



**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL
CURSO DE ENGENHARIA CIVIL**

GABRIEL TIBÉRIUS ANDRADE MOREIRA

**MODELAGEM DA PRODUÇÃO DE ÁCIDOS GRAXOS VOLÁTEIS E METANO
ATRAVÉS DE REGRESSÃO LINEAR MÚLTIPLA COM USO DE LINEARIZAÇÃO
DAS VARIÁVEIS E OTIMIZAÇÃO DO COEFICIENTE DE DETERMINAÇÃO**

**RECIFE
2023**

GABRIEL TIBÉRIUS ANDRADE MOREIRA

**MODELAGEM DA PRODUÇÃO DE ÁCIDOS GRAXOS VOLÁTEIS E METANO
ATRAVÉS DE REGRESSÃO LINEAR MÚLTIPLA COM USO DE LINEARIZAÇÃO
DAS VARIÁVEIS E OTIMIZAÇÃO DO COEFICIENTE DE DETERMINAÇÃO**

Trabalho de Conclusão de Curso apresentado
ao Curso de Graduação em Engenharia Civil da
Universidade Federal de Pernambuco, como
requisito parcial para obtenção do grau de
Bacharel em Engenharia Civil.

Área de concentração: Saneamento Ambiental

Orientador: Prof. D.Sc. Wanderli Rogério
Moreira Leite

RECIFE

2023

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Moreira, Gabriel Tibérius Andrade.

Modelagem da produção de ácidos graxos voláteis e metano através de regressão linear múltipla com uso de linearização das variáveis e otimização do coeficiente de determinação / Gabriel Tibérius Andrade Moreira. - Recife, 2023. 76 p. : il., tab.

Orientador(a): Wanderli Rogério Moreira Leite

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Engenharia Civil - Bacharelado, 2023.

Inclui referências, apêndices.

1. Modelagem da digestão anaeróbia. 2. Microalgas. 3. Metano. 4. Linearização de variáveis. 5. Otimização do coeficiente de determinação. 6. Ácidos graxos voláteis. I. Leite, Wanderli Rogério Moreira. (Orientação). II. Título.

620 CDD (22.ed.)

GABRIEL TIBÉRIUS ANDRADE MOREIRA

**MODELAGEM DA PRODUÇÃO DE ÁCIDOS GRAXOS VOLÁTEIS E METANO
ATRAVÉS DE REGRESSÃO LINEAR MÚLTIPLA COM USO DE LINEARIZAÇÃO
DAS VARIÁVEIS E OTIMIZAÇÃO DO COEFICIENTE DE DETERMINAÇÃO**

Trabalho de Conclusão de Curso apresentado
ao Curso de Graduação em Engenharia Civil da
Universidade Federal de Pernambuco, como
requisito parcial para obtenção do grau de
Bacharel em Engenharia Civil.

Aprovado em: 07 / 11 / 2023 com nota 9,7.

BANCA EXAMINADORA

Prof. D.Sc. Wanderli Rogério Moreira Leite (Orientador)
Universidade Federal de Pernambuco

Prof. D.Sc. Rubens Alves Dantas (Examinador Interno)
Universidade Federal de Pernambuco

M.Sc. Agnes Adam Duarte Pinheiro (Examinadora Externa)
Universidade Federal de Pernambuco

“Shemá Yisrael, Hashem Elohênu, Hashem Echad”

Devarim, 6:4

“Ouve Israel! O Eterno é nosso D’us! O Eterno é Um!”

Deuteronômio, 6:4

Dedico este trabalho à minha mãe,
Zenilda Andrade, pela dedicação
incondicional e exclusiva à nossa Família

AGRADECIMENTOS

Agradeço primeiramente ao Eterno, Bendito seja Ele, por me guiar, me orientar, me iluminar, me abençoar, me guardar e me fazer chegar até aqui. Agradeço ao Eterno também por eu ter tantas pessoas a quem agradecer.

Em seguida ao Eterno, agradeço à minha Mãe, Zenilda Andrade, por ser a principal responsável por quem eu sou e por tudo o que tenho conquistado. Agradeço ao meu pai, Alexandre César, por tornar possível a minha graduação e formação. Agradeço à minha irmã, Giselle Acácia, por ter sido tão importante na minha criação e educação. Agradeço à minha irmã, Mayara Gabrielle, pela sua influência em quem eu me tornei.

Agradeço ao Professor Wanderli Leite, por ter me aberto tantas oportunidades desde cedo no curso, por ter ido além dos ensinamentos como orientador, ter me instruído com habilidade em lições que vou levar para a vida, e pela paciência e compreensão. Agradeço ao Professor Rubens Dantas por ter fornecido a licença do SAB Comparativo gratuitamente e ter me ensinado toda a parte estatística que apliquei no presente trabalho. Agradeço à prof.a Lourdinha Florêncio, coorientadora na pesquisa e fundadora do LSA junto com o prof. Mario Kato.

Agradeço aos técnicos do LSA/LABIOTA Ronaldo, com toda a sua alegria contagiante, Danúbia, pela paciência e prestatividade, Iago, com sua disciplina e orientação, são indispensáveis para a realização da pesquisa que deu origem a este trabalho e que continua sendo realizada.

Agradeço à Álcool Química Nacional, do Grupo JBS, por receber o grupo de pesquisa sempre tão bem, manter a colaboração na troca de conhecimento e fornecimento de vinhaça. Agradeço à FACEPE por ter me concedido a bolsa para realização da pesquisa.

Agradeço à minha tutora da iniciação científica, Agnes Pinheiro, pela parceria na pesquisa. A Edilberto da Silva pelas ajudas no laboratório e na produção dos trabalhos, cumprindo o papel de um segundo tutor. À Dicla Cesário pela ajuda na formatação do TCC. Agradeço aos outros parceiros de pesquisa, Francielle Machado, Acson de Lima, Gisella Araújo, Alynne Soares e Sayonara de Araújo. E aos demais colegas do laboratório que sempre dão uma força, em especial à Manoella Candido, Luiz Pereira e Alfredo Quirino.

Agradeço aos amigos da Equipe Rocket: Débora Soares, Edinaldo França, Túlio Queiroz, Clauber Bezerra, Carlos Lamour, Alinne Kétally, Yasmin Bonfim, Wagner Isaucon, por toda a parceria, companheirismo, apoio e amizade desde que a turma de civil se formou. Foram prêmio ganho, noites viradas juntos fazendo trabalhos e provas, muita troca de conhecimento, materiais, estágios e empregos, fotos de formatura de madrugada, mas acima de tudo, muita amizade e companheirismo.

Agradeço a Hipólito Roberto pela parceria desde o primeiro período e à Twany Correia que junto com Hipólito me apoiou em um momento crítico do curso, além da ajuda em vários outros momentos. A Carlos Matheus França, por ser um irmão inestimável desde o começo de nossa amizade cedo no curso. À Maria Clara Candido de Lima por ter me dado forças, conforto, consolo em um momento muito difícil do curso e da vida, por continuar fazendo isso com sua bonita e cativante amizade diária. À Laura Gomes da Costa pela parceria nas monitorias, estágios, TCC's e pela forte amizade que surgiu a partir disso, com participações um do outro em momentos importantes das nossas vidas.

Agradeço aos coordenadores do curso de Engenharia Civil, À Liliane Fonseca pelo apoio e prestatividade na minha conclusão do curso, a Anderson Paiva, Tiago Ancelmo por todo o suporte em relação a eu ser guardador do sábado e demais demandas que tive ao longo do curso. Bem como aos professores da UFPE que foram tão importantes na minha formação profissional e acadêmica. Aos professores do IFPE, em especial Juliana Moraes, Anamaria Duarte, Clifford Ericson. A todos os mestres que tive ao longo da caminhada, em especial à professora que me sugeriu entrar no IFPE e ao professor Elvis.

Agradeço aos amigos da UFPE que me deram suas agradáveis amizades que tornaram a rotina mais leve, suporte na logística do curso e parcerias em trabalhos e estudos: Gabriel Carlini; Catarina Zovka; Paulo Ferreira; Natália Martins; Mayara Moraes; o casal João Vitor e Letícia Karine; Maria Andrade; o trio Mabelle, Levi e Lara Lis; a dupla Anderson Antônio e Beatriz Melo.

Agradeço aos amigos que confio e carrego comigo há tanto tempo que são como família: Diogo Santos, Arianne Gomes, Diego Barbosa.

A todos, eu agradeço muito. Que o Eterno retribua em dobro, os abençoe e guarde.

RESUMO

O desenvolvimento de modelos preditores da digestão anaeróbia permite prever a produção de biometano (bio-CH₄) e de ácidos graxos voláteis (AGV), economizando custos e tempo experimental para saber essas produções e guiando a realização de novos experimentos pelo conhecimento das variáveis mais significantes com seus valores ótimos. Assim sendo, desenvolveram-se modelos preditores através da regressão linear múltipla (RLM) pelo método dos mínimos quadrados, com transformações de linearização das variáveis e de otimização do coeficiente de determinação (R^2), para experimento de codigestão anaeróbia de biomassa de microalgas (BM) proveniente de lagoa de alta taxa (LAT) com vinhaça de cana-de-açúcar (VC). A codigestão se deu em sistema de dois estágios com reator acidogênico (RA) produtor de AGV no primeiro estágio, e reator metanogênico (RM) produtor de bio-CH₄ no segundo estágio. O experimento teve duração de 270 dias e foi dividido em três fases experimentais: *start-up*; codigestão de BM e VC; e codigestão de biomassa de microalgas pré-tratada termicamente (BM-PT) com VC. As transformações das variáveis para linearização e otimização do R^2 foram realizadas no *software* SAB Comparativo®, desenvolvido pela Dantas Engenharia, onde foram testados arranjos de 7 escalas: x , $1/(x)$, $\ln(x)$, $\text{Raiz}(x)$, $1/\text{Raiz}(x)$, $(x)^2$, $1/(x)^2$. Os testes iniciais de otimização dos modelos do RA foram com 4 variáveis independentes e a concentração de AGV no efluente como dependente, resultando em 2.520 arranjos testados. Foi gerado um modelo preditor para as duas últimas fases do RA, levando a 2 modelos. Os testes iniciais do RM foram com 13 variáveis independentes e o rendimento de metano (rend.CH_4) como variável dependente, resultando em 8.648.640 modelos testados. Os testes para o RM foram divididos em etapas. Foi gerado um modelo para as duas últimas fases do RM, com a fase 2 subdividida devido à mudança de aparelho de quantificação do biogás. Todos os modelos tiveram valor de R^2 acima de 0,89 e nível de significância global acima de 99,99%, com maior R^2 para o RA de 0,95 e maior R^2 para o RM de 0,96. Dentre as variáveis independentes que foram significantes para os modelos do RA, 3 foram selecionadas como mais importantes e suas melhores escalas indicadas, juntamente com a variável dependente. Para o RM, 5 variáveis foram mais relevantes e suas escalas indicadas, bem como para a variável dependente. O método de transformações das variáveis permitiu a modelagem em um caso que a RLM falhou e aumento do R^2 em todos os outros casos, chegando ao ganho de até 54 pontos percentuais.

Palavras-chave: Modelagem da digestão anaeróbia, microalgas, metano, linearização de variáveis, otimização do coeficiente de determinação, ácidos graxos voláteis.

