Water Science & Technology

IWA PUBLISHING

Water Science & Technology Vol 86 No 8, 1887 doi: 10.2166/wst.2022.315

Design methodologies to determine optimal staging of membrane-aerated biofilm reactors for mainstream treatment with anammox

Brett M. Wagner*, Glen T. Daigger and Nancy. G. Love

Department of Civil and Environmental Engineering, University of Michigan, 1351 Beal Avenue, Ann Arbor, MI 48109, USA
*Corresponding author. E-mail: brettwag@umich.edu

BMW, 0000-0003-2963-8326; GTD, 0000-0003-3857-8345; NGL, 0000-0002-9184-2451

ABSTRACT

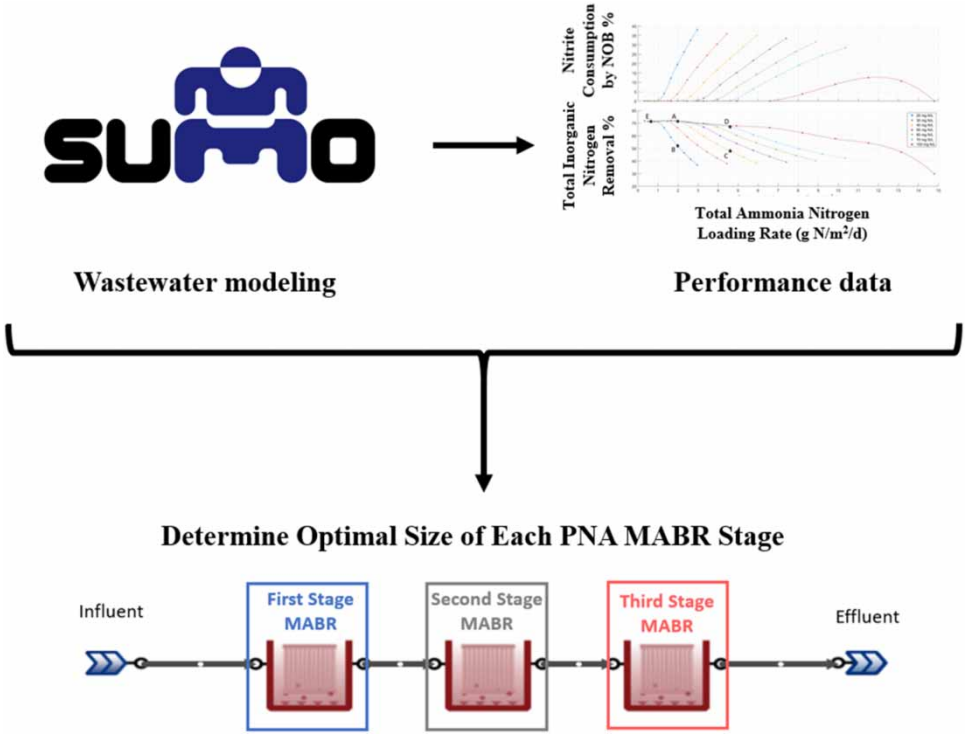
Partial nitritation anammox (PNA) membrane-aerated biofilm reactors (MABRs) can be used in mainstream nitrogen removal to help facilities reduce their energy consumption. Previous PNA MABR research has not investigated the impacts of staging, i.e. arraying MABRs in series, on their nitrogen removal performance, operation, and ability to suppress nitrite oxidizing bacteria. In this paper, a mathematical model simulated PNA MABR performance at different influent total ammonia concentrations and loadings. A design methodology for staging PNA MABRs was created and found that the amount of membrane surface area is dependent upon the total ammonia-nitrogen concentration and loading, and the air loading to the membrane must be proportional to the total ammonia-nitrogen loading to maximize the total inorganic nitrogen (TIN) removal rate. This led to approximately equal-sized stages that each had a TIN removal percentage of 71% of the influent total ammonia nitrogen. Staging a treatment train resulted in 9.8% larger total ammonia and 9.3% larger total nitrogen removal rates when compared with an un-staged reactor. The un-staged reactor also was not able to produce an effluent total ammonia concentration below 5 mg N/L which would be necessary for many facilities' permits.

Key words: membrane-aerated biofilm reactor (MABR), partial nitritation anammox, staging

HIGHLIGHTS

- Staging PNA MABRs requires a methodology that suppresses NOB and prevents O_2 inhibition of *AMX*.
- The size of each stage depends on the NH_x-N loading, NH_x-N concentration, and the membrane air loading to the reactor.
- This method will typically result in equal-sized stages.
- Staging a treatment train increases the NH_x-N removal rate by 10% and the total nitrogen removal rate by 9% when compared with an un-staged reactor.

GRAPHICAL ABSTRACT



INTRODUCTION

Membrane-aerated biofilm reactors (MABRs) incorporating partial nitritation plus anammox (PNA) can potentially perform resource efficient, carbon-free nitrogen removal in mainstream wastewater treatment. The anammox process relies on the oxidation of a fraction of the ammonia to nitrite (partial nitritation) through aerobic ammonia oxidizing organisms (AOO) and the anaerobic oxidation of the remaining ammonia with the produced nitrite to nitrogen gas through anammox bacteria (AMX). This combined PNA process is resource efficient when compared with traditional nitrification and heterotrophic denitrification, as it requires approximately 60% less oxygen and eliminates the need for exogenous organic carbon (Strous *et al.* 1998). The reduction in oxygen directly correlates to energy savings. Furthermore, since anammox is an autotrophic process, the organic carbon in the wastewater can be captured in an upstream high-rate process and used in resource recovery, such as energy production through anaerobic digestion.

MABRs, commercialized in the USA in 2016, present a new opportunity for water resource recovery facilities (WRRFs) to significantly decrease their energy consumption. MABRs consist of gas-permeable membranes that provide oxygen to biofilms growing on the membranes' exterior surfaces. These membranes provide oxygen directly to the biofilm and are up to four times more energy efficient than traditional fine bubble aerators (Peeters *et al.* 2017). The biofilm found in an MABR is counter-diffusional with oxygen diffusing through the membrane and other essential substrates diffusing into the biofilm from the bulk liquid. This allows the aeration of the biofilm to be controlled independently of the bulk liquid conditions and enables MABR operators to intentionally create aerobic and anaerobic regions within the biofilm (Terada *et al.* 2007). Therefore, integrating the PNA process via the MABR technology in mainstream nitrogen removal makes it possible to reduce a WRRF's energy consumption by up to 76% and allows the WRRF to further increase its energy generation ability through anaerobic digestion of the diverted organic carbon.

To transition PNA MABRs to full scale, research is needed to guide the proper design of PNA MABRs in series, referred to here as a PNA MABR treatment train. To create a treatment train, reactors are staged to increase the overall substrate removal rate for the entire treatment train while simultaneously achieving lower effluent substrate concentrations that are not possible in an un-staged reactor. The increase in substrate removal rate can be explained by differences in the reactors' bulk liquid substrate concentrations which control the flux of substrate into their biomass and the resulting substrate removal rate.

With this setup, the bulk liquid substrate concentration in the first reactor will be higher than the concentration in subsequent reactors, so the first reactor will have the largest flux of substrate into its biofilm and the largest substrate removal rate. In the subsequent reactors, the substrate concentration will decrease until it has complied with its effluent objective, but the lower substrate concentrations will also decrease the substrate removal rates in these reactors. When examining the treatment train as a whole, the increase in the first stage's removal rate compensates for the decreased removal rates in the later reactors, resulting in a net increase in the substrate removal rate for the overall process when compared to an un-staged reactor. In addition to an increase in the overall removal rate, staged treatment trains can also achieve lower effluent concentrations than un-staged reactors. For PNA MABRs, this is especially advantageous as our previous modeling work showed that it is difficult to decrease the total inorganic nitrogen (TIN) concentration of an un-staged PNA MABR by increasing the membrane surface area and the air loading to the membrane because of the difficulty in suppressing nitrite oxidizing bacteria (NOB) (Wagner *et al.* 2022). Thus, staging will be essential for PNA MABRs to achieve low effluent substrate concentrations.

Staging can also improve process performance by decreasing the competition between heterotrophs and autotrophs in the biofilm. Heterotrophic bacteria will outcompete autotrophic AOO in MABRs for oxygen while the ratio of influent organic carbon to nitrogen (C:N) is greater than 7, and they will outcompete AMX for nitrite at C:N ratios greater than 2. In a treatment train with a C:N ratio greater than 2, the first reactor will remove most of the organic carbon, since it has the largest organic carbon concentration. This will decrease the C:N ratio in subsequent stages and enable slower growing autotrophic bacteria to grow (Matsumoto *et al.* 2007; Lackner *et al.* 2008). Despite these benefits, PNA MABRs in the final stages of a treatment train will have the lowest effluent ammonia concentration, as ammonia is consumed in the previous stages. This ammonia concentration may not be sufficient to maintain an ammonia concentration large enough for stable and reliable NOB suppression, so the impacts of staging on PNA MABRs must be investigated.

