

A universal empirical equation to estimate the abundance of carbapenem-resistant genes during aerobic digestion of wastewater sludge

Eskandar Poorasgari ^{a,*} and Banu Örmeci ^b

^a Department Civil, Geological and Environmental Engineering, 57 Campus Drive, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada

^b Department of Civil and Environmental Engineering, 3438 Mackenzie, Carleton University, 1125 Colonel By Drive, Ottawa, ON, K1S 5B6, Canada

*Corresponding author. E-mail: arf867@mail.usask.ca

 EP, 0009-0008-6260-8085

ABSTRACT

Carbapenem-resistant genes (CRGs) exist in wastewater and accumulate in wastewater sludge. Due to the potential threat posed by the CRGs, it is important to quantify CRGs and predict their removal and discharge concentrations during aerobic sludge digestion. Nonetheless, gene quantification is tedious, error-prone and expensive. This study aims to develop multiple regression models to estimate CRGs from sludge parameters that are routinely measured for the monitoring and design of aerobic sludge digesters. Batch reactors were operated at mesophilic and thermophilic temperatures for 20–35 days. Sludge samples were periodically taken during aerobic digestion. Three CRGs (*bla*_{GES}, *bla*_{OXA-48} and *bla*_{IMP-27}) together with 16S rRNA and integron class 1 genes were quantified. Aerobic digestion reduced the abundance of all target genes. Multiple regression modelling was conducted in linear (LM) and non-linear (NLM) modes. Sums of squared errors of the LM models were 0–0.048, whereas those of the NLM models were 0–0.003. Adjusted R² ranges of the LM and NLM models were 0.774–0.931 and 0.986–1, respectively. Overall, the NLM models predicted the abundance of target genes more accurately than the LM models. NLM models may be used to modify the design and operational parameters of aerobic sludge digesters.

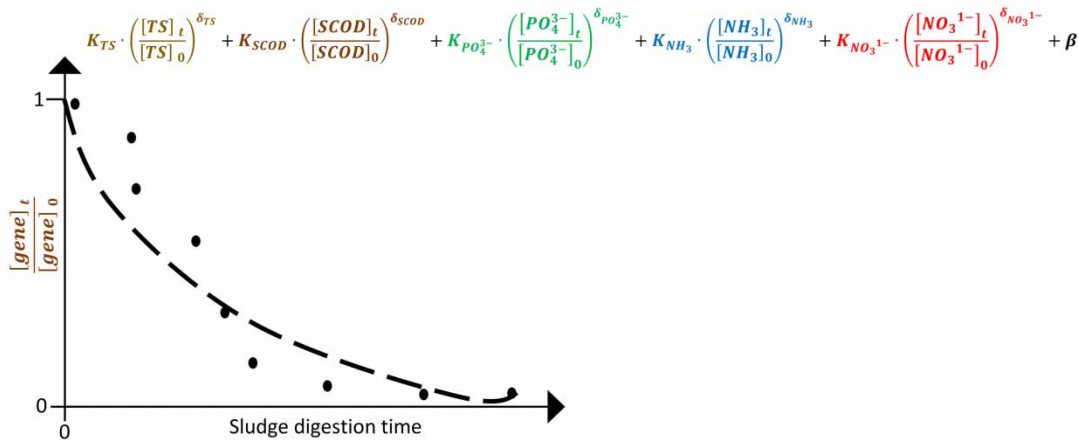
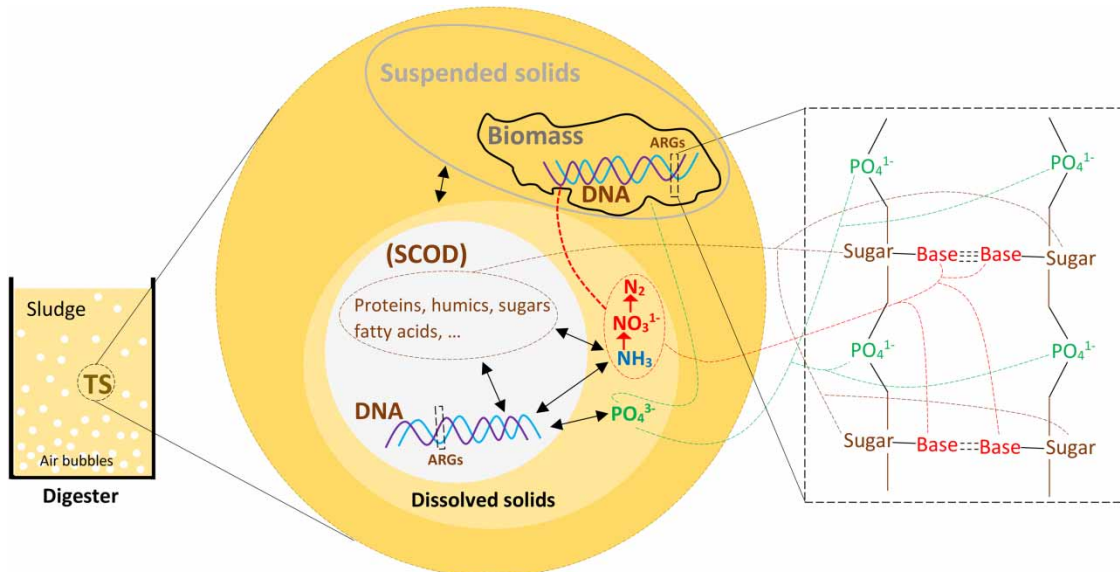
Key words: aerobic digestion, antibiotic resistance, carbapenem, regression modelling, sludge, wastewater

HIGHLIGHTS

- 16S rRNA, Int 1, *bla*_{GES} and *bla*_{OXA-48} were quantified during aerobic digestion.
- ARGs–solids–nutrients (ASN) nexus explained the abundance of the target genes.
- ASN nexus was mathematically expressed by regression modelling.
- Regression models included single and multiple, linear and non-linear: L_S , NL_S , $L_{S,U}$, $NL_{S,U}$, L_M and NL_M models.
- Multi-factor non-linear models (NLM) were more accurate than the other models.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



1. INTRODUCTION

Wastewater treatment produces sludge, which must be treated to reduce its pathogen and organic matter content before land application. Aerobic digestion is a common treatment method of sludge produced by wastewater treatment plants (WWTPs) up to a capacity of 2 m³/s, performed at mesophilic and thermophilic temperatures (Tchobanoglous *et al.* 2014). Antibiotic-resistant genes (ARGs), which are a part of DNA, encode resistance against bacterial infections (Davies & Davies 2010). Multiple studies demonstrated that antibiotic-resistant bacterial infections are becoming a major cause of death in North America and Europe (Bancroft 2007; Spellberg *et al.* 2011; Cassini *et al.* 2019). ARGs are an emerging group of contaminants in sludge (Mao *et al.* 2015). Aerobic digestion affects the quantity of ARGs, some of which survive aerobic digestion (Zhang *et al.* 2021). Therefore, safe design and monitoring of aerobic sludge digesters in terms of ARG removal would need engineers and operators to quantify ARGs.

Carbapenem-resistant genes (CRGs) are an important group of ARGs because carbapenem antibiotics are one of the few remaining options to treat bacterial infections (Sekyere 2016). There are many CRGs (Nordmann *et al.* 2012), among which bla_{GES}, bla_{OXA} and bla_{IMP} (Liang *et al.* 2018; Ekwanzala *et al.* 2020; Wang *et al.* 2022) have been detected in wastewater and sludge. Therefore, these CRGs are important and need to be the focus of further studies on wastewater sludge treatment. Nevertheless, it has not been investigated how the quantity of these genes changes during the aerobic digestion of wastewater sludge. Moreover, some bacteria are capable of horizontal gene transfer (HGT) using a gene that code for integrase class 1 (Int 1 gene). This gene can help to transmit CRGs (Xiong *et al.* 2013) and exists in wastewaters (An *et al.* 2018). Apart from that, 16S rRNA gene is a part of the bacterial genome, and it can be used for analyzing bacterial biomass (Smets *et al.* 2016). Hence, researchers

usually study Int 1 and 16S rRNA variations with other genes in ARG studies (Nordmann *et al.* 2012; Wang *et al.* 2022).

To enumerate genes, DNA isolation and quantitative polymerase chain reaction (qPCR) are required (Xiong *et al.* 2013; Paulus *et al.* 2019), which are expensive, tedious and error-prone (Milobedzka *et al.* 2022). In addition, practitioners working with wastewater sludge do not normally have the background and training for such molecular techniques. Thus, a reasonable method of gene estimation is needed that can be employed by the personnel who design and operate the aerobic digesters of wastewater sludge. Such estimation can be based on the sludge characteristics. Wastewater sludge consists of solids and water, where the former contains DNA. This compound can be chemically oxidized and consists of hydrogen, oxygen, carbon, nitrogen and phosphorus (Brezonik & Arnold 2011). Chemical oxidizability means that DNA is partly responsible for the chemical oxygen demand (COD) of the sludge. With this perspective, it is not surprising that the relationship between ARGs quantity and various sludge parameters has been scrutinized frequently (Tong *et al.* 2018; Tong *et al.* 2019a; Jang *et al.* 2020). The parameters include total solids, volatile solids (VSs) and volatile suspended solids (VSSs) (Jang *et al.* 2020; Poorasgari & Örmeci 2022a), phosphorus, COD and nitrogenous compounds (Tong *et al.* 2018, 2019a). Note that these parameters are normally measured for operation and monitoring purposes. VS and VSS encompass intercellular biomolecules such as proteins, lipids and carbohydrates. Also, the mentioned intercellular compounds are partly responsible for COD. A fraction of COD is soluble (SCOD), whose concentration changes during aerobic digestion (Kavitha *et al.* 2016), and a part of it is regarded as food for bacteria (Tchobanoglous *et al.* 2014). Ammonia is a nitrogenous compound in sludge whose concentration varies during aerobic digestion (Liu *et al.* 2012). This compound is a nutrient, but it becomes toxic when its concentration goes beyond certain limits (Gray 2004; Tong *et al.* 2019b). Associations between ammonia and ARGs have been demonstrated (Tong *et al.* 2019b; Goetsch *et al.* 2020). Soluble orthophosphate (SOP) is another nutrient (Tchobanoglous *et al.* 2014; Baird *et al.* 2017) whose concentration also changes during aerobic digestion (Poorasgari & Örmeci 2022c).