ABSTRACT

The development of predictive models for anaerobic digestion makes it possible to predict the production of biomethane (bio-CH₄) and volatile fatty acids (VFA), saving costs and experimental time to know these productions and guiding the carrying out of new experiments by knowing the most significant variables with its optimal values. Therefore, predictive models were developed through multiple linear regression (RLM) using the least squares method, with linearization transformations of variables and optimization of the coefficient of determination (R^2), for an experiment on anaerobic co-digestion of microalgae biomass (BM) from a high rate lagoon (LAT) with sugar cane vinasse (VC). Co-digestion took place in a two-stage system with an acidogenic reactor (RA) producing AGV in the first stage, and a methanogenic reactor (RM) producing bio-CH₄ in the second stage. The experiment lasted 270 days and was divided into three experimental phases: start-up; co-digestion of BM and VC; and co-digestion of thermally pretreated microalgae biomass (BM-PT) with VC. The transformations of the variables for linearization and optimization of R^2 were carried out in the SAB Comparativo® software, developed by Dantas Engenharia, where arrangements of 7 scales were tested: x , $1/(x)$, $\ln(x)$, $\text{Root}(x)$, $1/\text{Root}(x)$, $(x)^2$, $1/(x)^2$. The initial optimization tests of the RA models were with 4 independent variables and the VFA concentration in the effluent as dependent, resulting in 2,520 tested arrangements. A predictor model was generated for the last two phases of the RA, leading to 2 models. The initial RM tests were with 13 independent variables and methane yield (CH₄ yield) as the dependent variable, resulting in 8,648,640 tested models. The tests for RM were divided into stages. A model was generated for the last two phases of the RM, with phase 2 subdivided due to the change in the biogas quantification device. All models had an R^2 value above 0.89 and an overall significance level above 99.99%, with the highest R^2 for the RA of 0.95 and the highest R^2 for the RM of 0.96. Among the independent variables that were significant for the RA models, 3 were selected as most important and their best scales indicated, together with the dependent variable. For the RM, 5 variables were most relevant and their scales indicated, as well as for the dependent variable. The variable transformation method allowed modeling in a case where the RLM failed and an increase in R^2 in all other cases, reaching a gain of up to 54 percentage points.

Keywords: Modeling of anaerobic digestion, microalgae, methane, linearization of variables, optimization of the coefficient of determination, volatile fatty acids.

LISTA DE FIGURAS

Figura 1 – Evolução da concentração de ácidos graxos voláteis (AGV) no efluente do RA e afluente do RM	29
Figura 2 – Evolução do rendimento de metano do RM	30
Figura 3 – Distâncias de Cook <i>versus</i> valores ajustados do modelo preditor da fase II do reator acidogênico.....	44
Figura 4 – Histograma dos resíduos do modelo preditor da fase II do reator acidogênico	45
Figura 5 – Valores ajustados <i>versus</i> valores observados do modelo preditor da fase II do reator acidogênico.....	45
Figura 6 – Resíduos padronizados <i>versus</i> valores ajustados do modelo preditor da fase II do reator acidogênico.....	47
Figura 7 – Distâncias de Cook <i>versus</i> valores ajustados do modelo preditor da fase III do reator acidogênico.....	49
Figura 8 – Resíduos padronizados <i>versus</i> valores ajustados do modelo preditor da fase III do reator acidogênico.....	50
Figura 9 – Valores ajustados <i>versus</i> valores observados do modelo preditor da fase III do reator acidogênico.....	50
Figura 10 – Histograma dos resíduos do modelo preditor da fase III do reator acidogênico..	51
Figura 11 – Distâncias de Cook <i>versus</i> valores ajustados do modelo preditor da fase II-1 do reator metanogênico	54
Figura 12 – Resíduos padronizados <i>versus</i> valores ajustados do modelo preditor da fase II-1 do reator metanogênico	54
Figura 13 – Valores ajustados <i>versus</i> valores observados do modelo preditor da fase II-1 do reator metanogênico	55
Figura 14 – Histograma dos resíduos do modelo preditor da fase II-1 do reator metanogênico	55
Figura 15 – Distâncias de Cook <i>versus</i> valores ajustados do modelo preditor da fase II-2 do reator metanogênico	58
Figura 16 – Resíduos padronizados <i>versus</i> valores ajustados do modelo preditor da fase II-2 do reator metanogênico	58
Figura 17 – Valores ajustados <i>versus</i> valores observados do modelo preditor da fase II-2 do reator metanogênico	59

Figura 18 – Histograma dos resíduos do modelo preditor da fase II-2 do reator metanogênico	59
Figura 19 – Distâncias de Cook <i>versus</i> valores ajustados do modelo preditor da fase III do reator metanogênico	62
Figura 20 – Resíduos padronizados <i>versus</i> valores ajustados do modelo preditor da fase III do reator metanogênico	62
Figura 21 – Valores ajustados <i>versus</i> valores observados do modelo preditor da fase III do reator metanogênico	63
Figura 22 – Histograma dos resíduos do modelo preditor da fase III do reator metanogênico	63

LISTA DE QUADROS

Quadro 1 - Valores de R^2 alcançados de acordo com as fases experimentais com e sem transformações de linearização das variáveis	65
Quadro 2 - Equações dos modelos de acordo com o reator e com a fase experimental	66
Quadro 3 - Valores de R^2 alcançados através de diferentes métodos de geração de modelos preditores	66

LISTA DE TABELAS

Tabela 1 – Exemplo de dados para modelo de regressão linear múltipla	31
Tabela 2 – Classificação da correlação em relação ao nível de dependência linear	35
Tabela 3 – Escalas de transformação de linearização das variáveis e otimização do coeficiente de determinação.	38
Tabela 4 – Variáveis utilizadas na geração dos modelos de regressão para o reator acidogênico	39
Tabela 5 – Variáveis utilizadas na geração dos modelos de regressão para o reator metanogênico.....	41
Tabela 6 - indicadores globais do ajuste do modelo preditor da fase II do reator acidogênico	43
Tabela 7 - <i>Outliers</i> e pontos influentes retirados do modelo preditor da fase II do reator acidogênico	43
Tabela 7 - indicadores individuais da qualidade de ajuste das variáveis independentes do modelo preditor da fase II do reator acidogênico	43
Tabela 9 – Matriz de correlação do modelo preditor da fase II do reator acidogênico	46
Tabela 10 - Indicadores globais do ajuste do modelo de regressão sem transformação das escalas fase II do reator acidogênico	47
Tabela 11 - Indicadores globais do ajuste do modelo preditor da fase III do reator acidogênico	48
Tabela 12 - Ponto influenciante retirado do modelo preditor da fase III do reator acidogênico	48
Tabela 13 - Indicadores individuais da qualidade de ajuste das variáveis independentes do modelo preditor da fase III do reator acidogênico.....	48
Tabela 14 – Matriz de correlação do modelo preditor da fase III do reator acidogênico	51
Tabela 15 - Indicadores globais do ajuste do modelo de regressão sem transformação das escalas da fase III do reator acidogênico	52
Tabela 16 - Indicadores globais do ajuste do modelo preditor da fase II-1 do reator metanogênico.....	53
Tabela 17 - <i>Outliers</i> e ponto influenciante retirados do modelo preditor da fase II-1 do reator metanogênico.....	53
Tabela 18 - Indicadores individuais do ajuste das variáveis do modelo preditor da fase II-1 do reator metanogênico	53

Tabela 19 – Matriz de correlação do modelo preditor da fase II-1 do reator metanogênico ...	56
Tabela 20 - Indicadores globais da qualidade de ajuste do modelo preditor da fase II-1 do reator metanogênico	56
Tabela 21 - Indicadores globais do ajuste do modelo preditor da fase II-2 do reator metanogênico.....	57
Tabela 22 - Ponto influenciante retirado do modelo preditor da fase II-2 do reator metanogênico.....	57
Tabela 23 - Indicadores individuais do ajuste das variáveis do modelo preditor da fase II-2 do reator metanogênico	57
Tabela 24 – Matriz de correlação do modelo preditor da fase II-2 do reator metanogênico ...	60
Tabela 25 - Indicadores globais do ajuste do modelo de regressão sem transformação das escalas fase II-2 do reator metanogênico.....	60
Tabela 26 - Indicadores globais do ajuste do modelo preditor da fase III do reator metanogênico.....	61
Tabela 27 - <i>Outlier</i> retirado do modelo preditor da fase III do reator metanogênico	61
Tabela 28 - Indicadores individuais do ajuste das variáveis do modelo preditor da fase III do reator metanogênico	61
Tabela 29 – Matriz de correlação do modelo preditor da fase III do reator metanogênico	64
Tabela 30 - Indicadores globais do ajuste do modelo de regressão sem transformação das escalas da fase III do reator metanogênico	64

LISTA DE ABREVIATURAS E SIGLAS

%S	Percentual do substrato
AI/AP	Relação alcalinidade intermediária/alcalinidade parcial
AI/AT	alcalinidade intermediária/alcalinidade total
ABNT	Associação Brasileira de Normas Técnicas
ACE	Variável independente ácido acético
adc	adicionado
AGV	Ácidos graxos voláteis
AIAP, AI/AP	Variável independente relação alcalinidade intermediária/alcalinidade parcial
AIAT, AI/AT	Variável independente relação alcalinidade intermediária/alcalinidade total
ALCAINT	Variável independente alcalinidade intermediária
ALCAPAR	Variável independente alcalinidade parcial
ALCATOT	Variável independente alcalinidade total
AMON	Variável independente nitrogênio amoniacal
ANFIS	Sistema de inferência fuzzy baseado em rede adaptativa
ANN	Rede neural artificial
APHA	Associação Americana de Saúde Pública
bio-CH ₄	Biometano
BM	Biomassa de microalgas
BM-PT	Biomassa de microalgas pré-tratada
C/N	Relação carbono/nitrogênio
CaCO ₃	Carbonato de cálcio
CCD	Método de design de composto central
CH ₄	Gás metano
CNTP	Condições Normais de Temperatura e Pressão
CO ₂	Gás dióxido de carbono
COT	Carbono Orgânico Total
COV	Carga orgânica volumétrica
CSTR	<i>Continuous stirred tank reactor</i>
CTG	Centro de Tecnologia e Geociências
DA	Digestão anaeróbia
DQO	Variável independente carga de demanda química de oxigênio adicionada
DQOs	Demanda Química de Oxigênio solúvel

DQOt	Demanda Química de oxigênio total
e	Resíduo
$\bar{e}_i$	Resíduo padronizado
ETE	Estação de tratamento de esgoto
Fc	Estatística de Fischer-Snedecor
FORSU	Fração orgânica dos resíduos sólidos urbanos
GC-FID	Cromatógrafo gasoso com detector de ionização de chama
GC-TCD	Cromatógrafo gasoso com detector de condutividade térmica
H ₂	Gás hidrogênio
H ₂ S	Gás sulfídrico
H ₂ SO ₄	Ácido sulfúrico
HCl	Ácido clorídrico
HPLC	Cromatógrafo Líquido de Alta Eficiência
I/S	Relação inóculo/substrato
IS	Índice de Sinergia
LAT	Lagoa de alta taxa
MMQ	Método dos mínimos quadrados
MO	Matéria orgânica
NaHCO ₃	Bicarbonato de sódio
NaOH	Hidróxido de sódio
NH ₃	Amônia
NH ₄ ⁺	Amônio
N-NH ₃	Nitrogênio amoniacal
NTK	Nitrogênio Total Kjeldahl
O ₂	Gás oxigênio
OD	Oxigênio dissolvido
ODS	Objetivos do desenvolvimento sustentável
ONU	Organização das Nações Unidas
ORP	Potencial de oxi-redução
P total	Fósforo total
PBM	Potencial bioquímico de metano
PH	Variável independente potencial hidrogeniônico
pH	Potencial hidrogeniônico
PHA	Polihidroxialcanoatos

