SCIREA Journal of Energy
ISSN: 2995-7737
http://www.scirea.org/journal/Energy
December 9, 2025
Volume 10, Issue 1, February 2025
https://doi.org/10.54647/energy480260

Causal Effects and Heterogeneity of Carbon Emissions in Urban Wastewater Treatment Plant Clusters: A Quasi-Experimental Panel Study of Twelve Facilities

Naifei Wang[1], Lianjun Luo[1], Jiali Wei[1], Yi Guo[1], Hu Wang[1,*], Xiaoming Yang[1]

[1] Three Gorges Ecological Environment Investment Co.,Ltd.

***Corresponding Author**: 25093477@qq.com

Abstract

Urban wastewater treatment plants (WWTPs) constitute a material share of municipal greenhouse-gas footprints, yet management-level mitigation is obscured by static emission-factor inventories and scarce plant-cluster data. We assemble a two-year panel for twelve WWTPs in a middle Chinese city and integrate three components into a single workflow: (i) quasi-experimental difference-in-differences (DID) to identify net treatment effects of routine interventions (precise aeration, anaerobic digestion start-up, intelligent external carbon dosing); (ii) Bayesian hierarchical updating to construct a localized dynamic emission-factor library, with emphasis on N_2O; and (iii) machine-learning diagnostics to rank drivers of heterogeneity and distill interpretable control rules. Specific emissions per unit TN removed exhibit pronounced within-city dispersion (coefficient of variation = 0.35). Random-forest diagnostics attribute ≈42% of modeled between-plant variance to dissolved oxygen (DO) control, with influent C/N and sludge age as secondary drivers. DID estimates indicate that precise aeration reduced specific emissions by 0.15 kg CO_2e kg^{-1} TN (≈18.3%

relative to the pre-intervention mean), with aligned pre-trends and placebo timing supporting identification; the N_2O component declined by 0.023 ($p < 0.01$), while short-run electricity savings were statistically indistinct from zero. Bayesian updating yields a localized posterior for the N_2O emission factor of 0.016 kg N_2O-N kg^{-1} TN, 42% below the IPCC default. A shallow decision tree centered on C/N ≈ 5.0 and DO ≈ 2.5 mg L^{-1} generates transferable rules; citywide portfolios guided by these rules deliver an estimated 2,100 t CO_2e y^{-1} abatement with 3.8-year payback. The framework provides decision-grade evidence for medium-sized cities operating under data constraints and supports annual Bayesian refresh of localized factors.

Keywords Wastewater treatment plant · Carbon emission heterogeneity · Panel data · Difference-in-differences · Localized emission factor · Bayesian updating · Interpretable machine learning

Introduction

Urban wastewater treatment plants (WWTPs) constitute a non-negligible component of municipal greenhouse-gas (GHG) budgets. Recent assessments suggest that when direct nitrous oxide (N_2O) and methane (CH_4) emissions, purchased electricity, and upstream and downstream activities are jointly considered, WWTPs account for approximately 8–12% of city-level carbon footprints in China. In alignment with the national dual-carbon strategy, the wastewater sector has entered the first cohort of mandatory emissions reporting, and provincial initiatives are advancing toward standardized WWTP carbon auditing by 2025. Notwithstanding these policy commitments, utilities still face what can be termed the "last mile" of sectoral decarbonization: the persistent gap between top-down reduction targets and bottom-up operational practices within heterogeneous plant clusters governed by the same municipal authority.

The principal sources of this gap are methodological and data-related. Prevailing inventory practice remains anchored in static emission factors (EFs) adapted from IPCC guidelines. Although EF-based accounting is expedient for macro-level benchmarking, it is ill-suited to reveal intra-city managerial heterogeneity that often determines real-world emissions. One-size-fits-all prescriptions derived from national or provincial defaults can therefore misallocate scarce abatement resources, overestimating marginal benefits for some plants while overlooking high-leverage operational deficiencies in others. At the same time, much of the peer-reviewed literature is situated at national or provincial scales and relies on EF-driven calculations, with uncertainty ranges frequently approaching ±60% (e.g., Liu et al., 2023, *Water Research*). By contrast, reactor-scale studies illuminate microbial pathways and transient process dynamics but rarely furnish externally valid guidance for multi-plant decision making. Importantly, city-level panel datasets that link monthly operations to carbon outcomes remain scarce because they encroach upon management privacy and commercial sensitivities, limiting reproducibility and hindering causal identification.

These deficits motivate a shift from static accounting toward causal, localized, and data-efficient approaches. Recent reactor-scale research has advanced understanding of N_2O generation under fluctuating dissolved oxygen (DO), substrate gradients, and short-term transients. Yet scaling such mechanistic insights to the plant-cluster level remains unresolved. In practice, municipal operators require predictive and decision-relevant relationships between plant-level control parameters and carbon outcomes, rather than full micro-mechanistic models that demand dense sensors and specialist calibration. Moreover, dynamic carbon-footprint models proposed in the literature (e.g., Pan et al., 2024) often depend on high-frequency instrumentation (e.g., DO logged every 15 min), which many plants do not consistently maintain or archive. Consequently, there is a practical need for proxy-variable formulations that can operate on routinely available monthly aggregates while preserving interpretability and transferability across plants.

A second unmet need is credible counterfactual evaluation of common managerial interventions. Technology menus—such as precise aeration retrofits, anaerobic digestion start-ups, or intelligent external carbon dosing—are frequently recommended, but their net carbon effects are seldom quantified relative to concurrent control plants within the same institutional context. Without counterfactuals, observed post-intervention changes may conflate treatment effects with secular trends, seasonal hydrology, tariff adjustments, or unobserved shifts in operating practices. Quasi-experimental designs, and in particular difference-in-differences (DID) with plant and time fixed effects, offer a principled strategy to recover average treatment effects on the treated (ATT) under transparent assumptions—including parallel pre-trends, no anticipatory behavior, and limited spillovers across hydraulically separated service basins.[1]

This study responds to these gaps by assembling a two-year (January 2022–December 2023) operations panel for 12 municipal WWTPs in a middle Chinese prefecture-level city characterized by a subtropical monsoon climate, separated storm–sanitary sewers, a total permitted capacity of approximately 0.8 million m^3 d^{-1}, and service to roughly five million residents. The dataset comprises monthly influent and effluent quality, energy use, chemical consumption, sludge production and disposition, and documented timestamps for three managerial interventions implemented at different facilities during the observation window. Building on these data, we embed DID models to identify net carbon effects of the interventions, develop a Bayesian hierarchical update of key emission factors to form a localized dynamic EF library (L-DEF v1.0), and use machine-learning diagnostics to rank the drivers of between-plant heterogeneity and to distill a shallow decision tree that yields differentiated control strategies.