The associations described above imply that there is an ARGs–solids–nutrients (ASN) nexus in wastewater, which may be mathematically represented by regression modelling. A single regression model establishes a quantitative relationship between a dependent variable and a predictor, whereas a multiple regression model estimates the dependent variable from the values of more than one predictor. Several single and multiple regression models have been reported (Tong *et al.* 2018, 2019a, b). Nevertheless, the predictors of those models included, among other things, Chao and Shannon indices which are suitable for studying the biodiversity of a microbial community using specific software (Gihring *et al.* 2012; Thurkral 2017; Willis 2019). As well, quantification of particular parts of the 16S rRNA gene would be required to apply those models for estimating the genes (Větrovský & Baldrian 2013; Dueholm *et al.* 2022). This would make the models difficult to use by the wastewater practitioner. Based on the ASN nexus, Poorasgari & Örmeci (2022a) established a new modelling approach to predict the abundance of target genes during anaerobic sludge digestion. In that approach, all the predictors were parameters routinely measured during the operation of anaerobic digesters. However, it is not known whether the regression models developed by Poorasgari & Örmeci (2022a) would work for the aerobic digesters of wastewater sludge. Due to nitrification and denitrification during aerobic digestion (Tchobanoglous *et al.* 2014), nitrate concentration can change (Liu *et al.* 2012; Poorasgari & Örmeci 2022c). Therefore, factoring in the nitrate concentration may be important for building proper regression models of the aerobic sludge digesters.

In this investigation, it is hypothesized that the abundance of 16S rRNA, Int 1, bla_{GES}, bla_{OXA} and bla_{IMP} can be estimated from sludge total solids (TSS), SOP, SCOD, ammonia and nitrate concentrations. To test the hypothesis, the current study uses the ASN conceptual model to formulate a governing equation through different types of regressions. It is aimed to construct the equation in such a way that it is accurate and comprehensive, yet it can be understood and applied by the engineers and operators of aerobic sludge digesters. Such an equation is meant to estimate the abundance and removal of CRGs during the operation of aerobic digesters, while it could also be used for the safe design of such reactors.

2. METHODS

2.1. Digesters

Three batch aerobic sludge digesters were employed in the current investigation. The batch reactors were operated at 35, 45 and 55 °C, named AD.35, AD.45 and AD.55, respectively. The working volume of each reactor was

10 L. The reactor content consisted of 7.4 Kg primary sludge, 1.6 Kg T-WAS and 1 Kg inoculum. More information about the reactors is provided elsewhere (Poorasgari & Örmeci 2022b, 2022c). Sludge samples were taken from the digesters within time intervals of 5 days. The sludge samples were refrigerated at 1–3 °C until further analysis and DNA isolation.

2.2. Isolation and purification of DNA

DNA of the sludge samples was isolated using the Power Soil kit (Qiagen). About 0.25 g of sludge was used for each DNA isolation and three isolates were acquired for each sludge sample. A NucleoSpin gDNA clean-up kit (Macherey-Nagel GmbH & Co. KG) was used to improve the quality of the DNA isolates. After the clean-up, the DNA concentration of the isolates was measured using a NanoDrop 1000 spectrophotometer (NanoDrop, USA). Then, the isolates were stored at –20 °C until further analysis.

2.3. Genes quantification

Five target genes were enumerated: 16S rRNA, Int 1, *bla_{GES}*, *bla_{OXA-48}* and *bla_{IMP-27}*. Quantification was conducted using qPCR, where a Bio-Rad thermocycler (CC013200) and CFX Master software (Bio-Rad) were utilized. The reagent was SsoFast (Bio-Rad), whose ingredients were di-nucleotides, SYBR Green label, DNA polymerase enzyme and proprietary components. The volume of the reaction mix in each PCR tube was 25 µL. The mass of DNA for DNA isolates obtained from the sludge samples was 30 ng DNA per tube. The assay began with preheating at 95 °C for 180 s, and each cycle included denaturation at 95 °C for 10 s, annealing at primer-specific temperatures for 20 s, and polymerisation at 72 °C (see Table 1 for further information). The total cycle number of the CRGs was three more than that of the other two targets because preliminary tests had shown that the assays for the CRGs required more cycles to become quantifiable. Quantities with quantification cycle numbers more than those shown in Table 1 were considered zero.

The amplicon size and the specificity of primer annealing were first verified with the Primer-BLAST tool of the National Centre for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). To further confirm the annealing specificity of the primers and the size of the amplicons, gel electrophoresis (GE) was conducted (see Supplementary material S1 for details). Manufactured positive standards (Thermo Fisher) were used to prepare serial dilutions for calibrating the PCR machine. Standard solutions spanned 10²–10¹¹ copy numbers with an order of magnitude difference between each two subsequent dilutions. Calibration equations (Supplementary material, Table S2) were used to determine the absolute abundance of each target gene in the DNA isolate of the sludge samples in copy number per ng DNA. To determine the relative abundance of Int 1 genes and the CRGs, their absolute abundances were divided by the absolute abundance of their respective 16S rRNA genes.

2.4. Analytical methods

Standard methods were followed to measure TS concentration (Baird *et al.* 2017). To measure the predictors in the liquid phase, sludge liquid phase was obtained. To obtain the liquid phase, the sludge was first centrifuged at 14,000 g, and subsequently, the supernatant was filtered through a membrane filter with a pore diameter of 0.45 µm. The permeate of the filtration step was used for measuring the concentration of dissolved components. SCOD was measured using a high-range COD reagent (Hach colorimetric method 8000), and SOP was measured using a high-range molybdovanadate reagent (Hach, colorimetric methods 8114). Total ammonia nitrogen (TAN), nitrate (NO₃-N) and pH were measured using IntelliCAL™ (ISEH3181, Hach), IntelliCAL™ (ISENO 3181, Hach) and ORION91578 (Thermo Scientific) probes, respectively. The mass fraction of sludge water was calculated from TS concentration, and it was used to convert the concentrations of the dissolved parameters from g per kg of the liquid phase to g per kg of wet sludge. Further details about the analytical techniques are provided elsewhere (Poorasgari & Örmeci 2022b, 2022c). A fraction of TAN is free ammonia nitrogen (FAN), which is more toxic than ammonium ions (Gray 2004). The FAN concentration was calculated from TAN by Equation (1), which is based on NH₃-NH₄⁺ equilibrium in water-based solutions.

$$\text{FAN} = \frac{\text{TAN}}{1 + \left[\left(\frac{K_b}{K_w} \right) \times 10^{-\text{pH}} \right]} \quad (1)$$

where K_b and K_w are the equilibrium constants of ammonia basicity and water hydrolysis, respectively. The values of the constants depend on temperature. The K_b and K_w values for the reactor temperatures were

Table 1 | Details of primer pairs, PCR steps and amplicons

Target gene	Primer sequence	Primer concentration (μM)	Annealing T ($^{\circ}\text{C}$)	Cycle number	Polymerization time (s)	Amplicon size (bp)	Reference
16S rRNA	F: TGCCGCGTGTATGAAGAAGG R: TGCGGGTAACGTCAATTGCT	0.1	60.5	35	20	96	Poorasgari & Örmeci (2022a)
Int 1	F: CCT CCC GCA CGA TGA TC R: TCC ACG CAT CGT CAG GC	0.24	60.2	35	20	280	Diehl & LaPara (2010)
bla _{GES}	F: GAAACCAAACGGGAGACGC R: CTTGACCGACAGAGGCAACT	0.1	60	38	20	207	Mlynarcik <i>et al.</i> (2016)
bla _{OXA-48}	F: AGGCACGTATGAGCAAGATG R: TGGCTTGTTTGACAATACGC	0.1	58.5	38	15	189	Subirats <i>et al.</i> (2017)
bla _{IMP-27}	F: AAAGGCACTGTTTCCTCACA R: TCGCCAGCCAATAACTAACC	0.1	57.3	38	15	169	Mlynarcik <i>et al.</i> (2016)

T, temperature; μM , micro-mole per litre; bp, base pair.

estimated using the data and trends provided in the literature (Bates & Pinching 1949; Bandura & Lvov 2006). pH values that were used to calculate FAN concentrations were presented previously (Poorasgari & Örmeci 2022c).

2.5. Kinetics of target genes abundance

Equations (2)–(4) represent first-order, second-order and refined Collins–Selleck (CS) kinetics, respectively (Tchobanoglous *et al.* 2014). These equations were used to analyze the variation rate of gene abundance during aerobic digestion.

$$\ln \frac{[\text{gene}]_t}{[\text{gene}]_0} = -K_1 t \quad (2)$$

$$\frac{1}{[\text{gene}]_t} = \frac{1}{[\text{gene}]_0} + K_2 \quad (3)$$

$$\log_{10} \frac{[\text{gene}]_t}{[\text{gene}]_0} = \Lambda_{\text{CS}} (\log_{10} t - \log_{10} b) \quad (4)$$

where the absolute abundance of the target gene is denoted by $[\text{gene}]$, K_1 is the first-order rate constant in d^{-1} and K_2 is the second-order rate constant in $(\text{ng DNA}) \cdot (\text{copy number})^{-1} \cdot \text{d}^{-1}$. Λ_{CS} is the slope of the curve in the refined Collins–Selleck (CS) model.

2.6. Development of regression models

The predictors of regression modelling included TS, SCOD, SOP, FAN and $\text{NO}_3\text{-N}$. SOP, FAN, $\text{NO}_3\text{-N}$ and compounds responsible for SCOD constitute total dissolved solids (TDS), which is a part of TS. Nevertheless, our solids measurements made it clear that TDS accounted for less than 20% of TS. It was very likely that the remaining 80% contained a large portion of the total DNA, and hence, the target genes. That is why TS was considered a predictor, despite being partly represented by the other four predictors. The overall trends of FAN concentration (Supplementary material, Figure S3) were more meaningful than those of TAN (Poorasgari & Örmeci 2022c). This is the reason for choosing FAN for the ammonia-gene link in the model.

