Unraveling operational drivers of nitrous oxide emissions in biological wastewater treatment systems through machine learning analysis of multi-decadal datasets

Gnanaraj Augustine, Kartik Chandran*

*Department of Environmental Engineering, Columbia University, New York, 10027, USA

ABSTRACT: This study focused on the development of machine-learning- (ML-) based strategies for mitigating nitrous oxide (N_2O) emissions from various wastewater treatment systems in the United States measured using a benchmark USEPA-endorsed protocol. Results revealed that in general, poor process performance correlated with higher N_2O emissions. Specifically, local variables including zone-specific dissolved oxygen, ammonia, and nitrite concentrations and global variables including effluent nitrite and nitrate concentrations contributed positively towards N_2O emissions from both aerobic and anoxic zones of the process bioreactors. The optimal operational conditions identified for minimizing N_2O emissions included, operation of aerobic and anoxic zones at DO > 4 mg O_2/L and < 1 mg O_2/L, respectively, coupled with appropriate solids retention times (SRTs) that maximize process performance. Accordingly, our results strongly underscore the utility of ML models in combination with bioprocess fundamentals for predicting and mitigating N_2O emissions, while concomitantly optimizing wastewater treatment operations.

KEYWORDS: *nitrous oxide, activated sludge, greenhouse gas, machine learning*

Introduction

Nitrous oxide (N_2O) is a potent greenhouse gas (GHG) and ozone depleting substance [1,2]. Wastewater treatment plants (WWTPs) have been increasingly recognized as substantial anthropogenic sources of N_2O emissions[3–7]. On one hand, increasingly stringent effluent nitrogen limits from wastewater treatment plants have spurred a global effort to develop and implement engineered biological nutrient removal (BNR) processes to improve water quality[4]. On the other hand, suboptimal operating conditions and process configurations associated with BNR operations can exacerbate N_2O production[8].

Biological pathways and drivers for chemolithoautotrophic and chemoorganoheterotrophic N_2O production during BNR operations. During nitrification, ammonia-oxidizing bacteria (AOB)

and nitrite-oxidizing bacteria (NOB) sequentially oxidize ammonia (NH_3) to nitrite (NO_2^-) and nitrate (NO_3^-), respectively, using oxygen as the preferred terminal electron acceptor[9]. During denitrification, NO_3^- is reduced to dinitrogen (N_2) gas, coupled with the utilization of organic or inorganic electron donors[9]. In nitrifying WWTPs, AOB produce nitric oxide (NO) and N_2O via nitrifier denitrification under limiting dissolved oxygen (DO) and non-limiting NO_2^- concentrations [6,10–15]. Also, under non-limiting DO and NH_3 concentrations, AOB produce N_2O during the oxidation of hydroxylamine (NH_2OH) [10,16]. N_2O can also be produced through chemical oxidation of NH_2OH abiotically [10,16]. During typical engineered denitrification, chemoorganoheterotrophic bacteria produce N_2O under operating conditions that favor inhibition of nitrous oxide reductase (Nos) [16], for instance, associated with inhibition by oxygen and NO_2^- [12,17,18]. Despite the contribution of both nitrification and denitrification in N_2O production, it has now been established that both *production* and *emission* of N_2O from aerated zones are higher than from non-aerated zones at WWTPs [4].

The characterization and minimization of N_2O emissions from WWTPs has assumed widespread and critical importance as wastewater utilities aim to minimize the overall carbon footprint of treatment operations, while adhering to increasingly stringent discharge regulations[19]. Previous N_2O footprinting studies performed at full-scale BNR and non-BNR WWTPs around the United States reported N_2O emission fractions (EFs) in the range 0.01-1.8% relative to the influent TKN load[4]. Other full-scale studies have reported widely varying EFs in the range 0.001-25.4% of the influent load[4,20].

N_2O production is directly correlated with nitrogen conversion dynamics in WWTPs further indicating the potential of N_2O emissions as BNR performance indicators [4–6,20]. Various operational and environmental factors such as solids retention time (SRT), specific AOB activity, temperature, and influent loading patterns contribute to both N_2O emissions and overall process performance in BNR systems [4–7,20]. BNR operations at sub-optimal conditions, such as excessively high or low DO concentrations combined with inadequate SRT, can lead to transient or permanent NH_3 or NO_2^-

accumulation. Such accumulation is symptomatic of inadequate nitrification and denitrification performance and can be correlated, in turn, to NO and N_2O production and emission[5–10]. Mechanistic models have been used to predict and reduce N_2O emissions in the BNR systems from WWTPs, ranging from models based on the activated sludge framework [21] to models based on gene- and protein-expression[15,22–25]. Other data-driven models have also been proposed, which when trained on robust and representative datasets, offer a balance between the mathematical and metabolic intricacies relating to N_2O production and their practical end-application for process control[4,26]. Data-driven models such as unsupervised Principal Component Analysis (PCA) and supervised machine learning (ML) models have shown greater insights into improving process performance and identifying factors contributing to N_2O emissions[27–39]. Supervised ML models have demonstrated the potential for predicting N_2O emissions from full-scale BNR operations, and studies have shown the usage of ML models focused on predicting effluent water quality[32–36] or separate N_2O emissions.[40–43], with limited dataset and using different models and hyperparameter optimization (HPO)[43].

In this study, we subjected both historical and currently obtained datasets to unsupervised PCA and supervised ML modeling approaches to uncover the complex interplay between operational strategies, N_2O emissions and effluent water quality and to develop actionable insights into WWTP operating strategies that minimize both *gaseous* and *aqueous* nitrogenous discharges. Through this research, we offer new insights into concurrently maximizing BNR process efficiency and achieving decarbonization of BNR operations, thereby promoting a more sustainable wastewater sector.

Results

Updated N_2O emissions from advanced mainstream and sidestream activated sludge processes

We first present the results obtained from recently conducted (2023-2024) N_2O measurement campaigns at five different next-generation mainstream and sidestream wastewater treatment processes in the Northeastern and Mid-Atlantic regions of the United States (**Table 1**). For the

78 *mainstream processes* tested, the average emission factor (EF), representing the fraction of the
79 influent nitrogen load emitted as N_2O, was statistically lower than the IPCC EF (1.6±1.2%) ($p=0.03$,
80 $\alpha=0.05$)[3]. Within the mainstream processes evaluated, the EF for the four-pass step-feed BNR process
81 tested (0.20±0.03%) was statistically higher ($p=0.01$, $\alpha=0.05$) than for the processes based on
82 ammonia-based aeration control (ABAC) (0.004±0.00%), partial denitratation-anammox (PdNA)
83 (0.006±0.00%), and fermentate-fed partial denitratation (0.003±0.00%). In contrast, the *sidestream*
84 *process* exhibited EFs in the range 6.3 -14 %, which were statistically higher ($p<0.05$) than both the
85 mainstream and IPCC EFs. Compared to our previous study EF**(Table 1),** our current study EF from
86 next-generation mainstream systems shows lower EF and better N-removal (**Figure S1**), showcasing
87 our recent development in improved process operation and reduced N_2O emissions.

88 However, some process-specific nuances could also be inferred based on the current datasets. For
89 the mainstream PdNA process, operating temperature did not statistically impact either the BNR
90 performance (% TN removal, which averaged 72±9 % across temperatures) or the N_2O EF (**Table**
91 **1**). In contrast, the emissions from the step-feed BNR process were higher during the lower
92 temperature monitoring campaign (**Table 1**). On the other hand, ABAC-based process represents
93 distinct operating characteristics with uncoupled organic carbon, nitrification, and post-
94 denitrification, and displayed a low EF. A specific N-removal efficiency for the ABAC process could
95 not be correlated with the measured N_2O emissions, which were restricted to the aerated zones in the
96 process (**Table 1**). Therefore, N_2O emissions are subjective to their process operations within the
97 designed process configuration, and are not truly influenced by operating temperatures.

98 The SHARON process tested treats post-anaerobic digestion centrate and operates via a combination
99 of nitritation (oxidation of influent NH_3 to NO_2^-) followed by glycerol-mediated denitritation
100 (reduction of NO_2^- to N_2). The SHARON process displayed low ammonia conversion (46±15% and
101 54±11%) and low nitrogen removal (36±17% and 38±16%)**(Table 1)**. These observed process
102 performance measures corresponded with a higher N_2O EF (6.3-14%), relative to the other processes

tested. Conceptually, sidestream treatment processes, including those operating in partial nitritation mode can display EFs as low as 0.5%[4,44], suggesting considerable room for reduction in emissions with concomitant performance improvements in the system tested herein.