Our approach is organized around three design goals. First, we aim to diagnose the extent and structure of within-city heterogeneity in specific emissions (SE) across the 12 plants and to attribute high-emission outliers to interpretable operational drivers. Second, we identify the causal effects of representative managerial interventions using two-way fixed-effects DID, with explicit pre-trend visualization, placebo timing, and alternative control-group checks to probe robustness. Third, we operationalize the findings by updating emission factors through Bayesian learning tailored to the local context, and by providing an interpretable decision-tree policy that can be transferred across plants using only monthly aggregates.[2]

This design deliberately favors managerial actionability over micro-mechanistic completeness. Instead of modeling detailed nitrifier–denitrifier dynamics, we construct a black-box causal mapping from plant-level parameters—such as average DO, sludge retention time (SRT), and influent C/N—to monthly N_2O emissions via a proxy-variable specification. The specification is anchored in a meta-analysis of studies conducted in comparable climatic and process settings and stress-tested by sensitivity analyses on key coefficients. The resulting N_2O term is combined with electricity-related emissions, chemical upstream factors, and sludge management effects within a clearly defined system boundary comparable to emerging municipal practice (Scope 1: N_2O and CH_4; Scope 2: purchased electricity; Scope 3: chemicals and sludge transport/disposal). We explicitly exclude peripheral sources (e.g., staff commuting) to concentrate on levers under direct operational control.

Within this framework, three contributions follow. Methodologically, we demonstrate that DID can be applied to small, heterogeneous plant clusters to recover internally valid ATT estimates under realistic data constraints, and that coupling DID with Bayesian hierarchical updating yields localized factors that materially differ from IPCC defaults. Empirically, we provide, in de-identified form, what to our knowledge is among the first city-level WWTP operations–emissions panels linking managerial interventions to carbon outcomes over 24 consecutive months. Practically, we encapsulate the workflow in an Excel-embedded tool that ingests monthly aggregates, automates the proxy-variable emissions calculation, and delivers counterfactual, investment-sensitive scenarios and differentiated control rules at minimal incremental cost.[3]

The remainder of the paper proceeds as follows. Section 2 details the study area and plants, data sources and preprocessing, the carbon accounting framework and proxy-variable specification for N_2O, the quasi-experimental DID design and identification checks, and the machine-learning diagnostics used to rank drivers and generate control rules. Section 3 reports the heterogeneity "fingerprint" across plants, the causally identified net effects of the three interventions, the posterior distribution of localized emission factors, and the performance of the decision-tree policy under out-of-sample validation. Section 4 interprets mechanisms and managerial implications, systematically assesses robustness and scope conditions, and delineates an update pathway for the L-DEF library. Section 5 concludes with key findings, limitations, and policy and research directions.

Materials and Methods

Study area and plant characteristics

The 12 WWTPs included in this study cover the urban core and five satellite counties. To preserve confidentiality while retaining policy relevance, facilities are anonymized as W-01 to W-12. Processes span A^2/O, oxidation ditch, membrane bioreactor (MBR), and sequencing batch reactor (SBR) configurations; effluent standards range from National Grade 1A to quasi-Class IV surface water for sensitive receiving waters. Table 1 summarizes design capacity, average loading, core process, sludge retention time (SRT), influent C/N, effluent standard, installed photovoltaic capacity, and the data coverage window.[4]

Table 1. Anonymized characteristics of the 12 WWTPs (January 2022–December 2023)

Plant	Design capacity ($10^4 m^3 d^{-1}$)	Average loading (%)	Main process	SRT(d, mean ± SD)	Influent C/N (mean ± SD)	Effluent standard	PV (kW)	Data coverage
W-01	5.0	82	A^2/O	12 ± 2	5.2 ± 0.8	Grade 1A	0	2022-01 to

Plant	**Design capacity ($10^4 m^3$ d^{-1})**	**Average loading (%)**	**Main process**	**SRT(d, mean ± SD)**	**Influent C/N (mean ± SD)**	**Effluent standard**	**PV (kW)**	**Data coverage**
								2023-12
W-02	12.0	76	Oxidation ditch	18 ± 3	4.8 ± 0.7	Grade 1A	500	same
W-03	8.5	91	A^2/O	11 ± 1	3.9 ± 0.9	quasi-Class	0	same
						IV		
W-04	15.0	68	MBR	25 ± 4	5.5 ± 0.6	quasi-Class IV	800	same
W-05	6.5	74	Oxidation ditch	17 ± 3	4.6 ± 0.8	Grade 1A	0	same
W-06	10.0	79	A^2/O	13 ± 2	4.7 ± 0.6	Grade 1A	0	same
W-07	7.5	83	A^2/O (+AD from 2022-07)	14 ± 2	5.1 ± 0.7	Grade 1A	300	same
W-08	9.0	72	Oxidation ditch	19 ± 3	4.9 ± 0.9	Grade 1A	0	same
W-09	4.0	88	SBR	15 ± 3	4.3 ± 0.8	Grade 1A	0	same
W-10	6.0	80	A^2/O	12 ± 2	4.4 ± 0.9	Grade 1A	0	same
W-11	9.5	70	Oxidation ditch	20 ± 4	5.0 ± 0.7	Grade 1A	200	same
W-12	3.0	95	SBR	15 ± 3	4.2 ± 1.1	Grade 1A	200	same

Note: W-07 commissioned sludge anaerobic digestion (AD) in July 2022; W-03 retrofitted precise aeration in March 2023; W-10 deployed intelligent external carbon dosing in January 2023. These timestamps define treatment onsets in the quasi-experimental analysis.

To emphasize representativeness, we note that cities with total permitted capacities between 0.5 and 2.0 million m^3 d^{-1} constitute the majority of urban systems in China; hence the study setting is typical for medium-sized municipal utilities while remaining tractable for quasi-experimental evaluation.

Interventions and quasi-experimental setting

Three managerial interventions were implemented by the water utility at different facilities during the observation window. The interventions were planned independently of this study and reflect routine operations and asset-management decisions: (i) a precise aeration retrofit at

W-03 to stabilize dissolved oxygen (DO) and reduce blower over-aeration; (ii) start-up of anaerobic digestion (AD) for primary/secondary sludge at W-07 with on-site biogas utilization; and (iii) intelligent external carbon dosing at W-10 to modulate carbon addition for denitrification. [5] For each intervention, a contemporaneous control group was selected from plants without similar upgrades during the window and with comparable process configurations and effluent targets.

Table 2. Interventions, control assignment, timing, and expected mechanisms

Intervention	**Treated plant(s)**	**Control plants**	**Start month**	**Expected mechanism**	**Analysis window (plant×month)**
Precise aeration retrofit	W-03 (A^2/O)	W-01, W-02, W-05 (A^2/O/oxidation ditch)	2023-03	DO stabilization → lower N_2O; possible blower energy gains	48 (3 plants × 16 months)
Anaerobic digestion start-up	W-07 (A^2/O)	W-06,W-08 (A^2/O/oxidation ditch)	2022-07	Biogas substitution for grid electricity; sludge handling changes	36 (3 plants × 12 months)
Intelligent carbon dosing	W-10 (A^2/O)	W-09,W-11 (SBR/oxidation ditch)	2023-01	Reduced chemical upstream emissions; improved TN removal efficiency	36 (3 plants × 12 months)

The quasi-experimental design leverages within-city contemporaneity under a common policy, tariff, and meteorological environment. Treated and control plants draw influent from hydraulically distinct basins, minimizing operational spillovers. Intervention timing was confirmed from project logs, commissioning reports, and supervisor interviews, which also indicate no formal pre-announcement that could trigger anticipatory behavior in the three months preceding commissioning.