To reduce the effects of potential inaccuracies by measurement errors, the ratios of the genes and predictors were calculated and used in model development. The ratios were computed by dividing the value of the parameter at any given time by that at time zero. Equation (5) was used as the governing formula of single regression modelling:

$$\text{Gene ratio} = K_{\text{predictor}} \cdot (\text{Predictor ratio})^{\delta_{\text{predictor}}} + \beta \quad (5)$$

where the predictor-specific coefficient and exponent are signified by $K_{\text{predictor}}$ and $\delta_{\text{predictor}}$, respectively. The model intercept is represented by β . The goodness of the single regression models was assessed by the sum of squared errors (SSE) and the coefficient of determination (R^2). For the linear mode (L_S), where the index S signifies single regression, $\delta_{\text{predictor}}$ was set equal to one. Because of the toxicity of such compounds as ammonia and the competitive growth of bacteria (Gray 2004; Tchobanoglous *et al.* 2014), a portion of which carry ARGs, the relationship between the ARGs and the predictors could become non-linear. Moreover, the growth of ARGs-harboring bacteria in sludge digesters might follow Monod-type trends (Tchobanoglous *et al.* 2014). This would further strengthen the idea of non-linearity of the relationship between the number of bacteria, and hence, ARGs and bacterial food (i.e., biodegradable SCOD). Therefore, non-linear regression modelling may be more reasonable than linear regression modelling in some situations. For the non-linear mode of single regression (NL_S), $\delta_{\text{predictor}}$ was considered to be smaller or greater than one, and β was set at zero.

Equation (6) was the governing expression considered for multiple regression models:

$$\text{Gene ratio} = \sum \{K_{\text{predictor}} \cdot (\text{Predictor ratio})^{\delta_{\text{predictor}}}\} + \beta \quad (6)$$

Equation (7) is the expanded form of Equation (6):

$$\text{Gene ratio} = K_{\text{TS}} \cdot (\text{TS ratio})^{\delta_{\text{TS}}} + K_{\text{SCOD}} \cdot (\text{SCOD ratio})^{\delta_{\text{SCOD}}} + K_{\text{SOP}} \cdot (\text{SOP ratio})^{\delta_{\text{SOP}}} + K_{\text{FAN}} \cdot (\text{FAN ratio})^{\delta_{\text{FAN}}} + K_{\text{NO}_3-\text{N}} \cdot (\text{NO}_3 - \text{N ratio})^{\delta_{\text{NO}_3-\text{N}}} + \beta \quad (7)$$

Equation (7) was also used in both linear and non-linear modes. For the linear mode, all the exponents of the linear mode (L_M) were set to 1. The index M denotes multiple regression. To apply Equation (7) in the non-linear mode (NL_M), exponents differing from one were taken from NL_S models with the R^2 value higher than 0.1 and the R^2 value of their respective L_S models. There were two reasons for doing this as follows. The exponents of the NL_S models with $R^2 < 0.1$ would be unreliable. Additionally, using exponents of the NL_S models with R^2 values smaller than those of the respective L_S models would not be logical. Once the predictor-specific powers were applied to the predictor ratios, the multiple regression model was fitted to the experimental data. The goodness of the multiple regression models was investigated by their SSE, R^2 , adjusted R^2 (R^2_{Adj}) and P -values. Equation (8) was used to compute the R^2_{Adj} values. In this equation, n_p is the number of predictors and $n_{\text{data points}}$ is the number of data points.

$$R^2_{\text{Adj}} = 1 - \frac{(1 - R^2)(n_{\text{data points}} - 1)}{(n_{\text{data points}} - n_p - 1)} \quad (8)$$

2.7. Statistical analysis

To determine the significance of the differences among the time series data, the two-factor analysis of variance (ANOVA) was conducted. Excel software was used to perform the ANOVA and regression modelling. The significance level of all statistical calculations of this study was 0.05.

3. RESULTS AND DISCUSSION

3.1. Absolute abundance of target genes

Variations of absolute abundance over digestion time are shown in Figures 1(a)–1(d). Initially, the highest and lowest absolute abundances belonged to Int 1 (Figure 1(b)) and $\text{bla}_{\text{OXA-48}}$ (Figure 1(d)), respectively. All target genes declined as a result of aerobic digestion. AD.55 achieved the highest initial reductions in the absolute abundance of 16S rRNA, Int 1 and $\text{bla}_{\text{OXA-48}}$, among all reactors. Within the first five days, AD.55 removed 70, 89, 88 and 91% of the absolute abundance of 16S rRNA, Int 1, bla_{GES} and $\text{bla}_{\text{OXA-48}}$ genes, respectively (Figures 1(a)–1(d)). There was a slight increase in the abundance of Int 1 in AD.35 and 45 after 30 days. This could be due to the adaptation of Int 1-carrying bacteria to the environment of the aerobic digesters. As Int 1 is associated with various types of ARGs, solid retention times (SRTs) longer or lesser than 30 days are not recommended for aerobic digesters. The reactors were able to completely remove bla_{GES} and $\text{bla}_{\text{OXA-48}}$ in 20 and 15 days, respectively (Figures 1(c) and 1(d)). Practically, CRGs were eliminated within 15 days of aerobic digestion, which may be considered as the optimum SRT for future aerobic digesters. Burch *et al.* (2013) observed 90% decrease in the abundance of 16S rRNA and Int 1, when they conducted batch aerobic digestion of wastewater sludge at room temperature for 20 days. Their reduction rate was comparable with those observed for AD.35. Nevertheless, the reduction trends over time were linear. It is worth mentioning that the abundance unit used by Burch *et al.* (2013) was copy number per mL, which could explain the difference between the decline patterns of their study and the current investigation. Jang *et al.* (2019) carried out thermophilic aerobic digestion on anaerobically digested sludge for 12 days. They observed an initial increase of the 16S rRNA by 3.68 folds, but an overall reduction of 80% was realized at the end of the aerobic digestion. They also demonstrated that Int 1 was reduced by 97% in a non-linear manner. On the other hand, Ma *et al.* (2011) observed an increase in Int 1 expressed as copy number per g dry solids after the aerobic digestion of a thermally hydrolysed and anaerobically digested sludge. Therefore, the effect of aerobic digestion on the abundance of target genes may depend on sludge preconditions and the reported unit of abundance. The trends of variation for 16S rRNA between AD.35 and 55, and for bla_{GES} between AD.35 and 45, were not significantly different (P -value > 0.05). All the other trends for each target gene between the reactors were significantly different (P -value < 0.05). The absolute abundance of $\text{bla}_{\text{IMP-27}}$ was below the quantifiable level.

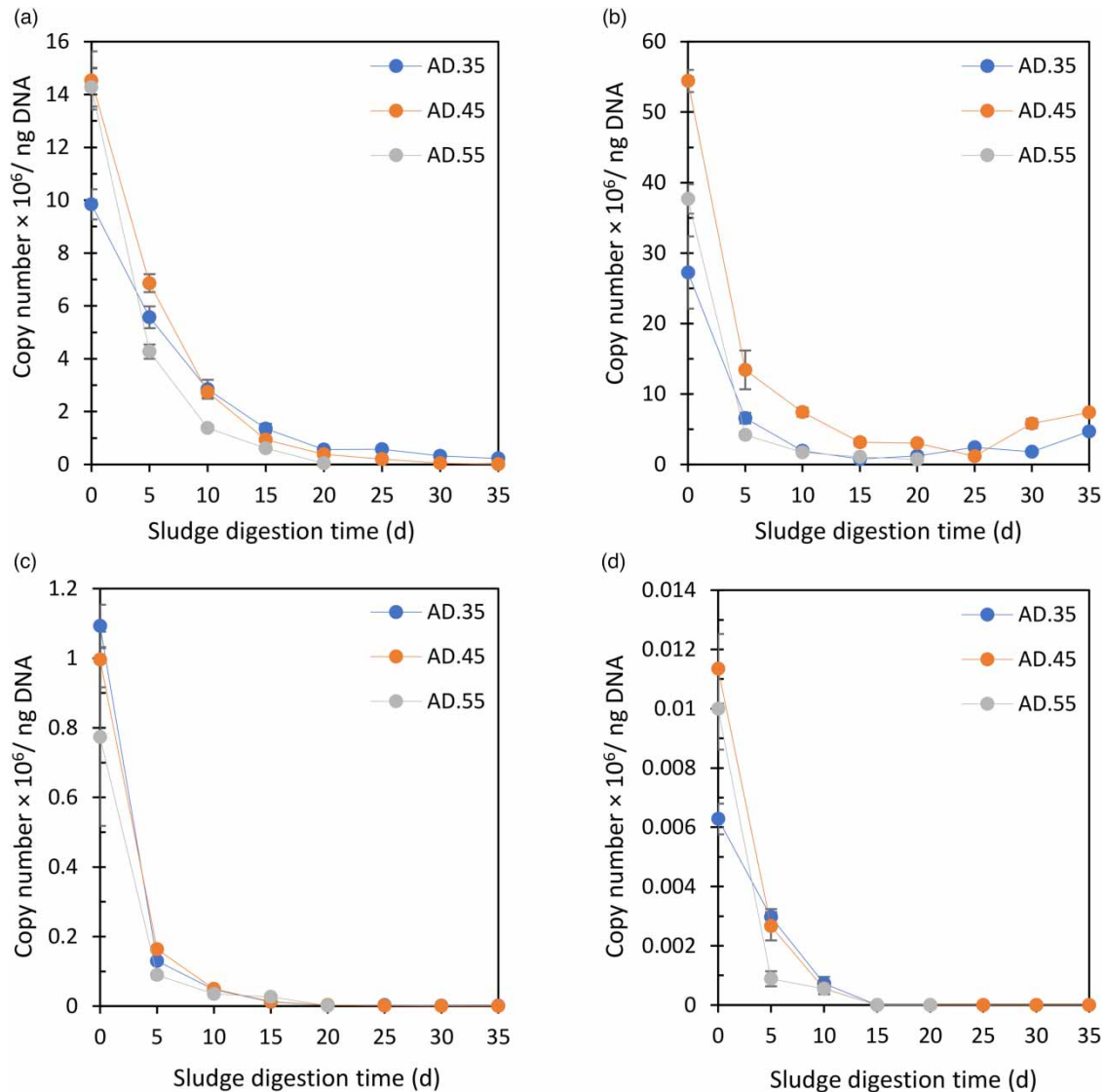


Figure 1 | Absolute abundance of the target genes: (a) 16S rRNA, (b) Int 1, (c) bla_{GES}, and (d) bla_{OXA-48} for batch reactors.