Such findings reinforce the previously demonstrated parallels between process performance and N_2O emissions from WWTPs[7,44–46]. Nevertheless, a robust approach capable of unraveling the complex interplay between process configurations, operating conditions, process performance and corresponding emissions is needed to explore the development of process optimization and decarbonization strategies, and is described next.

Unraveling the Interdependencies of N_2O Flux and Process Parameters Using a Multi-Decadal Wastewater Treatment Operations Dataset

Advanced unsupervised and supervised machine learning models were employed using our present and past wastewater process operational data spanning more than a decade of process monitoring (**Table 1**), with the goal of improving both *aqueous* and *gaseous* nitrogen discharges.

First, using unsupervised principal component analysis (PCA), we revealed the interdependencies between nitrification and denitrification process variables (nitrogen species, organic compounds, suspended solids), operating conditions (SRT, influent flow rate) and measured N_2O fluxes. In the presented PCA biplots, (**Figures 1A and 1B)** the length of the arrow represents the magnitude of the variable contribution. Positively correlated variables are clustered on the same side of the axis of a reference variable (for instance, N_2O fluxes), while the negatively and uncorrelated variables are positioned opposite and perpendicular, respectively. We further classified the different influential factors into two categories. Specifically, zone-specific variables including NH_3, DO, NO_2^- and SRT were classified as *local variables*, since they reflect spatially discrete sampling locations within the WWTP process bioreactors. Influent and effluent water quality parameters, influent flowrate, and temperature were considered as *global variables*.

127 **PCA-based interpretation of aerated zone emissions.** Based on aerated zone data from current and
128 previous sampling campaigns, PC1 and PC2 captured 35.7% and 19.1% of the variance, respectively
129 (**Figure.1A**). Among the local variables, NH_3, DO, NO_2^- and among the global variables, influent
130 TKN, effluent NO_2^-, effluent TSS, effluent ffCOD, and effluent tCOD were strongly and positively
131 correlated with N_2O fluxes (**Figures 1, and S2A**). One fundamental interpretation of such
132 correlations between the local variables and observed N_2O emissions from the aerated zones points
133 to the role of nitrifier nitrification and nitrifier denitrification pathways. During nitrifier-nitrification,
134 non-limiting DO and NH_3 concentrations in aerated zones can lead to incomplete NH_2OH oxidation
135 by AOB, which results in NO and N_2O production therein [4,10,16]. During nitrifier-denitrification, AOB
136 produce N_2O primarily under conditions that result in non-limiting NO_2^- concentrations, potentially
137 exacerbated by limiting DO and non-limiting NH_3 concentrations [4,10]. Of the global variables, high
138 influent TKN and effluent NO_2^- concentrations could indirectly correspond with higher N_2O fluxes
139 depending on SRT and the corresponding extant local variables (**Figure 1A**). The observed positive
140 correlation of measured N_2O fluxes with different effluent TSS and organic carbon (ffCOD and
141 tCOD) concentrations for a given degree of nitrification cannot be interpreted in isolation and likely
142 needs to be considered using more complex models (such as the ML models considered below).

143 Negative correlations of SRT, zone-specific NO_3^-, and influent tCOD concentrations with measured
144 N_2O fluxes were also observed (**Figures 1A, and S2A**), suggesting overall that conditions that
145 promote full nitrification (higher zone-specific NO_3^- concentrations) at WWTPs (higher SRTs)
146 correspond to lower N_2O fluxes (in keeping with previous results) [47]. In general, the link between
147 lower N_2O emissions and lower influent tCOD concentrations is indirect and points to increased
148 oxygen availability for complete nitrification compared to organic carbon oxidation[9].

149 Of the remaining global variables, effluent NO_2^- concentrations correlated positively with aerobic
150 N_2O fluxes, and likely represent the link inadequate nitrogen removal and higher N_2O emissions[4]
151 (**Figures.1A and S2A**). The global variables captured by both PC1 and PC2 included temperature,

influent flow rate, influent TIN, influent NH_3, effluent VSS, effluent NH_3, and did not statistically correlate with aerobic-zone N_2O fluxes (**Figures 1A, and S2A**).

PCA-based interpretation of anoxic-zone emissions. For data obtained from all the previously and currently sampled anoxic zones, PC1 and PC2 captured 32.4% and 15.3% of the variance, respectively. Local variables including zone-specific DO and NO_2^- concentrations were positively correlated with measured N_2O fluxes (**Figures 1B, and S2B**). Notably, zone-specific DO and NO_2^- concentrations clustered closely together, emphasizing their potential significant interactive impacts on N_2O production again linked with poor process operation and incomplete denitrification, as also shown previously [4].

Global variables such as effluent NO_3^-, effluent NO_2^-, influent TKN, influent NH_3, and influent TIN concentrations, exhibited a positive correlation with N_2O fluxes (**Figures 1B, and S2B**). In contrast, zone-specific NO_3^- concentrations and influent flow rate were negatively correlated with N_2O fluxes (**Figure 1B**). Effluent TKN and effluent NH_3 concentrations were not statistically correlated with N_2O fluxes (**Figures 1B, and S2B**).

Hyperparameter Optimization of ML models for predicting N_2O flux under different wastewater treatment parameters

We further developed ML models with combinations of different process variables and parameters from our past and present data to predict and mitigate N_2O emissions from WWTP operations. The ML models were trained to predict N_2O flux (g-N/min-m^2), N_2O emission rate (g-N/d), and N_2O concentration (ppm) using three machine learning algorithms namely Random forest (RF), Gradient Boosting (GBM), and XG Boost. Each model was optimized under different hyperparameter tuning to identify the optimum parameter for achieving the best prediction performance (**Table S1**). Among these, XG boost consistently demonstrated superior performance, achieving an R^2 of 0.99 for N_2O flux prediction in both aerobic and anoxic conditions. The Root Mean Square Error (RMSE) for XG

Boost was significantly lower compared to RF and GB, with values of 0.053 and 0.022 for aerobic and anoxic conditions (Table S2). Furthermore, N_2O emission rates and concentrations also showed higher prediction accuracy with XG Boost compared to RF and GB models (**Table S2**).

Machine Learning Insights into the direct impact of wastewater effluent quality on N_2O flux

From the trained ML models, permutation importance values were extracted for both aerobic and anoxic zone conditions (**Figure 2A-C**). The most important variables influencing aerated-zone N_2O emission fluxes included influent flow rate, local NH_3, DO, and NO_2^- concentrations, SRT, and Effluent NO_2^- concentrations (**Figure 2B**). For non-aerated zones, the most important variables influencing N_2O emission fluxes were temperature, zone-specific DO, and NO_2^- concentrations, SRT, and Influent NH_3, Effluent NO_3^-, Effluent TIN and Effluent sCOD concentrations **(Figure 2C)**. These variables identified by the ML models as being influential on N_2O emissions aligned with the PCA analysis and underscored the now well-established link between process performance and emissions [4]. Our results also aligned with findings from a similar ML study using AdaBoost, which pointed to temperature, DO, NH_3, NO_2^-, and NO_3^- as being influential factors for aerobic zone N_2O emissions, without corresponding insights into anoxic-zone emissions [27]. Our study highlights the critical need to consider quantifying and predicting N_2O emissions from anoxic zones, as their contribution to N_2O production and emissions are inevitably under poor operating conditions[4,5].

We further integrated SHAP analysis to uncover the magnitude and directionality influence of the variables towards predicted N_2O flux, bridging the gap between model predictions and actionable process insights. *SHAP analysis for aerobic* N_2O flux indicates that influent flow rate is the most significant variable compared to others (**Figure 3A**). An increase in influent flow rate positively influences higher N_2O flux, as it is on the positive side of the SHAP values. NH_3 and DO demonstrate considerable influence as elevated NH_3 SHAP value and its concentrations explain increased N_2O flux, while higher DO levels can result in both higher and lower N_2O flux conditions,

200 depending on the condition. The results show that operations under non-limiting NH_3 concentrations
201 can lead to N_2O production, regardless of whether DO levels are high or low. Therefore, optimizing
202 DO levels *in conjunction with* appropriate NH_3 concentrations is essential for effective nitrification.