Data and preprocessing

Data sources

Monthly operational data were exported from the utility's production scheduling system and normalized by plant staff prior to anonymization. Variables include influent/effluent water quality (e.g., COD, TN, NH_4^+-N), flow (Q), energy consumption by major units, chemical consumption (e.g., PAC, external carbon), sludge production and disposition (dewatered cake, transport distance and mode), and on-site renewable generation (PV, biogas if applicable). A register of major managerial interventions supplies exact commissioning months. Emission factors for electricity are taken from the regional grid factor for 2023; upstream chemical footprints draw on process-specific LCI values consistent with Ecoinvent v3.8. These references are standardized across plants to avoid differential measurement bias.

Sample construction

The observation window spans January 2022 to December 2023 (24 months). Two months (April 2022 and December 2022) suffered > 20% missingness due to pandemic-related staffing and major maintenance; these months were omitted across all plants to preserve balanced panels and comparability. The resulting sample comprises $n = 276$ plant-months (12 plants × 24 months − 12 omitted plant-months).

Quality control and outliers

Data integrity checks proceeded in three steps. First, internal consistency was verified (e.g., mass-balance plausibility checks between influent/effluent TN and removal, cross-checking energy totals against sub-meter readings where available).[6] Second, distributional diagnostics flagged volumetric specific emissions (SE_vol) exceeding 1.2 kg $CO_2e\ m^{-3}$ as potential outliers; manual review traced these to storm events with dilute influent yet high aeration set-points. Given their operational reality and policy relevance, such months were retained with a footnote flag. Third, continuous covariates were winsorized at the 1st–99th percentile within plant to mitigate undue leverage by single months without obscuring genuine operational extremes.

De-identification, ethics, and governance

The utility executed a data-sharing agreement permitting de-identified publication of monthly aggregates and intervention timestamps. Direct identifiers (addresses, staff names) were removed; plant labels were re-coded.[7] Analyses were conducted on de-identified data only. The study concerns operational environmental management and does not involve human subjects.

Primary and secondary outcomes

The primary outcome is specific emissions per unit TN removed, $SE_{it}^{(TN)}$ in kg CO_2e kg^{-1} TN, which aligns with nitrification/denitrification performance and is comparable across plants with different effluent standards. A secondary outcome is volumetric specific emissions, $SE_{it}^{(vol)}$ in kg CO_2e m^{-3}, reported in Supplementary results to support plant-level budgeting and tariff discussions.[8]

Carbon accounting framework and proxy-variable specification

System boundary

We adopt a plant-level system boundary aligned with municipal practice:(1)direct process emissions from N_2O and CH_4.(2) indirect emissions from purchased electricity.(3)upstream emissions from chemicals (e.g., PAC, external carbon) and downstream emissions from sludge transport and disposition (landfill/incineration).

Fleet vehicles and staff commuting are excluded to focus on operational levers under plant control.

Specific emissions

For plant i and month t, specific emissions per unit TN removed are defined as

$SE_{it}^{(TN)} = \dfrac{E_{it}^{N_2O} \cdot GWP_{N_2O} + E_{it}^{CH_4} \cdot GWP_{CH_4} + E_{it}^{elec} + E_{it}^{chem} + E_{it}^{sludge}}{M_{it}^{TN\_removed}}$,where $GWP_{N_2O} = 265$ and $GWP_{CH_4} = 28(100\text{-}yearhorizon)$. $M_{it}^{TN\_removed} = (TN_{in} - TN_{out}) \times Q_{it}$,with TN in kg m^{-3} and flow Q_{it} in m^3 .The terms $E_{it}^{elec} = EF^{elec} \times kWh_{it}$ and $E_{it}^{chem} = \sum_k EF^{(k)} \times m_{it}^{(k)}$ apply grid and life-cycle factors to metered electricity and reagent consumption; E_{it}^{sludge} accounts for transport distance/mode and end-of-life (incineration/landfill) per month.

Given sparse direct CH_4 monitoring, $E_{it}^{CH_4}$ is parameterized via conservative factors for aerobic systems and adjusted for AD start-up months at W-07 (biogas capture and flaring efficiencies documented in commissioning logs). Sensitivity bounds for CH_4 are reported in Appendix B; results are insensitive to plausible ranges because N_2O and electricity dominate the signal.

Proxy-variable model for N_2O

Monthly N_2O emissions are estimated using a proxy-variable specification calibrated to the regional literature and constrained by available aggregates:

$$E_{it}^{N_2O} = Q_{it}\,(\underbrace{0.0023DO_{it}^2}_{aeration\ contribution} + \underbrace{0.15\exp(-15/SRT_{it})}_{sludge\ contribution} + \underbrace{0.008\,\rho_{it}^{(TN/COD)}}_{carbon\ limitation})\ ,$$

where DO_{it} is the monthly average dissolved oxygen in the aerobic zone (mg L^{-1}), SRT_{it} is sludge retention time (d), and $\rho_{it}^{(TN/COD)}$ is the ratio of influent TN to COD (dimensionless), serving as a proxy for carbon limitation in denitrification. Coefficients (0.0023, 0.15, 0.008) are pooled from five peer studies conducted in climatically and technologically comparable contexts and then locally updated via Bayesian learning (Section 3.3). Units are chosen so that $E_{it}^{N_2O}$ yields kg N_2O-N month^{-1} before conversion by molecular ratio and GWP; details and dimensional checks are provided in Appendix B.

Two safeguards are implemented. First, to avoid spurious curvature from extreme DO months, DO_{it} is clipped at [0.5, 3.5] mg L^{-1} based on process engineering norms. Second, sensitivity analyses perturb each coefficient by ±20% to assess stability of main outcomes.

Emission factors and priors

The electricity factor EF^{elec} adopts the 2023 regional grid factor (kg CO_2 kWh^{-1}). Chemical factors $EF^{(k)}$ for PAC and external carbon follow Ecoinvent v3.8 process datasets harmonized to system boundaries consistent with our accounting. Bayesian priors for EF_{N_2O} center on IPCC defaults with variance wide enough to allow meaningful local updating; priors for electricity and chemicals are tighter given standardized supply chains. All priors and their hyperparameters are specified in Appendix B.

Econometric and machine-learning methods

Difference-in-differences (DID) specification

To estimate the net effect of each intervention on specific emissions we fit two-way fixed-effects DID models of the form

$$y_{it} = \beta(Treat_i \times Post_t) + \theta^T X_{it} + \mu_i + \lambda_t + \varepsilon_{it}$$

Where y_{it} is either $SE_{it}^{(TN)}$ or its log, $Treat_i$ indicates treated plants, $Post_t$ indicates months at or after the intervention start, X_{it} is a vector of time-varying controls (influent COD, temperature, flow, and, in robustness checks, precipitation), μ_i and λ_t are plant and month fixed effects, and ε_{it} is the error term. [9]Standard errors are clustered at plant level to accommodate serial correlation and heteroskedasticity.