3.2. Correlations between target genes

Similar trends of the absolute abundance of the target genes (Figures 1(a)–1(f)) suggested the existence of correlations between the CRGs. Linear correlations between the target genes were established, and their R^2 and P -values are shown in Table 2. All the correlations were statistically significant. The overall strength of the correlations in terms of R^2 can be put in the following order: AD.55 > AD.45 > AD.35. With these correlations, linear equations can be readily built. Such equations are helpful for the operation of sludge digesters because they enable us to estimate the abundance and removal of all target genes by quantifying only one of the target genes. Correlations between 16S rRNA, ARGs and Int 1 have been reported for influents of WWTPs (Chen & Zhang 2013). In the current investigation, the high abundance of Int 1 and its correlation with bla_{GES} and bla_{OXA-48} imply that sludge bacteria were capable of horizontal transfer. Jang *et al.* (2019) analyzed the correlations between Int 1 and multiple ARGs by the Pearson correlation coefficient (r or square root of R^2). The reported r values were between 0.021 and 0.983, which translates to an R^2 range of 4×10^{-4} – 0.966. Also, the same study reported both positive and negative correlations. In the present research, R^2 values were greater than 0.8, and no negative correlation was observed. The R^2 values obtained from correlating Int 1 and the other target genes in the current investigation are more in line with the range 0.37–0.99 reported by Jang *et al.* (2020). The discrepancies could be due to the differences in the process, reactor feed and types of correlated ARGs.

Table 2 | R^2 and P -values (subscripts) for linear inter-gene correlations

Reactor	Target gene	16S rRNA	Int 1	bla _{GES}
AD.35	16S rRNA			
	Int 1	0.808 _{0.002}		
	bla _{GES}	0.816 _{0.002}	0.975 _{0.000}	
	bla _{OXA-48}	0.974 _{0.000}	0.895 _{0.000}	0.880 _{0.001}
AD.45	16S rRNA			
	Int 1	0.906 _{0.000}		
	bla _{GES}	0.908 _{0.000}	0.987 _{0.000}	
	bla _{OXA-48}	0.941 _{0.000}	0.986 _{0.000}	0.994 _{0.000}
AD.55	16S rRNA			
	Int 1	0.963 _{0.003}		
	bla _{GES}	0.967 _{0.003}	1 _{0.000}	
	bla _{OXA-48}	0.961 _{0.003}	0.999 _{0.000}	0.998 _{0.000}

3.3. Kinetics of target genes

Table 3 displays kinetic parameters and respective coefficients of determination. The rate constants were specific for each target gene. The magnitude of the rate constants followed the order bla_{OXA-48} > bla_{GES} > 16S rRNA. In addition, there was an overall increase in the absolute values of K_1 with increasing temperature. The R^2 values of the first-order fit for Int 1 in AD.35, AD.45 and AD.55 were 0.212, 0.368 and 0.870, respectively. This range is considerably higher than the range 0.01–0.54 reported by Diehl & LaPara (2010), where wastewater solids were digested at mesophilic and thermophilic temperatures. Notwithstanding, the R^2 value for the first-order fit of Int 1 in AD.35 was 2.55 times smaller than the corresponding value obtained by Diehl & LaPara (2010) from an aerobic digester operated at 37 °C. On the other hand, Jang *et al.* (2019) fitted the first-order kinetic equation smoothly to the abundance of Int 1 measured during aerobic digestion at 55 °C, similar to the current investigation. Second-order and CS kinetic fittings were also carried out. The second-order kinetic equation has previously been used to describe temporal variations of the target genes during physical and chemical processes only (Yu *et al.* 2017; Gao *et al.* 2020). The current study, however, shows that the second-order kinetics represented a variation rate of Int 1 and bla_{OXA-48} in AD.55 more acceptably, i.e., higher R^2 values than the first-order kinetics did. CS kinetics could not be performed for bla_{OXA-48} due to a lack of sufficient data. In the case of the other targets, the CS model yielded more accurate fits than the second-order model.

Table 3 | Kinetic parameters and R^2 for kinetics of the target genes

Reactor	Target gene	K_1	$R^2 - 1^{st}$ order	K_2	$R^2 - 2^{nd}$ order	Λ_{CS}	b	$R^2 - CS$
AD.35	16S rRNA	-0.110	0.970	-0.120	0.887	1.680	4.086	0.966
	Int 1	-0.042	0.212	-0.006	0.03	0.733	0.002	0.065
	bla _{GES}	-0.205	0.948	-39.537	0.775	2.823	2.723	0.969
	bla _{OXA-48}	-0.215	0.970	-120.98	0.857			
AD.45	16S rRNA	-0.205	0.980	-2.444	0.438	3.166	5.164	0.863
	Int 1	-0.057	0.368	-0.008	0.146	0.584	0.250	0.248
	bla _{GES}	-0.272	0.995	-33.399	0.681	3.127	3.229	0.944
	bla _{OXA-48}	-0.303	0.999	-173.04	0.871			
AD.55	16S rRNA	-0.269	0.954	-0.907	0.560	2.920	3.814	0.825
	Int 1	-0.186	0.870	-0.069	0.981	1.2627	0.879	0.999
	bla _{GES}	-0.283	0.916	-33.737	0.539	2.606	2.600	0.703
	bla _{OXA-48}	-0.291	0.872	-173.32	0.989			

It is important to note that the type of kinetics can change the design equations of continuous flow reactors including continuous flow aerobic digesters. Equations for the completely stirred mixed flow reactor (CMFR) and the plug-flow reactor (PFR) are provided below, where τ is SRT. Mathematical operations and assumptions used for deriving Equations (9)–(14) are available in Supplementary material S4.

CMFR:

$$\text{First – order: } [\text{gene}]_{\text{out}} = \frac{[\text{gene}]_{\text{in}}}{1 + K_1 \tau} \quad (9)$$

$$\text{Second – order: } [\text{gene}]_{\text{out}} = \frac{[\text{gene}]_{\text{in}}}{1 + [\text{gene}]_{\text{out}} K_2 \tau} \quad (10)$$

$$\text{CS: } [\text{gene}]_{\text{out}} = [\text{gene}]_{\text{in}} \times \left[1 - \frac{\Lambda_{\text{CS}}}{b^{\Lambda_{\text{CS}}}} \tau^{\Lambda_{\text{CS}}} \right] \quad (11)$$

PFR:

$$\text{First – order: } [\text{gene}]_{\text{out}} = [\text{gene}]_{\text{in}} \times \exp\{-K_1 \tau\} \quad (12)$$

$$\text{Second – order: } [\text{gene}]_{\text{out}} = \frac{[\text{gene}]_{\text{in}}}{1 + [\text{gene}]_{\text{in}} K_2 \tau} \quad (13)$$

$$\text{CS: } [\text{gene}]_{\text{out}} = [\text{gene}]_{\text{in}} \times \exp\left\{-\frac{\Lambda_{\text{CS}}}{b^{\Lambda_{\text{CS}}}} \tau^{\Lambda_{\text{CS}}}\right\} \quad (14)$$

3.4. Relative abundance of target genes

Relative abundances of the target genes are shown in Figures 2(a)–2(c). Except for bla_{GES} of AD.35 and AD.45, the trends of the relative abundance were different from those of the absolute abundance. In AD.55, the relative abundance of the targets fluctuated during digestion. In that reactor, the relative abundance levels showed peaks on days 10 and 15. The exponential increase in the relative abundance of Int 1 gene observed in AD.34 and 45 (Figure 2(a)) suggests that a much higher proportion of Int 1-carrying bacteria than the total bacterial community adapted themselves with the conditions of the digesters. In AD.55, this happened much earlier, and it was followed by a sharp decline (i.e., a peak that reached 38 on day 10). The sharp increase could be due to the dominance of thermophilic bacteria carrying Int 1 and the effects of the higher temperature on biokinetics. The sharp decline might be explained by the significant amount of water loss due to the evaporation at 55 °C. We discussed the water evaporation rate and its effects on biokinetics at various temperatures, in a previous study (Poorasgari & Örmeci 2022b). The fluctuations in the relative abundance of Int 1 are in line with the observations of Diehl & LaPara (2010). In AD.35 and AD.45, the relative abundance of the 16S rRNA increased in a logarithmic manner. The relative abundance of $\text{bla}_{\text{OXA-48}}$ experienced a sharp decline in a rather linear manner within 15 days, after which the target gene became un-quantifiable.

3.5. Regression models

The predictor ratios of the regression models were obtained from the concentrations of TS, SCOD, SOP, TAN and $\text{NO}_3\text{-N}$, which have been provided elsewhere (Poorasgari & Örmeci 2022b, c). The results of L_S and NL_S modelling have been summarised in Supplementary material, Table S5. R^2 values of up to 0.658 were observed for the L_S models. Also, the SOP concentration non-linearly correlated with bla_{GES} in AD.55 with an R^2 value of 0.999. Association has been observed between ARGs and phosphate (Ping *et al.* 2022). Initial nitrate concentrations were almost zero just before starting the aerobic digestion experiment. Nitrate concentration increased with time, as the reactor was being constantly aerated (Poorasgari & Örmeci 2022c). Chemical oxidation of ammonia requires specific conditions such as catalysts (Johnston *et al.* 2022). Therefore, the increase in nitrate concentration would have to be due to biological nitrification (Tchobanoglous *et al.* 2014). After 25 days of operation, there was a significant drop of nitrate concentration in AD.45. This suggests that there was denitrification, which was likely to occur because of the low concentration of dissolved oxygen (Poorasgari 2023) and probable localized anoxic islands deep inside the sludge aggregates. Thus, the observed increase and decrease in nitrate concentration (Poorasgari & Örmeci 2022c) signified the occurrence of nitrification and denitrification, respectively (Tchobanoglous *et al.* 2014). Bacteria involved in the biological conversion of inorganic nitrogen species may be able to carry ARGs (León & Roy 2003; Agersø & Sandvang 2005; Zhang *et al.* 2009). In the current investigation, there were strong negative correlations between the nitrate