Effective PNA implementation relies on successful suppression of NOB activity relative to AOO activity. NOB aerobically metabolize nitrite formed by AOOs to nitrate. Nitrate in turn, requires organic carbon for its reduction to nitrogen gas through heterotrophic denitrification, which decreases the energy and carbon efficiency of the process. Because NOB compete with AOO for oxygen and AMX for nitrite, achieving stable NOB suppression represents a key challenge to PNA process performance. Multiple operational strategies for NOB suppression have been investigated in MABRs including high temperatures (Gong *et al.* 2007, 2008; Pellicer-Nàcher *et al.* 2010; Sun *et al.* 2010; Terada *et al.* 2010; Pellicer-Nàcher & Smets 2012; Liu *et al.* 2016; Ribeiro Augusto *et al.* 2018), sequential aeration (Pellicer-Nàcher *et al.* 2010; Bunse *et al.* 2020), and bioaugmentation (Pellicer-Nàcher *et al.* 2010; Sun *et al.* 2010). Many of these strategies result in a residual ammonia concentration large enough that the AOO and AMX growth rates are large enough to outcompete NOB for their shared substrates, oxygen for AOO and nitrite for AMX. Consequently, PNA reactors must be operated to have a residual ammonia concentration that does not limit AOO and AMX growth rates, and they must be loaded with enough oxygen to allow AOO to produce nitrite for AMX but not enough for NOB to produce nitrate or for AMX to be inhibited by oxygen. The importance of using a residual ammonia concentration to prevent the limitation of AOO growth rates and to suppress NOB in partial nitritation reactors has been well documented in suspended growth reactors (Wett *et al.* 2013; Pérez *et al.* 2014, 2015; Regmi *et al.* 2014; Isanta *et al.* 2015; Lotti *et al.* 2015; Poot *et al.* 2016; Bekele *et al.* 2020). Furthermore, control processes using a residual ammonia concentration create oxygen-limiting conditions which prevent oxygen inhibition of AMX. This is essential for a PNA process where the aerobic and anaerobic metabolisms are housed within one reactor. The ammonia residual required to prevent oxygen inhibition of AMX and to suppress NOB is dependent on the amount of oxygen supplied to the process and is, therefore, widely varied in literature, with bulk liquid ammonia concentrations ranging from 2–5 mg N/L (Bekele *et al.* 2020) to 32 mg N/L (Isanta *et al.* 2015).

The residual ammonia concentration within the biofilm of a PNA MABR has not been widely studied because measuring the biofilm's ammonia concentration is difficult and intrusive. Instead, a reasonable approach involves measuring the bulk liquid ammonia concentration and calculating its diffusion throughout the biofilm depth. However, this is not typically done for process control. Instead, controlling the amount of air loaded to the membrane proportionally to the influent ammonia-nitrogen loading achieves a residual ammonia concentration in the biofilm and is common for PNA MABRs (Terada *et al.* 2007; Lackner *et al.* 2008; Pellicer-Nàcher *et al.* 2010; Sun *et al.* 2010; Gilmore *et al.* 2013; Acevedo Alonso & Lackner 2019; Bunse *et al.* 2020; Wagner *et al.* 2022). This control strategy has been compared with the theoretical oxygen demand required to oxidize 1 g of ammonia nitrogen, based on AMX and AOO (PNA) stoichiometry, which equals 1.75 g O_2/g N (Strous *et al.* 1998) to ensure that the biofilm is oxygen limited and has a sufficient residual ammonia concentration. The recommended ratio of oxygen to nitrogen loadings has been as low as 1 g O_2/g N (Wagner 2022), slightly lower than 1.75 (Terada *et al.*

2007; Acevedo Lackner *et al.* 2008; Alonso & Lackner 2019), slightly higher 1.75, (Pellicer-Nàcher *et al.* 2010; Gilmore *et al.* 2013), and as high as 3–5 g O_2/g N (Bunse *et al.* 2020). The wide range of ideal ratios from the modeling and experimental studies can be explained by differences in their chemical oxygen demand (COD) loadings and by a hypothesis proposed in (Sun *et al.* 2010) which states that this ratio can be highly impacted by the biokinetics of the biofilm and reactor configuration. These studies also varied in biofilm thicknesses which, when fixed in a model or in an experimental reactor with air scouring, can limit the ability of a biofilm to adjust its aerobic and anaerobic bacterial population abundances to the varied nitrogen and oxygen loadings. For example, if the oxygen loading to a thin 100 μm biofilm was adjusted to treat a large nitrogen load, the proportionally large oxygen load may cause the biofilm to be fully penetrated with oxygen, preventing the formation of an anaerobic region for AMX. Thus, the ratio of oxygen and nitrogen loadings must be viewed as an initial starting point in PNA MABR operation to ensure that AOO and AMX have a residual ammonia concentration, enough air is supplied to AOO to produce nitrite, and that the air loading is not too high to allow NOB growth or AMX inhibition by oxygen. The ratio can be further modified as operation of the reactor continues and effluent quality is analysed.

This study investigated the design and operation of PNA MABRs treating mainstream wastewater at various nitrogen loading rates and concentrations. In objective 1, a mathematical model of the PNA MABR process was used to generate performance data required for this investigation because the number of cases that needed to be examined greatly exceed the capacity to experimentally produce them. The PNA MABR model can be viewed as a digital twin for the subject process. Given the performance of the MABRs under these conditions, a design methodology and staging process is proposed and calculated in objective 2. This methodology is unique for PNA MABRs because design methodologies used for other biofilm-based reactors, which size a reactor based off the relationship between the substrate loading rate and the removal rate (Grady *et al.* 2011), resulted in poor inorganic total nitrogen removal from excessive NOB growth (data not shown). In objective 3, the staging process was tested against a hypothetical distribution of membrane surface area in three staged reactors to confirm its validity. Finally in objective 4, the impact of operational parameters such as the ratio of oxygen and nitrogen loadings and the exogenous COD loading on the staging design are presented.

METHODS

Objective 1: Create a PNA MABR model to quantify total inorganic nitrogen removal performance for different ammonia concentrations and loadings

A mathematical model created in an established wastewater modeling software, SUMO 19.2 (Dynamita 2019), was used to simulate a PNA MABR. The SUMO software combines a one-dimensional biofilm model, a biokinetic model, and various reactor models to create a complete wastewater modeling package.

The multispecies one-dimensional biofilm model in SUMO applies user inputs of the biofilm specific mass, density, thickness, and number of layers to set a fixed biofilm thickness that is divided into distinct layers. In the biofilm model, each biofilm layer acts as a completely mixed reactor. Substrates diffuse from the bulk liquid through the user specified hydrodynamic boundary layer into the outermost biofilm layer. The substrates then diffuse into adjacent biofilm layers at half the rate of diffusion that would occur in the bulk liquid. Oxygen substrate from the membrane in the MABR is supplied to the innermost biofilm layer at a fixed loading rate which diffuses outward towards the bulk liquid. Particulate solids within each layer transfer between the biofilm layers at a specified rate. Finally, attachment and detachment occur at fixed, specified rates in the outermost biofilm layer.

The SUMO2 biokinetic model library with two-step nitrification and denitrification was used, as it was the only library that included AMX. The library accounted for AOO, NOB, AMX, ordinary heterotrophic organisms, glycogen accumulating organisms, phosphorus accumulating organisms, anoxic methanol utilizers, acetoclastic methanogens, and hydrogenotrophic methanogens. Given the conditions analysed in this study, glycogen accumulating organisms, phosphorus accumulating organisms, anoxic methanol utilizers, acetoclastic methanogens, and hydrogenotrophic methanogens were quickly outcompeted and did not significantly affect the results. A temperature of 20°C was used, so temperature corrections were unnecessary. Half saturation coefficients of organisms within the biofilm were reduced by a factor of 0.1 to account for the fact that diffusion limitations are explicitly incorporated into the biofilm model. The model did not consider the effects of pH, metabolic intermediates for nitrogen removal beyond nitrite, and variations in biochemical kinetic parameters. A table listing all SUMO default parameters, including biokinetic and stoichiometric parameters, is found in the Supplemental Information.

The simulated influent wastewater to the MABR was assumed to come from the effluent of an upstream high-rate bioabsorption process that captured particulate and colloidal carbon in the wastewater. The MABR influent had different flow rates, influent total ammonia concentrations (representing the sum of ammonia and ammonium nitrogen with no organic nitrogen), and readily biodegradable COD concentrations depending on the model objective listed below. The biofilm was modeled with five layers with a total thickness of 700 μm, and a hydraulic boundary layer thickness of 100 μm was used. The reactor size and flow parameters are modeled after a laboratory-scale reactor, but removal rates and influent fluxes are normalized to the membrane surface area to make them scale independent. Simulations were generally run for 360 days with a calculation step size of 1 day to reach steady-state conditions, defined here as the moment when the daily difference in the effluent total ammonia/ammonium, nitrite, and nitrate was less than 0.001 mg N/L, but some simulations needed additional time to reach a steady-state effluent concentration. A list of all modified SUMO parameters is provided in Table S-1.