PHB	Polihidroxibutirato
PHV	Polihidroxivalerato
PROP	Variável independente ácido propiônico
r	Coeficiente de correlação
R'	Coeficiente de determinação ajustado
R ²	Coeficiente de determinação
RA	Reator acidogênico
REDOX	Variável independente potencial de oxi-redução
Rend.CH ₄	Rendimento de metano
RENDCH ₄	Variável independente rendimento de metano
RLM	Regressão linear múltipla
RM	Reator metanogênico
RSM	Método da superfície de resposta
S(b)	Desvio padrão da variável
S/E	Relação substrato/efluente
SD	Grau de solubilização
Se	Desvio padrão do modelo
SF	Sólidos fixos
SM	Standard Methods
SQR	Somatório do quadrado dos resíduos explicados
SQTO	Somatório dos quadrados dos resíduos totais
ST	Sólidos Totais
SV	Sólidos Voláteis
TDH	Tempo de detenção hidráulica
t _j	Estatística t
UASB	Reator Anaeróbico de Fluxo Ascendente
UFPE	Universidade Federal de Pernambuco
VALE	Variável independente ácido valérico
VC	Vinhaça de cana-de-açúcar
Vexp	Variância explicada
VFA	<i>Volatile fatty acids</i>
Vnexp	Variância não explicada

SUMÁRIO

1	INTRODUÇÃO E JUSTIFICATIVA.....	19
2	OBJETIVOS	23
2.1	Objetivo geral	23
2.2	Objetivos específicos.....	23
3	REVISÃO BIBLIOGRÁFICA	24
3.1	Modelos preditores aplicados à digestão anaeróbia gerados a partir de regressão linear.....	24
3.2	Digestão Anaeróbia	25
3.3	Lagoas de alta taxa	26
3.4	biomassa de Microalgas	27
4	MATERIAL E MÉTODOS	29
4.1	DADOS EXPERIMENTAIS USADOS NA PESQUISA	29
4.2	Inferência estatística e regressão linear	30
4.3	Regressão linear múltipla pelo método dos mínimos quadrados (MMQ).....	31
4.4	Teste de significância global do modelo pela análise da variância (Fc) e p-valor	32
4.5	Teste de significância individual PARA SELEÇÃO das variáveis independentes – estatística t (t_j).....	33
4.6	Coeficiente de correlação e multicolinearidade (r)	34
4.7	Coeficiente de determinação (R^2) e coeficiente de determinação ajustado (R').....	35
4.8	Parcimônia.....	36
4.9	Desvio padrão global do modelo (S_e).....	36
4.10	Transformações das variáveis para linearização dos dados, otimização do R^2 e escolha das variáveis nos modelos.....	37
4.11	Pressupostos básicos do método clássico de regressão	38
4.12	Geração do modelo estatístico para o reator acidogênico	39
4.13	Geração do modelo estatístico para o reator metanogênico	40
4.14	fases experimentais	42
5	RESULTADOS E DISCUSSÃO.....	43
5.1	Modelo preditor do reator acidogênico – fase II	43
5.2	Modelo preditor do reator acidogênico – fase III.....	48
5.3	Modelo preditor do reator metanogênico – fase II-1	52
5.4	Modelo preditor do reator metanogênico – fase II-2.....	57

5.5	Modelo preditor do reator metanogênico – fase III.....	61
5.6	Resumo das equações e comparação com outros trabalhos com diferentes métodos	65
5.7	Avaliação das variáveis mais importantes para cada reator	66
6	CONCLUSÕES	68
7	SUGESTÕES DE TRABALHOS FUTUROS.....	69
	REFERÊNCIAS	70
	APÊNDICE A – Dados observados no experimento.....	73

1 INTRODUÇÃO E JUSTIFICATIVA

O esgotamento dos combustíveis fósseis devido à sua disponibilidade limitada e à crescente demanda por energia causada pelo aumento populacional, industrialização e elevação do padrão de vida, tem levado à busca por fontes alternativas de energia. Grande destaque se tem dado aos biocombustíveis, fontes de energia sustentáveis graças a serem renováveis e diminuírem de maneira levando o impacto ambiental. No mesmo contexto, o esgotamento dos combustíveis fósseis leva ao esgotamento da matéria-prima utilizada em diferentes produtos industrializados, como por exemplo os plásticos, originados principalmente do petróleo. Isso tem levado à busca por fontes alternativas de matéria-prima. Os bioprodutos de valor agregado são uma dessas fontes alternativas, dando origem por exemplo aos bioplásticos (Nazarpour *et al.* 2022).

Assim sendo, muitas pesquisas têm sido desenvolvidas para melhorar as tecnologias de produção de biocombustíveis e bioprodutos de valor agregado, com foco em melhorar os rendimentos, os balanços energéticos, reduzir custos e tempo de produção para tornar essas tecnologias economicamente viáveis e competitivas com os combustíveis fósseis em escala industrial (Vishtal e Kraslawski, 2011; Tepari *et al.* 2020).

O biometano (bio-CH₄) e os ácidos graxos voláteis (AGV) têm se mostrado muito promissores como biocombustível e bioproduto de valor agregado, respectivamente, sendo amplamente estudados há algumas décadas. O bio-CH₄ é um gás combustível que pode ser queimado para a geração de energia elétrica e energia térmica, resultando em CO₂ e água (Chen, 2023). Enquanto que os AGV podem ser insumo para a produção de bioplásticos como os variados polihidroxialcanoatos (PHA), neles incluídos o polihidroxibutirato (PHB) e o polihidroxivalerato (PHV), entre outros. Os AGV e o bio-CH₄ podem ser produzidos a partir da digestão anaeróbia (Nazarpour *et al.* 2022).

A digestão anaeróbia (DA) consiste na decomposição e estabilização da matéria orgânica (MO) por micro-organismos na ausência de oxigênio (O₂). A DA permite a geração de bio-CH₄ e AGV a partir de resíduos orgânicos, contribuindo para a economia circular, reaproveitamento de recursos, diminuição do volume de resíduos e minimização do seu potencial poluidor. Exemplos de resíduos que podem ser usados são os industriais, de estações de tratamento de esgoto (ETE) e mesmo a fração orgânica dos resíduos sólidos urbanos (FORSU) (Chen, Wang e Liu, 2023).

Essa matéria orgânica usada na produção de energia e de produtos de valor agregado é chamada de biomassa. A biomassa pode ser classificada em relação à sua geração de acordo com o seu estágio de uso. Quando é usada na produção de energia assim que é extraída da sua fonte primária, que pode ser a colheita de uma plantação, por exemplo, é uma biomassa de primeira geração. Se é

um resíduo de um primeiro processo produtivo, é biomassa de segunda geração. Se é um resíduo de processo produtivo que já utilizou um resíduo anterior, é biomassa de terceira geração, e assim sucessivamente. O uso dos resíduos de ETE na produção de bioenergia classifica essa biomassa como de terceira geração e alia o conceito de estações de recuperação de recursos às ETE, estando em perfeita sintonia com a prática da economia circular (Vishtal e Kraslawski, 2011).

O uso de biomassa de primeira geração causa a competição com o ramo alimentício, inclusive em relação às áreas agrícolas, disponibilidade de água e fertilizantes, agravando o problema da escassez de alimentos. Porém, o estudo das algas e microalgas como biomassa tem crescido bastante nos últimos anos graças à sua alta concentração de lipídios, proteínas e carboidratos, disponibilidade abundante, facilidade de coleta, dispensa de terras, baixo custo, adaptação a todos os países do mundo, mitigação e remediação de CO₂ ao crescer usando gases de combustão (Nazarpour *et al.* 2022; Shobana *et al.* 2017). Além de a biomassa de microalgas mitigar o problema da competição com o ramo alimentício, ela pode ser advinda de resíduos de ETE que possuem lagoas de estabilização que envolvam processos aeróbios, como a lagoa aeróbia, lagoa facultativa ou lagoa de alta taxa (LAT). Nesses reatores, as microalgas podem desempenhar o papel de redução da matéria orgânica, melhora do oxigênio dissolvido (OD) e diminuição dos nutrientes nitrogênio e fósforo, funcionando bem em cooperação com as bactérias aeróbias que apresentam maiores taxas de degradação da MO, usam o OD produzido na fotossíntese das algas e as microalgas usam o CO₂ produzido na respiração das bactérias aeróbias.

Contudo, as microalgas possuem parede celular de baixa biodegradabilidade (recalcitrante) e concentração de proteínas muito elevada, que pode gerar toxicidade por amônia com baixa relação carbono/nitrogênio (C/N) (Machado, 2012). Soluções que têm sido adotadas para aumentar a biodegradabilidade da parede celular das microalgas são a codigestão com substratos mais biodegradáveis, o que aumenta a atividade enzimática hidrolítica dos micro-organismos, favorecendo a quebra da parede celular das microalgas, e os pré-tratamentos, que podem ser térmicos, químicos, mecânicos e enzimáticos, aumentando a solubilização das microalgas. A codigestão permite ainda a solução da baixa relação C/N através da mistura com resíduos ricos em carbono e pobres em nitrogênio (Shobana *et al.*, 2017; Sialve *et al.*, 2009).

No processo de DA estão incluídas quatro etapas. As duas primeiras, hidrólise e acidogênese, são favorecidas por níveis de potencial hidrogeniônico (pH) mais baixos (entre 5 e 6,5) e por maiores (10 a 15 gSV/L) cargas orgânicas volumétricas (COV), enquanto que as duas últimas, acetogênese e metanogênese, são favorecidas por maiores valores de pH o (em torno de 7) e COV mais baixas (2 a 5 gSV/L). Dessa maneira é comum otimizar a DA pela separação das etapas em dois sub-reatores,

caracterizando um sistema de dois estágios. O primeiro estágio, reator acidogênico (RA) ou fermentador, compreende as duas primeiras etapas, e o segundo estágio, reator metanogênico (RM), compreende as duas últimas etapas. Os AGV são gerados no RA, sendo um produto intermediário da DA, e o metano é gerado no RM, sendo um produto final da DA, juntamente com gás carbônico e água (Wang *et al.*, 2009). Além de operar em faixas diferentes de pH e COV, a separação em estágios permite outras estratégias de operação diferenciada entre os reatores, como por exemplo, temperatura, velocidade de agitação, adição de substâncias de controle das condições ambientais, como alcalinizantes ou inibidores, em cada reator.

A inferência estatística consiste na extrapolação para uma população inteira das características, chamadas de variáveis, estudadas em uma amostra, como por exemplo estimar a produção de AGV e bio-CH₄ em reatores com características operacionais e biomassas semelhantes a partir de uma amostra dessas produções. Uma das ferramentas utilizadas para estudar as variáveis da amostra é a regressão linear, que consiste em um método estatístico de trabalhar um conjunto de dados e encontrar um modelo, definido a princípio por uma equação linear, que descreve o comportamento do fenômeno estudado, relacionando variáveis explicativas com a variável explicada. Com isso é possível prever os valores da variável explicada que serão obtidos a partir de diferentes valores de entrada das variáveis explicativas (Dantas, 2005).

Esses modelos são chamados de preditores. Os modelos preditores da DA permitem prever a produção de bio-CH₄ e de AGV, economizando tempo experimental para saber essas produções, e guiando a realização de novos experimentos pelo conhecimento das variáveis mais significantes e seus valores ótimos (Niquini *et al.*, 2019; Raposo, Borja e Ibelli-Bianco, 2020). Além disso, a análise estatística necessária à geração dos modelos indica pontos atípicos entre os dados experimentais, indicando os dados com possíveis erros analíticos ou operacionais. Os pontos atípicos são classificados em *outliers*, pontos com grandes resíduos, e pontos influentes, pontos distantes da massa homogênea de dados e com pequena dispersão que mudam a direção da reta de regressão (Dantas, 2005).