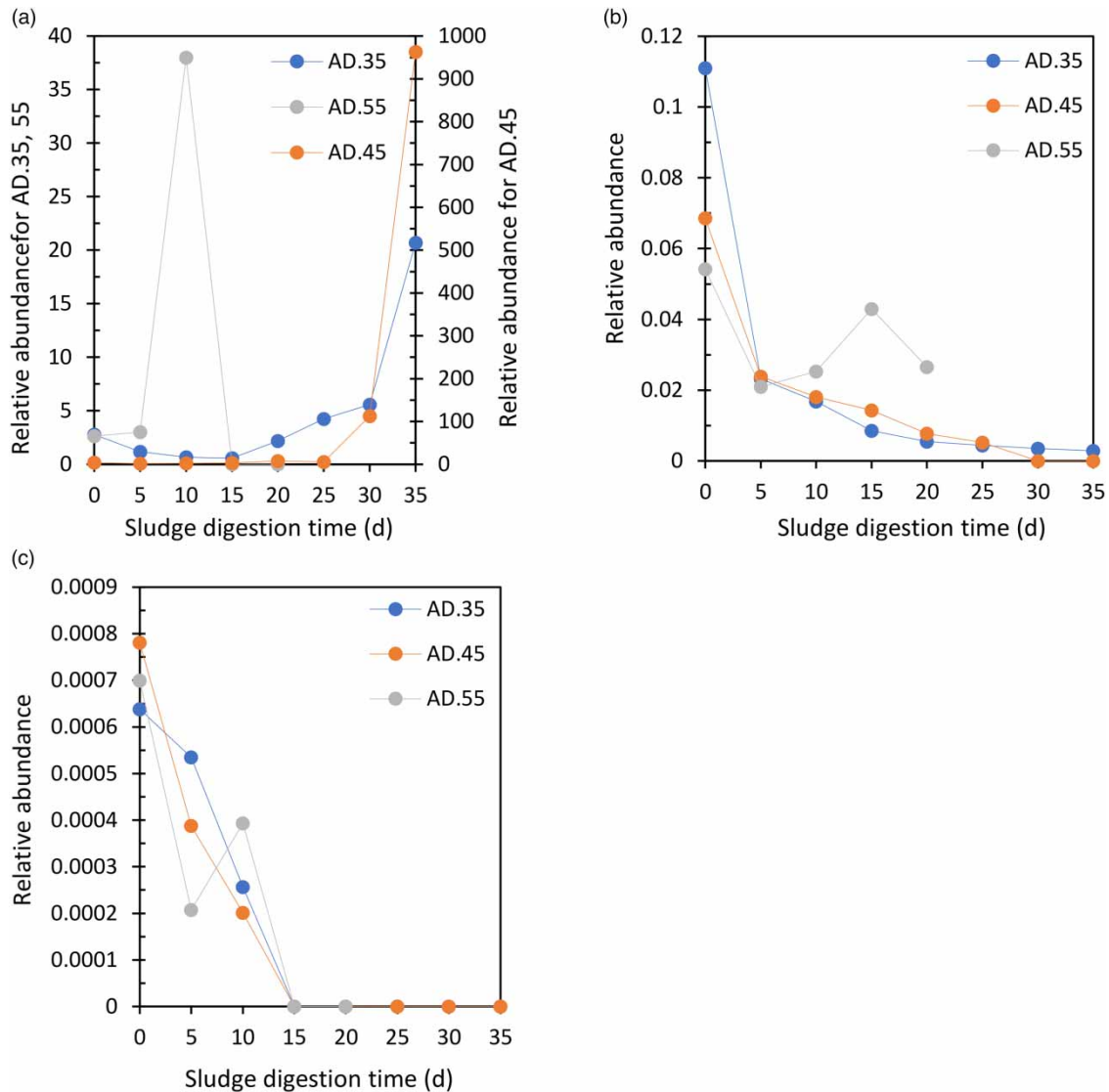


Figure 2 | Relative abundance of the target genes: (a) Int 1, (b) bla_{GES} , and (c) bla_{OXA-48} .

concentration and the target genes ($R^2 = 0.81-1$) in a non-linear manner. This is interesting because non-linear negative correlations have also been observed between ARGs and denitrification genes (Sun *et al.* 2017).

The results of L_M and NL_M modelling are shown in Figures 3(a)–3(h) and in Table 4. The L_M models (Figures 3(a), 3(c), 3(e) and 3(g)) resulted in more fluctuations than the NL_M models (Figures 3(b), 3(d), 3(f) and 3(h)). The L_M models of AD.35 and 45 estimated some negative ratios of the target genes (Figures 3(a), 3(c), 3(e) and 3(g)). The NL_M models, on the other hand, did not have this issue except for the bla_{OXA-48} ratios during digestion in AD.35 (Figure 3(h)).

Table 4 shows further details of calculations for L_M and NL_M . The coefficients and exponents of the obtained models were different for each case. An explanation would be the dissimilar interactions between parameters affecting the abundance of genes in various kinds of sludge bacteria during aerobic digestion. Int 1 and CRGs could be nonexistent in some bacteria. Also, not all bacteria have the same copy number of the 16S rRNA (Větrovský & Baldrian 2013). In addition, the temperature at which anaerobic digesters operate impacts the composition of the microbial population (Gagliano *et al.* 2015). Similar temperature effects could be expected during aerobic digestion. Consumption of substrates and secretion of extracellular material may differ among microbial communities. As expected, the coefficients of the predictors in the L_M models were different from those in the NL_M models. The highest and lowest coefficients were 6.753 for the SOP ratio and -5.316 for the TS ratio, respectively. Both mentioned values belonged to the NL_M model for 16S rRNA ratios during digestion in AD.45. The lowest exponent was -8.262 , which belonged to the TS ratio of the NL_M model for bla_{OXA-48}

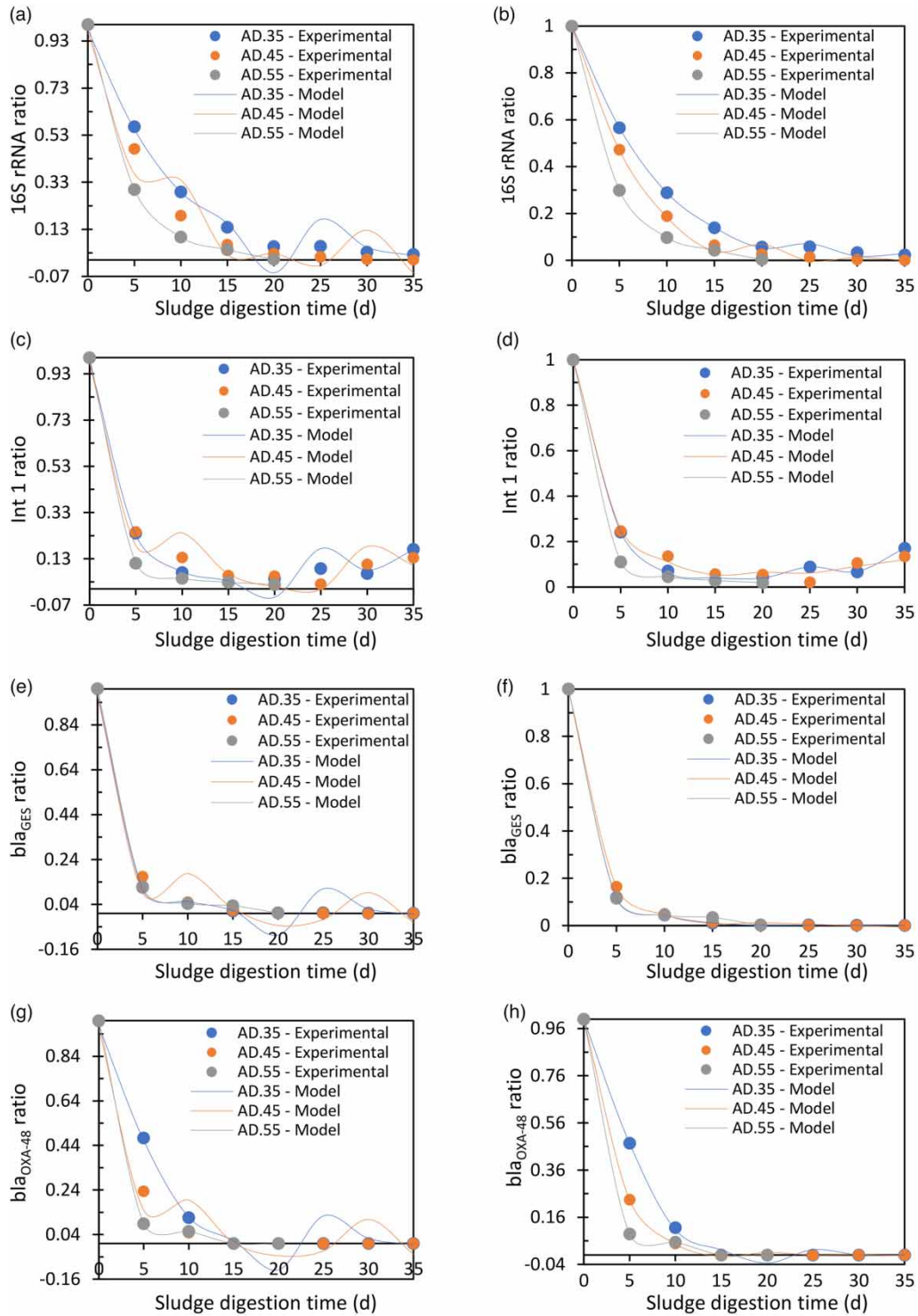


Figure 3 | Experimental and modelled ratios of the target genes using multiple regression modelling: (a), (c), (e) and (g) show linear model fittings; (b), (d), (f) and (h) show non-linear model fittings.

ratios during digestion in AD.55. The weight of the parameters can be assessed by the absolute values of the coefficients and the exponents. Overall, the lowest absolute values of the coefficients and the exponents occurred in the FAN ratio in both L_M and NL_M models. This suggests that FAN concentration was not as important as the other predictors. FAN can be toxic at concentrations above 150 mg N/L (Gray 2004). FAN concentrations of AD.45 and AD.55 remained below the mentioned level. However, AD.35 exceeded the toxic level on day 35 (Supplementary material, Figure S3). Thus, FAN should have contributed as a nutrient to the abundance of the target