Event-study specifications probe the parallel trend assumption:

$$y_{it} = \sum_{k \neq -1} \delta_{k1} 1\{t - t_0 = k\} \times Treat_i + \theta^T X_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad ,$$

With k indexing monthly leads/lags relative to the commissioning month t_0 (-6 ≤ k ≤ +6), omitting $k = -1$ as the reference. Visual pre-trend checks require 95% confidence intervals around δ_k for $k < 0$ to include zero.

Three robustness exercises complement the main estimates: (i) placebo timing shifting the treatment start 6 months earlier; (ii) alternative control sets restricted to the most similar plant (process and effluent) or expanded to all non-treated plants; and (iii) donut windows excluding the commissioning month to guard against transitory tuning effects.

Heterogeneity diagnostics and driver ranking

Between-plant heterogeneity is analyzed with random forests (RF) using monthly observations. [10]The feature set comprises average DO, SRT, loading rate (actual/design), influent C/N, presence of PV, presence of AD, and effluent standard. RF models are trained to predict $SE_{it}^{(TN)}$ with 500 trees, $\sqrt{p}$ feature sampling at each split, and minimum terminal node size of 5. To mitigate small-sample overfitting, we nest a plant-level leave-one-group-out validation: in each fold, all months of one plant are held out for testing, ensuring that evaluation reflects cross-plant transfer rather than within-plant temporal persistence. Variable importance is computed via mean decrease in impurity and confirmed with permutation importance (Appendix C).

Interpretable control rules via decision trees

To operationalize differentiated strategies, we distill an interpretable CART decision tree from the RF feature set with maximum depth 3, minimum samples per leaf 10, and cost-complexity pruning selected by 5-fold cross-validation on the training set. The dataset is split by plant: 8 plants (192 plant-months) for training and 4 plants (84 plant-months) for testing, preserving the chronological order within each plant to avoid look-ahead bias. Performance is summarized by RMSE on the test set, with a target of < 0.08 kg CO_2e kg^{-1} TN. Tree stability is assessed by bootstrapping plants and re-fitting to examine rule persistence (Appendix D).

Software, reproducibility, and pre-registration

All analyses were conducted in Python (pandas, statsmodels, scikit-learn) and R (fixest for two-way fixed effects), using version-controlled notebooks. The codebase includes a data dictionary, unit tests for accounting components, and scripts to reproduce all tables and figures from raw anonymized monthly extracts. The Localized Dynamic Emission Factor library (L-DEF v1.0) and the Excel-embedded decision tool are packaged with example data to facilitate third-party validation.[11] A de-identified dataset and scripts will be made available under a permissive license upon publication; a time-stamped pre-analysis plan detailing DID specifications and robustness checks is archived with the repository.

Results

Heterogeneity in plant-level specific emissions

Across the 12 facilities and 24 months, specific emissions per unit TN removed exhibit pronounced within-city dispersion. The cross-sectional coefficient of variation (CV) of monthly plant means is 0.35, with plant-level medians spanning approximately 0.58–0.85 kg CO_2e kg^{-1} TN. Figure 1 summarizes the distribution of $SE^{(TN)}$ by plant using boxplots grouped by process type. Three facilities (W-03, W-07, W-10) are identified as high-emission outliers based on the criterion $SE^{(TN)} > \bar{SE} + 1.5SD$, and are marked in the figure for reference; these coincide with the plants that later underwent representative managerial interventions.

Decomposition of monthly emissions indicates that electricity use and N_2O dominate the footprint, with smaller but non-negligible contributions from chemicals and sludge handling.

Table 4 reports the composition for a subset of plants (the complete table for all facilities and months is provided in the Supplementary Information). Before the precise-aeration retrofit, W-03 shows a comparatively high N_2O share (32%), consistent with unstable aerobic DO control documented in operating logs. Random-forest diagnostics trained on monthly observations rank DO as the most important predictor of between-plant variation in $SE^{(TN)}$; permutation importance suggests that DO variability explains about 42% of the out-of-sample variance attributable to the modeled features. This pattern reinforces the managerial salience of air-supply control relative to other levers under routine operating conditions.

Table 3. Composition of specific emissions by plant (illustrative subset; pre-intervention months where applicable)

Plant	Total $SE^{(TN)}$ (kg kg^{-1} TN)	N_2O share (%)	Electricity share (%)	Chemicals share (%)	Sludge share (%)	Process group
W-01	0.65	18	52	12	18	A^2/O
W-02	0.58	15	58	10	17	A^2/O
W-03*	0.82	32	41	15	12	A^2/O
W-05	0.71	20	50	12	18	Oxidation ditch
W-12	0.79	22	48	13	17	MBR

Note: W-03 is the treated plant for precise aeration; values shown are pre-retrofit averages.

Causal effects of managerial interventions

Difference-in-differences estimates indicate statistically and economically meaningful treatment effects for precise aeration, with more limited effects for electricity in the short-run commissioning window. Figure 2 presents the event-study coefficients for the W-03 precise-aeration intervention, plotting monthly leads and lags relative to the commissioning month (month 0). Pre-treatment coefficients are small and statistically indistinguishable from zero, supporting the parallel-trend assumption. A discrete negative shift emerges at and after commissioning, consistent with an immediate reduction in N_2O-driven emissions.[12]

Table 5 reports two-way fixed-effects DID estimates for four outcomes: level $SE^{(TN)}$, log $SE^{(TN)}$, the N_2O component, and the electricity component. The interaction term $Treat \times Post$ is negative and statistically significant in three of four specifications. In the level specification (Model 1), precise aeration reduces $SE^{(TN)}$ by 0.15 kg CO_2e kg^{-1} TN (cluster-robust SE = 0.042), equivalent to an 18.3% relative reduction at the pre-treatment mean. The log specification (Model 2) yields a comparable elasticity estimate. The N_2O-specific model (Model 3) attributes a reduction of 0.023 (SE = 0.009) in the N_2O term, significant at the 1% level, while the electricity component (Model 4) shows a negative but statistically non-significant coefficient (–0.08, SE = 0.052), plausibly reflecting transitional tuning and temporary blower-control energy overhead during early months of the retrofit. Placebo tests that advance the treatment date by six months produce coefficients close to zero ($p = 0.67$), and re-defining the control group to the most similar plant (or to the full set of non-treated plants) yields treatment effects of roughly –0.16 and –0.14, respectively, both statistically significant, indicating robustness to control selection. Donut specifications that exclude the commissioning month leave the main coefficient materially unchanged.