Objective 2: Create a design methodology for a PNA MABR without influent exogenous COD and propose a staging process for a PNA MABR treatment train

The methodology for designing the reactor staging was established by simulating a 2.5 L MABR with varied total ammonia fluxes for influent total ammonia concentrations ranging from 10–70 mg N/L and 100 mg N/L, which represents a high nitrogen concentration mainstream wastewater scenario. The influent nitrogen flux was modified for each influent total ammonia concentration by altering the influent flow rate from 1 to 9 L/d which corresponded to influent nitrogen fluxes of 0.3–14.8 g $N/m^2/d$, and the membrane surface area was kept constant at 0.06 m^2. The air loading to the membrane was initially kept constant regardless of the influent total ammonia-nitrogen loading, but this created excessive oxygen-limited conditions and restricted the TIN removal rate (Figure S-1). The membrane air loading was then modified to a 1:1 ratio of oxygen to influent total ammonia which corresponded to oxygen fluxes of 0.3–14.8 g $O_2/m^2/d$ to create an oxygen-limited biofilm that balanced AOO and AMX activity. The TIN removal percentage was then calculated for each simulation. It is important to note that, since the membrane air loading was proportional to the influent total ammonia-nitrogen loading, increasing the nitrogen flux led to a corresponding increase in the membrane air loading.

Using the curves created for the diverse total ammonia fluxes, total ammonia concentrations, and TIN removal rates, a process train can be designed when given an influent flow rate, total ammonia concentration, and a desired effluent concentration. In a real-world design process with fixed flows and concentrations, the membrane surface area and the oxygen loading to the membrane in each stage are critical parameters that an engineer would modify in designing a staged MABR.

To demonstrate the staging methodology that is described later in the results section, an example treatment train was designed and calculated, with the requirements and design process presented here for clarity. The example treatment train was required to treat an influent flow rate of 5 L/d with a total ammonia concentration of 50 mg N/L. A total nitrogen effluent concentration below 5 mg N/L and a total ammonia concentration below 1 mg N/L were selected as the design effluent quality objectives to simulate the effluent requirements in a strict discharge permit. For the design, three reactors in series with a combined membrane surface area of 0.29 m^2 were required to reach the desired effluent limits. A TIN removal of 71% was chosen for the first two reactors, since it was the largest removal percentage for every influent nitrogen concentration tested. The membrane surface area required in those stages was determined by finding the largest influent total ammonia loading rate for the given influent total ammonia concentration for which the TIN removal was 71%. This occurred before a precipitous drop in the TIN removal percentage and maximized the mass of nitrogen removed in each stage. With the influent total ammonia loading rate determined, the membrane surface area was calculated by dividing that loading rate by the influent flow rate and influent total ammonia-nitrogen concentration. For the last stage, a smaller TIN removal percentage of 69% was used since it used less membrane surface area while still meeting the required effluent TIN and total ammonia concentration limits. The oxygen loading to each reactor was calculated at a 1:1 ratio to the influent ammonia-nitrogen loading. This design was then simulated in SUMO in three 0.83 L MABR units (2.5 L total) using the influent flow rate, total ammonia concentration, and the calculated membrane surface areas and air loadings for each reactor. The simulated effluent of each reactor was then compared to the calculated values from the design methodology to evaluate its validity.

Objective 3: Test the staging design against a diverse distribution of membrane surface areas in three staged reactors

To determine if the calculated design produced the lowest effluent total nitrogen concentration for a given quantity of MABR units, i.e. if the MABRs were being used most effectively, the same total membrane surface area as the calculated design in

objective 2 (0.29 m^2) was distributed across three reactors in different proportions, as shown in Figure 1. Each reactor in series had its oxygen loading modified from 0.1 to 4.1 g $O_2/m^2/d$ to minimize the effluent total nitrogen concentration and was not controlled proportionally to the influent total ammonia concentration to verify the design air loading used in the design methodology. The volume of each reactor was kept constant at 0.83 L (2.5 L total), regardless of the membrane surface area, to ensure that the hydraulic residence time (HRT) for the process was equivalent to the HRT of the reactor that produced the original design. The TIN removal percentage and the ratio of influent oxygen and total ammonia were then calculated for each set of surface area distributions.

Objective 4: Determine how exogenous COD affects the design and operation

The previous design methodology and operation optimization did not consider the impacts of exogenous COD in the influent. To ascertain how the TIN removal percentage and methodology is impacted by exogenous COD, the simulations conducted in objective 1 with different total ammonia fluxes for influent total ammonia concentrations ranging from 10–70 mg N/L and 100 mg N/L were re-simulated with readily biodegradable COD (rbCOD) concentrations of 25, 50, and 100 mg COD/L added to the influent. The membrane air loading was increased by ratios of 0.5, and 1.0 g O_2/g COD to adapt to the increased oxygen demand. The TIN and rbCOD removal rates were then calculated for every simulation. The simulations with the largest TIN removal rate for a given influent rbCOD concentration, total ammonia concentration, and air loading were kept for analysis, and the other simulations were discarded, as they resulted from non-optimal operation of the MABR's air loading. Additional simulations were discarded if they produced effluent nitrate concentrations greater than 1 mg N/L, which resulted from excessive NOB activity without significant heterotrophic denitrification. To analyse if exogenous COD affected the TIN removal rate, simulations were binned based on their influent rbCOD to influent total ammonia loadings, and the TIN removal rate was averaged in each bin to simplify and illustrate observed performance trends. In the 0.0 C:N bin, only

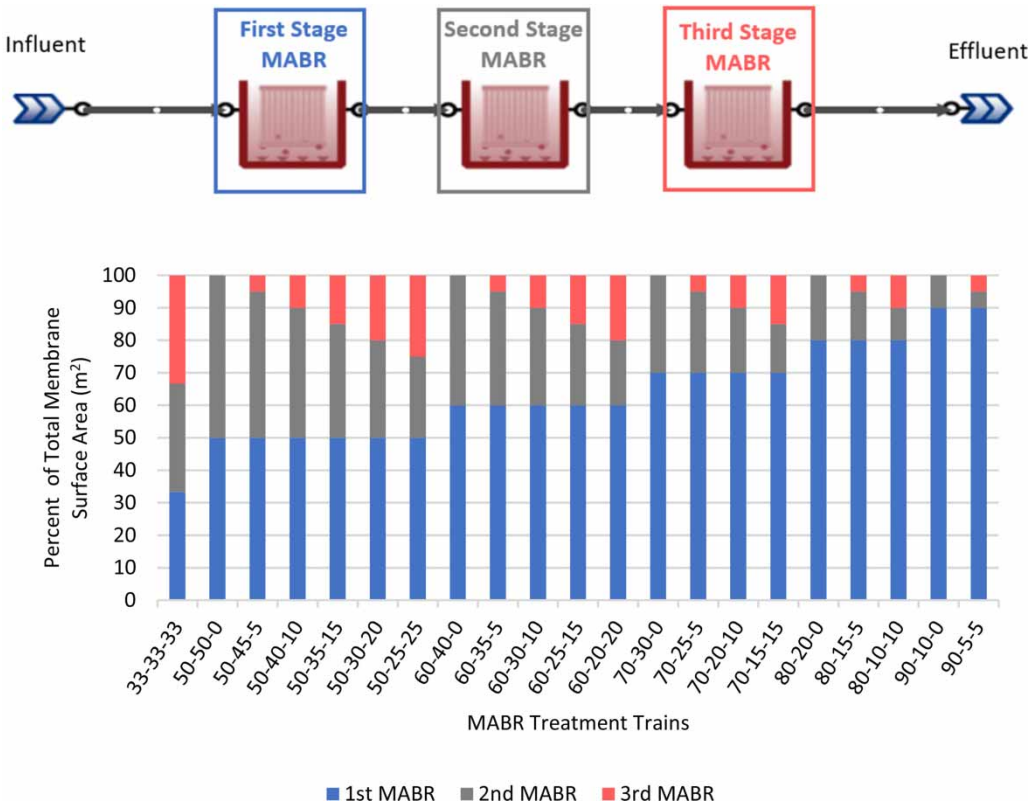


Figure 1 | Schematic of MABR treatment train (top) and the distribution of MABR surface areas in the tested treatment trains (bottom). The values listed in the x-axis on the bottom figure correspond to the percentage of total MABR surface area in the first, second, and third reactors, respectively, in an MABR treatment train.

simulations with NOB consumption of nitrite less than 1% of the total nitrite consumption were included, and other simulations were discarded.

Similarly, the impact of exogenous COD on the optimal operation of staged PNA MABRs must be considered. To quantify this, the treatment train created from the design methodology in objective 2 with an influent flow rate of 5 L/d, a total ammonia concentration of 50 mg N/L, and a total membrane surface of 0.29 m^2 distributed into three 0.83 L MABRs had 50 mg rbCOD/L added to its influent. An initial simulation was run without altering any operational parameters, but each reactor was under-loaded with oxygen, resulting in poor total nitrogen and COD removal rates (data not shown). The membrane air loading in each reactor was then increased by 0.5 g O_2/g COD of the influent COD to each reactor in the initial simulation to adapt to the increased oxygen demand, and the treatment train was modeled in SUMO. The results of this simulation were then compared with the results without exogenous COD.