Table 4 | Coefficients, exponents and goodness parameters of the models

Reactor	Gene	Model	K_{TS}	δ_{TS}	K_{SCOD}	δ_{SCOD}	K_{SOP}	δ_{SOP}	K_{FAN}	δ_{FAN}	K_{NO3-N}	δ_{NO3-N}	β	SSE	R^2	R^2_{Adj}	P-value	
AD.35	16S rRNA	L_M	0.075	1	-0.306	1	0.332	1	-0.000	1	-0.147	1	1.023	0.027	0.968	0.889	0.078	
		NL_M	0.546	1	-2.452	-0.675	0.256	1	-0.281	-0.439	2.281	-1.584	0.649	0.000	1	0.998	0.001	
	Int 1	L_M	-0.250	1	-0.369	1	0.160	1	0.000	1	-0.109	1	1.551	0.015	0.980	0.931	0.048	
		NL_M	-0.274	-2.241	0.530	-3.593	-0.016	1	0.000	1	0.724	-1.118	0.035	0.000	1	0.999	0.001	
	bla _{GES}	L_M	-0.401	1	-0.316	1	0.233	1	0.000	1	-0.116	1	1.579	0.022	0.973	0.907	0.065	
		NL_M	-0.394	-1.105	-0.137	-2.377	0.030	1	-0.033	-0.783	1.255	-3.098	0.280	0.000	1	1	0.000	
	bla _{OXA-48}	L_M	0.105	1	-0.491	1	0.218	1	0.000	1	-0.129	1	1.272	0.031	0.967	0.883	0.081	
		NL_M	-1.144	-1.578	0.686	-2.495	-0.008	-3.303	0.000	1	1.137	-1.172	0.328	0.002	0.998	0.992	0.006	
	AD.45	16S rRNA	L_M	-0.126	1	-0.378	1	0.043	1	0.006	1	-0.081	1	1.505	0.056	0.935	0.774	0.154
			NL_M	-5.316	-5.708	0.098	-2.675	6.753	-2.856	-0.745	-1.091	0.216	-2.009	-0.005	0.003	0.997	0.989	0.008
		Int 1	L_M	-0.298	1	-0.371	1	0.060	1	0.003	1	-0.049	1	1.629	0.025	0.965	0.879	0.084
			NL_M	1.556	-1.836	-0.026	-2.769	-1.966	-0.825	-0.552	-0.546	1.930	-0.87	0.058	0.003	0.996	0.986	0.010
bla _{GES}		L_M	-0.353	1	-0.380	1	0.052	1	0.002	1	-0.038	1	1.686	0.037	0.956	0.847	0.106	
		NL_M	2.584	-5.760	0.040	-4.856	-2.432	-3.652	-0.889	-1.162	1.705	-2.165	-0.007	0.000	1	0.999	0.001	
bla _{OXA-48}		L_M	-0.280	1	-0.402	1	0.048	1	0.003	1	-0.048	1	1.644	0.048	0.944	0.803	0.135	
		NL_M	-0.019	-2.427	0.038	-2.807	0.108	-2.422	-0.353	-0.752	1.227	-1.822	-0.001	0.000	1	0.999	0.001	
AD.55		16S rRNA	L_M	0.000	1	-0.973	1	0.318	1	-0.003	1	-0.048	1	1.706	0.000	1	1	0.000
			NL_M	0.339	-4.258	-0.875	-4.769	0.000	-2.367	1.624	-0.839	-0.059	-1.276	-0.029	0.000	1	1	0.000
		Int 1	L_M	0.000	1	-1.354	1	0.420	1	-0.003	1	-0.052	1	1.988	0.000	1	1	0.000
			NL_M	-0.062	-2.476	0.582	-3.527	0.000	-1.756	0.001	0.775	0.504	-1.041	-0.025	0.000	1	1	0.000
	bla _{GES}	L_M	0.000	1	-1.329	1	0.410	1	-0.003	1	-0.050	1	1.972	0.000	1	1	0.000	
		NL_M	0.217	-4.326	-1.193	-5.344	0.000	-2.561	2.517	-1.006	-0.524	-1.371	-0.016	0.000	1	1	0.000	
	bla _{OXA-48}	L_M	0.000	1	-1.432	1	0.431	1	-0.002	1	-0.052	1	2.055	0.000	1	1	0.000	
		NL_M	-0.216	-8.262	1.071	-4.069	0.000	-2.151	-0.897	-0.653	1.041	-1.008	0.000	0.000	1	1	0.000	

genes in AD.45 and AD.55 during the whole digestion process. In AD.35, on the other hand, FAN may have acted as an inhibitor on the abundance of the target genes. Notwithstanding, the negligible and, in some cases, no contribution of FAN to the target gene ratio suggests that ammonia could have behaved as both nutrient and inhibitor in AD.35 and 45. That is, the antagonistic effects of nourishment and inhibition would have resulted in the diminished impact of the FAN ratio.

The absolute values of β for the NL_M models were smaller than those for the L_M models. Since the intercept is meant to correct the constant error of the model, smaller absolute values of β are more acceptable. In other words, estimations of the model are more reliable when the absolute value of β is smaller. Given the maximum value of 1 for the target gene ratio, an intercept with the absolute value equal to or greater than 1 means that the error compensation largely contributed to the estimated target gene ratio. An extreme case was the L_M model for bla_{OXA-48} ratios of AD.55, where β was 2.055. Evidently, applying the exponents of the NL_S models (Supplementary material, Table S5) to Equation (7) improved the reliability of the model by reducing the absolute value of β . The intercept of the NL_M model went down to zero for bla_{OXA-48} of AD.55. As well, the exponents' applications reduced the SSE of multiple regression modelling for AD.35 and AD.45. The most noticeable one was observed in the case of the 16S rRNA of AD.45, where the SSE of the NL_M model was 0.053 below that of its corresponding L_M model. The overall R^2 and R^2_{Adj} values of the NL_M models of AD.35 and AD.45 were larger than those of the respective L_M models. The most noticeable case was the 16S rRNA of AD.45. In that case, R^2 and R^2_{Adj} of the NL_M model were 1.07 and 1.28 times higher than those parameters in the L_M model. The multiple regression models of AD.55 had the highest R^2 and R^2_{Adj} values.

The NL_M models of AD.35 and 45 had lower overall P -values than the respective L_M models. The most striking of all occurred for bla_{OXA-48} of AD.45, where the P -value of the NL_M model was 135 times smaller than that of the corresponding L_M model. Also, a 106-fold decrease was observed for bla_{GES} of the same reactor. The L_M models developed for 16S rRNA, bla_{GES} and bla_{OXA-48} of AD.35, and for all target genes of AD.45 were insignificant (P -value > 0.05). On the other hand, the P -value was less than 0.05 for all multiple regression models of AD.55, and for all NL_M models of AD.35 and AD.45. The good fit at 55 °C is, at least partly, due to the fact that the number of observations and the number of model predictors were the same (Izenmann 2008). Therefore, the L_M and NL_M models for AD.55 need to be tested with other datasets with more observations, before they are considered by engineers in charge of aerobic digesters. Compared to the multiple regression models reported (Tong *et al.* 2018; Tong *et al.* 2019a; Jang *et al.* 2020; Poorasgari & Örmeci 2022a), the NL_M models of the current investigation resulted in a lower range for the P -value and a higher range for the R^2 value. Nevertheless, the cited investigations developed those models for processes other than the aerobic digestion of wastewater sludge.

It would be interesting to see how the goodness parameters of single and multiple regression models compare. To do so, the SSE values of the multiple regression models were divided by the minimum SSE values of their respective linear regression models, and the R^2_{Adj} values of the multiple regression models were divided by the

maximum R^2 values of their respective single regression models. The results of the former and latter divisions are termed the SSE ratio and the R^2_{Adj} ratio, respectively. The multiple regression model is more accurate than the single regression model when the SSE ratio < 1 and the R^2_{Adj} ratio > 1 . On the other hand, the single regression model is more accurate when the SSE ratio > 1 and the R^2 ratio < 1 . The computed SSE and R^2_{Adj} ratios are presented in Table 5. All target genes and reactors had SSE ratios smaller than 1 and zero in several cases. Also, the R^2_{Adj} ratio of AD.35 and AD.45 for all target genes was greater than 1. In the case of AD.35 for bla_{GES} , the R^2_{Adj} ratio of the L_M model (Table 4) was 1.95 times larger than the highest R^2 value of the respective L_S model (Supplementary material, Table S5). Note that these calculations were carried out using all the decimals produced by Excel for the goodness parameters of the involved models but the values in Tables 4 and Supplementary material, Table S5 are shown with three decimals. If the R^2 values of the single (Supplementary material, Table S5) and multiple (Table 4) regression models are compared, the multiple regression models seem even more precise and accurate. From this perspective, multiple regression modelling with the general formula shown in Equation (7) was a stronger method of estimating the target gene ratio than single regression modelling. In the current study, the superiority of non-linear models to linear models is obvious. This may be explained by the effects of toxicity and biokinetics on the relationship between ARG abundances and model predictors (see Section 2.6). The effects of non-linearity on the decisions made by the designers and operators of aerobic sludge digesters would depend on the predictor ratios and the exponent of the predictor ratios. A worked example of using the non-linear form of Equation (7) is provided in Supplementary material S6.

Table 5 | Comparison between single (L_S and NL_S) and multiple (L_M and NL_M) regression models

Reactor	Target gene	Model	SSE ratio	R^2_{Adj} ratio	
AD.35	16S rRNA	L	0.087	1.409	
		NL	0.001	1.214	
	Int 1	L	0.048	1.616	
		NL	0.002	1.045	
	bla_{GES}	L	0.049	1.946	
		NL	0.000	1.008	
	bla_{OXA-48}	L	0.196	1.544	
		NL	0.043	1.054	
	AD.45	16S rRNA	L	0.145	1.392
			NL	0.001	1.08
Int 1		L	0.124	1.216	
		NL	0.028	1.048	
bla_{GES}		L	0.148	1.205	
		NL	0.001	1.001	
bla_{OXA-48}		L	0.615	1.165	
		NL	0.579	1	
AD.55		16S rRNA	L	0.000	1
			NL	0.000	1
	Int 1	L	0.000	1	
		NL	0.000	1	
	bla_{GES}	L	0.000	1	
		NL	0.000	1	
	bla_{OXA-48}	L	0.000	1	
		NL	0.000	1	

A challenge to comparisons between the multiple and single regression models is the argument of overfitting (von Neumann *et al.* 2016). Increasing the number of predictors in regression models can make the estimated datapoints closer to their experimental counterparts for a specific set of data, while the model might not fit a

new set of data. To tackle this issue, a new model was constructed in which the ratios of all the predictors were unified (i.e., added together). Also, a proportionality factor or a unified constant (K_U) was incorporated into the equation. This way, only one predictor is involved from a mathematical point of view, which settles the concern about overfitting. This unified single regression model is expressed by Equation (15).

$$\text{Gene ratio} = K_U \cdot [(\text{TS ratio})^{\delta_{\text{TS}}} + (\text{SCOD ratio})^{\delta_{\text{SCOD}}} + (\text{SOP ratio})^{\delta_{\text{SOP}}} + (\text{FAN ratio})^{\delta_{\text{FAN}}} + (\text{NO}_3 - \text{N ratio})^{\delta_{\text{NO}_3 - \text{N}}}] + \beta \quad (15)$$

The single predictor in the new model, indeed, represents all predictors of the NL_M model (Equation (7)). The unified model was applied in both linear ($L_{S,U}$) and non-linear ($NL_{S,U}$) fashions, in accordance with the explanations in Section 2.6. Basically, the unified model would be more practical for engineering purposes, as its equation is simpler than Equation (7). The unified model was first tested with the experimental data obtained from anaerobic sludge digesters operated at the same temperature and with the same feed. The test was relatively successful as the experimental data fitted the unified models well (Poorasgari 2023). However, the unified single regression performed poorly for the observations of the current study conducted on aerobic digestion. This was quite obvious in the case of the reactor AD.35 with both $L_{S,U}$ and $NL_{S,U}$ models (Supplementary material, Figure S7 and Table S8). Finally, parameters such as pH could be incorporated into the regression models. However, including additional parameters would further complicate the models and veer them off the ASN conceptual model.