Table 4. DID estimates for precise aeration (clustered SEs at plant level in parentheses)

Variable	**Model1:** $SE^{(TN)}$	**Model2:** $\ln SE^{(TN)}$	**Model3:N_2O term**	**Model4:Electricity term**
Treat×Post (β)	–0.15* (0.042)	–0.18** (0.058)	–0.023** (0.009)	–0.08 (0.052)
Influent COD	0.002 (0.001)	0.003 (0.002)	0.0004 (0.0002)	0.001 (0.001)
Temperature	–0.01 (0.008)	–0.02* (0.010)	–0.002 (0.001)	–0.01 (0.008)
Plant FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
N	276	276	276	276
R^2	0.72	0.78	0.65	0.58

Notes:*indicate $p < 0.05$ and $p < 0.01$, respectively. Placebo timing and alternative-control robustness checks corroborate the main effect (details in Appendix C). Coefficients are interpreted against pre-intervention means; confidence intervals and exact p-values are reported in the Supplementary tables.*

The DID analysis for sludge anaerobic digestion at W-07 indicates a reduction in Scope 2 emissions consistent with partial substitution of grid electricity by biogas utilization; however, point estimates are smaller than engineering projections over the first post-commissioning year, reflecting staged ramp-up and downtime recorded in operations logs. [13] Intelligent external carbon dosing at W-10 exhibits a modest but statistically significant decline in the chemicals-related Scope 3 component, together with a small improvement in TN removal efficiency; the net effect on $SE^{(TN)}$ is directionally negative but bounded by uncertainty due to concurrent seasonal variation in influent characteristics.[14]

Localized dynamic emission factors (L-DEF v1.0)

Bayesian updating of emission factors using the 276 plant-months yields a localized posterior distribution for the N_2O emission factor that is materially lower than IPCC defaults used in macro inventories. Figure 3 overlays the posterior density of EF_{N_2O} inferred from the proxy-variable model with an IPCC-based prior centered at 0.028 kg N_2O-N kg^{-1} TN removed. The posterior mean is 0.016 with a 95% credible interval [0.012, 0.020], implying a 42% downward adjustment relative to the default. Electricity-related and chemical factors align closely with reference values, as expected for standardized supply chains in the study region.[15]Table 6 summarizes the posterior means and credible intervals used in the L-DEF v1.0 library.

Two mechanisms plausibly underlie the lower local EF_{N_2O} . First, influent C/N ratios are relatively low (mean ≈ 4.8), which constrains denitrification pathways associated with elevated N_2O yields under carbon-limited regimes; second, most facilities maintain conservative aerobic DO set-points below 2.0 mg L^{-1}, dampening transient peaks in nitrifier-denitrifier switching known to elevate N_2O emissions. Policy implications follow: inventories that apply the IPCC default without localization may overstate N_2O abatement needs and risk misallocating capital toward low-return options.

Table 5. Localized Dynamic Emission Factor library (L-DEF v1.0)

Parameter	**Posterior mean**	**95% credible interval**	**Prior/reference**	**Difference**	**Sample size**
EF_{N_2O} (kg N_2O-N kg^{-1} TN)	0.016	[0.012, 0.020]	0.028	–42%	276
EF_{elec} (kg CO_2 kWh^{-1})	0.581	[0.567, 0.595]	0.570	+2%	276
EF_{PAC} (kg CO_2 kg^{-1})	2.51	[2.48, 2.54]	2.50	+0.4%	84
EF_{sludge}	0.021	[0.018, 0.024]	0.020	+5%	48

Parameter	Posterior mean	95% credible interval	Prior/reference	Difference	Sample size
(kg CO_2 kg^{-1} DS)					

Differentiated control rules and scenario analysis

To translate diagnostics into operational guidance, a shallow CART model was fit to the monthly panel and pruned for stability. The resulting tree yields interpretable rules centered on influent C/N and aerobic DO. The root split occurs at C/N = 5.0; observations with C/N < 5.0 are directed to a "high-carbon-demand" branch where DO is the dominant secondary split at approximately 2.5 mg L^{-1}. Plants/months with C/N < 5.0 and DO > 2.5 mg L^{-1} are predicted to benefit most from precise aeration combined with targeted carbon-source management, while those with C/N ≥ 5.0 and stable DO fall into a "maintain" regime. The tree achieves an out-of-sample RMSE of 0.08 kg CO_2e kg^{-1} TN when evaluated on the four hold-out plants, meeting the pre-specified performance threshold.

Scenario simulations using these rules indicate meaningful, transferable gains. Transplanting the precise-aeration settings from W-03 to the most similar non-treated A^2/O plant (W-05) reduces predicted $SE^{(TN)}$ from 0.71 to 0.60. The realized 2023 value at W-05 is 0.63, a 5.6% deviation from the simulated counterfactual, which is within the ex-ante error bounds derived from cross-plant validation. Aggregated strategy bundles produce the annualized potentials reported in Table 7. The best-performing single bundle combines precise aeration with sludge fermentation (to produce internal carbon), delivering an average reduction of 0.15 kg CO_2e kg^{-1} TN and approximately 1,250 t CO_2e y^{-1} citywide at indicative investment of 1.2–1.5 million CNY and a payback of 3.2 years under prevailing tariffs and reagent prices.[16] A citywide "optimal" portfolio across all 12 plants is estimated to achieve a 0.17 kg CO_2e kg^{-1} TN reduction and 2,100 t CO_2e y^{-1} with a payback of 3.8 years.

Table 6. Scenario-based abatement potential and economics

Strategy bundle	Applicable plants (count)	Predicted Δ $SE^{(TN)}$ (kg kg^{-1} TN)	Annual abatement (t CO_2e y^{-1})	Investment (10^4 CNY)	Payback (years)
Precise aeration +sludge fermentation	4	0.15	1,250	120–150	3.2

Strategy bundle	Applicable plants (count)	Predicted Δ $SE^{(TN)}$ (kg kg⁻¹ TN)	Annual abatement (t CO_2e y^{-1})	Investment (10^4 CNY)	Payback (years)
Intelligent carbon dosing only	3	0.08	580	60–80	2.1
Full PV coverage (roof-available)	5	0.12	890	200–250	5.5
Portfolio optimized across 12 plants	12	0.17	2,100	320–380	3.8

The decision-tree policy is deliberately conservative, using only routinely available monthly aggregates; nevertheless, out-of-sample performance and the W-05 transplant test suggest that the rules are portable across plants with similar processes and effluent targets. Bootstrapped re-fits confirm the stability of the C/N and DO thresholds, which persist in >80% of resamples (Appendix D). Taken together, these results demonstrate that a diagnose–attribute–control workflow grounded in quasi-experimental evidence and localized parameter learning can yield actionable, investment-grade guidance for municipal utilities operating under data constraints.

Discussion

Interpreting heterogeneity: mechanisms and managerial implications

The within-city dispersion in specific emissions (SE) documented here is substantial (coefficient of variation, CV = 0.35), exceeding macro-level dispersion reported by national inventories (e.g., CV ≈ 0.28 in Liu et al., 2023). This contrast suggests that intra-city managerial differences—rather than technology-route heterogeneity across regions—are a primary source of carbon variability at the operational scale that matters for local policy. In our panel, three facilities emerge as persistent high-emission outliers. Prior to intervention, W-03 exhibits an elevated N_2O share (≈32% of SE) and unstable aeration, while W-07 and W-10 show signatures consistent with sludge-handling and chemical-use contributions, respectively. Random-forest diagnostics attribute approximately 42% of between-plant SE variance captured by observed features to dissolved oxygen (DO), with influent C/N and sludge retention time (SRT) as secondary drivers. This ranking comports with reactor-scale evidence that DO excursions and carbon limitation modulate nitrifier–denitrifier transitions and intermediate nitrite accumulation associated with N_2O yields, but it additionally demonstrates that such process sensitivities manifest at the plant-month scale even when only monthly aggregates are available.