203 Our results also suggest potential interactive effects of operating SRT with DO and NH_3
204 concentrations (**Figure 3A**), which has a firm biochemical basis. For instance, operation at elevated
205 SRT can lead to reduced N_2O fluxes, when coupled with optimization of DO concentrations to
206 achieve a given NH_3 conversion. Based on our data, NO_2^- concentrations had a lower impact on N_2O
207 fluxes compared to that of NH_3, DO, and SRT. Again, the production of NO_2^- can only be controlled
208 by concurrently optimizing DO and SRT along with NH_3 concentration.

209 Based on *SHAP analysis for anoxic* zones, the N_2O fluxes therein were highly influenced by
210 DO concentrations (**Figure 3B**), which could be linked to DO-mediated inhibition of N_2O reduction
211 to N_2[4,5,18]. A biphasic response of N_2O fluxes to SRT values was observed once again pointing to
212 possible interactions with extant DO and NO_2^- concentrations. NO_3^- on the other hand, influenced
213 anoxic-zone N_2O to a lower extent. Higher NO_3^- feature value representing higher SHAP value
214 (**Figure 3B**) does not directly imply its involvement in producing N_2O. The global variables, effluent
215 NO_2^- and NO_3^- have higher feature values in the positive side of the SHAP (Figure 3B), reflecting
216 that higher effluent NO_2^- and NO_3^- correlate with increased N_2O flux as it represents an ineffective
217 denitrification process.

218 **Optimization of process performance from the ML predictive outcome**

219 ***i. Aerobic zone optimization***

220 The process optimization approach entailed identifying operating ranges for SRT, which is the
221 principal process parameter coupled with process performance responses that concurrently minimize
222 off-gas N_2O fluxes, via PDP analysis. **Figure 4A** suggests that NH_3 concentrations < 3 mg L^{-1} at a
223 SRT < 10 days show less N_2O flux. To study the relationship between DO and NH_3, **Figure 4B**

224 suggests that operating conditions that achieve NH_3 concentrations < 3 mg L^{-1}, with DO < 3 mg L^{-1},
225 are effective in minimizing N_2O production. In contrast, operating conditions that foster concomitant
226 non-limiting DO and NH_3 concentrations, correspond to high N_2O fluxes **(Figure 4B).** Such
227 predictions are in excellent alignment with fundamental and mechanistic understanding of drivers for
228 N_2O emissions under aerobic conditions at WWTPs, via a combination of oxidative and reductive
229 pathways primarily in AOB[4–10,44–48].

230 ***ii. Anoxic zone optimization***

231 Similarly, conditions that minimize N_2O fluxes from anoxic zones included operation at DO
232 concentrations < 1 mg/L, however, coupled with a total process SRT < 20 days (**Figure 4C**). On the
233 other hand, N_2O flux increases when operated at DO concentration > 1 mg/L, under SRT < 20
234 days(**Figure 4C**). **Figure 4D** shows that when operated at DO concentrations > 1 mg/L, results in
235 NO_2^- accumulation in the anoxic zone, which subsequently increases N_2O flux. Higher DO and NO_2^-
236 concentrations promote high N_2O flux even under increased SRT. Also, we acknowledge that
237 operation at such high SRT values may be an option for some but not all WWTPs. Therefore,
238 understanding *process control vs process operation* in a dynamic environment is essential to mitigate
239 N_2O fluxes and improve reactor performance.

240 ## Discussion

241 **Poor process operation leads to high N_2O emissions irrespective of process configuration and**
242 **temperature variations.**

243 Based on a combination of current measurements and extensive historic datasets, we determined that
244 less than optimal process operations can manifest themselves in combinations of both local and global
245 variables that in turn are correlated with higher N_2O production and emission across aerobic and
246 anoxic zones. Among the identified variables, SRT and DO emerged as the critical control factors
247 governing emissions. The key local and global variables identified from the unsupervised PCA

248 followed by supervised ML permutation feature importance and SHAP align well with the established
249 understanding of the nitrification and denitrification processes[4–7,29–31], thus providing consistent
250 insights into the contribution and significance linking to the process operation, N_2O emission, and
251 effluent water quality. The PDP analysis conducted for aerobic zones illustrates the relationship of
252 NH_3 concentration with key process control variables, DO and SRT, identifying the optimal operating
253 conditions for effective nitrification and reduced N_2O flux (**Figure 4A, B**). Similarly, in anoxic zones,
254 the PDP analysis highlights the interactions among DO, SRT, and NO_2^- pinpointing optimal
255 conditions to achieve effective denitrification and reduced N_2O emissions (**Figure 4C, D**). Thus, the
256 findings from ML-based analysis reveal the interplay between the operational variables and their
257 dynamic environment, that contribute to N_2O fluxes. Using interpretation tools, like PDP, coupled
258 with biochemical understanding, can provide real-time insights to optimize wastewater treatment
259 processes to minimize both *aqueous* and *gaseous* nitrogenous discharges.

260 **Integration of real-time process control based on ML framework**

261 This study demonstrates the immense potential of data-driven models for process control and
262 optimization. Upon further integrating such models into Supervisory Control And Data Acquisition
263 (SCADA) systems, they continuously process data generated from real-time or non-real-time
264 measurements, facilitating the development of digital twin architectures based on model-based
265 systems engineering (MBSE)[49]. Data-driven models can provide early fault detection, and advanced
266 process control capabilities, which mechanistic models typically cannot deliver using real-time data[49].
267 Indeed, previous studies have successfully applied data-driven approaches in real-time reactor
268 operations, enhancing reactor performance and effluent water quality [50–53]. The emphasis on data-
269 driven modeling for GHG emissions, particularly N_2O emissions, continues to expand, driven by
270 ongoing efforts to understand their process dynamics better and improve forecasting accuracy[27,40–43].
271 Our current work demonstrates the capabilities of ML models and their useful insights into the
272 dynamic interplay and link between process operations, N_2O emissions and effluent water quality.

The advanced interpretation tools, PDP, within the supervised ML models allows for better process control setpoints and leverages the scope for digital and hybrid twin systems for optimized prediction and management of WWTP operations.

Materials and Methods

Field Data Collection using an USEPA Approved Protocol to Measure N_2O Emissions from BNR Processes

The currently reported overall N_2O measurement campaign was performed at three different BNR wastewater treatment plants around the US. **Table S3.** provides an overview of the sampling techniques performed across both mainstream and sidestream WWTPs. The gaseous N_2O emission concentrations from the BNR processes were quantified using an USEPA endorsed protocol[54]. The Surface Emission Isolated Flux Chamber (SEIFC), the only floating body approved by the Environmental Protection Agency (EPA) ($\Delta P=0$), has a circular cross-sectional area of 0.13 m^2 and is used to measure gaseous nitrogen fluxes from activated sludge tanks[54,55]. The SEIFC consists of a floating enclosed space through which exhaust gas were collected in real time fashion. The SEIFC was employed to obtain headspace N_2O concentration(ppm) and fluxes at distinct locations (aerated and nonaerated) in a treatment process. At each location, online measurements of gaseous N_2O concentration and gas flow were conducted over a 24-hour period to address the diurnal variability of gas emissions. The gaseous N_2O concentration was measured via infrared gas filter correlation (Teledyne API, San Diego, California) at a frequency of 1/min. For emission measurements from the closed-surface, sidestream SHARON process, the N_2O analyzer was directly connected to the off-gas stacks. Mainstream N_2O measurements were conducted at both aerobic and anoxic zones in the target WWTPs. Measurement campaigns in both mainstream and sidestream processes were conducted over multiple days to determine the spatial and temporal variability. The use of GC-TCD in past studies to measure the off-gas flow has provided valuable data at different WWTPs around US[4,55].

We used an anemometer (Hotwire thermo anemometer) to measure the air flow from the off-gas of the flux chamber, which helps to identify the N_2O flux from the aerated and non-aerated zones without requiring heavy-weight GC and other supplies. In other studies, a major drawback when monitoring flux from the aerated and anoxic zone is to rely on the blower diffuser data and bulk airflow measured in a vent from a closed tank[56]. In an aerobic zone, inconsistent blower supply due to aging infrastructure and lack of continuous blower data limits flux measurements from the tank, representing substantial spatial variable emissions. Anoxic zones do not have blower data, limiting the ability to monitor and measure N_2O flux contribution from their AST systems.