RESULTS AND DISCUSSION

The main objective of this paper was to create a design methodology for staged PNA MABRs. This methodology is necessary, as design procedures used for other biofilm-based reactors resulted in poor inorganic total nitrogen removal, described later. This is due to the fact that traditional biofilm reactors are simpler than PNA MABRs since they typically rely on conventional oxidation metabolisms such as nitrification and COD oxidation rather than partial nitritation and anammox which require NOB to be suppressed. The relationship between the substrate loading rate and the surficial removal rate in traditional biofilm reactors is also not dependent on the concentrations of ammonia and COD because reactor performance is less impacted by competition for shared substrates between different metabolisms. For example in a PNA reactor, nitrite consumption by NOB will decrease the nitrogen removal rate, but in a traditional nitrification–heterotrophic denitrification reactor, nitrite consumption by denitrifying heterotrophs will decrease the carbon and oxygen requirement and improve the resource efficiency of the process.

Design of PNA MABRs is especially complex, as operators must balance the trade-off between total ammonia and total nitrogen removal while suppressing NOBs (Bunse *et al.* 2020). The air loading to the membrane needs to be adjusted according to the influent total ammonia loading to guarantee that the biofilm is oxygen limited which limits NOB growth and oxygen inhibition of AMX. However, adjusting air loading to the membrane according to the influent total ammonia loading is not sufficient by itself. For example, at large nitrogen loadings, the air loading to the membrane can be large enough that NOB begin to outcompete AMX for nitrite, resulting in nitrate production. Thus, appropriate methodologies must size a stage that limits the total ammonia loading, and therefore the air loading, to provide enough oxygen to the membrane to balance nitrite production by AOO but hinder nitrate production by NOB and oxygen inhibition of AMX. The methodology must also consider the concentration of ammonia, as the ammonia concentration controls the specific growth rates of AOO and AMX which are essential for NOB suppression. Before creating the methodology, the performance of a PNA MABR must be established under a variety of conditions.

Objective 1: NOB suppression in a PNA MABR is dependent on the influent total ammonia concentration, the influent total ammonia loading, and the air loading to the reactor

By simulating a PNA MABR under different influent flow rates and total ammonia concentrations, the performance of the reactor can be demonstrated under numerous nitrogen loadings and concentrations which is essential for determining how many reactors are needed to treat a given wastewater flow. From these simulations, we observed that increasing the flow rate to the MABR decreased its HRT and increased the nitrogen loading rate for each influent total ammonia concentration tested. As the nitrogen loading rate increased for every influent ammonia-nitrogen concentration, the TIN removal percentage decreased and never increased, Figure 2 (bottom). This decrease in TIN removal percentage was abrupt for lower influent total ammonia concentrations, occurring at loading rates less than 3 g N/m^2/d for influent ammonia concentration less than 50 mg N/L, and more gradual for larger influent concentrations. When comparing the reactor's performance against the stoichiometric nitrogen removal for AOO and AMX, the TIN removal percentage never reached the theoretical removal percentage of 89% for any nitrogen loading rate or influent concentration because the biofilm was loaded with oxygen below the stoichiometric oxygen demand for PNA, 1.75 g O_2/g N (Strous *et al.* 1998), to create oxygen-limiting conditions to prevent NOB growth. The effluent TIN concentration also increased with increased nitrogen loading rates (Figure S-2). These findings are unsurprising and can be observed in other biofilm-based reactors such as rotating biological contactors and packed towers.

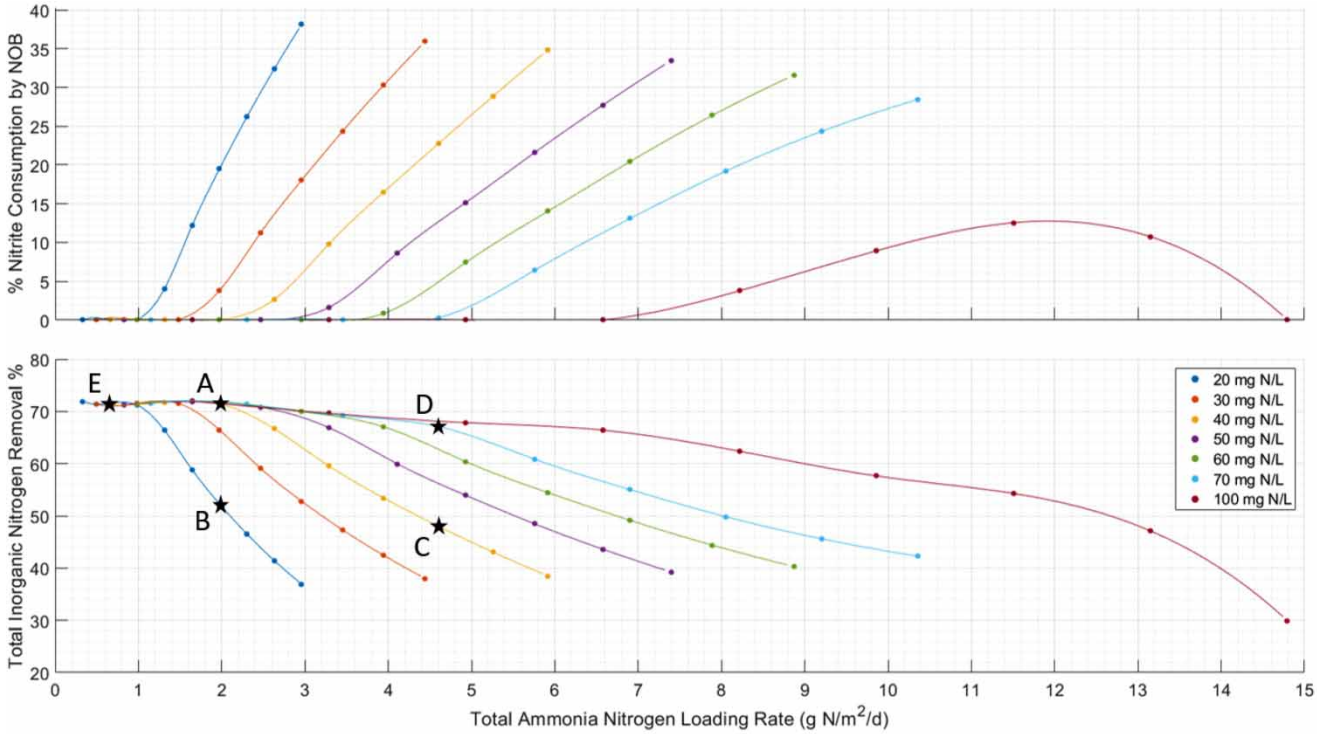


Figure 2 | Comparison of TIN removal percentages for varied total ammonia loading rates and influent ammonia-nitrogen concentrations (bottom) and comparison of the percentage of nitrite consumed by NOB for varied total ammonia loading rates and influent ammonia-nitrogen concentrations (top). Starred points correspond to the scenarios described in Table 1. The air loading to the reactor in each figure was maintained at a 1:1 ratio of influent oxygen to total ammonia nitrogen.

Notably, Figure 2 (bottom) shows the impact of the influent total ammonia concentration loaded to the reactor. In MABRs, activity within the biofilm is dependent upon the substrate concentration in the bulk liquid and its flux into the biofilm. By comparing multiple influent total ammonia concentrations at a constant loading rate, it can be observed that the TIN removal percentage increases as the influent concentration increases. Using larger influent total ammonia concentrations elevates the specific growth rates of AOO and AMX within the biofilm which increases the nitrogen removal performance. This increased performance is importantly nonlinear and is dependent upon the influent total ammonia-nitrogen concentration and loading rate. For example, at a nitrogen loading rate of 2 g N/m^2/d, the TIN removal percentage increased from 52 to 66% when the influent total ammonia concentration increased from 20 to 30 mg N/L. However, when the influent total ammonia concentration increased from 30 to 40 mg N/L, the TIN removal percentage only increased from 66 to 71%, so an equivalent increase in the influent total ammonia concentration resulted in a smaller increase in the removal percentage. Because the TIN removal is nonlinear and dependent upon the total ammonia concentration, traditional design methods used for simpler metabolisms will not apply.

It is logical to think that a PNA MABR should be fed a large nitrogen loading rate to maximize the mass of nitrogen removed from the reactor. However, increasing the nitrogen loading rate would require an equivalent increase to the air loading to the reactor, due to the 1:1 ratio of oxygen to nitrogen loadings, to ensure that the additional nitrogen can be oxidized by AOO and AMX. This can lead to excessive NOB growth which can be observed in Figure 2 (top). The sharp increase in the nitrite consumption by NOB corresponds to the point at which the TIN removal percentage begins to significantly decrease for each influent total ammonia concentration in Figure 2 (bottom).