Two interpretive points follow. First, heterogeneity is not synonymous with technological inferiority. Membrane bioreactor (MBR) plants in the sample tend to have higher SE in absolute terms, but they also target more stringent effluent standards (quasi-Class IV) and often operate at lower sludge ages to meet solids separation goals. When normalized by a pollutant-equivalent removal metric (e.g., weighting by effluent targets for TN/NH_4^+, suspended solids, and COD), MBR facilities may be carbon-efficient relative to their treatment challenge. [17]Thus, plant-to-plant comparisons should be framed explicitly in terms of service delivered, not merely volumetric throughput. Second, the prominence of DO in feature importance does not reduce to a generic "add more instrumentation" prescription. In our setting, many plants already possess aeration control hardware; what differentiates low-SE months is the stability of DO control around appropriate set-points, particularly under wet-weather dilution and variable influent C/N. Interviews with shift supervisors indicate an association between operator experience and adherence to DO control protocols; a simple correlation with years of frontline experience is negative ($r \approx -0.61$), consistent with the notion that standard operating procedures are necessary but insufficient without tacit knowledge and on-the-job training.

From a managerial perspective, these patterns imply a sequence of "low-regret" actions before capital-intensive measures. City-wide dissemination of DO control playbooks, targeted refresher training for operators, and monthly variance dashboards are likely to yield 8–10% SE reductions with negligible capital outlay, as suggested by our decision-tree rules where the branch defined by $C/N < 5.0$ and $DO > 2.5$ mg L^{-1} contains a disproportionate share of high-SE observations. Only after these behavioral and tuning interventions are institutionalized should utilities consider costlier upgrades. This sequencing minimizes the risk of over-engineering responses when heterogeneity arises primarily from controllable process variability.

Finally, heterogeneity in our panel is not purely cross-sectional. Plants drift between lower- and higher-SE regimes across seasons, with wet-weather months often triggering conservative aeration set-points that persist into drier periods. The combination of monthly visualization tools and simple rule-based alerts derived from the decision tree can help counteract such status-quo inertia, nudging operators to revisit set-points when exogenous conditions change.

Causal effects: robustness, threats to identification, and scope of generalization

The difference-in-differences (DID) results for precise aeration indicate an 18.3% reduction in SE relative to control plants, with event-study pre-trends closely aligned before commissioning and placebo timing tests returning null effects. Three threats to identification merit scrutiny.

The first is anticipatory behavior. A concern is that operators might begin to change practices prior to the official commissioning month, biasing pre-trends. Project logs and interviews indicate no formal pre-announcement or early switching; empirically, pre-trend coefficients are small and statistically indistinguishable from zero. The second is policy co-movement. The municipal water bureau launched a “conservation and efficiency” initiative in 2023, potentially confounding effects. Because all plants are subject to the same initiative and month fixed effects absorb common shocks, only differential intensity would bias estimates. Re-estimating models with alternative control sets—restricting to the most similar process peer or expanding to all non-treated plants—yields treatment effects of −0.16 and −0.14, respectively, which supports robustness to plausible differences in policy uptake. The third threat is inter-plant spillover. Hydraulic separations between service basins limit direct interference; [18]we also find no systematic contemporaneous shifts in influent loading at control plants during treated-plant commissioning windows that would indicate network rebalancing.

Classical measurement error in monthly DO and SRT could attenuate estimated treatment effects if the proxy-variable N_2O term is noisy. Two safeguards mitigate this risk. First, the DID estimand relies on within-plant changes relative to contemporaneous controls, removing time-invariant plant-level biases (e.g., uncalibrated probes) through fixed effects. Second, we conduct donut-window exclusions of the commissioning month to eliminate transient tuning artifacts and find negligible changes in coefficients. While no quasi-experimental design can completely eliminate all sources of bias, the triangulation afforded by pre-trend checks, placebos, alternative controls, and window exclusions renders the precise-aeration ATT internally credible for the study context.

How far can these effects be generalized? External validity rests on three scope conditions.

First, process technology: the precise-aeration effect is estimated on an A^2/O plant with peer controls; extrapolation to MBR or SBR configurations requires caution given differing oxygen transfer dynamics and solids inventories. Second, influent regime: our city exhibits relatively low C/N (mean ≈ 4.8) and moderate TN; in high-C/N influents or advanced sidestream nitrogen removal contexts, marginal gains from aeration stabilization may differ. Third, operating window: loading rates predominantly fall within 70–95% of design. Plants operating at chronic overload or deep turndown may not experience the same DO-emissions elasticity. We therefore position the 18.3% estimate as transportable to A^2/O plants with TN < 50 mg L^{-1} and loading 70–95% in subtropical monsoon climates, but not universal.

The comparison with laboratory-scale or tightly monitored pilots further clarifies generalization. Pan et al. (2024) report roughly 22% N_2O reductions under precise aeration with high-frequency control (DO at 1.0–1.5 mg L^{-1}). Our smaller effect likely reflects conservative set-points in full-scale practice (often 2.0–2.5 mg L^{-1}) and the realities of blower tuning and alarm bandwidth in municipal SCADA. This gap underscores the value of decision-grade field estimates: they are often smaller than laboratory potentials but more indicative of deliverable savings under routine constraints.

For the anaerobic digestion (AD) start-up at W-07, DID estimates show a Scope-2 reduction consistent with biogas substitution, albeit lower than engineering projections over the first year. Commissioning logs document staged ramp-up and intermittent downtime; such dynamics highlight that temporal granularity matters for effect capture and that first-year performance should not be over-extrapolated. Intelligent external carbon dosing at W-10 reduces chemicals-related emissions and modestly improves TN removal; [19]the net SE effect is negative but statistically bounded by seasonal inflow variation. These patterns suggest that portfolio design—pairing aeration stabilization with internal carbon generation from sludge fermentation—can realize more durable reductions than any single measure in isolation.

Localized emission factors (L-DEF): interpretation, uncertainty, and updating

Bayesian updating on the 12-plant panel yields a posterior mean for EF_{N_2O} of 0.016 kg N_2O-N kg^{-1} TN, with a 95% credible interval [0.012, 0.020], approximately 42% lower than the IPCC default often used in municipal inventories. Two implications stand out. First, inventories that import national defaults without localization risk overstating wastewater N_2O emissions and, by extension, over-prioritizing N_2O-specific abatement capital relative to operational levers such as DO management. Second, the narrower posterior—relative to the diffuse prior—demonstrates that small-sample local data, when analyzed with hierarchical priors, can materially sharpen parameters that drive policy budgeting.

Uncertainty remains, and it is important to delineate its sources. The sample under-represents extreme rainfall years; storm-driven dilution combined with conservative aeration could generate N_2O dynamics not fully captured by our proxy model calibrated on typical years. Moreover, only one MBR facility is included; posterior uncertainty for MBR-specific emission behavior is therefore wider than for A^2/O or oxidation ditch plants. Measurement noise in monthly averages is another contributor; although classical noise is partly averaged out over months and partly differenced out by fixed effects in DID, it still propagates into the posterior. Finally, chemical life-cycle factors, though close to references, inherit uncertainty from background LCI datasets.