Figure. S7 and supplementary section 7 explains the drawbacks of using blower data to measure aerobic zone emissions and neglecting the anoxic zone, which can significantly eliminate the consideration of N_2O contribution from the anoxic zone. We have also calculated the off-gas estimations comparing blower and air flow data. **Figure. S8** represents the difference between air flow and blower off-gas flow rates for every aerobic zone during our campaign in step-feed BNR process. The results show that consideration of blower data can overestimate emission rates from WWTPs. In addition to our comparative study on blower vs air flow data, we also identified the optimum sweep air introduction to the flux chamber in the anoxic zone to quantify fluxes. **Figure. S9** discusses the introduction of sweep air into the flux chamber in an anoxic zone, where the optimum sweep air introduction are determined based on the off-gas saturation flow rate and gaseous N_2O dilution.

Concomitantly, during each monitoring campaign, wastewater nitrogen species concentrations, including influent and effluent bioreactor samples were obtained using an autosampler every 1 hour to measure NH_3, NO_2^-, and NO_3^- concentrations using Lachat Quickchem Flow injection analysis system [57]. Activated sludge mixed-liquor temperature, pH, DO are acquired from the SCADA system and soluble chemical oxygen (sCOD) and total chemical oxygen demand (tCOD) were measured with Hach TNTplus kits and a Hach DR2800 spectrophotometer (HACH

Loveland, CO). Mixed-liquor suspended solids (MLSS) and volatile suspended solids (VSS) concentrations were analyzed using Standard Methods 2504D and 2540E [57].

Bioreactor profile samples were collected from nitrifying and denitrifying tanks once a day to measure NH_3, NO_2^-, and NO_3^- concentrations using Hach TNTplus kits and a Hach DR2800 spectrophotometer (HACH Loveland, CO). DO concentrations, temperature (℃), and pH were measured using HACH handheld sensors (HACH Loveland, CO). Process air flow rate (standard cubic feet per minute, SCFM), SRT, the surface area of the tanks (m^2) and influent flow rates (million gallons per day, MGD) were obtained from the plant operation records.

The N_2O fluxes were calculated from both aerated and non-aerated zone using Eq.1 and the emissions were calculated using Eq.2.

$$N_2O_{Flux}\left(\frac{g-N}{min-m^2}\right) = \left(\frac{(V_{air}*A_t*1000*60)+Q_{N2O_inst}}{A_{SEIFC(}m^2)}\right) * C_{N_2O}(\frac{g-N}{L}) \qquad \textbf{Eq.1}$$

$$N_2O_{ER}\left(\frac{g-N}{d}\right) = N_2O_{Flux}\left(\frac{g-N}{min-m^2}\right) * A_{tank}\,(m^2) * 1440 \qquad \textbf{Eq.2}$$

$$EF\ \% = \frac{N_2O_{ER}\left(g-\frac{N}{d}\right)}{\text{Influent TKN Load}\left(g-\frac{N}{d}\right)} * 100 \qquad \textbf{Eq.3}$$

Where N_2O_{flux} is the measure of N_2O flux(g-N/min.m^2) from the activated sludge tank, V_{air} is the air velocity ($m.s^{-1}$) measured in the anemometer from the SEIFC parallel to the N_2O concentration measurements, A_t is the Area of the tube(m^2), Q_{N2O_inst} is the off-gas flow recorded on the N_2O analyzer ($L.min^{-1}$) from the SEIFC, A_{SEIFC} is the area of the flux chamber (m^2) and C_{N2O} is the average concentration of the gaseous N_2O concentration (g-$N.L^{-1}$). N_2O_{ER} is the N_2O emission rate (g-N/d) from the overall tank, and A_{tank} is the area of the activated sludge tank (m^2). To determine the Emission Factor (EF), the Q_{N2O} (N_2O emission rate (g-N/d)) was normalized to the influent Total Kjeldahl Nitrogen (TKN) (g-N/d) load using Eq.3.

Data Preprocessing for Calculating EF and Implementing unsupervised PCA analysis

The measured data were grouped into zone-specific aerobic and anoxic zone measurements, which were then used to calculate N_2O emission rates and EF across the system **(Figure 2A)**. The dataset throughout this study are expressed to two significant figures. Furthermore, past and present datasets were processed to identify missing and non-numeric values before initiating the analysis in machine learning algorithms. The missing values were filled with "*missForest,"* a package in the R studio that successfully handles the missing values for different variables, even under non-linear relations[58].

To perform the unsupervised PCA analysis, processed datasets are arranged according to aerobic and anoxic zone conditions. We used "*prcomp"* package with "*factoextra"* visualization packages in R studio to advance the PCA visualization. For Pearson correlation, "*corrplot"* and "*ggcorrplot"* packages were used to develop the correlation heatmap.

Supervised Machine Learning Approach to Predict N_2O Fluxes from Different Operating Conditions

To develop the supervised ML models, the datasets were log-transformed and min-max normalized [0,1]. **Figure S10** provides the graphical representation of the preprocessed dataset in a boxplot for both aerobic and anoxic datasets. The pre-processed dataset were randomly split into 75% for training and 25% for testing. The dataset split for the training data were divided into k-fold cross-validation (cv) splits to optimize the model hyperparameters effectively. Outlier detection was performed to understand the data sets variability for all the independent variables associated with N_2O fluxes.

Supervised ML comprising an ensemble based on decision trees were employed. They are Random Forest, Gradient Boosting, and Xtreme Gradient Boost. The model hyperparameter tuning of the selected models was optimized using a grid search algorithm. The grid search explores a pre-define space of parameters to find the set of values that maximizes the model mean performance on the k-

fold CV subsets. The hyperparameters and optimized hyperparameter values for each model are provided in **Supplementary Section 4.** All the model evaluations are implemented in R-studio.

Model evaluation and comparison

The model accuracy with N_2O flux estimates was evaluated in both statistical metrics and dynamic process evaluations. **Table S2** shows regression-based metrics, including the coefficients of determination (R^2), which are used to estimate the proportion of the variance the model explains. Mean Square Error (MSE) and Root Mean Square Error (RMSE) were also employed to understand the magnitude of errors that occur across different ranges of N_2O flux prediction for both aerobic and anoxic prediction. This provides detailed residual analysis to provide deeper insight into the model performance and detect potential areas of improvement.

Tools utilized for N_2O flux prediction from machine learning models

We employed SHAP and PDP analyses to interpret the ML model predictions and to investigate the directionality and relative influence of process variables on the predicted N_2O flux. For this purpose, we used the *iml* package to perform model-agnostic interpretation, *pdp* to generate partial dependence plots, *ggbeeswarm* to create beeswarm plots for visualizing SHAP values, and *shapviz* for efficient and scalable SHAP-based visualizations. The *viridis* package was used to apply perceptually uniform, colorblind-friendly palettes, and *plotly* was utilized to develop interactive visualizations of SHAP and PDP outputs.

References

1. Ravishankara, A. R., Daniel, J. S. & Portmann, R. W. Nitrous Oxide (N_20): The Dominant Ozone-Depleting Substance Emitted in the 21st Century. *Science* **326**, 123–125 (2009).
2. Stocker, T. *Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. (Cambridge University Press, 2014).
3. 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories — IPCC. https://www.ipcc.ch/report/2019-refinement-to-the-2006-ipcc-guidelines-for-national-greenhouse-gas-inventories/.
4. Ahn, J. H. *et al.* N2O emissions from activated sludge processes, 2008-2009: Results of a national monitoring survey in the united states. *Environmental Science and Technology* **44**, 4505–4511 (2010).
5. Rassamee, V., Sattayatewa, C., Pagilla, K. & Chandran, K. Effect of oxic and anoxic conditions on nitrous oxide emissions from nitrification and denitrification processes. *Biotechnology and Bioengineering* **108**, 2036–2045 (2011).
6. Yu, R., Kampschreur, M. J., Loosdrecht, M. C. M. van & Chandran, K. Mechanisms and Specific Directionality of Autotrophic Nitrous Oxide and Nitric Oxide Generation during Transient Anoxia. *Environ. Sci. Technol.* **44**, 1313–1319 (2010).
7. Brotto, A. C. *et al.* Characterization and mitigation of nitrous oxide (N2O) emissions from partial and full-nitrification BNR processes based on post-anoxic aeration control. *Biotechnology and Bioengineering* **112**, 2241–2247 (2015).
8. Kampschreur, M. J., Temmink, H., Kleerebezem, R., Jetten, M. S. M. & van Loosdrecht, M. C. M. Nitrous oxide emission during wastewater treatment. *Water Research* **43**, 4093–4103 (2009).
9. Grady, C. P. L. *Biological Wastewater Treatment*. (New York : Marcel Dekker, 1999).