Modelos de regressão linear simples, com uma única variável explicativa, não têm demonstrado bom desempenho em explicar a DA devido aos dados de produção de bio-CH₄ não serem lineares e a DA ter muitas variáveis intervenientes. A regressão linear múltipla, com mais de uma variável explicativa, tem tido poder explicação, representado pelo coeficiente de determinação R², com valores maiores. Contudo, modelos preditores gerados com inversas das variáveis, termos de ordem superior e interações entre as variáveis têm alcançados valores ainda melhores de R, motivando

a avaliação de transformações de linearização das variáveis e otimização do R^2 em escalas diferentes das já testadas (Niquini *et al.*, 2019; Xu *et al.*, 2014).

Assim sendo, o presente trabalho desenvolve modelos de regressão linear múltipla, com transformações para linearização das variáveis e para otimização do R^2 , preditores da produção de ácidos graxos voláteis e de metano a partir da codigestão anaeróbia de biomassa de microalgas e vinhaça de cana-de-açúcar em sistema de dois estágios. Também é realizada a seleção das variáveis mais importantes para a modelagem de cada reator, bem como a identificação de pontos atípicos e a validação do método de linearização das variáveis e otimização do R^2 . O experimento analisado, além de empregar a codigestão, também utilizou o pré-tratamento térmico.

Este trabalho está perfeitamente alinhado com alguns dos 17 objetivos de desenvolvimento sustentável (ODS) estabelecidos na agenda 2030 pela Organização das Nações Unidas (ONU). Esses ODS estabelecem metas do desenvolvimento sustentável nas áreas econômica, social e ambiental. A pesquisa de produção de energia proposta no presente trabalho faz parte do plano para o objetivo 7 – energia limpa e acessível, uma vez que o metano será aproveitado ao invés de ser lançado na atmosfera, e tem potencial para ser uma energia mais barata, muito menos poluidora e plenamente renovável em comparação com a advinda dos combustíveis fósseis. O uso de resíduos de ETE (BM), e indústrias (VC), faz parte do objetivo 11 – cidades e comunidades sustentáveis, devido a diminuir o impacto poluidor, o volume em aterros e à recuperação de recursos. O enquadramento no conceito de economia circular, a produção de matéria-prima para bioplásticos, levam ao objetivo 12 – assegurar padrões de produção e consumo sustentáveis. A associação de ETE como estações de recuperação de recursos e o próprio tratamento dos resíduos no processo estudado faz parte do objetivo 6 – assegurar a disponibilidade e gestão sustentável da água e saneamento para todas e todos.

Da mesma maneira, o presente trabalho está em concordância com a atualização brasileira do marco legal do saneamento através da lei nº 14.026 de 2020 que estabelece a universalização do saneamento básico no Brasil, inclusive por meio do acesso de pelo menos 90% da população a sistema de coleta e tratamento de esgotos até o dia 31 de dezembro de 2033, o que levará ao aumento do número de ETE, quantidade de esgoto tratado e consequentemente aumento de lagoas de estabilização com processos aeróbios que têm como resíduo a BM que pode ser tratada e recuperada de acordo com a presente pesquisa.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Modelar através da regressão multilinear com linearização das variáveis e com otimização do R^2 a produção de ácidos graxos voláteis e de metano a partir da codigestão anaeróbia de biomassa de microalgas e vinhaça de cana-de-açúcar em sistema de dois estágios.

2.2 OBJETIVOS ESPECÍFICOS

- Desenvolver e avaliar modelos de regressão multilinear pelo método dos mínimos quadrados dos dados de produção de ácidos graxos voláteis do reator acidogênico;
- Desenvolver e avaliar de modelos de regressão multilinear pelo método dos mínimos quadrados dos dados de produção de metano do reator metanogênico;
- Avaliar as variáveis independentes e as escalas de transformações de linearização das variáveis mais significativas para a geração dos modelos;
- Validar a linearização das variáveis por meio de transformações, juntamente com a otimização do valor do R^2 dos modelos de regressão multilinear dos dados de codigestão anaeróbia;

3 REVISÃO BIBLIOGRÁFICA

3.1 MODELOS PREDITORES APLICADOS À DIGESTÃO ANAERÓBIA GERADOS A PARTIR DE REGRESSÃO LINEAR

Segundo Niquini *et al.* (2019), o desempenho da DA na produção de metano geralmente é avaliado através de testes de potencial bioquímico de metano (PBM). Porém, os testes despendem longos períodos e trabalho intensivo para a realização, em torno de 30 dias, levando à busca por alternativas de avaliação da DA. Os modelos matemáticos têm se mostrado atrativos devido a gerarem respostas rapidamente e poderem ser continuamente usados após a sua geração, reduzindo a necessidade de análises laboratoriais.

Desde 2007, Gunaseelan constatou que os modelos de regressão linear simples, com uma única variável independente, são ineficazes em modelar a DA. A regressão linear múltipla (RLM) foi mais eficaz em prever a produção de metano. Só em 2014, Xu *et al.* desenvolveram modelos de RLM considerando interações entre variáveis (através de multiplicação e divisão entre variáveis) e termos quadráticos e cúbicos, chegando a maiores valores de R^2 . Mas as interações e os termos de ordem maior (potência acima de 1) foram aplicados apenas para a variável operacional substrato/efluente (S/E), podendo ter desconsiderado a provável relação não-linear entre outras variáveis com a produção de metano e as interações entre outras variáveis.

Niquini *et al.* (2019) realizaram os testes mais completos de RLM em relação à manipulação das variáveis encontrados na presente pesquisa. Eles consideraram 11 variáveis independentes, classificadas como relativas ao substrato: sólidos voláteis, extrativos, lignina, celulose e hemicelulose; relativas ao inóculo: alcalinidade e nitrogênio amoniacal total; operacionais: relação substrato/efluente, tamanho de partícula, razão carbono/nitrogênio (C/N) e sólidos totais. Os modelos foram testados considerando interações, inversas, e ordens superiores até 7 para todas as variáveis. Como resultado, alcançaram R^2 de 0,83 para uma amostra de 86 dados de diferentes experimentos de PBM de variadas biomassas lignocelulósicas de mais de um laboratório. Ou seja, conseguiram boa capacidade de explicação do modelo para uma amostra bastante diversa, como será explicado a seguir, indicando a vantagem de modelos lineares que consideram termos não lineares, fundamentando o método de transformações das variáveis aqui empregado.

Raposo, Borja e Ibelli-Bianco (2020) compilaram 26 trabalhos de geração de modelos preditores da produção de metano em testes PBM avaliando a possibilidade de um modelo global que se aplique a todos os tipos de amostras. Esta avaliação concluiu que é difícil chegar a tal modelo

global devido à grande variabilidade dos dados causada pela diversidade de metodologias e falta de precisão dos laboratórios.

Destaca-se que não foram encontrados trabalhos utilizando a geração de modelos preditores de RLM com três das escalas usadas no presente trabalho ($\ln(x)$, $\text{Raiz}(x)$, $1/\text{Raiz}(x)$), sendo, portanto, um acréscimo inovador para modelagem da DA.

3.2 DIGESTÃO ANAERÓBIA

A digestão anaeróbia é um dos mecanismos de degradação da matéria orgânica que ocorre em baixas concentrações de oxigênio, ou mesmo na anaerobiose, tendo como produtos finais metano, gás carbônico, água, gás sulfídrico, amônia, dentre outros. Esse mecanismo se dá em quatro fases: hidrólise, acidogênese, acetogênese e metanogênese. Cada fase ocorre pela atividade de diferentes micro-organismos: bactérias fermentativas, hidrolíticas e acidogênicas, nas fases de hidrólise e acidogênese; bactérias acetogênicas; na fase de acetogênese; *archaea* metanogênicas na fase de metanogênese (Chernicharo, 1997).

A seguir são detalhadas as fases da digestão anaeróbia e os fatores intervenientes em cada uma delas.

1. **Hidrólise:** fase de quebra de compostos orgânicos complexos (proteínas, carboidratos, lipídeos) a compostos orgânicos mais simples (aminoácidos, monossacarídeos, peptídeos) realizada pelas bactérias fermentativas hidrolíticas. As bactérias hidrolíticas secretam enzimas que fazem essa conversão. Os compostos orgânicos mais simples são melhor aproveitados pelos micro-organismos da fase seguinte. O sucesso da hidrólise depende múltiplos fatores, dentre eles: a temperatura no reator, tempo de residência do substrato no reator, composição do substrato, pH do meio, concentração de íons amônio, concentração de produtos da hidrólise, bem como o contato eficiente entre biomassa e o substrato (Machado, 2012; Chernicharo, 1997).
2. **Acidogênese:** fase de conversão dos compostos orgânicos mais simples a produtos fermentativos variados, sobretudo ácidos graxos voláteis (AGV), executada pelas bactérias fermentativas acidogênicas. Os AGV são ácidos orgânicos de cadeia curta (2 a 6 carbonos, etanóico, propanóico, butanóico, isobutanóico, pentanóico, isopentanóico, hexanóico), sendo o principal alimento das bactérias da fase seguinte. Além dos AGV, temos succinato, lactato e álcoois (Machado, 2012; Chernicharo, 1997).

3. **Acetogênese:** as bactérias acetogênicas oxidam compostos intermediários provenientes da acidogênese a substratos assimiláveis pelas *archaea* metanogênicas, especialmente acetato e hidrogênio. Dentre esses compostos intermediários, o acetato e o propionato são os compostos que geralmente se encontram em maior quantidade nos biodigestores anaeróbios (Machado, 2012; Vavilin *et al.*, 1995).
4. **Metanogênese:** fase final, que tem como produto de interesse o gás metano. Se dá pela ação das *archaea* metanogênicas de dois grupos: as acetoclásticas, que se alimentam de acetato, e as hidrogenotróficas, que se alimentam de hidrogênio. Ambas pertencem ao domínio Archea, porém as acetoclásticas são responsáveis pela maior parte da conversão a metano. Por serem os organismos mais sensíveis de todo o processo, são limitantes da eficiência da digestão anaeróbia, exigindo condições muito favoráveis. Elevadas concentrações de compostos tais como amônia, sais e AGV podem inibir as *archaea* metanogênicas, principalmente as acetoclásticas (Machado, 2012).

O uso da digestão anaeróbia para o tratamento de efluentes no Brasil é menos frequente do que o uso da digestão aeróbia devido à maior eficiência desta última. No entanto, ainda assim a digestão anaeróbia é muito utilizada e tem grande vantagem na recuperação de recursos, graças à produção do metano, biocombustível, e na preservação ambiental, uma vez que comporta grandes cargas orgânicas, maiores do que as comportadas pela degradação aeróbia (Machado, 2012).

3.3 LAGOAS DE ALTA TAXA

As lagoas de estabilização são reatores de tratamento de águas residuárias simples e de funcionamento natural que se caracterizam como grandes tanques de pequena profundidade, definidos geralmente por diques de terra. Nesses tanques há um processo simbiótico entre algas e bactérias. Entretanto, há diversos níveis de simplicidade operacional e requisitos de área para os diferentes tipos de lagoas de estabilização (Barroso Júnior, 2015; Machado, 2012).

As LAT são um tipo de lagoa de estabilização configurada normalmente em canais abertos com profundidade reduzida (0,2 a 1,0 m), que utiliza um sistema de mistura simples e controladores de fluxo (baffles) para otimização da produção de oxigênio através da fotossíntese. O misturador é usado para mover a cultura e deve fornecer velocidade média da água de 0,15 a 0,3 m/s, causando vórtices turbulentos que formam a mistura vertical. A mistura vertical aumenta o crescimento de microalgas, assegurando frequente exposição das células à luz solar, evitando a sedimentação das microalgas/flocos bacterianos e aumentando a difusão de nutrientes em torno da célula. Esta

profundidade, juntamente com elevadas cargas de nitrogênio e de fósforo das águas residuárias, com a agitação permite que as microalgas se proliferem a altas concentrações, com valores máximos de clorofila-a que podem ultrapassar 3.000 mg/m³ (Barroso Júnior, 2015).