4. CONCLUSIONS

Experiments and data analyses were conducted to mathematically describe the ASN nexus during the aerobic digestion of wastewater sludge. The aerobic digestion process reduced the absolute abundance of the target genes quantified in the current study. The absolute value of K_1 followed the order $\text{bla}_{\text{OXA-48}} > \text{bla}_{\text{GES}} > 16\text{S rRNA}$. Single and multiple regression models were established to predict CRGs as a function of TS, SCOD, SOP, FAN and $\text{NO}_3\text{-N}$. Calculations of the SSE and R^2 values indicated that the multiple regression models, including L_M and NL_M , were more accurate than the single regression models. In multiple regression models, FAN was the lowest impact predictor among the other four. The absolute value of β ranged 1.023–2.055 and 0–0.649 for the L_M and NL_M models, respectively. Consistent with this, the highest SSE value of NL_M modelling was up to 18.7 times, smaller than that of L_M modelling. Also, the NL_M models of AD.35 and AD.45 were superior to their L_M counterparts in terms of R^2 , R^2_{Adj} and P -values. This study, for the first time, demonstrated that the non-linear multiple regression models can estimate the abundance of target genes during the aerobic digestion of sludge. These models can help engineers and operators determine the SRT and digestion temperature required for removing the target genes of interest at a given rate during the aerobic digestion of wastewater sludge.

ACKNOWLEDGEMENTS

This research was funded by the Natural Sciences and Engineering Research Council of Canada, Discovery Grant program and Jarislowsky Chair Funds awarded to Professor Banu Örmeci.

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

- Agersø, Y. & Sandvang, D. (2005) Class 1 integrons and tetracycline resistance genes in *alcaligenes*, *arthrobacter*, and *pseudomonas* spp. isolated from pigsties and manured soil, *Appl. Environ. Microb.*, **71**, 7941–7947. <https://doi.org/10.1128/AEM.71.12.7941-7947.2005>.
- An, X.-L., Chen, Q.-L., Zhu, D., Zhu, Y.-G., Gillings, M. R. & Su, J.-Q. (2018) Impact of wastewater treatment on the prevalence of integrons and the genetic diversity of integron gene cassettes, *Appl. Environ. Microb.*, **84**, e02766–17. <https://doi.org/10.1128/AEM.02766-17>.

- Baird, R., Eaton, A. D. & Rice, E. W. (2017) *Standard Methods for the Examination of Water and Wastewater*, 23rd edn. Washington, DC, USA: APHA, AWWA, WEF.
- Bancroft, E. A. (2007) Antimicrobial resistance: It's not just for hospitals, *JAMA*, **298**, 1803–1804. <https://doi.org/10.1001/jama.298.15.1803>.
- Bandura, A. V. & Lvov, S. N. (2006) The ionization constant of water over wide ranges of temperature and density, *J. Phys. Chem. Ref. Data*, **35**, 15–30.
- Bates, R. G. & Pinching, G. D. (1949) Acidic dissociation constant of ammonium ion at 0 ° to 50 °C, and the base strength of ammonia, *J. Res. Nat. Bur. Stan.*, **42**, 419–430.
- Brezonik, P. L. & Arnold, W. A. (2011) *Water Chemistry*. New York, NY, USA: Oxford University Press.
- Burch, P. R., Sadowsky, M. J. & LaPara, T. M. (2013) Aerobic digestion reduces the quantity of antibiotic resistance genes in residual municipal wastewater solids, *Front. Microbiol.*, **4**, 17. <https://doi.org/10.3389/fmicb.2013.00017>.
- Cassini, A., Högberg, L. D., Plachouras, D., Quattrocchi, A., Hoxha, A., Simonsen, G. S., Colomb-Cotinat, M., Kretzschmar, M. E., Devleeschauwer, B., Cecchini, M., Ouakrim, D. A., Oliveira, T. C., Struelens, M. J., Suetens, C. & Monnet, D. L. (2019) Burden of AMR collaborative group. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: A population-level modelling analysis, *Lancet Infect. Dis.*, **19**, 56–66. [http://dx.doi.org/10.1016/S1473-3099\(18\)30605-4](http://dx.doi.org/10.1016/S1473-3099(18)30605-4).
- Chen, H. & Zhang, M. (2013) Occurrence and removal of antibiotic resistance genes in municipal wastewater and domestic sewage treatment systems in eastern China, *Environ. Int.*, **55**, 9–14. <https://doi.org/10.1016/j.envint.2013.01.019>.
- Davies, J. & Davies, D. (2010) Origins of evolution and antibiotic resistance, *Microbiol. Mol. Biol. Rev.*, **74**, 417–433. <https://doi.org/10.1128/MMBR.00016-10>.
- Diehl, D. L. & LaPara, T. M. (2010) Effect of temperature on the fate of genes encoding tetracycline resistance and the integrase of class 1 Integrons within anaerobic and aerobic digesters treating municipal wastewater solids, *Environ. Sci. Technol.*, **44**, 9128–9133. <https://doi.org/10.1021/es102765a>.
- Dueholm, M. K. D., Nierychlo, M., Andersen, K. S., Rudkjøbing, V., Knutsson, S., Consortium, M. G., Albertsen, M. & Nielsen, P. H. (2022) MiDAS 4: A global catalogue of full-length 16S rRNA gene sequences and taxonomy for studies of bacterial communities in wastewater treatment plants, *Nat. Commun.*, **13**, 1908. <https://doi.org/10.1038/s41467-022-29438-7>.
- Ekwanzala, M. D., Dewar, J. B., Kamika, I. & Momba, M. N. B. (2020) Genome sequence of carbapenem-resistant *Citrobacter koseri* carrying blaOXA-181 isolated from sewage sludge, *J. Glob. Antimicrob. Resist.*, **20**, 94–97. <https://doi.org/10.1016/j.jgar.2019.07.011>.
- Gagliano, M. C., Braguglia, C. M., Gallipoli, A., Gianico, A. & Rossetti, S. (2015) Microbial diversity in innovative mesophilic/thermophilic temperature-phased anaerobic digestion of sludge, *Environ. Sci. Pollut. R.*, **22**, 7339–7348. <https://doi.org/10.1007/s11356-014-3061-y>.
- Gao, J.-F., Duan, W.-J., Zhang, W.-Z. & Wu, Z.-L. (2020) Effects of persulfate treatment on antibiotic resistance genes abundance and the bacterial community in secondary effluent, *Chem. Eng. J.*, **382**, 121860. <https://doi.org/10.1016/j.cej.2019.05.221>.
- Gihring, T. M., Green, S. J. & Schadt, C. W. (2012) Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes, *Environ. Microbiol.*, **14**, 285–290. <https://doi.org/10.1111/j.1462-2920.2011.02550.x>.
- Goetsch, H. E., Love, N. G. & Wigginton, K. R. (2020) Fate of extracellular DNA in production of fertilizers from source-separated urine, *Environ. Sci. Technol.*, **54**, 1808–1815. <https://doi.org/10.1021/acs.est.9b04263>.
- Gray, N. F. (2004) *Biology of Wastewater Treatment. Second Edition*. London, UK: Imperial College Press.
- Izenmann, A. (2008) *Modern Multivariate Statistical Techniques*. New York, NY, USA: Springer.
- Jang, H. M., Choi, S., Shin, J., Kan, E. & Kim, Y. M. (2019) Additional reduction of antibiotic resistance genes and human bacterial pathogens via thermophilic aerobic digestion of anaerobically digested sludge, *Bioresource Technol.*, **273**, 259–268. <https://doi.org/10.1016/j.biortech.2018.11.027>.
- Jang, H. M., Lee, J., Shin, S. G., Shin, J. & Kim, Y. M. (2020) Comparing the fate of antibiotic resistance genes in two full-scale thermophilic anaerobic digestion plants treating food wastewater, *Bioresource Technol.*, **312**, 123577. <https://doi.org/10.1016/j.biortech.2020.123577>.
- Johnston, S., Cohen, S., Nguyen, C. K., Dinh, K. N., Nguen, T. D., Giddey, S., Munnings, C., Simonov, A. A. & MacFarlane, D. R. (2022) A survey of catalytic materials for ammonia electrooxidation to nitrite and nitrate, *ChemSusChem*, **15**, e202200614. <https://doi.org/10.1002/cssc.202200614>.
- Kavitha, S., Brindha, G. M. J., Rajashankar, K., Yeom, I. T. & Banu, J. R. (2016) Effect of enzyme secreting bacterial pretreatment on enhancement of aerobic digestion potential of waste activated sludge interceded through EDTA, *Bioresource Technol.*, **200**, 161–169. <http://dx.doi.org/10.1016/j.biortech.2015.10.026>.
- León, G. & Roy, P. H. (2003) Excision and integration of cassettes by an integron integrase of *Nitrosomonas europaea*, *J. Bacteriol.*, **185**, 2036–2041. <https://doi.org/10.1128/JB.185.6.2036-2041.2003>.
- Liang, W., Liu, H., Duan, G.-C., Zhao, Y., Chen, S., Yang, H.-Y. & Xi, Y.-L. (2018) Emergence and mechanism of carbapenem-resistant *Escherichia coli* in Henan, China, 2014, *J. Infect. Public Heal.*, **11**, 347–351. <https://doi.org/10.1016/j.jiph.2017.09.020>.
- Liu, S., Zhu, N. & Li, L. Y. (2012) The one-stage autothermal thermophilic aerobic digestion for sewage sludge treatment: Stabilization process and mechanism, *Bioresource Technol.*, **104**, 266–273. <https://doi.org/10.1016/j.biortech.2011.11.041>.