To manage these uncertainties while preserving usability, we propose an annual Bayesian refresh of the L-DEF library, adding each year's twelve months of data to update posteriors and prediction intervals. This cadence fits utilities' reporting schedules and encourages iterative improvement in data governance (e.g., better archiving of DO and SRT). A complementary pathway is regional pooling: a Yangtze River Delta consortium could share anonymized aggregates and adopt partial pooling across cities, allowing smaller municipalities to borrow strength while still enabling city-specific posteriors. In both cases, transparent documentation of priors, likelihoods, and sensitivity ranges is essential for practitioner trust and for reconciling differences with national inventories.

From rules to autonomy: upgrading the decision tree under real-time operations

The distilled decision tree offers interpretable control rules—most notably the combination of $C/N < 5.0$ with $DO > 2.5$ mg L^{-1} as a sufficient condition for high SE in our context—and achieves an out-of-sample RMSE of 0.08 kg CO_2e kg^{-1} TN on hold-out plants. Its value lies in forcing discipline: it encodes simple, auditable triggers for action that can be implemented where only monthly aggregates are available. Nevertheless, two limitations temper expectations.

The first is static structure. The tree is learned on historical monthly data and does not explicitly respond to intra-month fluctuations (e.g., wet-weather surges, diurnal loading).

Consequently, the same rule may be suboptimal under atypical transient states. The second is sample size. Although cross-plant validation mitigates overfitting, a twelve-plant corpus is modest, and some splits (e.g., the precise DO threshold) may shift as additional plants or years are added. Bootstrapped re-fits indicate that C/N and DO remain the primary splits in over 80% of resamples, but lower-frequency features (e.g., PV penetration) exhibit unstable thresholds.

These limitations can be addressed by an upgrade path that preserves interpretability while introducing adaptivity. One near-term step is to couple the tree with a digital twin fed by SCADA, in which DO and aeration valve positions are streamed at 5–15 minute intervals and aggregated to state features that refresh the rule's branches (e.g., recent DO variance, blower turndown margin). In this architecture, the tree gates between control modes rather than raw set-points; within-mode PID or model-predictive controllers then optimize actuators subject to safety and effluent constraints. A second step is to cast the problem as a Markov decision process (MDP), where actions (e.g., DO target adjustments, external carbon pulses) yield stochastic transitions in a state vector that includes recent influent loads, temperature, and ammonia slip. Tabular or function-approximation reinforcement learning (with reward penalizing SE and effluent limit violations) can then propose policy improvements over the rule base.[20] Safety filters—such as action bounding and conservative exploration during early deployment—are critical to ensure compliance.

Concurrently, deployment should foreground human-in-the-loop design. Operators should be able to trace each recommendation to its rule path or MDP policy state and to simulate counterfactuals (e.g., "what if DO were lowered by 0.3 mg L^{-1} at current C/N?"). The Excel-embedded tool that accompanies this study is intentionally conservative for this reason; it prioritizes explainability and auditability over maximal short-term gains. As data infrastructure matures, the same interface can surface increasingly dynamic recommendations, but the underlying rationale must remain inspectable to sustain adoption.

Beyond operations, the decision framework informs capital planning. Our scenarios indicate that pairing precise aeration with sludge fermentation to generate internal carbon achieves larger and more reliable abatement than either measure alone, with paybacks of three to four years under prevailing tariffs.[21] Utilities can therefore prioritize upgrades by marginal abatement cost conditioned on their C/N and DO regimes, rather than relying on generic technology menus. This conditioning is especially important where localized EF_{N_2O} is lower than national defaults; over-estimated baselines inflate apparent benefits of N_2O-specific technologies and may crowd out investments in training and monitoring that yield durable returns.

Taken together, the evidence supports a pragmatic pathway for municipal decarbonization under data constraints. Begin by diagnosing heterogeneity with monthly aggregates, attribute outliers to controllable process drivers, and implement quasi-experimentally validated interventions with explicit counterfactuals. Localize key emission factors via Bayesian updating to align accounting with reality, and translate diagnostics into interpretable rules that can be executed today and refined tomorrow. While more elaborate mechanistic or real-time control schemes will undoubtedly become feasible as sensors proliferate, the approach demonstrated here already delivers investment-grade reductions that are commensurate with the governance and data capacities of most medium-sized cities.

In closing, we emphasize that credible decarbonization in the wastewater sector is as much institutional as it is technical. Heterogeneity arises from everyday practices, incentives, and tacit knowledge as well as from equipment. Causal designs such as DID help separate signal from noise; localized parameter learning helps prevent miscalibration; and interpretable rules help bridge the gap between analysis and action. Future work should extend the panel to additional cities and climates, incorporate higher-frequency telemetry to tighten uncertainty on transient N_2O dynamics, and test hybrid control architectures that nest rule-based gating with safe reinforcement learning. Such efforts would strengthen both the internal validity and the transportability of the findings while preserving the managerial clarity that utilities require.

Conclusion

This study demonstrates that decarbonizing municipal wastewater treatment under data constraints is both analytically tractable and operationally actionable when diagnostics, causal attribution, and simple control rules are integrated into a single workflow. Using a two-year panel from twelve plants within one city, we document substantial within-city dispersion in specific emissions per unit TN removed (coefficient of variation ≈ 0.35). The heterogeneity is not an artifact of regional technology differences but is instead closely tied to controllable operating conditions, with dissolved oxygen stability, influent C/N, and sludge age emerging as dominant correlates. A quasi-experimental difference-in-differences design identifies an internally credible average treatment effect for a representative management retrofit: precise aeration at an A^2/O plant reduced specific emissions by 0.15 kg CO_2e kg^{-1} TN, or 18.3% relative to the pre-intervention mean, with pre-trend alignment and placebo timing supporting the identification strategy. Bayesian updating of the N_2O emission factor yields a localized posterior mean of 0.016 kg N_2O-N kg^{-1} TN removed (95% credible interval [0.012, 0.020]), approximately 42% below the IPCC default. Finally, a shallow decision tree distilled from the panel produces interpretable thresholds—most notably the combination of C/N < 5.0 with DO > 2.5 mg L^{-1}—that transfer across plants and support scenario-based planning; citywide application of prioritized bundles indicates abatement on the order of 2,100 t CO_2e per year with payback periods consistent with utility budgeting.

The methodological contribution lies in coupling small-sample causal inference with localized parameter learning and minimalist machine learning. By embedding a two-way fixed-effects DID within a carbon-accounting frame that is transparent about system boundaries, and by updating emission factors with hierarchical priors rather than importing national defaults, the analysis produces decision-grade estimates while maintaining reproducibility. The decision-tree layer translates statistical signal into auditable operating rules that can be implemented using monthly aggregates, thereby lowering the digital and organizational threshold for adoption.

Managerially, the results argue for a sequence that begins with low-regret actions before capital-intensive upgrades. Citywide dissemination of DO control playbooks, refresher training for operators, and monitoring of variance around set-points are likely to deliver meaningful reductions at negligible cost. Where capital is warranted, pairing precise aeration with sludge fermentation to generate internal carbon appears more durable than single-measure deployments, particularly in influent regimes characterized by low C/N. Because localized N_2O factors are materially below national defaults in this context, utilities should recalibrate baselines to avoid over-prioritizing technologies whose benefits were computed against overstated emissions.