Desta forma, este sistema é considerado um processo de tratamento de baixo custo que maximiza a utilização da energia solar e possibilita maior absorção de nutrientes dos esgotos, se comparado ao sistema de lagoa convencional. Além disso, há o benefício de aproveitamento do CO₂ (gás do efeito estufa) o que indica seu potencial de aplicação para o futuro (Barroso Júnior, 2015).

Por fim, o consumo do CO₂ pelas microalgas causa o aumento do pH a valores acima de 9,0, com pico de 11,00 ao meio-dia. A ocorrência de pH de 9,2 durante 24 horas condiciona a redução de 100% de *Escherichia Coli* e presumidamente outros patógenos. Uma vez que a LAT seja configurada para tempo de detenção hidráulica de 3 a 5 dias, com fluxo contínuo de esgoto e mecanismo de mistura completa, é possível obter elevada taxa de desinfecção no efluente da LAT, acentuada pela radiação solar que também age no decaimento dos organismos patogênicos (Barroso Júnior, 2015).

3.4 BIOMASSA DE MICROALGAS

A biomassa proveniente de lagoas de alta taxa é composta majoritariamente por microalgas, que agem na remoção de nutrientes e liberação de oxigênio e, em seguida, por bactérias aeróbias, responsáveis pela degradação da matéria orgânica. Essa relação entre as bactérias aeróbias e as microalgas é simbiótica, onde as bactérias liberam o CO₂ como produto da respiração, que é captado pelas microalgas, e as microalgas liberam O₂ aproveitado na respiração das bactérias. No entanto, a concentração de nutrientes presente nos esgotos domésticos e em grande parte dos esgotos industriais, a baixa profundidade das LAT, favorecendo a disponibilidade de energia luminosa, e o constante aporte de CO₂ da respiração das bactérias favorecem a superpopulação de microalgas na biomassa proveniente de LAT (Machado, 2012; Barroso Júnior, 2015).

As microalgas pertencem a um grupo heterogêneo de organismos, predominantemente aquáticos e geralmente microscópicos unicelulares, que podem formar colônias ou filamentos, bastante diversificados quanto a sua morfologia e estrutura celular. São caracterizadas pela presença de pigmentos, entre eles a clorofila, responsáveis por coloração variada e por mecanismo fotoautotrófico, sendo capazes de utilizar CO₂ para produzir oxigênio na presença de luz. Filogeneticamente, as microalgas são compostas de espécies procarióticas ou eucarióticas. A literatura traz informações a respeito de que existem dezenas de milhares de espécies de microalgas, algumas centenas são estudadas quanto ao seu conteúdo químico e apenas umas poucas espécies têm sido cultivadas em quantidades industriais (Machado, 2012; Goularte, 2014).

Segundo Machado (2012), o uso das microalgas para a produção de biocombustíveis tem sido um assunto de grande interesse e destaque em pesquisas acadêmicas e industriais, motivado por preocupações tais como a esgotabilidade dos combustíveis fósseis, a utilização de terras áridas e culturas alimentares para geração de bioenergia alternativa (biodiesel). Entre esses biocombustíveis está o metano, produto da digestão anaeróbia. Ornelas (2021) explica ainda que o cultivo de microalgas tem sido bastante estudado no campo da biotecnologia, uma vez que são uma fonte rica em carboidratos, lipídios, proteínas, pigmentos e vitaminas. Porém, quando produzidas para o tratamento de efluentes, é mais a sua valorização em fontes de biocombustíveis, como o biogás, biodiesel e bioetanol que é estudada.

A composição celular das microalgas atende os requisitos de nutrientes da microflora anaeróbia, uma vez que apresentam carbono, nitrogênio e fósforo, além de outros micronutrientes importantes como ferro, zinco e cobalto. A lignina contida nas microalgas é pouca, ou quase ausente, contribuindo para uma boa estabilidade e conversão de alta eficiência na DA, se comparada a outros substratos. Mas a composição em relação aos compostos orgânicos é muito variável, dependendo da espécie, sendo as composições para proteínas (6 – 52%), carboidratos (5– 23%) e lipídeos (7 – 23%). O alto teor de proteína que pode ser encontrado resulta na liberação de amônia, gerando toxicidade para os micro-organismos da DA e baixa relação C/N, em média de 10,2 para microalgas de água doce, enquanto que o valor ideal é em torno de 60. Outro problema para a eficiência da DA das microalgas é a parede celular ser recalcitrante, bloqueando a atividade enzimática do consórcio anaeróbio e reduzindo a sua biodegradabilidade (Machado, 2012).

Uma das soluções para a baixa relação C/N, que ocasiona desequilíbrio para a DA em valores abaixo de 20, é a codigestão das microalgas com um cosubstrato pobre em nitrogênio. (Sialve *et al.*, 2009). Além de o aumento da atividade enzimática dos micro-organismos, gerado pela presença de um cosubstrato de fácil degradação, contribuir para a quebra da parede celular das microalgas.

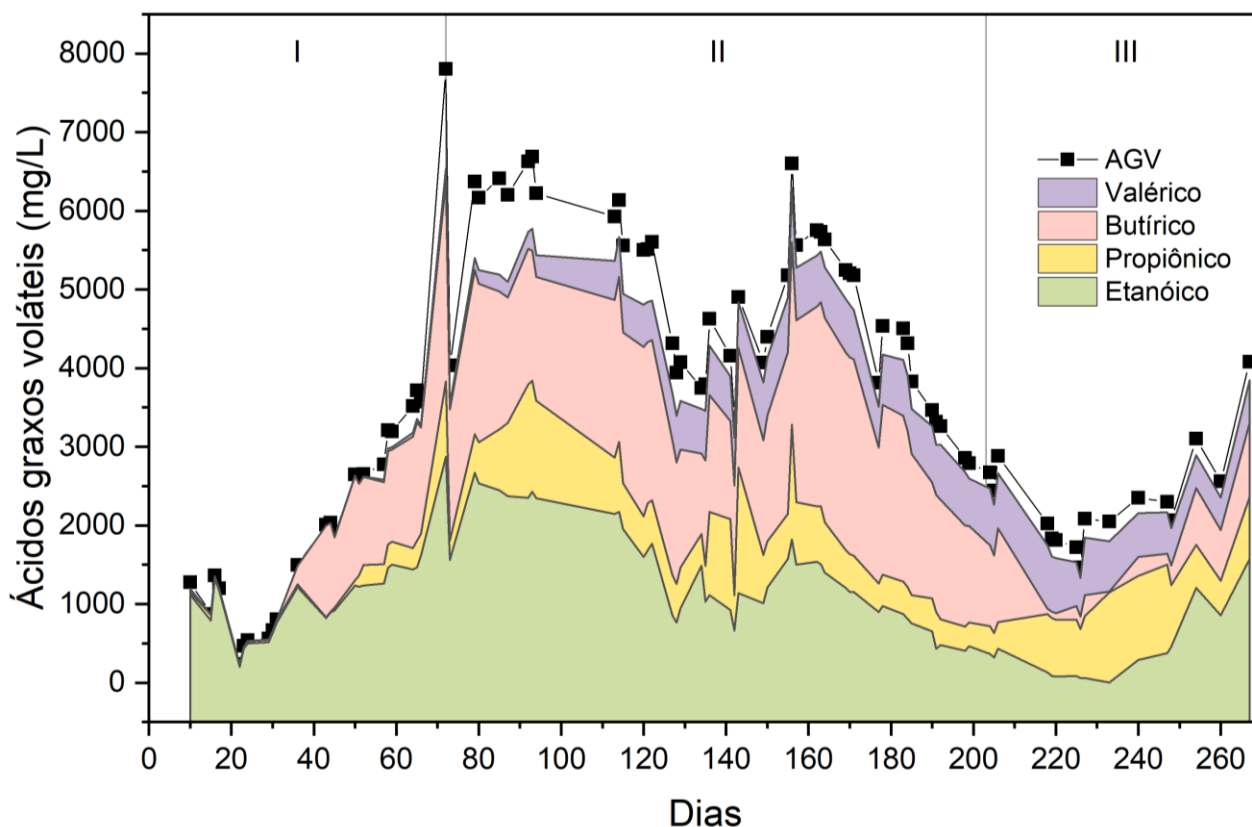
Goularte (2014) realizou experimentos de codigestão de microalgas da espécie spirulina sp. LEB 18 com glicerol e obteve ganhos na produção de biogás para a codigestão (0,16 – 0,24 L/d) quando comparada com a digestão de apenas spirulina sp. LEB 18 (0,05 L/d). Da mesma maneira, a produção específica de biometano foi maior na codigestão, sendo de 0,53 L.CH₄/gSV para 50% de spirulina, 50% de glicerol e concentração total de 10 gSV/L; contra 0,03 L.CH₄/gSV para 100% de spirulina e concentração total de 10 gSV/L.

4 MATERIAL E MÉTODOS

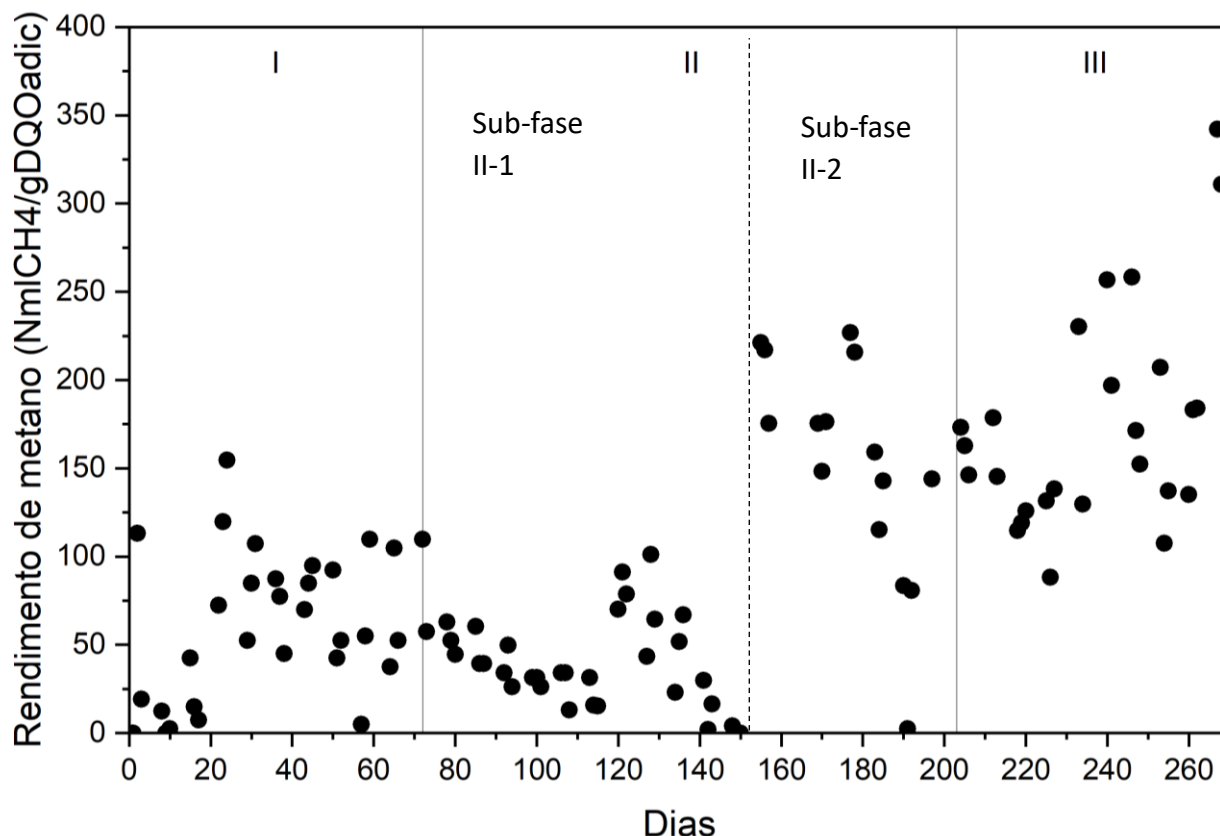
4.1 DADOS EXPERIMENTAIS USADOS NA PESQUISA