- Ma, Y., Wilson, C. A., Novak, J. T., Riffat, R., Aynur, S., Murthy, S. & Pruden, A. (2011) Effect of various sludge digestion conditions on sulfonamide, macrolide, and tetracycline resistance genes and class I integrons, *Environ. Sci. Technol.*, **45**, 7855–7861. <https://doi.org/10.1021/es200827t>.
- Mao, D., Yu, S., Rysz, M., Luo, Y., Yang, F., Li, F., Hou, J., Mu, Q. & Alvarez, P. J. J. (2015) Prevalence and proliferation of antibiotic resistance genes in two municipal wastewater treatment plants, *Water Res.*, **85**, 458–466. <https://doi.org/10.1016/j.watres.2015.09.010>.
- Milobedzka, A., Ferreira, C., Vaz-Moreira, I., Calderón-Franco, D., Gorecki, A., Purkrtova, S., Bartacek J, J., Dziewit, L., Singleton, C. M., Nielsen, P. H., Weissbrodt, D. G. & Manaia, C. M. (2022) Monitoring antibiotic resistance genes in wastewater environments: The challenges of filling a gap in the One-Health cycle, *J. Hazard. Mater.*, **424**, 127407. <https://doi.org/10.1016/j.jhazmat.2021.127407>.
- Mlynarcik, P., Roderova, M. & Kolar, M. (2016) Primer evaluation for PCR and its application for detection of carbapenemases in enterobacteriaceae, *Jundishapur J. Microb.*, **9**, e29314. <https://doi.org/10.5812/jjm.29314>.
- Nordmann, P., Dortet, L. & Poirel, L. (2012) Carbapenem resistance in Enterobacteriaceae: Here is the storm!, *Trends Mol. Med.*, **18**, 263–272. <https://doi.org/10.1016/j.molmed.2012.03.003>.
- Paulus, G. K., Hornstra, L. M., Alygizakis, N., Slobodnik, J., Thomaidis, N. & Medema, G. (2019) The impact of on-site hospital wastewater treatment on the downstream communal wastewater system in terms of antibiotics and antibiotic resistance genes, *In. J. Hyg. Envir. Heal.*, **222**, 635–644. <https://doi.org/10.1016/j.ijheh.2019.01.004>.
- Ping, Q., Zhang, Z., Ma, L., Yan, T., Wang, L. & Li, Y. (2022) The prevalence and removal of antibiotic resistance genes in full-scale wastewater treatment plants: Bacterial host, influencing factors and correlation with nitrogen metabolic pathway, *Sci. Total Environ.*, **827**, 154154. <http://dx.doi.org/10.1016/j.scitotenv.2022.154154>.
- Poorasgari, E. (2023) Effects of anaerobic and aerobic digestion at mesophilic and thermophilic temperatures on the quantity of carbapenem resistance genes in wastewater sludge: PhD Dissertation 2. Ottawa, ON, Canada: Carleton University. https://scholar.google.com/citations?view_op=view_citation&hl=en&user=o8Ldm8MAAAAJ&citation_for_view=o8Ldm8MAAAAJ:UeHWp8X0CEIC.
- Poorasgari, E. & Örmeci, B. (2022a) Development of non-linear empirical models to estimate the abundance of carbapenem resistance genes during anaerobic digestion of wastewater sludge at mesophilic and thermophilic temperatures, *Chem Eng. J.*, **450**, 138290. <https://doi.org/10.1016/j.cej.2022.138290>.
- Poorasgari, E. & Örmeci, B. (2022b) A mathematical model for the kinetics of concurrent volatile solids destruction and water evaporation during aerobic and anaerobic digestion of wastewater sludge under mesophilic and thermophilic temperatures, *J. Water Process Eng.*, **46**, 102540. <https://doi.org/10.1016/j.jwpe.2021.102540>.
- Poorasgari, E. & Örmeci, B. (2022c) Relationship between ultraviolet-visible spectra and soluble species in the liquid phase of wastewater sludge during biological digestion under mesophilic and thermophilic conditions. *J. Biochem. Eng. J.*, **179**, 108318. <https://doi.org/10.1016/j.bej.2021.108318>.
- Sekyere, J. O. (2016) Current state of resistance to antibiotics of last-resort in South Africa: A review from a public health perspective, *Front. Public Health*, **4**. <https://doi.org/10.3389/fpubh.2016.00209>.
- Smets, W., Leff, J. W., Bradford, M. A., McCulley, R. L., Lebeer, S. & Fierer, N. (2016) A method for simultaneous measurement of soil bacterial abundances and community composition via 16S rRNA gene sequencing, *Soil Biol. Biochem.*, **96**, 145–151.
- Spellberg, B., Blaser, M., Guidos, R. J., Boucher, H. W., Bradley, J. S., Eisenstein, B. I., Gerding, D., Lynfield, R., Reller, L. B., Rex, J., Schwartz, D., Septimus, E., Tenover, F. C. & Gilbert, D. N. (2011) Combating antimicrobial resistance: Policy recommendations to save lives, *Clin. Infect. Dis.*, **52**, S397–S428.
- Subirats, J., RoyoJosé, E., Balcázar, L. & Borrego, C. M. (2017) Real-time PCR assays for the detection and quantification of carbapenemase genes (bla KPC, bla NDM, and bla OXA-48) in environmental samples, *Environmental Sci. Pollut. R*, **24**, 6710–6714. <https://doi.org/10.1007/s11356-017-8426-6>.
- Sun, M., Ye, M., Liu, K., Schwab, A. P., Liu, M., Jiao, J., Feng, Y., Wan, J., Tian, D., Wu, J., Li, H., Hu, F. & Jiang, X. (2017) Dynamic interplay between microbial denitrification and antibiotic resistance under enhanced anoxic denitrification condition in soil, *Environ. Pollut.*, **222**, 583–591. <http://dx.doi.org/10.1016/j.envpol.2016.10.015>.
- Tchobanoglous, G., Stensel, H. D., Tsuchihashi, R. & Burton, F. (2014) *Wastewater Engineering: Treatment and Resource Recovery. 5th Edition*. New York, NY, USA: McGraw-Hill Education.
- Thukral, A. K. (2017) A review on measurement of alpha diversity biology, *Agric Res J*, **54** (1), 1–10. doi:10.5958/2395-146X.2017.00001.1.
- Tong, J., Lu, X., Zhang, J., Angelidakis, I. & Wei, Y. (2018) Factors influencing the fate of antibiotic resistance genes during thermochemical pretreatment and anaerobic digestion of pharmaceutical waste sludge, *Environ. Pollut.*, **243**, 1403–1413. <https://doi.org/10.1016/j.envpol.2018.09.096>.
- Tong, J., Tang, A., Wang, H., Liu, X., Huang, Z., Wang, Z., Zhang, J., Wei, Y., Su, Y. & Zhang, Y. (2019a) Microbial community evolution and fate of antibiotic resistance genes along six different full-scale municipal wastewater treatment processes, *Bioresour. Technol.*, **272**, 489–500.
- Tong, J., Fang, P., Zhang, J., Wei, Y., Su, Y. & Zhang, Y. (2019b) Microbial community evolution and fate of antibiotic resistance genes during sludge treatment in two full-scale anaerobic digestion plants with thermal hydrolysis pretreatment, *Bioresour. Technol.*, **288**, 121575. <https://doi.org/10.1016/j.biortech.2019.121575>.
- Větrovský, T. & Baldrian, P. (2013) The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses, *PLoS One*, **8**, e57923. <https://doi.org/10.1371/journal.pone.0057923>.

- von Neumann, J. (2016) Model selection and overfitting, *Nature Methods*, **13**, 703–704.
- Wang, J., Li, M., Guan, A., Liu, R., Qi, W., Liu, H. & Qu, J. (2022) Can radicals-orientated chemical oxidation improve the reduction of antibiotic resistance genes (ARGs) by mesophilic anaerobic digestion of sludge?, *J. Hazard. Mater.*, **426**, 128001. <https://doi.org/10.1016/j.jhazmat.2021.128001>.
- Willis, A. D. (2019) Rarefaction, alpha diversity, and statistics, *Front. Microbiol.*, **10**, 2407. <https://doi.org/10.3389/fmicb.2019.02407>.
- Xiong, J., Alexander, D. A., Ma, J. H., Déraspe, M., Low, D. E., Jamieson, F. B. & Roy, P. H. (2013) Complete sequence of pOZ176, a 500-Kilobase IncP-2 plasmid encoding IMP-9-Mediated carbapenem resistance, from Outbreak Isolate *Pseudomonas aeruginosa* 96, *Antimicrob. Agents Ch.*, **57**, 3775–3782. <https://doi.org/10.1128/AAC.00423-13>.
- Yu, W., Zhan, S., Shen, Z., Zhou, Q. & Yang, D. (2017) Efficient removal mechanism for antibiotic resistance genes from aquatic environments by graphene oxide nanosheet, *Chem. Eng. J.*, **313**, 836–846. <https://doi.org/10.1016/j.cej.2016.10.107>.
- Zhang, Y., Marrs, C. F., Simon, C. & Xi, C. (2009) Wastewater treatment contributes to selective increase of antibiotic resistance among *Acinetobacter* spp, *Sci. Total Environ.*, **407**, 3702–3706. <https://doi.org/10.1016/j.scitotenv.2009.02.013>.
- Zhang, Z., Liu, H., Wen, H., Gao, L., Gong, Y., Guo, W., Wang, Z., Li, X. & Wang, Q. (2021) Microplastics deteriorate the removal efficiency of antibiotic resistance genes during aerobic sludge digestion, *Sci. Total Environ.*, **798**, 149344. <https://doi.org/10.1016/j.scitotenv.2021.149344>.

First received 21 July 2024; accepted in revised form 18 October 2024. Available online 4 November 2024