To summarize how the total ammonia loading, total ammonia concentration, oxygen loading, TIN removal percentage, and the NOB nitrite consumption impact the staging of PNA MABRs, five scenarios were identified in Figure 2 and are described in Table 1. Scenario A represents an ideal outcome that maximizes the TIN removal rate for an influent total ammonia concentration of 40 mg N/L. The total ammonia loading selected for this scenario was increased to the point before NOB would begin to consume nitrite which maximized the mass of nitrogen removed in the reactor at a TIN removal of 71%. Scenario B

Table 1 | Staging scenarios at different total ammonia loadings and concentrations

Scenario	NH_x loading rate ($g\ N/m^2/d$) NH_x concentration ($mg\ N/L$)	O_2 Loading ($g\ O_2/m^2/d$)	Concentration in the biofilm layer closest to the membrane			TIN removal % TIN removal rate ($g\ N/m^2/d$)	NO_2^- consumption by NOB %	Outcome
			O_2 ($mg\ O_2/L$)	NH_x ($mg\ N/L$)	NO_2^- ($mg\ N/L$)			
A	2.0 40	2.0	0.0050	1.4	0.15	71 1.4	0.023	Optimal amount of membrane surface area
B: Equivalent NH_x loading as Scenario A, but lower NH_x concentration	2.0 20	2.0	0.0057	0.29	0.24	52 1.0	20	Not enough membrane surface area: large NOB activity
C: Equivalent NH_x concentration as Scenario A, but larger NH_x loading	4.6 40	4.6	0.0064	0.34	2.9	48 2.2	23	Not enough membrane surface area: large NOB activity
D: Larger NH_x loading rate and concentration	4.6 70	4.6	0.0065	0.43	3.9	67 4.4	0.27	Acceptable amount of membrane surface area (given the large NH_X concentration), but AMX beginning to be outcompeted
E: Equivalent NH_x concentration as Scenario A, but smaller NH_x loading rate	0.66 40	0.66	0.0035	7.2	0.024	71 0.47	0.00	Too much membrane surface area: low removal rate

highlights the importance of the total ammonia concentration in staging. Although the reactor in this scenario has the same oxygen and total ammonia loading as Scenario A, its influent total ammonia concentration has been halved. This decreases the residual ammonia concentration in the bulk liquid which decreases the AOO and AMX activities, making NOB suppression more difficult and results in 20% nitrite consumption by NOB. Scenario C highlights the importance of the total ammonia loading rate in staging design. The reactor in this scenario has the same influent total ammonia concentration as Scenario A, but the total ammonia loading is increased. Since the oxygen loading is dependent on the influent total ammonia loading, Scenario C represents a scenario where the combined total ammonia and oxygen loadings are too large to suppress NOB for the given total ammonia concentration, resulting in a low TIN removal of 48% and a nitrite consumption by NOB of 23%. Scenario D highlights how large total ammonia concentrations and loadings affect staging. The reactor in this scenario has a TIN removal percentage of 67% which is 4% lower than the TIN removal in Scenario A, since AMX is beginning to be outcompeted for space in the biofilm. However, since its total ammonia concentration of 70 mg N/L is relatively high compared to the other scenarios, it can maintain a residual ammonia concentration and suppress NOB. This scenario emphasizes that total ammonia loadings greater than 4 g $N/m^2/d$ can only be treated by a reactor when its influent total ammonia concentration is greater than 70 mg N/L, if NOB suppression is desired. Finally, Scenario E highlights how a reactor can be oversized. The reactor in this scenario has the same influent ammonia concentration as Scenario A, but it has a 2.9 times smaller loading rate at equivalent TIN removal percentages. Thus, the TIN removal rate can be increased by increasing the total ammonia loading, without impacting NOB suppression.

To further examine how the total ammonia concentration impacts the NOB activity, the nitrite consumption by NOB was graphed versus the total ammonia concentration in the biofilm layer closest to the membrane in Figure S-3. In this layer, a

residual ammonia concentration within the layer of at least 0.5 mg N/L was found to suppress NOB growth, and when the concentration decreased below 0.5 mg N/L, the percentage of nitrite consumed by NOB significantly increased. This ammonia residual concentration is lower than the residual required for NOB suppression in the bulk liquid of suspended growth reactors. However, this concentration is averaged throughout the entire layer, so the concentration at the base of the membrane, where the oxygen is being supplied, will be larger than this average. Thus, in order to have optimal staging, each reactor must have a residual total ammonia concentration of at least 0.5 mg N/L in the biofilm layer closest to the membrane, regardless of the influent concentration.

Objective 2: The optimal amount of membrane surface area (the size of each stage) is determined by maximizing the total inorganic nitrogen removal percentage while minimizing NOB growth

Using the TIN removal rates from Figure 2 (bottom) and the NOB nitrite consumption data from Figure 2 (top), a staging design methodology was created. Various approaches were initially considered for the methodology such as using Figure S-4, which relates the TIN removal rate to the total ammonia loading rate similar to other biofilm-based reactor design methodologies, to select a total ammonia loading near its horizontal asymptote. This would maximize the TIN removal rate, minimize the amount of membrane surface area used in each stage, and would result in many small reactors in series (data not shown). However, loading the reactor with large ammonia loadings near their horizontal asymptotes in Figure S-4 would result in low TIN removal rates and large nitrite consumption by NOB, shown in Figure 2. Thus, a different methodology was required to maximize the TIN removal percentage and minimize the amount of membrane surface area.

The new methodology presented here was able to maximize the amount of nitrogen removed in each stage and minimize the amount of membrane surface area in each stage. To do this, the amount of membrane surface area was calculated by selecting the maximum TIN removal percentage achievable for most influent total ammonia nitrogen concentrations, 71%, and the largest nitrogen load for which the influent nitrogen concentration can achieve that TIN removal percentage. For the example design requirements listed in the methods (influent flow rate = 5 L/d, influent total ammonia concentration = 50 mg N/L, effluent total nitrogen effluent concentration below 5 mg N/L, and effluent total ammonia concentration below 1 mg N/L), three reactors were required. In a treatment scheme with different influent and effluent requirements than those in the example, the number of reactors can be determined by consecutively removing 71% of the TIN for the influent total ammonia flux in each stage until the desired effluent conditions are achieved. It should be stressed that this design method is formulated on the influent total ammonia load, not the influent TIN load. Similarly, the percentage of TIN removed is a percentage of the influent ammonia load. Thus, some nitrate will be generated and will accumulate in the subsequent reactors, as heterotrophic denitrification is minimal with no exogenous COD. This design methodology will typically result in equal amounts of membrane surface area in the beginning stages, with the last stage's size and membrane air loading modified to reach the desired effluent concentrations. Although the design methodology differs from that applied for simpler biological reactions, the result of having equally sized stages is seen in other biofilm reactors including rotating biological contactors (Grady *et al.* 2011). For this example, the first two reactors had a TIN removal percentage of 71%, and the last reactor had a smaller TIN removal percentage of 69% in order to use less membrane surface area while still meeting the effluent requirements. Sample calculations for the example treatment train are shown in the Supplemental Information under Figure S-5.

This design methodology did not include a safety factor in determining the amount of membrane surface area in each stage. Therefore when applying this design methodology to practice, each stage should have its membrane surface area increased and its influent total ammonia loading rate decreased. After implementing the safety factor, the new influent total ammonia loading rate for each stage on Figure 2 will be to the left of the point where the TIN removal significantly decreases and NOB begin to consume nitrite. This will decrease the surficial nitrogen removal rate in each reactor, but it will enable the staged treatment train to effectively treat wastewater flows with dynamic influent nitrogen loadings without enabling NOB growth.

The design methodology produced an example treatment train with 35.5% of its membrane surface area in each of the first two stages and 29% in the last stage. Effluent total nitrogen, total ammonia, nitrite, and nitrate concentrations for the designed treatment train are shown in Table 2. The design was calculated to have an effluent total nitrogen concentration of 3.5 mg N/L and a total ammonia concentration of 0.9 mg N/L. 76% of the total nitrogen calculated to be removed in the entire treatment train occurred in the first stage since it had the largest loading and largest effluent ammonia concentration, seen in Table 2, which promoted mass transfer into the biofilm and increased the specific growth rates of AOO and AMX. The first two stages were calculated to have a 71% TIN removal of their influent total ammonia, but the

Table 2 | Comparison of the calculated and modeled performance of the treatment train produced with the design methodology

	1st stage effluent		2nd stage effluent		3rd stage effluent	
	Calculated	SUMO	Calculated	SUMO	Calculated	SUMO
Total nitrogen mg N/L	15	15	5.6	5.5	3.5	3.7
Total ammonia mg N/L	13	12	3.1	3.0	0.93	0.86
Nitrite mg N/L	0.00	0.00	0.00	0.00	0.00	0.00
Nitrate mg N/L	2.0	2.2	2.5	2.4	2.5	2.8

second stage had a smaller mass removal due to the lower loading applied to the reactor and removed 19% of the train's total nitrogen. The last reactor had the smallest total ammonia loading and had an influent total ammonia concentration of only 3.1 mg N/L. Since the effluent total ammonia concentration was required to be below 1 mg N/L, the total ammonia concentration was the limiting factor for the design, and the last stage was sized just large enough to reach the effluent limit. As a result, the last stage only removed 5% of the train's total nitrogen loading and was used primarily for total ammonia oxidation.

The calculated performance of the treatment train using the design methodology paralleled the results from its performance when modeled in SUMO, Table 2. The total nitrogen concentrations differed only by 0.2 mg N/L, and the same total ammonia concentration was achieved. This supports the validity of the design methodology to create staged PNA MABRs, but experimental calibration with real-world reactors is required.