Os dados experimentais usados para a geração dos modelos de regressão foram obtidos em trabalho recentemente desenvolvido por Pinheiro (2023) no Laboratório de Saneamento Ambiental (LSA), no Centro de Tecnologia e Geociências (CTG) da Universidade Federal de Pernambuco (UFPE), em Recife, Pernambuco, Brasil. O experimento consistiu na codigestão anaeróbia operada em sistema de dois estágios do tipo *continuous stirred tank reactor* (CSTR), no qual o primeiro estágio caracterizou-se como reator acidogênico, produtor de AGV e o segundo estágio definiu-se como reator metanogênico, produtor de CH_4 . A separação em dois estágios deu-se, dentre as demais possibilidades de operação diferenciada entre os estágios, para garantir pH mais baixo (entre 5 e 6,5) e COV maior (10 a 15 gSV/L) necessário para o RA, e pH mais alto (em torno de 7) e COV menor (2 a 5 gSV/L) para o RM. O experimento foi dividido em três fases experimentais de acordo com os tipos de substratos digeridos, que foram vinhaça de cana-de-açúcar (VC) na fase I, para o *start-up* dos reatores; biomassa de microalgas (BM) e vinhaça na fase II; e biomassa de microalgas pré-tratadas (BM-PT) e vinhaça fase III. Essas mesmas fases experimentais dividiram a modelagem do experimento. Na Figura 1 pode ser visto o gráfico da produção de AGV e, na Figura 2, o gráfico do rendimento de metano, respectivamente, nos 270 dias do experimento.

Figura 1 – Evolução da concentração de ácidos graxos voláteis (AGV) no efluente do RA e afluente do RM



Fonte: adaptado de Pinheiro (2023).

Figura 2 – Evolução do rendimento de metano do RM

Fonte: Adaptado de Pinheiro (2023).

A mistura digerida foi previamente validada como sinérgica em trabalho desenvolvido por Silva (2022), também no LSA/UFPE. O referido trabalho se consistiu em testes de potencial bioquímico de metano (PBM) realizados após planejamento de misturas pelo método estatístico centróide-simplex para avaliar o sinergismo de misturas entre biomassa de microalgas e cosubstratos, com posterior otimização estatística dos resultados para encontrar a mistura ótima e o PBM ótimo. Entre essas misturas esteve a combinação de BM e VC, cujo índice de sinergia de $1,10 \pm 0,02$, maior do 1,0, indicou mistura sinérgica com 10% de produção de metano acima da esperada pelos resultados de PBM dos substratos separadamente.

4.2 INFERÊNCIA ESTATÍSTICA E REGRESSÃO LINEAR

A inferência estatística consiste em tirar conclusões sobre características de uma população através de métodos estatísticos aplicados em uma amostra extraída dessa população. O presente estudo analisa por meio da inferência estatística a produção de ácidos graxos voláteis (AGV) em reator acidogênico (RA) e de rendimento de metano em reator metanogênico (RM) a partir de amostras de experimento realizado em escala bancada.

4.3 REGRESSÃO LINEAR MÚLTIPLA PELO MÉTODO DOS MÍNIMOS QUADRADOS (MMQ)

No estudo corrente utilizou-se o MMQ, escolhido devido a ser o método clássico para regressões lineares, consolidado, com ampla disponibilidade de materiais para embasamento. O MMQ é definido como aquele que minimiza os quadrados dos desvios entre a reta, ou o plano, do modelo e os dados da amostra. Como consequência da minimização dos quadrados dos desvios, os próprios desvios são mínimos e assim a reta, ou o plano, gerada pelo método é a mais ajustada aos dados (Atkinson, 1985; Myers e Montgomery, 1997; Dantas, 2005).

Dado um conjunto de pontos, como o representado a seguir.

Tabela 1 – Exemplo de dados para modelo de regressão linear múltipla

X ₁	X ₂	...	X _k	y
X ₁₁	X ₁₂	...	X _{1k}	y ₁
X ₂₁	X ₂₂	...	X _{2k}	y ₂
.	.	...	.	.
.	.	...	.	.
.	.	...	.	.
X _{n1}	X _{n2}	...	X _{nk}	y _n

Fonte: o autor (2023).

A equação geral, na forma matricial, de um plano gerado pela regressão linear múltipla é definida, em termos das observações da Tabela 1, como na equação **Erro! Fonte de referência não encontrada.**:

$$y = X\beta + \varepsilon \quad (1)$$

Onde

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad X = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \quad e \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_k \end{pmatrix}$$

O vetor das observações y tem dimensão n, X é uma matriz (n × p) e o vetor dos níveis das variáveis independentes β tem dimensão (p = k + 1). O método dos MQ fornece o valor de β que minimiza a soma dos quadrados dos erros ε_i. A soma dos quadrados dos erros é

$$\begin{aligned} S(\beta) &= \sum_{i=1}^n \varepsilon_i^2 = \varepsilon' \varepsilon = (y - X\beta)'(y - X\beta) \\ &= y'y - \beta'X'y - y'X\beta + \beta'X'X\beta \\ &= y'y - 2\beta'X'y + \beta'X'X\beta \end{aligned} \quad (2)$$

Para minimizar o quadrado dos erros, derivamos parcialmente em relação aos coeficientes estimados e igualamos a zero, assim, os estimadores devem satisfazer a:

$$\left. \frac{\partial S}{\partial \beta} \right|_{\hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0 \quad \Rightarrow \quad X'X\hat{\beta} = X'y \quad (3)$$

A Equação (3) é denominada de equação normal dos mínimos quadrados na forma matricial. Desde que $X'X$ seja positiva definida, podemos resolver a Equação (3) multiplicando ambos os seus membros por $(X'X)^{-1}$. Portanto, os estimadores de mínimos quadrados de β são

$$\hat{\beta} = (X'X)^{-1}X'y \quad (4)$$

E o modelo de regressão ajustado é

$$\hat{y} = X\hat{\beta} \quad (5)$$

Na forma escalar, o modelo é $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_k x_{ik}$.

4.4 TESTE DE SIGNIFICÂNCIA GLOBAL DO MODELO PELA ANÁLISE DA VARIÂNCIA (Fc) E P-VALOR

As hipóteses testadas para aceitação ou rejeição dos modelos, onde β é o coeficiente de cada variável j constituinte do modelo, foram:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_K = 0 \text{ (hipótese nula)}$$

$$H_1: \beta_j \neq 0, \text{ para ao menos um } j$$

O teste usado para verificar a hipótese correta, foi a análise da variância através da estatística F_c , calculada conforme a equação (6), onde $\hat{Y}$ representa os pontos da curva gerada pelo modelo, $\bar{Y}$ representa a reta horizontal gerada pela média aritmética das n observações, Y representa os pontos dos dados observados no experimento, e k é o número de variáveis independentes do modelo. A diferença $n-k-1$ significa o número de graus de liberdade da variância não explicada e k também é o número de graus de liberdade da variância explicada (Myers *et al*, 2002; Dantas, 2005).

$$F_c = \frac{V_{exp}}{V_{nexp}} = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \cdot \frac{n-k-1}{k} \quad (6)$$

A estatística F_c é definida como a razão entre a variância explicada (equação 7) e a variância não explicada (equação 8). Buscou-se maior variância explicada e menor variância não explicada possíveis, maximizando F_c através de muitos graus de liberdade (muitos dados observados (n) e poucas variáveis independentes (k)). Isso levou ao cuidado em utilizar apenas as variáveis independentes necessárias e significantes para o modelo, em sintonia com o princípio da parcimônia.

$$V_{exp} = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{K} \quad (7)$$

$$V_{nexp} = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-k-1} \quad (8)$$

Os valores encontrados de F_c foram comparados com os valores tabelados ($F_{\alpha,k,n-k-1}$) em função do nível de significância pretendido para o modelo (α), graus de liberdade da variância explicada (k) e graus de liberdade da variância não explicada ($n-k-1$). Nos casos em que $F_c > F_{\alpha,k,n-k-1}$ aceitou-se a hipótese H_1 . Esses valores foram tabelados pelo estatístico Fischer através da distribuição de Snedecor, que rege o comportamento da estatística F_c .

Além da comparação entre F_c e $F_{\alpha,k,n-k-1}$, calculou-se o p-valor, que, para o teste de significância global do modelo, é a probabilidade de $F_{\alpha,k,n-k-1} > F_c$. Caso $\alpha > p$ -valor, aceitou-se H_1 . No presente trabalho, para o teste de significância global do modelo, utilizou-se $\alpha=1\%$. Para uma exposição mais intuitiva do nível de significância, muitas vezes foi expresso em porcentagem como $(1 - p\text{-valor}) \times 100$.

4.5 TESTE DE SIGNIFICÂNCIA INDIVIDUAL PARA SELEÇÃO DAS VARIÁVEIS INDEPENDENTES – ESTATÍSTICA t (t_j)

Para identificar quais variáveis foram significantes nos modelos, procedeu-se o teste das seguintes hipóteses, onde b é o coeficiente de cada variável j constituinte do modelo:

$$H_0: \beta_j = 0 \text{ (hipótese nula)}$$

$$H_1: \beta_j \neq 0$$

O teste usado para verificar a hipótese correta foi baseado na estatística t , dada pela equação (9), onde t_j é a estatística t para a variável j , β_j é o coeficiente da variável j na equação do modelo e $s(b_j)$ é o desvio padrão da variável testada. Vale salientar que o teste de significância individual também é realizado para a interseção da curva do modelo com o eixo das abscissas (Myers, Montgomery e Vining, 2002; Dantas, 2005).

$$t_j = \left| \frac{\beta_j}{s(\beta_j)} \right| \quad (9)$$

O desvio padrão $s(b_j)$ da variável j é calculado conforme a equação (14), em função do desvio padrão global do modelo (Se), dado pela equação (10), e dos desvios dos dados da variável j em relação à sua média ($x_i - \bar{x}$).

$$S(\beta_j) = S_e \cdot \frac{1}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (10)$$

Ressalta-se ainda que a estatística t mede o nível de dispersão da variável testada em relação à sua média, uma vez que um desvio padrão pequeno minimizará o denominador da equação e aumentará o valor da estatística t , indicando que os pontos estão concentrados próximos à média da variável e que a variável é significativa. Assim sendo, buscamos o maior valor possível para a estatística $|t|$, em módulo.

Os valores encontrados para t_j foram comparados com os valores tabelados ($t_{\alpha/2, n-k-1}$) em função do nível de significância ($\alpha/2$) e dos graus de liberdade do modelo ($n-k-1$). Nos casos em que $t_j > t_{\alpha/2, n-k-1}$ aceitou-se a hipótese H_1 , mantendo a variável avaliada no modelo, do contrário, a variável foi retirada do modelo. Os valores tabelados foram extraídos da distribuição T-student, que descreve a estatística t . Se utiliza o nível de significância $\alpha/2$ devido à distribuição ser bicaudal.