10. Chandran, K., Stein, L. Y., Klotz, M. G. & Van Loosdrecht, M. C. M. Nitrous oxide production by lithotrophic ammonia-oxidizing bacteria and implications for engineered nitrogen-removal systems. *Biochemical Society Transactions* **39**, 1832–1837 (2011).
11. Zumft, W. G. Cell biology and molecular basis of denitrification. *Microbiol Mol Biol Rev* **61**, 533–616 (1997).
12. Kester, R. A., De Boer, W. & Laanbroek, H. J. Production of NO and N(inf2)O by Pure Cultures of Nitrifying and Denitrifying Bacteria during Changes in Aeration. *Appl Environ Microbiol* **63**, 3872–3877 (1997).
13. Ritchie, G. A. & Nicholas, D. J. Identification of the sources of nitrous oxide produced by oxidative and reductive processes in Nitrosomonas europaea. *Biochem J* **126**, 1181–1191 (1972).
14. Stüven, R., Vollmer, M. & Bock, E. The impact of organic matter on nitric oxide formation by Nitrosomonas europaea. *Arch. Microbiol.* **158**, 439–443 (1992).
15. Yu, R., Perez-Garcia, O., Lu, H. & Chandran, K. *Nitrosomonas europaea* adaptation to anoxic-oxic cycling: Insights from transcription analysis, proteomics and metabolic network modeling. *Science of The Total Environment* **615**, 1566–1573 (2018).
16. Schreiber, F., Wunderlin, P., Udert, K. M. & Wells, G. F. Nitric oxide and nitrous oxide turnover in natural and engineered microbial communities: biological pathways, chemical reactions, and novel technologies. *Front Microbiol* **3**, 372 (2012).
17. Wunderlin, P., Mohn, J., Joss, A., Emmenegger, L. & Siegrist, H. Mechanisms of N2O production in biological wastewater treatment under nitrifying and denitrifying conditions. *Water Research* **46**, 1027–1037 (2012).
18. Lu, H. & Chandran, K. Factors promoting emissions of nitrous oxide and nitric oxide from denitrifying sequencing batch reactors operated with methanol and ethanol as electron donors. *Biotechnology and Bioengineering* **106**, 390–398 (2010).

19. US EPA, O. Effluent Guidelines. https://www.epa.gov/eg (2014).
20. Daelman, M. R. J., van Voorthuizen, E. M., van Dongen, U. G. J. M., Volcke, E. I. P. & van Loosdrecht, M. C. M. Seasonal and diurnal variability of N2O emissions from a full-scale municipal wastewater treatment plant. *Science of the Total Environment* **536**, 1–11 (2015).
21. Gernaey, K. V., van Loosdrecht, M. C. M., Henze, M., Lind, M. & Jørgensen, S. B. Activated sludge wastewater treatment plant modelling and simulation: state of the art. *Environmental Modelling & Software* **19**, 763–783 (2004).
22. Perez-Garcia, O., Mankelow, C., Chandran, K., Villas-Boas, S. G. & Singhal, N. Modulation of Nitrous Oxide (N2O) Accumulation by Primary Metabolites in Denitrifying Cultures Adapting to Changes in Environmental C and N. *Environ. Sci. Technol.* **51**, 13678–13688 (2017).
23. Frontiers | Metabolic Network Modeling of Microbial Interactions in Natural and Engineered Environmental Systems. https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2016.00673/full.
24. Perez-Garcia, O., Chandran, K., Villas-Boas, S. G. & Singhal, N. Assessment of nitric oxide (NO) redox reactions contribution to nitrous oxide (N2O) formation during nitrification using a multispecies metabolic network model. *Biotechnology and Bioengineering* **113**, 1124–1136 (2016).
25. Perez-Garcia, O., Villas-Boas, S. G., Swift, S., Chandran, K. & Singhal, N. Clarifying the regulation of NO/N2O production in *Nitrosomonas europaea* during anoxic–oxic transition via flux balance analysis of a metabolic network model. *Water Research* **60**, 267–277 (2014).
26. Behera, C. R., Srinivasan, B., Chandran, K. & Venkatasubramanian, V. Model based predictive control for energy efficient biological nitrification process with minimal nitrous oxide production. *Chemical Engineering Journal* **268**, 300–310 (2015).

27. Khalil, M., AlSayed, A., Liu, Y. & Vanrolleghem, P. A. Machine learning for modeling N2O emissions from wastewater treatment plants: Aligning model performance, complexity, and interpretability. *Water Research* **245**, 120667 (2023).

28. Bellandi, G., Weijers, S., Gori, R. & Nopens, I. Towards an online mitigation strategy for N2O emissions through principal components analysis and clustering techniques. *J Environ Manage* **261**, 110219 (2020).

29. Villez, K. *et al.* Combining multiway principal component analysis (MPCA) and clustering for efficient data mining of historical data sets of SBR processes. *Water Sci Technol* **57**, 1659–1666 (2008).

30. Sobczyk, M., Pajdak-Stós, A., Fiałkowska, E., Sobczyk, Ł. & Fyda, J. Multivariate analysis of activated sludge community in full-scale wastewater treatment plants. *Environ Sci Pollut Res* **28**, 3579–3589 (2021).

31. Cui, F. *et al.* Application of principal component analysis (PCA) to the assessment of parameter correlations in the partial-nitrification process using aerobic granular sludge. *Journal of Environmental Management* **288**, 112408 (2021).

32. Çimen Mesutoğlu, Ö. & Gök, O. Prediction of COD in industrial wastewater treatment plant using an artificial neural network. *Sci Rep* **14**, 13750 (2024).

33. Liu, T., Zhang, H., Wu, J., Liu, W. & Fang, Y. Wastewater treatment process enhancement based on multi-objective optimization and interpretable machine learning. *Journal of Environmental Management* **364**, 121430 (2024).

34. Mathur, R. *et al.* Modeling of two-stage anaerobic onsite wastewater sanitation system to predict effluent soluble chemical oxygen demand through machine learning. *Sci Rep* **14**, 1835 (2024).

35. Zhang, Y. *et al.* Machine learning modeling for the prediction of phosphorus and nitrogen removal efficiency and screening of crucial microorganisms in wastewater treatment plants. *Science of The Total Environment* **907**, 167730 (2024).
36. Sheik, A. G. *et al.* Prediction of wastewater quality parameters using adaptive and machine learning models: A South African case study. *Journal of Water Process Engineering* **67**, 106185 (2024).
37. Szeląg, B., Zaborowska, E. & Mąkinia, J. An algorithm for selecting a machine learning method for predicting nitrous oxide emissions in municipal wastewater treatment plants. *Journal of Water Process Engineering* **54**, 103939 (2023).
38. Yaqub, M., Asif, H., Kim, S. & Lee, W. Modeling of a full-scale sewage treatment plant to predict the nutrient removal efficiency using a long short-term memory (LSTM) neural network. *Journal of Water Process Engineering* **37**, 101388 (2020).
39. Hu, J., Yin, R., Wang, J. & Ren, H. Data driven multiple objective optimization of AAO process towards wastewater effluent biological toxicity reduction. *npj Clean Water* **7**, 1–11 (2024).
40. Seshan, S., Poinapen, J., Zandvoort, M. H., van Lier, J. B. & Kapelan, Z. Forecasting nitrous oxide emissions from a full-scale wastewater treatment plant using LSTM-based deep learning models. *Water Research* **268**, 122754 (2025).
41. Xu, X. *et al.* Prediction of nitrous oxide emission of a municipal wastewater treatment plant using LSTM-based deep learning models. *Environ Sci Pollut Res* **31**, 2167–2186 (2024).
42. Li, K. *et al.* An Integrated First Principal and Deep Learning Approach for Modeling Nitrous Oxide Emissions from Wastewater Treatment Plants. *Environmental Science and Technology* **56**, 2816–2826 (2022).
43. Khalil, M., AlSayed, A., Liu, Y. & Vanrolleghem, P. A. An integrated feature selection and hyperparameter optimization algorithm for balanced machine learning models predicting N2O

emissions from wastewater treatment plants. *Journal of Water Process Engineering* **63**, 105512 (2024).