Objective 3: Optimal PNA MABRs treatment trains will include approximately equal-sized reactors in series

To confirm whether the design methodology optimized the amount of nitrogen removed for a given quantity of membranes and to validate whether equal-sized reactors will result in the best TIN removal, the simulations described in objective 3 were performed. Treatment trains with an equivalent amount of membrane surface area had their membranes distributed into separate reactors at different proportions, seen in Figure 1, and the combined total nitrogen removal rate for the treatment trains are shown in Figure 3. From this figure, we can confirm that our proposed methodology, which generally created equally sized stages, resulted in the best total nitrogen removal rate. The total nitrogen removal rate for the train with equally distributed membrane surface area, 33-33-33, had the largest total nitrogen removal rate of 0.81 g $N/m^2/d$. Although the total nitrogen

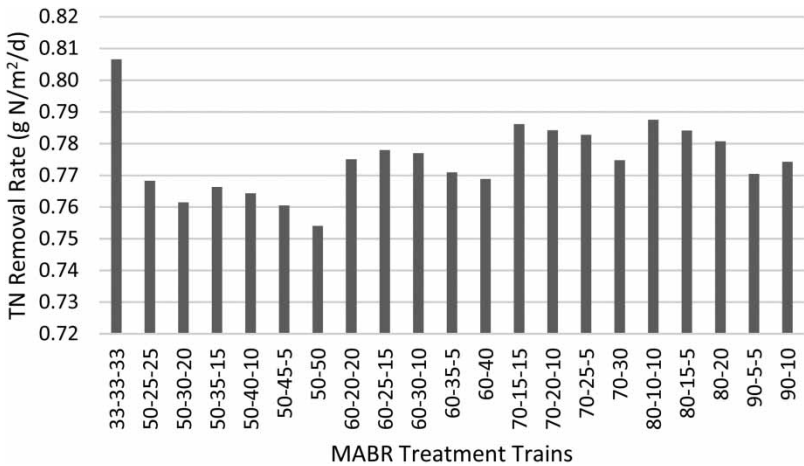


Figure 3 | Total nitrogen (TN) removal rate for the tested treatment trains. The values listed in the x-axis correspond to the percentage of total MABR surface area in the first, second, and third reactors in a treatment train, respectively. The HRT was constant for each reactor, and the removal rate represents the removal rate for the entire treatment train. The air loading in each staged reactor was optimized to minimize the effluent total nitrogen concentration in each stage and was not maintained at a 1:1 ratio of influent oxygen to nitrogen.

removal rates only differed by 7%, the effluent total nitrogen concentration in the treatment train with equally distributed membrane surface area had a 26% lower effluent total nitrogen concentration than the second-best treatment train.

Similar to the treatment train created using the design methodology, the total nitrogen removal rate in the first stage was significantly higher than the rates in the second and third stages with a maximum difference of 1.07 g $N/m^2/d$ between the first and second stages, shown in Figure S-7. The average total nitrogen removal rates in the first, second, and third stages were 1.12, 0.34, and 0.14 g $N/m^2/d$, respectively, for all the treatment trains tested. The first stage's removal rate compensated for the decreased removal rates in the later reactors, resulting in a net increase in the substrate removal rate for the overall process when compared to an un-staged reactor. In addition, the last reactor ensured that the treatment train met their effluent quality objectives.

When comparing the performance of the train with equally distributed membrane surface areas to the performance of an un-staged 2.5 L MABR with an equivalent surface area in Table S-2, the benefits of staging become evident. The staged treatment train had a total ammonia removal rate and a total nitrogen removal rate that were 9.8 and 9.3% larger than the un-staged reactor, respectively. The un-staged reactor also was not able to get its effluent total ammonia concentration below 5 mg N/L without causing excessive NOB growth which would be necessary for many facilities' permits. Despite further modifications to the un-staged reactor, it was not able to reach a comparable performance to the staged reactor after having its membrane surface area increased by a factor of four (data not shown). This observation of un-staged MABRs struggling to achieve low effluent concentrations has also been observed by Wagner *et al.* (2022). Ultimately, staging a PNA MABR will increase the reactor's ammonia and total nitrogen removal rates, enable it to oxidize total ammonia to lower limits, and limit the NOB specific growth rate by distributing the membrane air loading to the three reactors.

Effective operation of staged PNA MABRs requires each reactor to have its oxygen loading controlled

Optimal operation of PNA MABRs in a treatment train requires different oxygen loading setpoints for each reactor. Figure S-8 shows the ratios of influent oxygen to influent total ammonia fluxes for each reactor in the tested treatment trains. In these simulations, each reactor in the series had its oxygen loading modified to minimize the effluent total nitrogen concentration and was not controlled proportionally to the influent total ammonia concentration. The first stages' oxygen loading ratios were typically 1.2 g O_2/g N, but the configuration with equally distributed surface areas (33-33-33) had an optimal ratio of 1.0 g O_2/g N in its first stage. Since the design methodology was created using a ratio of 1.0 g O_2/g N, this ratio can be used for a treatment train with equal-sized reactors. However if the first reactor is proportionally larger than the other reactors, the design methodology may need to be slightly modified to account for the marginal increase in the optimal oxygen loading ratio. Interestingly, the optimal oxygen loading ratio decreases for the second and third stages when the reactors are not equally sized. On average, the optimal oxygen loadings to the second and third reactors are 22 and 45% lower than the loading applied to the first reactor, respectively. This suggests that if unequal stages are used, the design methodology may oversupply the second and third reactors with air, potentially decreasing their TIN removal rates. Nevertheless, our methodology appears to be a valid procedure for designing PNA MABRs.

Objective 4: The addition of exogenous COD to the influent increases the MABR's TIN removal rate for low C:N ratios

The previous results did not incorporate the consequences of having exogenous COD in the MABR's influent. In theory, PNA MABRs would be placed after a high-rate carbon capture process to redirect exogenous COD for resource recovery and to protect the MABR from large COD loadings. However, high-rate systems are not able to completely capture all the COD that they are fed with, so the COD's impact on the TIN removal rate and the ability of the MABR to remove the residual COD is important to include in the design methodology.

The TIN removal rates and rbCOD removal rates of simulations, which added exogenous rbCOD at 25, 50, and 100 mg rbCOD/L to the MABR influent, are shown in Figures 4 and 5, respectively. In Figure 4, the addition of exogenous COD almost doubled the average TIN removal rate from 1.5 g N/m^2/day at a C:N ratio of 0 to an average TIN removal of 2.94 and 2.86 g N/m^2/day for a C:N ratio range of 0.01–0.50 and 0.51–1.00, respectively. This increase in the TIN removal rate was primarily driven by an increase in heterotrophic denitrification. With exogenous COD, the percentage of nitrogen removed by heterotrophic denitrification increased from 3 to 5% when exogenous COD was not added to up to 80% (data not shown). At low C:N ratios below 1, heterotrophic denitrifiers did not outcompete AMX and were able to reduce the nitrate produced through anammox to nitrogen gas, significantly increasing the TIN removal rate. As the C:N ratio increased beyond 1.0, the average TIN removal rate decreased, and at C:N ratios greater than 2.0, the average TIN removal rate

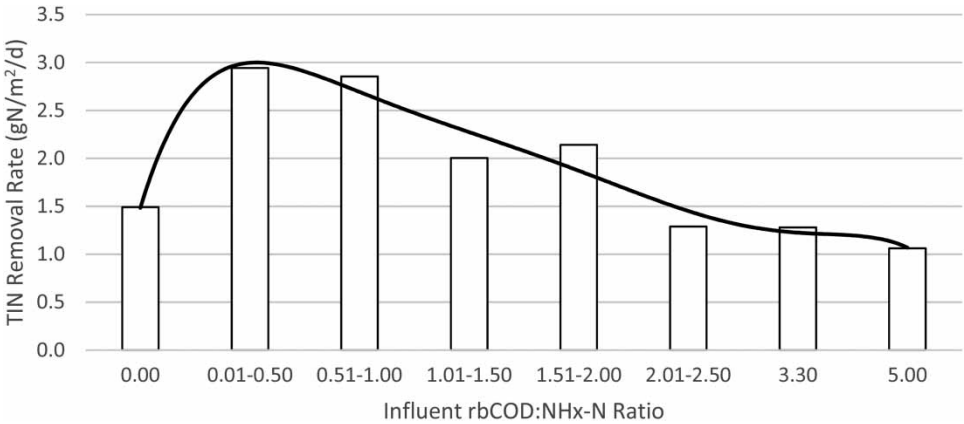


Figure 4 | Impact of the influent C:N ratio on the TIN removal rate.