Além da comparação entre t_j e $t_{\alpha/2, n-k-1}$, calculou-se o p -valor, que, para o teste de significância individual de cada variável, é a probabilidade de $t_{\alpha/2, n-k-1} > t_j$. Caso $\alpha > p$ -valor, aceitou-se H_1 . No presente trabalho, para o teste de significância individual das variáveis, utilizou-se $\alpha=5\%$. Para uma exposição mais intuitiva do nível de significância, muitas vezes foi expresso em porcentagem como $(1 - p\text{-valor}) \times 100$. As variáveis que não foram significativas em cada modelo foram retiradas e o processo de geração do modelo recommençado.

4.6 COEFICIENTE DE CORRELAÇÃO E MULTICOLINEARIDADE (R)

Para medir a dependência linear entre as variáveis utilizadas em cada reator (acidogênico e metanogênico), empregou-se o coeficiente de correlação r , que verifica duas variáveis por vez, dado pela equação (11). Onde x_i representa o dado experimental i observados para a variável x e $\bar{x}$ representa a média dos dados experimentais para a variável x . Da mesma forma, y_i representa o dado experimental i observados para a variável y e $\bar{y}$ representa a média dos dados experimentais para a variável y .

$$r = \frac{\sum_{i=1}^n [(x_i - \bar{x}) \cdot (y_i - \bar{y})]}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} \quad -1 \leq r \leq 1 \quad (11)$$

A correlação pode ser direta, quando $r > 0$, correspondendo a uma reta inclinada crescente; nula, quando $r = 0$, correspondendo a uma reta horizontal; ou inversa, quando $r < 0$, correspondendo a uma reta inclinada decrescente.

A dependência linear foi classificada ainda em relação ao seu grau de correlação, conforme a Tabela 2.

Tabela 2 – Classificação da correlação em relação ao nível de dependência linear

Valor de r	Nível da correlação
$ r = 0,00$	Nula
$0,00 < r \leq 0,30$	Fraca
$0,30 < r \leq 0,70$	Média
$0,70 < r \leq 0,90$	Forte
$0,90 < r \leq 0,99$	Fortíssima
$ r = 1,00$	Perfeita

Fonte: adaptado de Dantas (2005).

O ideal é obter um modelo isento de correlação entre as variáveis independentes, no qual as variáveis sejam de fato linearmente independentes entre si, uma vez que uma nova variável colinear traz informações semelhantes às já existentes, podendo ser pouco significativas para o modelo. Isso é um dos pressupostos básicos da regressão linear pelo método dos mínimos quadrados. No entanto, o efeito químico de algumas variáveis independentes utilizadas no experimento influencia os fenômenos observados nas demais variáveis. Isto é chamado de multicolinearidade. A multicolinearidade foi usada para indicar restrições do modelo em função das faixas de valores empregadas para as variáveis correlacionadas. Foram consideradas multicolinearidades com $r \geq 0,70$ como restritoras do uso do modelo. O coeficiente de correlação foi utilizado apenas para identificar restrições no uso dos modelos, a escolha das variáveis independentes presentes nos modelos foi feita através do teste de significância individual das variáveis, já explicado na seção anterior (4.5).

4.7 COEFICIENTE DE DETERMINAÇÃO (R^2) E COEFICIENTE DE DETERMINAÇÃO AJUSTADO (R'^2)

O coeficiente de determinação (R^2) foi calculado segundo a equação (12), em que r é o coeficiente de correlação, SQR é o somatório dos quadrados dos resíduos explicados, $SQTO$ é o somatório dos quadrados dos resíduos totais, $\hat{Y}_i$ representa os pontos da curva gerada pelo modelo, $\bar{Y}$ representa a média aritmética das observações, e Y_i representa os pontos dos dados observados no experimento.

$$R^2 = r^2 = \frac{SQR}{SQTO} = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (12)$$

Ao relacionar a variabilidade explicada (SQR) com a variabilidade total ($SQTO$), o coeficiente de determinação mede quanto da variabilidade total do fenômeno é explicada pelo modelo, ou seja,

é uma medida do poder de explicação do modelo. Esse coeficiente varia entre zero, caso em que nada do fenômeno é explicado pelo modelo, e 1, caso em que a totalidade do fenômeno é explicada pelo modelo.

No entanto, R^2 sempre aumenta com a inclusão de uma nova variável, o que o torna menos adequado na avaliação de modelos com múltiplas variáveis independentes. Deste modo, o coeficiente de determinação ajustado (R') se torna mais adequado. O R' leva em consideração a quantidade de graus de liberdade do modelo, que sempre diminui com a inclusão de uma nova variável. Assim, R' só aumenta com a inclusão de uma nova variável se o ganho pela inclusão da variável supera a penalização da perda de um grau de liberdade. O limite do valor de R' é o próprio valor de R^2 e valores iguais de R^2 e R' indicariam que as informações trazidas pela inclusão de todas as variáveis compensariam a perda de graus de liberdade de maneira perfeita.

O R' foi calculado segundo a equação (13), onde R^2 é o coeficiente de determinação, n é a quantidade de dados experimentais usados na geração do modelo e k é o número de variáveis independentes. Vale ressaltar que $n-1$ compreende o número de graus de liberdade da variância total e $n-k-1$ é da variância não explicada.

$$R' = 1 - (1 - R) \cdot \frac{(n-1)}{(n-k-1)} \quad (13)$$

4.8 PARCIMÔNIA

Aplicou-se o princípio da parcimônia de Guilherme de Ockham (Occam's Razor). Este princípio, na estatística, enuncia que deve ser dada preferência ao melhor modelo mais simples possível, ou seja, o modelo que possui o maior poder de predição com o menor número de variáveis. Este princípio foi apoiado, inclusive, nos fatos anteriormente vistos: um menor número de variáveis independentes aumenta o valor da estatística F_c pelo aumento do número de graus de liberdade, variáveis colineares basicamente trazem as mesmas informações para o modelo, a seleção das variáveis que de fato vale a pena acrescentar de acordo com o R' .

4.9 DESVIO PADRÃO GLOBAL DO MODELO (S_e)

O desvio padrão do modelo (S_e) foi dado pela equação (14), em que Y_i representa o dado experimental do ponto i , $\hat{Y}_i$ representa o valor calculado pelo modelo para o ponto i , n significa a quantidade de dados utilizados na modelagem e k o número de variáveis independentes. A quantidade de graus de liberdade do modelo é dada pela diferença $n-k-1$, em que um grau de liberdade é perdido

para cada variável estimada pelo modelo (variáveis independentes) e mais um grau de liberdade é perdido na estimação da interseção do modelo.

$$Se = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-k-1}} \quad (14)$$

O desvio padrão global do modelo foi um dos critérios de seleção do melhor modelo ajustado, uma vez que o desvio padrão é uma medida da dispersão dos dados e que um menor desvio padrão indica um modelo mais preciso.

4.10 TRANSFORMAÇÕES DAS VARIÁVEIS PARA LINEARIZAÇÃO DOS DADOS, OTIMIZAÇÃO DO R² E ESCOLHA DAS VARIÁVEIS NOS MODELOS

As transformações das variáveis consistem em calcular os valores correspondentes a cada dado na nova escala, que pode ser quadrática ou logarítmica, por exemplo. De posse dos valores calculados, analisa-se a correlação linear entre a variável transformada e a variável dependente. Caso a correlação seja linear, calcula-se o R² referente à regressão linear com os dados naquela escala, caso a correlação não esteja linear, passa-se para a escala de teste seguinte. A escala eleita para cada variável é a que apresenta melhores R² e R', otimizando a capacidade de explicação do modelo além de linearizar a variável. Após a regressão linear e geração do modelo, as variáveis são devolvidas à escala original, realizando a operação inversa da que foi feita na transformação. Esse método encontra equações que não são lineares e se adequam melhor à curva da variável dependente em função das variáveis independentes, mesmo utilizando o método de regressão linear, seus pressupostos e seus indicadores da qualidade de ajuste.

No caso da regressão linear múltipla, um arranjo das variáveis nas escalas testadas é verificado por vez. Até que todos os arranjos possíveis tenham sido testados e o que apresentou melhor R² e melhor R' é escolhido. Mas esse modelo escolhido foi gerado com todas as variáveis, então retira-se do modelo aquela que apresentou menor significância de acordo com a estatística t e p-valor e repete-se a operação iterativamente até ter um modelo apenas com variáveis significantes, respeitando o princípio da parcimônia e a análise da contribuição trazida pelas variáveis através do R'. Este foi o método de seleção das variáveis empregado no presente estudo.

O método de transformações das variáveis foi descrito por Dantas (2005). Para os arranjos das variáveis nas escalas e suas respectivas regressões, utilizou-se o *software* SAB Comparativo®, desenvolvido por Dantas (2005) e pela Dantas Engenharia, que realiza este trabalho de maneira automática. As 7 escalas utilizadas nos testes de transformações foram as que seguem na Tabela 3.

Tabela 3 – Escalas de transformação de linearização das variáveis e otimização do coeficiente de determinação.

Escala
X
1/(x)
Ln(x)
Raiz(x)
1/Raiz(x)
(x) ²
1/(x) ²

Fonte: o autor (2023).

4.11 PRESSUPOSTOS BÁSICOS DO MÉTODO CLÁSSICO DE REGRESSÃO

Procedeu-se a padronização dos resíduos segundo a equação (15), onde $\hat{e}_i$ é o i-ésimo resíduo padronizado, “e” é o i-ésimo resíduo, S_e é o desvio padrão do modelo.

$$\hat{e}_i = \frac{e}{S_e} \quad (15)$$

A padronização leva os resíduos a terem média 0 e desvio padrão igual a 1, permitindo a análise do enquadramento dos dados e do modelo em alguns dos pressupostos básicos do método clássico de regressão linear.

A homoscedasticidade, independência, não autocorrelação e ausência de formação de *clusters* foram verificadas no gráfico de resíduos padronizados *versus* valores ajustados.

A normalidade dos dados foi verificada através da comparação do histograma dos resíduos com o valor típico de sino da distribuição normal.

Os *outliers* foram verificados pela presença de todos os resíduos padronizados no intervalo [-2;+2] no gráfico de resíduos padronizados *versus* valores ajustados (Myers, Montgomery e Vining, 2002; Dantas, 2005), os pontos com resíduo padronizado fora do intervalo foram classificados como outliers. No atual trabalho os pontos influentes são associados a erro experimental, erro analítico, eventos adversos no funcionamento dos reatores, ou à fase de *start-up* dos reatores.

A presença de pontos influentes foi verificada pelo teste formal da distância de Cook de maneira gráfica no gráfico das distâncias de Cook *versus* valores ajustados. A distância foi calculada pela equação (16), onde D_i é a distância de Cook para a i-ésima observação; n é o número de observações, $\hat{Y}_j$ são os ponto do modelo com todas as observações; $\hat{Y}_{j(i)}$ são os pontos do modelo sem a i-ésima observação; k é o número de variáveis independentes; o termo +1 considera a estimação do

intersecto e EQM é o erro quadrático médio, dado pela equação (17). Estes pontos exercem influência atípica sobre o modelo, fazendo com que haja uma diferença estatisticamente significativa entre o modelo gerado sem esse ponto e o modelo com esse ponto, sendo indicada a análise da remoção do ponto. No atual trabalho os pontos influentes são associados a erro experimental, erro analítico, eventos adversos no funcionamento dos reatores, ou à fase de *start-up* dos reatores (Cook e Weisberg, 1999; Dantas, 2005).

$$D_i = \frac{\sum_{i=1}^n (\hat{Y}_j - \hat{Y}_{j(i)})^2}{(k+1) \cdot EQM} \quad (16)$$

$$EQM = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n} \quad (17)$$

Na equação (17) EQM é o erro quadrático médio; Y_i é o i -ésimo dado observado; $\hat{Y}_i$ é o i -ésimo valor estimado pelo modelo; n é a quantidade de observações.