44. Ahn, J. H., Kwan, T. & Chandran, K. Comparison of Partial and Full Nitrification Processes Applied for Treating High-Strength Nitrogen Wastewaters: Microbial Ecology through Nitrous Oxide Production. *Environ. Sci. Technol.* **45**, 2734–2740 (2011).
45. Kim, Y. M., Park, H. & Chandran, K. Nitrification inhibition by hexavalent chromium Cr(VI)--Microbial ecology, gene expression and off-gas emissions. *Water Res* **92**, 254–261 (2016).
46. Brotto, A. C., Annavajhala, M. K. & Chandran, K. Metatranscriptomic Investigation of Adaptation in NO and N2O Production From a Lab-Scale Nitrification Process Upon Repeated Exposure to Anoxic–Aerobic Cycling. *Front Microbiol* **9**, 3012 (2018).
47. Li, B. & Wu, G. Effects of Sludge Retention Times on Nutrient Removal and Nitrous Oxide Emission in Biological Nutrient Removal Processes. *Int J Environ Res Public Health* **11**, 3553–3569 (2014).
48. Brotto, A. C. *et al.* Factors controlling nitrous oxide emissions from a full-scale activated sludge system in the tropics. *Environ Sci Pollut Res* **22**, 11840–11849 (2015).
49. Wang, A.-J. *et al.* Digital Twins for Wastewater Treatment: A Technical Review. *Engineering* **36**, 21–35 (2024).
50. Zhou, J. *et al.* Sub-period division strategies combined with multiway principle component analysis for fault diagnosis on sequence batch reactor of wastewater treatment process in paper mill. *Process Safety and Environmental Protection* **146**, 9–19 (2021).
51. AL-Kordy, S. U. & Khudair, D. B. H. Effluent quality assessment of sewage treatment plant using principal component analysis and cluster analysis. *Journal of Engineering* **27**, 79–95 (2021).

52. Liu, H., Yang, J., Zhang, Y. & Yang, C. Monitoring of wastewater treatment processes using dynamic concurrent kernel partial least squares. *Process Safety and Environmental Protection* **147**, 274–282 (2021).
53. Yang, C., Zhang, Y., Huang, M. & Liu, H. Adaptive dynamic prediction of effluent quality in wastewater treatment processes using partial least squares embedded with relevance vector machine. *Journal of Cleaner Production* **314**, 128076 (2021).
54. Chandran, K., Pagilla, K. & Katehis, D. Field Protocol With Quality Assurance Plan - Characterization of Nitrogen Greenhouse Gas Emissions From Wastewater Treatment Bnr Operations. 1–47 (2009).
55. Chandran, K. *Protocol for the Measurement of Nitrous Oxide Fluxes from Biological Wastewater Treatment Plants*. *Methods in Enzymology* vol. 486 (Elsevier Inc., 2011).
56. Gruber, W. *et al.* Estimation of countrywide N2O emissions from wastewater treatment in Switzerland using long-term monitoring data. *Water Research X* **13**, (2021).
57. Andrew D. Eaton, American Water Works Association, Water Environment Federation. *Standard Methods for the Examination of Water and Wastewater*. (APHA-AWWA-WEF, Washington, D.C., 2005).
58. Stekhoven, D. J. & Bühlmann, P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* **28**, 112–118 (2012).

Tables and Figures

Table 1. Summary of N_2O emission fraction at a mainstream and sidestream process configuration in a Full-Scale BNR WWTPs

Plant configuration	Temp (°C)	Q (MGD)	Influent TKN load (g-N/day)	Removed TN load (g-N/day)	% Nitrogen removed	% influent TKN emitted as N_2O (EF)	Reference
Step feed BNR	18±3	21	1.6E+06	1.0E+06	62±21	0.20±0.03	This study
	23±3	23	1.6E+06	8.5E+05	56±16	0.03±0.006	
PdNA	18±1.7	15	1.8E+06	1.2E+06	68±7	0.006±0.00	
	25±2.8	12	2.1E+06	1.7E+06	77±8	0.005±0.00	
ABAC	26±2.2	23	1.4E+06	NA	NA	0.004±0.00	
Fermentate-fed partial denitratation	25±0.7	2.0	4.7E+01	3.3E+01	71±20	0.003±0.00	
SHARON Train 1	30±0.56	1.0	2.5E+06	9.5E+05	36±17	6.3±1.8	
SHARON Train 2		0.79	2.0E+06	7.0E+05	38±16	14±2.0	
Separate-stage BNR	15±0.48	23	1.8E+06	3.6E+05	20±1	0.03±0.00	Ahn.et al[4]
	23±0.28	27	2.3E+06	4.3E+05	19±4	0.01±0.00	
Four-stage Bardenpho	14±0.26	7.8	8.6E+05	1.7E+05	20±1	0.16±0.10	
	23±0.20	8.1	7.4E+05	7.6E+04	10±4	0.60±0.29	
Step-feed BNR 1	19±0.22	29	3.1E+06	1.4E+06	45±8	1.6±0.83	
	25±0.28	30	2.9E+06	9.4E+05	32±10	0.62±0.27	
Step-feed non-BNR	17±0.12	71	8.6E+06	4.4E+06	51	0.18±0.18	
	26±0.81	93	8.9E+06	4.2E+06	47	1.8±0.79	
Separate centrate	30±2.3	2.0	8.8E+06	5.5E+06	63	0.24±0.02	
	34±0.32	1.6	8.5E+06	4.2E+06	49	0.54±0.16	
Plug-flow 1	11±0.20	18	1.8E+06	1.0E+06	56	0.40±0.14	
	23±0.46	15	1.8E+06	7.3E+05	41	0.41±0.14	
Plug-flow 2	11±0.41	8.7	6.3E+05	4.0E+05	63	0.62±0.15	
	22±0.58	6.6	6.6E+05	4.0E+05	61	0.09±0.03	
MLE 1	26±1.8	4.0	6.8E+05	1.9E+05	28	0.07±0.04	
MLE 2	26±0.17	4.1	6.9E+05	1.5E+05	22	0.06±0.02	
Step-feed BNR 2	29±0.18	14	2.2E+06	2.9E+05	13	1.5±0.02	
Oxidation ditch	19±0.58	3.4	3.9E+05	4.3E+04	11±1	0.03±0.01	
Step-feed BNR 3	24±0.78	57	7.8E+06	8.6E+05	11±2	0.05±0.03	

Legend of Figures

Figure 1. PCA analysis on N_2O flux and wastewater quality parameters. A,B, N_2O flux (g-N d^{-1}-m^2) and their multivariate relationship between local and global variables in aerobic **(A)** and anoxic **(B)** zones.

Figure 2. Data collection and processing in ML with permutation importance score from the XG Boost algorithm. A, pipeline representing quantifying real-time N_2O emission data, N-profile and plant operational data, integrated into ML models provides mitigation strategies to reduce N_2O emissions along with clean water and better sanitation system. **B,C,** permutation importance and their contributing towards predicted N_2O flux in aerobic **(B)** and anoxic **(C)** zone

Figure 3. The Beeswarm SHAP analysis of predicted N_2O flux with wastewater parameters. A,B, Magnitudina and directionality insights from ML interpretable SHAP across both local and global variables in aerobic **(A)** and anoxic **(B)** zones predicted N_2O flux

Figure 4. 2D PDP on predicted N_2O flux with selective key features. Predicted N_2O flux and its dynamic relationship towards key process variables in PDP, across aerobic **(A,B)** and anoxic **(C,D)** zones. **(A)** predicted N_2O flux with NH_3 and SRT, **(B)** predicted N_2O flux with NH_3 and DO, **(C)** predicted N_2O flux with DO and SRT, **(D)** predicted N_2O flux with DO and NO_2^-