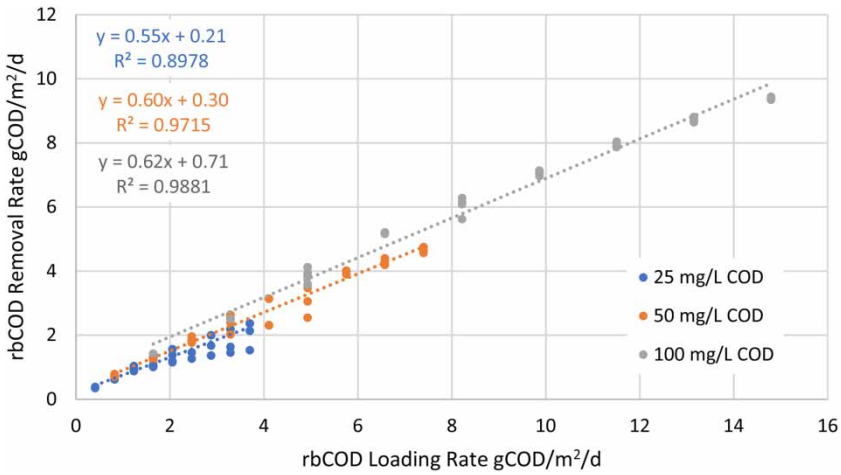


Figure 5 | Comparison between the rbCOD loading rate and the rbCOD removal rate.

decreased below the average TIN removal rate from the simulations without exogenous COD at a C:N ratio of 0.0. In fact, at C:N ratios of 3.3 and 5.0, PNA accounted for less than 1% of the nitrogen removal (data not shown). This decrease in the TIN removal rate highlights the sensitivity of PNA MABRs to the influent C:N ratio which has been observed in other studies (Lackner *et al.* 2008; Wagner *et al.* 2022). It also highlights the relative nitrogen removal efficiency of anammox versus heterotrophic denitrification. For municipal wastewater treated by a high-rate carbon capture process, mainstream nitrogen removal using staged MABRs is feasible, and the nitrogen removing metabolisms in the MABRs, either heterotrophic denitrification or anammox, will be influenced by the wastewater characteristics and the performance of the carbon capture process.

In addition to nitrogen removal, the rbCOD removal rate needs to be quantified for PNA MABRs to ensure that effluent COD limits are met. The rbCOD removal rate, displayed in Figure 5, correlated well with the rbCOD loading rate, and the rbCOD removal rate increased when the influent rbCOD increased for all simulations. Linear correlations with R^2 values of 0.90, 0.97, and 0.98 were observed for simulations with influent rbCOD concentrations of 25, 50, and 100 mg rbCOD/L, respectively. Simulations with influent rbCOD concentrations of 25 mg rbCOD/L had the lowest R^2 value, presumably because the nitrogen loading had a larger influence on the microbial community present within their biofilms. When the influent rbCOD concentration increased, the influence of the nitrogen loading on the microbial community within the biofilm decreased. The R^2 value also increased, emphasizing the increased importance of the influent rbCOD concentration on its removal rate. Simulations with 100 mg rbCOD/L in their influent also had the largest removal rates. This is unsurprising,

as they would have the largest effluent concentration which would increase the flux of rbCOD into the biofilm. These rbCOD removal rates may slightly overestimate the rbCOD removal rate seen in experimental reactors. In laboratory-scale PNA MABRs fed with municipal wastewater, (Bunse *et al.* 2020) observed a soluble COD (sCOD) removal of 50% with an average removal rate of 2.3 g $O_2/m^2/d$ after being fed with a surface loading ranging from 1.9 to 13.3 g $O_2/m^2/d$ at a constant HRT. However after switching to an operational strategy that supplied a constant nitrogen load rather than a constant HRT, the sCOD removal decreased and had a wider distribution of performance. The sCOD loading decreased to 0.4–7.2 g $O_2/m^2/d$ which resulted in an sCOD removal of 20–60% at an average rate of 1.0 g $O_2/m^2/d$ for an intermittently aerated reactor and 0.5 g $O_2/m^2/d$ for a continuously aerated reactor. Thus, additional experimentation, especially under dynamic flows and loadings, will be required to effectively validate the rbCOD removal rate.

To test how the addition of rbCOD would affect staging, the treatment train created from the design methodology had 50 mg rbCOD/L added to its influent. Because of the increase in oxygen demand, each reactor in the treatment train with 50 mg rbCOD/L had its membrane air increased by 0.5 g O_2/g COD. This design was simulated, and its results are compared to the design without exogenous COD in Table 3. In the first stage, the addition of COD and the increased air loading rate decreased the total ammonia concentration by 1.13 mg N/L and total nitrogen effluent concentration by 2.62 mg N/L. At an influent C:N ratio of 1.0, the nitrate that was previously produced by NOB and AMX was completely reduced to nitrogen gas by heterotrophic denitrifying organisms, reinforcing the results observed in Figure 4, and the increased air loading to the membrane enabled increased nitrification of total ammonia. The first stage also accounted for 73% of the rbCOD removal in the treatment train. The second stage had a slightly higher C:N ratio of 1.13, as the total ammonia removal rate was larger than the rbCOD removal rate in the previous stage. The removal rates of total ammonia and rbCOD were lower in this stage due to having smaller effluent concentrations, and the net improvements in the total nitrogen and total ammonia concentrations between the two trains slightly decreased by 0.95 and 0.73 mg N/L, respectively. Finally, the third stage had the lowest rbCOD and total ammonia removal, resulting in a combined improvement of 0.55 and 0.17 mg N/L for the total nitrogen and total ammonia concentrations. 98% of the rbCOD was also removed from the influent. In summary, adding COD to a staged PNA MABR treatment train increased the total nitrogen and total ammonia removal, since the C:N ratio was below 2.0 in every reactor. Most of the benefits resulted from the first stage which had the largest AMX specific growth rate, and heterotrophic denitrification was able to reduce the produced nitrate from NOB and AMX to nitrogen gas.

PRACTICAL CONSIDERATIONS

To recreate the strategy used in this paper to control the membrane air loading, the following control strategy could be implemented to treat dynamic flows and loadings. Each reactor stage needs its own oxygen loading determined from the influent total ammonia and COD loadings. Separate total ammonia and COD sensors are required in the influent to each reactor. Next, oxygen sensors can be placed in the MABR's air inlet and offgas to calculate the oxygen transfer to the biofilm. Finally, a nitrate sensor placed in the effluent of each reactor stage may be helpful to communicate instances of complete nitrification and will aid in quantifying the amount of nitrogen removed through PNA. These sensors will have to be precise to achieve strict total nitrogen and total ammonia permit limits, as variations between the observed and actual concentrations may cause process upsets. The combined sensor data can be used to control the membrane air actuator through proportional–integral–derivative (PID) controllers for each reactor stage. Full-scale treatment also may require multiple treatment trains to achieve

Table 3 | Comparison of staged treatment train performance with and without COD

	Influent		1st stage effluent		2nd stage effluent		3rd stage effluent	
	0 COD	50 COD	0 COD	50 COD	0 COD	50 COD	0 COD	50 COD
Total nitrogen mg N/L	50	50	15	12	5.5	3.9	3.7	3.2
Total ammonia mg N/L	50	50	12	11	3.0	2.6	0.86	0.69
Nitrite mg N/L	0.00	0.00	0.00	0.00	0.00	0.010	0.00	0.010
Nitrate mg N/L	0.00	0.00	2.2	0.00	2.4	0.75	2.8	2.0
rbCOD mg COD/L	0.00	50	0.073	14	0.073	2.2	0.073	0.49
Influent rbCOD:NH_x-N ratio	N/A	N/A	0.00	1.0	0.010	1.2	0.024	0.84

adequate treatment or to enable redundancy which would further increase the complexity of this control scheme. Therefore, research on more efficient, simpler, and cheaper control schemes may be desirable to enable the implementation of PNA MABRs in mainstream nitrogen removal.

Results from experimental reactors with ammonia concentrations typical for mainstream wastewater confirms that NOB suppression is possible but may be the most challenging aspect of achieving PNA in MABRs. Further work developing simple and effective NOB suppression strategies will be helpful to advance the use of MABRs for PNA. (Ribeiro Augusto *et al.* 2018) observed an average total nitrogen removal of 78% with a maximum total nitrogen removal of 84% in an MABR treating synthetic wastewater with an influent ammonium concentration of 50 mg N/L and no exogenous COD, which is higher than the 71% TIN removal used in our design methodology. However, they used an elevated water temperature of 30 °C in their MABR to improve anammox activity and to aid in NOB suppression. A study by (Bunse *et al.* 2020) compared intermittently and continuously aerated MABRs that treated real municipal wastewater. When both reactors were operated at a constant nitrogen loading, the intermittently aerated MABR had an average total nitrogen removal of 54–87% with a maximum of 96.9% while the continuously aerated reactor had a lower average total nitrogen removal of 10–55% with a maximum of 96.2%. Although intermittent aeration provided more reliable NOB suppression and larger total nitrogen removal percentage, it decreased the total ammonia removal percentage from 80–92% to 74–89% for the continuous and intermittently aerated reactors, respectively. To compensate for this trade-off in TIN and total ammonia removal percentages, other researchers have begun alternating the direction of the membrane gas flow which can combine the benefit of NOB suppression from intermittent aeration with the increased total ammonia removal from continuous aeration (Houweling 2020). However, as highlighted in the study by (Bunse *et al.* 2020), operating PNA MABRs under dynamic conditions can result in a wide range of total nitrogen removal percentages, regardless of the aeration strategy, which will make their implementation more difficult if effective control strategies are not developed.

CONCLUSIONS

This study created a methodology to design staged partial nitritation anammox MABRs. The main findings from this study are:

- PNA MABRs have more complex metabolisms than traditional biofilm reactors, so an alternative staging design methodology must be used to suppress NOB while preventing oxygen inhibition of AMX.
- The size of each stage depends on the total ammonia nitrogen loading, total ammonia nitrogen concentration, and the membrane air loading to the reactor. The air loading to each reactor should be adjusted in proportion to the total ammonia nitrogen loading. However, if the total ammonia nitrogen loading is too high, NOB can proliferate in the biofilm even if the oxygen is carefully controlled.
- Stages should be designed with a nitrogen loading that maximizes the TIN removal while minimizing NOB growth. This will typically result in equally sized stages, with the last stage's size and aeration modified to reach the desired effluent limits. This study used a TIN removal percentage of 71% to design each stage.
- The total ammonia and total nitrogen removal rates for a staged treatment train were 10 and 9% larger, respectively, when compared with an un-staged system.
- A robust control strategy will be required to implement PNA MABRs in practice, and additional experimentation is required to validate the performance under dynamic conditions.