Policy design benefits from the same localization. Annual municipal benchmarking that reports specific emissions and flags outlier plants can target oversight without penalizing facilities that face stricter effluent obligations. Standard-setting bodies and city consortia should enable Bayesian pooling of anonymized plant-month aggregates to sustain dynamic updates of regional emission factors. Incentive schemes that reward verified, quasi-experimentally identified net effects—rather than technology adoption per se—can reduce the risk of misallocated investment.

The study's limitations map directly to future work. The panel covers one city and two typical years; extreme wet-weather conditions and deeper process diversity (e.g., additional MBRs or sidestream nitrogen removal) remain underrepresented. Proxy-variable estimation of N_2O relies on monthly aggregates; higher-frequency telemetry would sharpen transient dynamics and could support mode-switching policies that complement the present rule set. While the decision tree achieves out-of-sample accuracy at the plant level, its static structure should be interpreted as a baseline for progressive integration with digital-twin architectures and safe reinforcement learning that preserve explainability.

Data and code availability will support independent replication and extension. De-identified plant-month data, the accounting and DID scripts, the Bayesian updating notebooks for the L-DEF library, and the Excel-embedded decision tool will be released under a permissive license upon publication, together with a time-stamped pre-analysis plan. In aggregate, the evidence supports a pragmatic path for medium-sized cities: diagnose heterogeneity with what is already measured, identify net effects with credible counterfactuals, localize the parameters that matter, and operationalize the result through rules that are simple enough to use and strong enough to matter.

Funding: Science funding of Three Gorges Ecological Environment Investment Co.,Ltd. No:NBWL202300224

References

[1] Y. Zhu, Z. Chang, Z. Li, et al., "Application of voltage to enrich anaerobic ammonium-oxidizing bacteria from mixed cultures for the degradation of actual wastewater containing COD," *Environ Monit Assess*, vol. 197, no. 12, pp. 1336, 2025.

[2] L. Yang, Y. Qin, Y. Zhao, et al., "A study on the classification and distribution characteristics of riparian vegetation at reach scales in the Nanliu River, China," *Environmental Monitoring and Assessment*, vol. 197, no. 12, 2025.

[3] Y. Lei, X. Jiang, S. Fan, et al., "Ecohydrological responses of vegetation changes to artificial water allocation schemes in the Heihe River Basin over the past 40 years," *Environ Monit Assess*, vol. 197, no. 12, pp. 1338, 2025.

[4] Y. Wang, A. Zhao, L. Zou, et al., “Ecosystem health responses to impervious surface expansion along a dynamic urban-rural gradient in China's Loess Plateau,” *Environ Monit Assess*, vol. 197, no. 11, pp. 1273, 2025.

[5] F. Alavian, F. Khodabakhshi, “Integrating artificial intelligence with microbial biotechnology for sustainable environmental remediation,” *Environ Monit Assess*, vol. 197, no. 11, pp. 1183, 2025.

[6] T. de Alencar Candido, S. S. Porfirio, K. S. B. de Lima Silva, et al., “Analyzing water quality monitoring data of environmentally protected water bodies,” *Environ Monit Assess*, vol. 197, no. 11, pp. 1187, 2025.

[7] Z. Wang, Y. Yu, H. Zhu, et al., “Spatiotemporal dynamics of soil moisture in a watershed in the Loess Plateau during vegetation restoration,” *Environ Monit Assess*, vol. 197, no. 11, pp. 1267, 2025.

[8] B. L. Masocha, P. Mhangara, “Satellite-driven assessment of methane trends, seasonal variability, and emission hotspots in Botswana's Central and Ngamiland Regions,” *Environ Monit Assess*, vol. 197, no. 10, pp. 1143, 2025.

[9] J. Park, H. T. Bui, E. Lee, et al., “Evaluating particulate matter and heavy metal accumulation, and air pollution tolerance in roadside woody plants, Cheongju City, South Korea,” *Environ Monit Assess*, vol. 197, no. 10, pp. 1133, 2025.

[10] S. T. Mabutana, S. Molander, P. Klintenberg, “Charcoal production in the tropical woodlands of southern Mozambique leads to land cover changes-the case of Combomune,” *Environ Monit Assess*, vol. 197, no. 10, pp. 1117, 2025.

[11] S. Sharma, B. Gajjar, C. Desai, et al., “Metagenomic analysis reveals the influence of wastewater discharge on the microbial community structures and spread of antibiotic-resistant bacteria at Mohar river, Gujarat,” *Environ Monit Assess*, vol. 197, no. 10, pp. 1112, 2025.

[12] A. A. Mohsen, A. K. Abdulkareem, “Assessing urban CO(2) emissions and sequestration potential: a case study of median strip trees in Baghdad,” *Environ Monit Assess*, vol. 197, no. 11, pp. 1197, 2025.

[13] Y. Tang, L. Liu, H. Sun, et al., “Comprehensive insights into the occurrence, spatiotemporal variations and risk assessment of antibiotics and their driving factors in a

large drinking water reservoir in the upper Yangtze River," *Environ Monit Assess*, vol. 197, no. 11, pp. 1170, 2025.

[14] N. F. M. Ali, I. Mohamed, R. M. Yunus, et al., "Spatio-temporal patterns of river water quality in the klang river basin, malaysia: a functional data analysis approach to detect pre- and post-pandemic shifts," *Environ Monit Assess*, vol. 197, no. 11, pp. 1198, 2025.

[15] S. Platikanov, A. Palomas, M. C. Mata, et al., "Coupling electrochemical and spectroscopic methods for river water dissolved organic matter characterization," *Environ Monit Assess*, vol. 197, no. 9, pp. 1071, 2025.

[16] N. Werfelli, R. Slaimi, G. Tayh, et al., "The irrigation impacts of treated wastewater over 35 years on soil properties and crop production," *Environ Monit Assess*, vol. 197, no. 9, pp. 1068, 2025.

[17] B. S. Rasool, A. K. Abbas, R. Haddad, "Photocatalytic innovations in environmental remediation: mechanisms, materials, and challenges for persistent organic pollutant removal," *Environ Monit Assess*, vol. 197, no. 10, pp. 1086, 2025.

[18] L. M. Boroujerd, A. Shakerdonyavi, Z. Asadgol, et al., "Machine learning-based prediction of drinking water quality index in Western Tehran using KAN, MLP, and traditional models," *Environ Monit Assess*, vol. 197, no. 9, pp. 1065, 2025.

[19] P. Ndem, O. Biose, E. Effiong, et al., "Electrochemical sensor-based assessment of CO(2) and CH(4) levels in Benin City under extreme heat wave conditions in Nigeria," *Environ Monit Assess*, vol. 197, no. 8, pp. 925, 2025.

[20] Z. Roodsarabi, H. Farhadi, H. Ebadi, et al., "Machine learning-based burned area detection using Sentinel-2 imagery and spectral indices," *Environ Monit Assess*, vol. 197, no. 8, pp. 905, 2025.

[21] K. Khalkhal, A. Reyhanitabar, S. Oustan, et al., "Improving municipal solid waste co-composting performance using activated biochar, leonardite, and MgSO(4)-modified zeolite," *Environ Monit Assess*, vol. 197, no. 8, pp. 847, 2025.