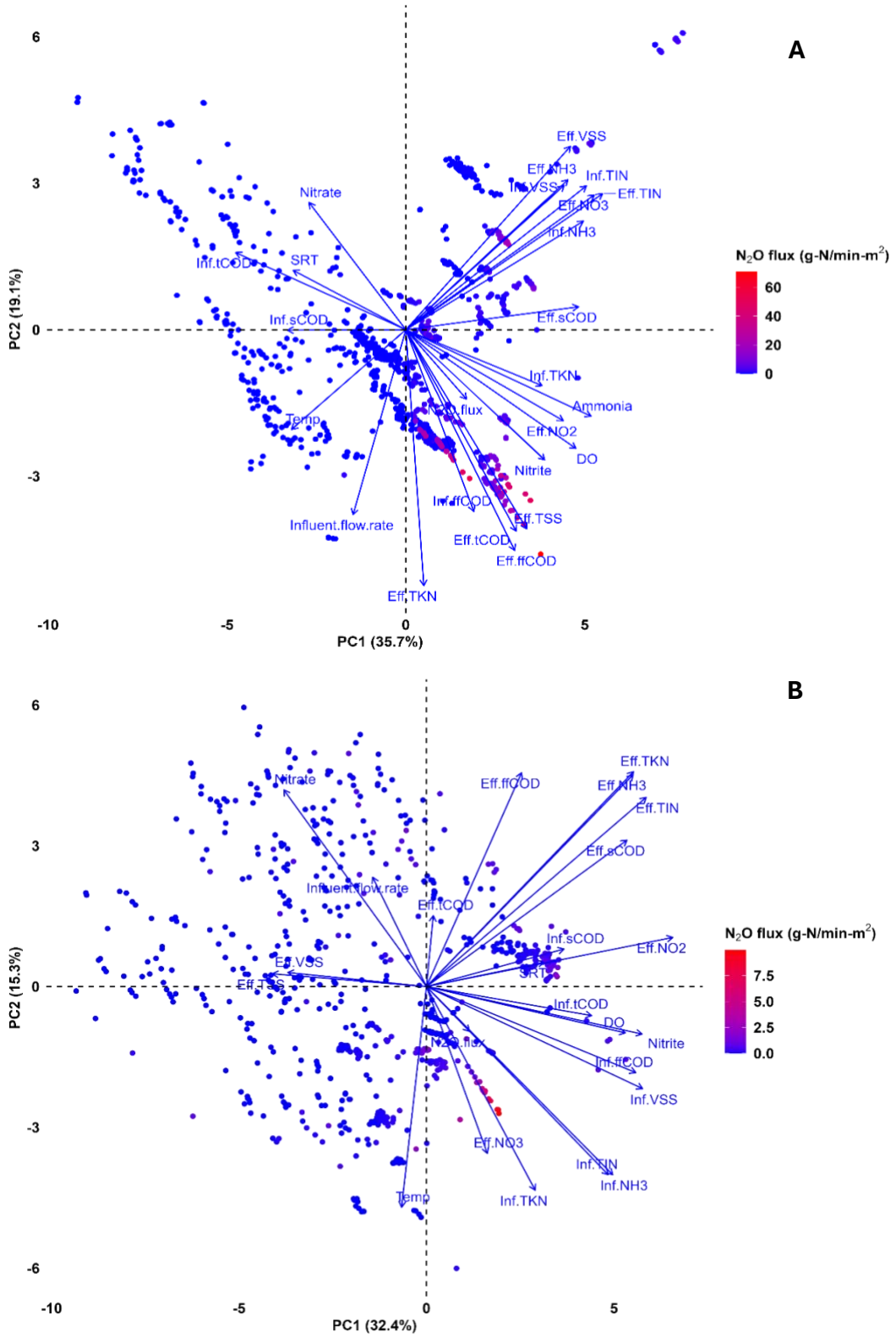

Figure 1. PCA analysis on N_2O flux and wastewater quality parameters. A,B, N_2O flux (g-N d^{-1}-m^2) and their multivariate relationship between local and global variables in aerobic **(A)** and anoxic **(B)** zones.

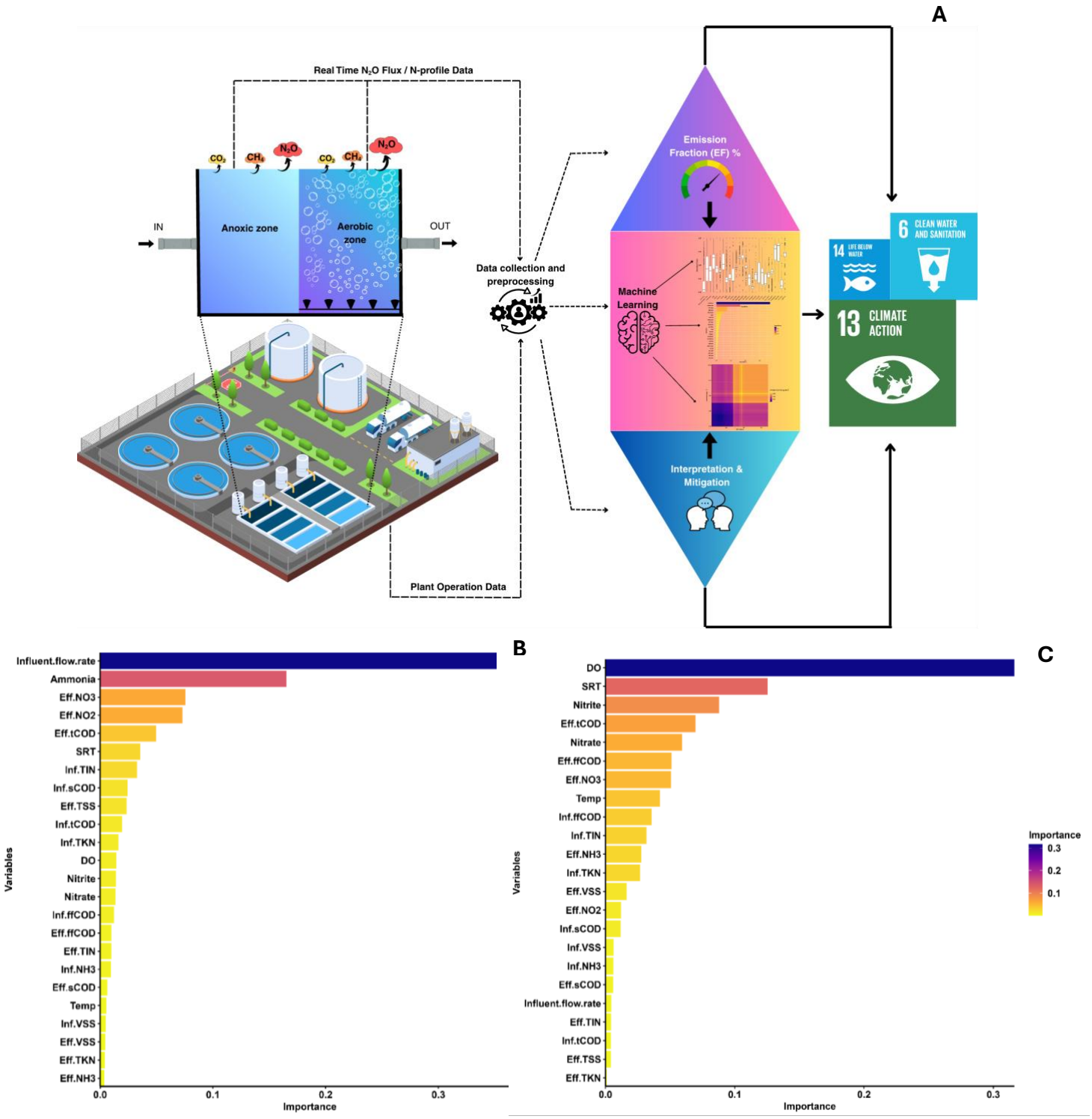

Figure 2. Data collection and processing in ML with permutation importance score from the XG Boost algorithm. A, pipeline representing quantifying real-time N_2O emission data, N-profile and plant operational data, integrated into ML models provides mitigation strategies to reduce N_2O emissions along with clean water and better sanitation system. **B,C,** permutation importance and their contributing towards N_2O flux in aerobic **(B)** and anoxic **(C)** zone