ACKNOWLEDGEMENTS

We would like to thank the United States National Science Foundation Graduate Research Fellowship (DGE 1841052) and the University of Michigan Rackham Predoctoral Fellowship for their financial support. We would also like to thank SUEZ Water Technologies & Solutions (Oakville, Ontario, Canada) for helpful discussions on MABRs and anammox.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Acevedo Alonso, V. & Lackner, S. 2019 Membrane aerated biofilm reactors – how longitudinal gradients influence nitrogen removal – a conceptual study. *Water Research* **166**, 115060. Available from: https://www.sciencedirect.com/science/article/pii/S0043135419308346.

Bekele, Z. A., Delgado Vela, J., Bott, C. B. & Love, N. G. 2020 Sensor-mediated granular sludge reactor for nitrogen removal and reduced aeration demand using a dilute wastewater. *Water Environment Research*, wer.1296. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/wer.1296.

Bunse, P., Orschler, L., Agrawal, S. & Lackner, S. 2020 Membrane aerated biofilm reactors for mainstream partial nitritation/anammox: experiences using real municipal wastewater. *Water Research X* **9**, 100066. https://doi.org/10.1016/j.wroa.2020.100066.

Dynamita 2019 *SUMO 19.2*. Sigale, France. Available from: http://www.dynamita.com/.

Gilmore, K. R., Terada, A., Smets, B. F., Love, N. G. & Garland, J. L. 2013 Autotrophic nitrogen removal in a membrane-aerated biofilm reactor under continuous aeration: a demonstration. *Environmental Engineering Science* **30** (1), 38–45. Available from: http://online.liebertpub.com/doi/abs/10.1089/ees.2012.0222.

Gong, Z., Yang, F., Liu, S., Bao, H., Hu, S. & Furukawa, K. 2007 Feasibility of a membrane-aerated biofilm reactor to achieve single-stage autotrophic nitrogen removal based on anammox. *Chemosphere* **69** (5), 776–784.

Gong, Z., Liu, S., Yang, F., Bao, H. & Furukawa, K. 2008 Characterization of functional microbial community in a membrane-aerated biofilm reactor operated for completely autotrophic nitrogen removal. *Bioresource Technology* **99** (8), 2749–2756.

Grady, C. P. L. J., Daigger, G. T., Love, N. G. & Filipe, C. D. M. 2011 *Biological Wastewater Treatment*. CRC Press, Boca Raton.

Houweling, D. 2020 ZeeNAMMOX deammonification with MABR – no carbon, no alkalinity and no bubbles. In: *WEFTEC 2020 Proceedings*.

Isanta, E., Reino, C., Carrera, J. & Pérez, J. 2015 Stable partial nitritation for low-strength wastewater at low temperature in an aerobic granular reactor. *Water Research* **80**, 149–158.

Lackner, S., Terada, A. & Smets, B. F. 2008 Heterotrophic activity compromises autotrophic nitrogen removal in membrane-aerated biofilms: results of a modeling study. *Water Research* **42** (4–5), 1102–1112.

Liu, Y., Ngo, H. H., Guo, W., Peng, L., Pan, Y., Guo, J., Chen, X. & Ni, B. J. 2016 Autotrophic nitrogen removal in membrane-aerated biofilms: archaeal ammonia oxidation versus bacterial ammonia oxidation. *Chemical Engineering Journal* **302**, 535–544. http://dx.doi.org/10.1016/j.cej.2016.05.078.

Lotti, T., Kleerebezem, R., Hu, Z., Kartal, B., De Kreuk, M. K., Kip, V. E. T., Kruit, C., Hendrickx, J., G, T. L. & Van Loosdrecht, M. C. M. 2015 Pilot-scale evaluation of anammox-based mainstream nitrogen removal from municipal wastewater. *Environmental Technology (United Kingdom)* **36** (9), 1167–1177. https://doi.org/10.1080/09593330.2014.982722.

Matsumoto, S., Terada, A. & Tsuneda, S. 2007 Modeling of membrane-aerated biofilm: effects of C/N ratio, biofilm thickness and surface loading of oxygen on feasibility of simultaneous nitrification and denitrification. *Biochemical Engineering Journal* **37** (1), 98–107.

Peeters, J., Adams, N., Long, Z., Côté, P. & Kunetz, T. 2017 Demonstration of innovative MABR low-energy nutrient removal technology at Chicago MWRD. *Water Practice and Technology* **12** (4), 927–936.

Pellicer-Nàcher, C. & Smets, B. F. 2012 Redox stratified biofilms to support completely autotrophic nitrogen removal: principles and results. In *SIDISA 2012 - Sustainable Technology for Environmental Protection*, Milan Italy. Available from: http://orbit.dtu.dk/en/publications/redox-stratified-biofilms-to-support-completely-autotrophic-nitrogen-removal-principles-and-results(f828fa2e-9e81-4862-b888-4d06b4e36c4c).html.

Pellicer-Nàcher, C., Sun, S., Lackner, S., Terada, A., Schreiber, F., Zhou, Q. & Smets, B. F. 2010 Sequential aeration of membrane-aerated biofilm reactors for high-rate autotrophic nitrogen removal: experimental demonstration. *Environmental Science and Technology* **44** (19), 7628–7634.

Pérez, J., Lotti, T., Kleerebezem, R., Picioreanu, C. & van Loosdrecht, M. C. M. 2014 Outcompeting nitrite-oxidizing bacteria in single-stage nitrogen removal in sewage treatment plants: a model-based study. *Water Research* **66** (ii), 208–218.

Pérez, J., Isanta, E. & Carrera, J. 2015 Would a two-stage N-removal be a suitable technology to implement at full scale the use of anammox for sewage treatment? *Water Science and Technology* **72** (6), 858–864.

Poot, V., Hoekstra, M., Geleijnse, M. A. A., van Loosdrecht, M. C. M. & Pérez, J. 2016 Effects of the residual ammonium concentration on NOB repression during partial nitritation with granular sludge. *Water Research* **106**, 518–530.

Regmi, P., Miller, M. W., Holgate, B., Bunce, R., Park, H., Chandran, K., Wett, B., Murthy, S. & Bott, C. B. 2014 Control of aeration, aerobic SRT and COD input for mainstream nitritation/denitritation. *Water Research* **57**, 162–171. http://dx.doi.org/10.1016/j.watres.2014.03.035.

Ribeiro Augusto, M., Rosseto, P. & Souza, T. S. O. D. 2018 Fast start-up of the single-stage nitrogen removal using anammox and partial nitritation (SNAP) from conventional activated sludge in a membrane-aerated biofilm reactor. *Bioresource Technology* **266**, 151–157. https://doi.org/10.1016/j.biortech.2018.06.068.

Strous, M., Heijnen, J. J., Kuenen, J. G. & Jetten, M. S. M. 1998 The sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microorganisms. *Applied Microbiology and Biotechnology* **50** (5), 589–596.

Sun, S., Pellicer-Nàcher, C., Terada, A., Lackner, S. & Smets, B. F. 2010 Start-up strategies of membrane-aerated biofilm reactors (MABR) for completely autotrophic nitrogen removal. In: *Proceedings of the 2nd IWA Specialized Conference on Nutrient Management in Wastewater Treatment Processes*. pp. 1–8.

Terada, A., Lackner, S., Tsuneda, S. & Smets, B. F. 2007 Redox-stratification controlled biofilm (ReSCoBi) for completely autotrophic nitrogen removal: the effect of co- versus counter-diffusion on reactor performance. *Biotechnology and Bioengineering* **97** (1), 40–51.

Terada, A., Lackner, S., Kristensen, K. & Smets, B. F. 2010 Inoculum effects on community composition and nitritation performance of autotrophic nitrifying biofilm reactors with counter-diffusion geometry. *Environmental Microbiology* **12** (10), 2858–2872.

Wagner, B. 2022 *Achieving Mainstream Partial Nitritation/Anammox Nitrogen Removal via a Membrane Aerated Biofilm Reactor (MABR)*.

Wagner, B. M., Daigger, G. T. & Love, N. G. 2022 Assessing membrane aerated biofilm reactor configurations in mainstream anammox applications. *Water Science & Technology* **85** (3), 943–960. https://doi.org/10.2166/wst.2022.013.

Wett, B., Omari, A., Podmirseg, S. M., Han, M., Akintayo, O., Gómez Brandón, M., Murthy, S., Bott, C., Hell, M., Takács, I., Nyhuis, G. & O'Shaughnessy, M. 2013 Going for mainstream deammonification from bench to full scale for maximized resource efficiency. *Water Science and Technology* **68** (2), 283–289.

First received 5 June 2022; accepted in revised form 26 September 2022. Available online 7 October 2022