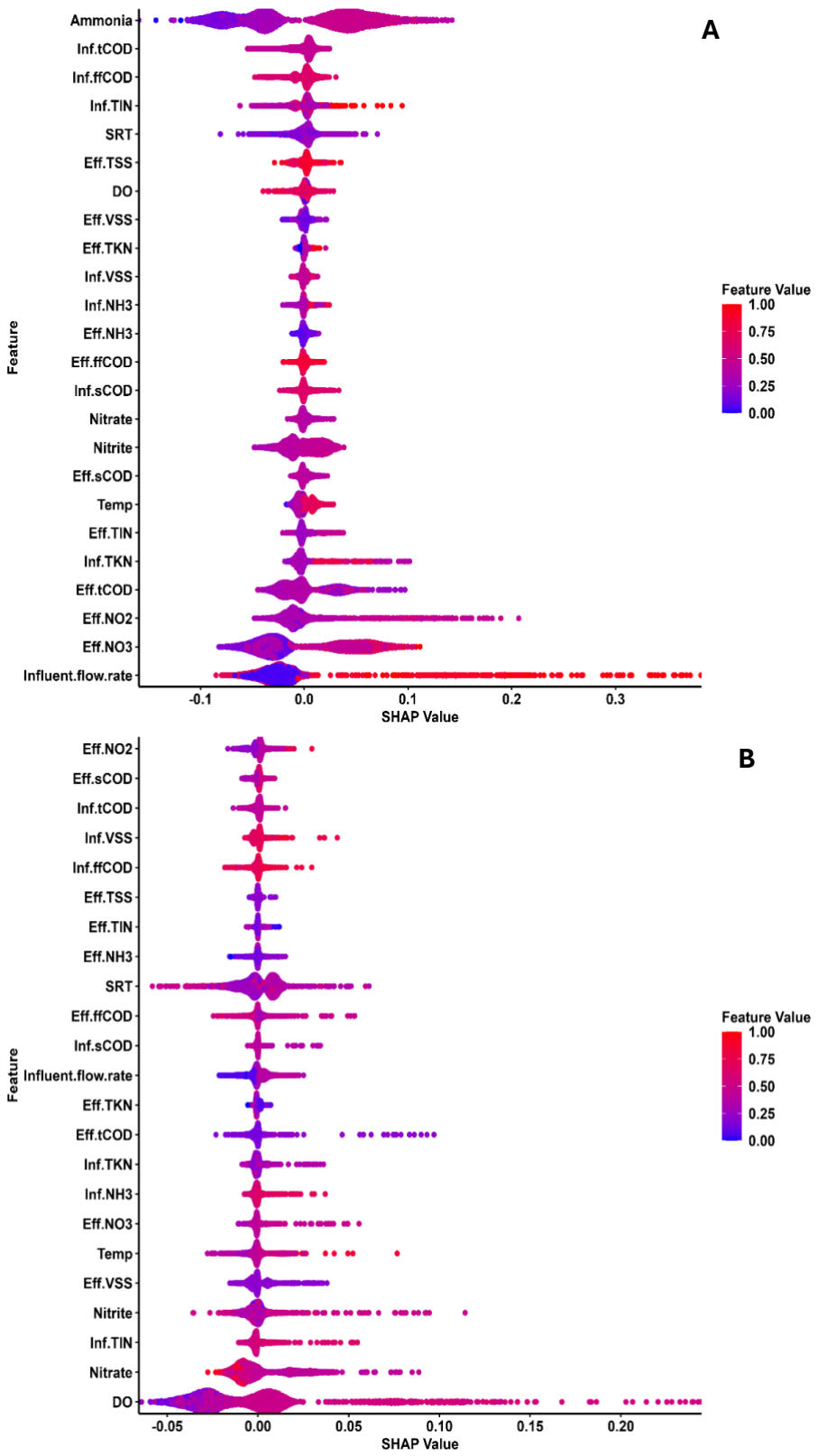

Figure 3. The Beeswarm SHAP analysis of predicted N_2O flux with wastewater parameters. A,B, Magnitudinal and directionality insights from ML interpretable SHAP across both local and global variables in aerobic **(A)** and anoxic **(B)** zones predicted N_2O flux .

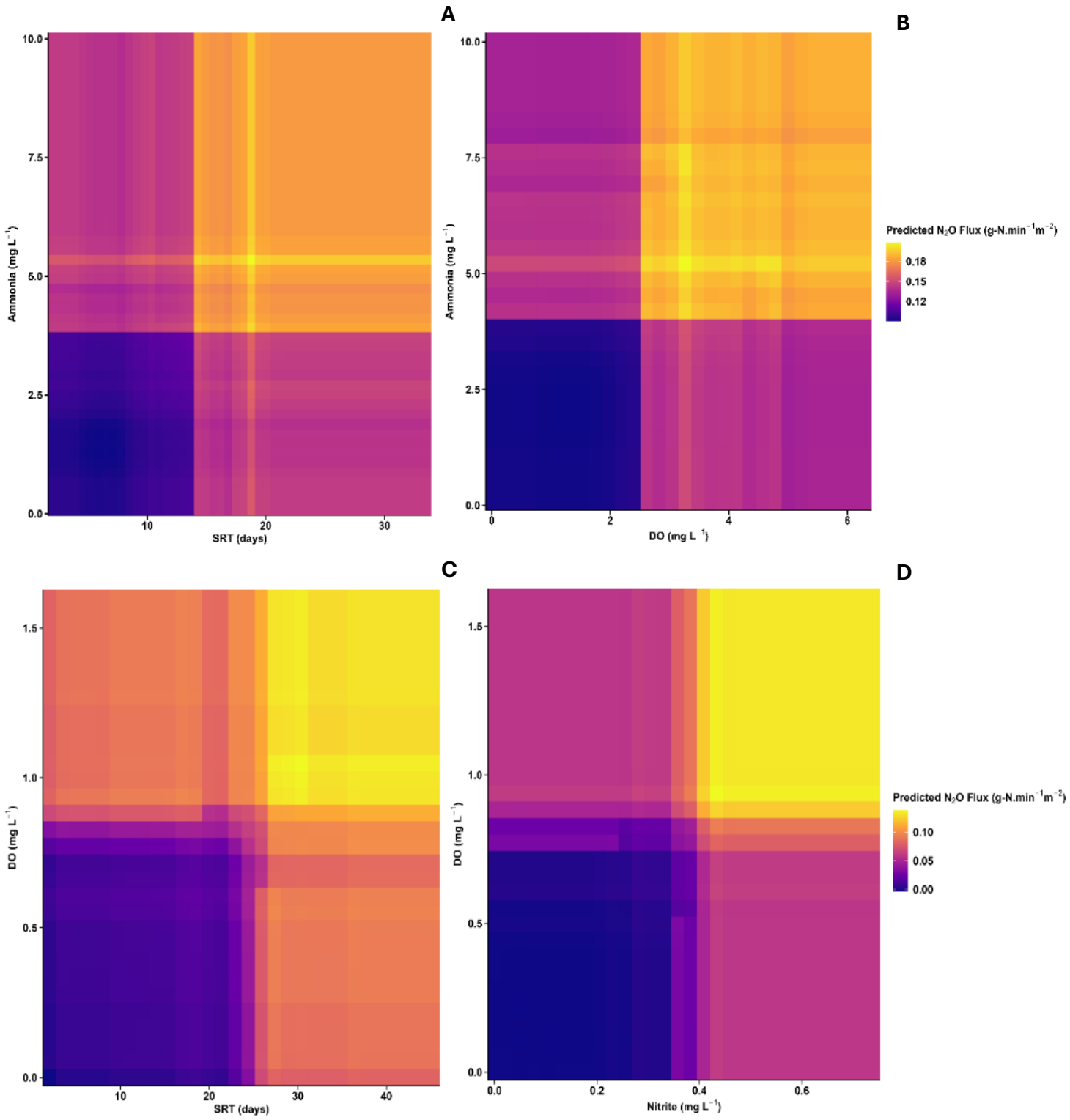

Figure 4. 2D PDP on predicted N_2O flux with selective key features. Predicted N_2O flux and its dynamic relationship towards key process variables in PDP, across aerobic **(A,B)** and anoxic **(C,D)** zones. **(A)** predicted N_2O flux with NH_3 and SRT, **(B)** predicted N_2O flux with NH_3 and DO, **(C)** predicted N_2O flux with DO and SRT, **(D)** predicted N_2O flux with DO and NO_2^-