Para evitar a micronumerosidade, também chamada de subamostragem, foi seguida a regra definida pela equação (18), onde k é a quantidade de variáveis independentes, n é a quantidade de dados observado e o termo $+1$ considera a estimação do intersecto no modelo.

$$n \geq 3 \cdot (k+1) \quad (18)$$

4.12 GERAÇÃO DO MODELO ESTATÍSTICO PARA O REATOR ACIDOGÊNICO

Para o reator acidogênico (RA), iniciou-se a linearização e otimização dos modelos para as 3 fases experimentais com as 5 variáveis, Tabela 4, de acompanhamento da estabilidade e desempenho do reator.

Tabela 4 – Variáveis utilizadas na geração dos modelos de regressão para o reator acidogênico

Variáveis independentes	Codificação	Variável dependente	Codificação
Potencial hidrogeniônico (adimensional)	PH		
Potencial de oxi-redução (mV)	REDOX		
Alcalinidade total (mg CaCO ₃ /L)	ALCATOT	Concentração de ácidos graxos voláteis (mg/L)	AGV
Nitrogênio Amoniacal (mg NH ₃ -N/L)	AMON		

Fonte: o autor (2023).

Uma vez que os arranjos de verificação da linearização e otimização foram de 5 variáveis em 7 escalas, o total de arranjos de variáveis transformadas na linearização e otimização verificadas pelo

software foram de 2.520 testes. Estiveram disponíveis 84 dados de produção de AGV e 71 dados das variáveis independentes.

Procedeu-se a análise exploratória através do gráfico de dispersão dos pontos da variável dependente *versus* a variável independente, para todas as variáveis, verificando a relação entre elas. AGV e pH não tiveram correlação linear, apenas 0,053 demonstrada na linha de tendência do gráfico, com correta indicação de maior produção de AGV com pH entre 5,5 e 5,7, corroborando o resultado de Pinheiro (2023). Entre REDOX e AGV também não houve linearidade, -0,13, com sinal negativo correto, uma vez que quanto mais anaeróbio melhor. ALCATOT teve linearidade satisfatória com AGV, 0,76, crescente, conforme esperado. AMON também teve linearidade satisfatória com AGV, 0,67, crescente. O crescimento de AGV com AMON concorda com o resultado encontrado por Pinheiro (2023) e foi explicado pela decomposição eficiente da BM, substrato rico em proteína, nos momentos de maior concentração de AGV, aumentando a concentração de nitrogênio amoniacal, mas sem ultrapassar o limite de inibição da DA.

Apesar da linearidade satisfatória para duas variáveis sem as transformações, o objetivo de otimização do R^2 tornou os testes de transformação das variáveis mais relevantes, levando inclusive à opção de não apresentar os gráficos da análise exploratória. Assim, a finalidade principal foi verificar se o comportamento das variáveis esteve correto, em segundo plano o auxílio na escolha das escalas.

4.13 GERAÇÃO DO MODELO ESTATÍSTICO PARA O REATOR METANOGENICO

Para o reator metanogênico (RM), percebeu-se a necessidade de realizar a modelagem dividida em 4 fases diferentes, ao invés de 3 fases conforme a divisão experimental, uma vez que o método de quantificação do biogás, no dia 151, durante a fase II, foi mudado de micromedidor de gás para gravimetria por deslocamento de Hidróxido de Sódio a 3% m/v, seguindo Aquino *et al* (2007). Assim sendo, a fase II foi subdividida em Fase II-1 e Fase II-2. A mudança do aparelho de quantificação refletiu de maneira muito evidente no gráfico do rendimento de metano da fase II (Figura 2), em que houve um salto no rendimento de metano de 43 ± 28 , antes do dia 151, para 147 ± 61 , após 151, e pela comparação do coeficiente de variação sem a separação dos dois períodos, 84%, com o coeficiente de variação separando os dois períodos, 64% no período II-1 e 42% no período II-2. Ressalta-se que os dados estatísticos apresentados no corrente parágrafo foram calculados após a eliminação dos *outliers* indicados pelo modelo de regressão.

Assim, iniciou-se a linearização e otimização dos modelos para as 3 fases experimentais, com a fase II dividida em 2 sub-fases, com as 13 variáveis, Tabela 5, de acompanhamento da estabilidade e desempenho do reator.

Tabela 5 – Variáveis utilizadas na geração dos modelos de regressão para o reator metanogênico

Variáveis independentes	Codificação	Variável dependente	Codificação
Potencial hidrogeniônico (adimensional)	PH	Rendimento de metano (Nml CH ₄ /gDQO ad)	RENDCH4
Potencial de oxi-redução (mV)	REDOX		
Alcalinidade total (mg CaCO ₃ /L)	ALCATOT		
Alcalinidade parcial (mg CaCO ₃ /L)	ALCAPAR		
Alcalinidade intermediária (mg CaCO ₃ /L)	ALCAINT		
Relação alcalinidade intermediária/alcalinidade parcial, AI/AP (proporção)	AIAP		
Relação alcalinidade intermediária/alcalinidade total, AI/AP (proporção)	AIAT		
Nitrogênio amoniacal (mg NH ₃ -N/L)	AMON		
Concentração de ácido acético no afluente (mg/L)	ACE		
Concentração de ácido propanóico no afluente (mg/L)	PROP		
Concentração de ácido valérico no afluente (mg/L)	VALE		
Concentração de ácidos graxos voláteis no afluente (mg/L)	AGV		
Carga de demanda química de oxigênio afluente ao reator (mgO ₂)	DQO		

Fonte: o autor (2023).

Uma vez que os arranjos de verificação da linearização e otimização foram de 13 variáveis em 7 escalas, o total de arranjos de variáveis transformadas na linearização e otimização verificadas pelo *software* foram de 8.648.640 testes. Devido ao tempo necessário para o *software* realizar todos os cálculos, seria demorado demais realizar a otimização em única vez, então a modelagem foi realizada por etapas, incluindo novas variáveis aos modelos à medida em que as variáveis não significativas foram retiradas. Foram utilizados 106 dados. Ressalta-se que houve uma menor disponibilidade de dados de concentração de AGV afluente em comparação com a disponibilidade de dados das outras variáveis devido à maior complexidade e demora na análise cromatográfica de AGV.

A análise exploratória das variáveis indicou correlação de -0,76 entre pH e RENDCH₄, maiores rendimentos de CH₄ com pH entre 6,5 e 6,9, reforçando o valor de referência de 6,5 de Pinheiro (2023). Entre REDOX e RENDCH₄, correlação de -0,05, negativa em conformidade com a anaerobiose. As correlações negativas entre ALCATOT, ALCAPAR, ALCAINT com RENDCH₄ (-0,66, -0,65, -0,6, respectivamente) são inesperadas, uma vez que o tamponamento do sistema para manutenção do pH na faixa ideal é essencial para a DA. No entanto, Pinheiro (2023) concluiu que o consumo da alcalinidade nos momentos de maior rendimento se deu para controlar o pH no aumento de AGV do afluente na fase II e pelo consumo de bicarbonato na produção de CH₄ por *archaeas* metanogênicas hidrogenotróficas na fase III. AMON e RENDCH₄ tiveram correlação de -0,45, refletindo corretamente o malefício da inibição por amônia. DQO e RENDCH₄ com correlação de -0,4, sugerindo inibição da metanogênese por excesso de MO em relação à capacidade do reator com o aumento da DQO afluente.

Por fim, as variáveis ligadas aos ácidos graxos voláteis, ACE (-0,30), PROP (0,14), VALE (0,34), AGV (-0,07) tiveram correlação com sentido inesperado, uma vez que o acético é diretamente convertido a acetato, sendo o mais importante AGV para a metanogênese, enquanto o propiônico é o AGV com efeito mais inibitório e o valérico necessita ser primeiramente convertido a acético, para só então passar a acetato. Contudo, a observação dos gráficos do anexo C permite inferir que o momento de maior concentração de acético no afluente coincidiu com a baixa do rendimento de CH₄ causada pelo período de adaptação da microbiota ao novo substrato. Bem como o ganho no rendimento de metano provocado pelo pré-tratamento coincidiu com o período de menor concentração de acético no afluente, além das maiores concentrações de propiônico e valérico, que sempre estiveram abaixo do limite de inibição.

Ressalta-se que, da mesma maneira que o RA, o objetivo de otimização do R² dá maior relevância aos testes de transformação das variáveis do que à escolha das escalas pela análise exploratória. Assim sendo, foi preferido não apresentar os gráficos da análise exploratória, onde a finalidade principal foi verificar se o comportamento das variáveis esteve correto.

4.14 FASES EXPERIMENTAIS

Para acelerar o processo de *start-up*, a fase experimental I teve como único substrato a VC, devido à maior biodegradabilidade e adaptação do inóculo utilizado no RM, lodo anaeróbio proveniente de reator de circulação interna tratando efluente de cervejaria. Assim sendo, optou-se por apresentar apenas os resultados dos modelos das fases experimentais II e III, que tiveram a codigestão do substrato de interesse no estudo realizado por Pinheiro (2023), que é a biomassa de microalgas, bruta na fase II e pré-tratada na fase III.

5 RESULTADOS E DISCUSSÃO

5.1 MODELO PREDITOR DO REATOR ACIDOGÊNICO – FASE II

O modelo preditor para a fase II (dias 72 a 203) do reator acidogênico foi gerado a partir de 37 dados experimentais (dados 19 a 55), que podem ser vistos no apêndice A. O poder de explicação do modelo da fase experimental II chegou a 89%, conforme R^2 de 0,89, Tabela 6; R' de 0,88, quase o mesmo valor do R^2 , demonstrando que a inclusão das 2 variáveis independentes foi significativa para o modelo; alta estatística Fc (116,81), alta aceitação pelo nível de significância (99,999%) e baixo desvio padrão (0,0004). Os 4 *outliers* e pontos influentes ignorados no modelo são indicados na Tabela 7.

Tabela 6 - indicadores globais do ajuste do modelo preditor da fase II do reator acidogênico

Modelo	Clássico de Regressão
Desvio Padrão	0,0004
Estatística Fc	116,8125
Nível de Significância do Modelo	0,0001
Coefficiente de determinação	0,8862
Coefficiente de determinação ajustado	0,8786
Observações	33

Fonte: o autor (2023).

Tabela 7 - *Outliers* e pontos influentes retirados do modelo preditor da fase II do reator acidogênico

Dado	Dias	PH	REDOX	ALCATOT	AMON	AGV
2	73	5,72	-182,4	1558,3645	86,94	4031,95
20	135	6,12	-345,4	1898,8475	16,67	3791,16
23	142	6,08	-363,3	1797,655	6,03	3375,586
36	177	5,81	-327	1811,941	72,6	3815,066

Fonte: o autor (2023).

Seguindo a metodologia de seleção das variáveis através do teste de significância individual, descrita na seção 4.5, as variáveis que permaneceram no modelo definitivo foram PH e ALCATOT, de acordo com a Tabela 8. A variável ALCATOT ficou 0,43% abaixo do nível mínimo aceitável de significância, porém, devido à dificuldade em encontrar um modelo que atendesse aos critérios exigidos, essa pequena diferença foi tolerada.

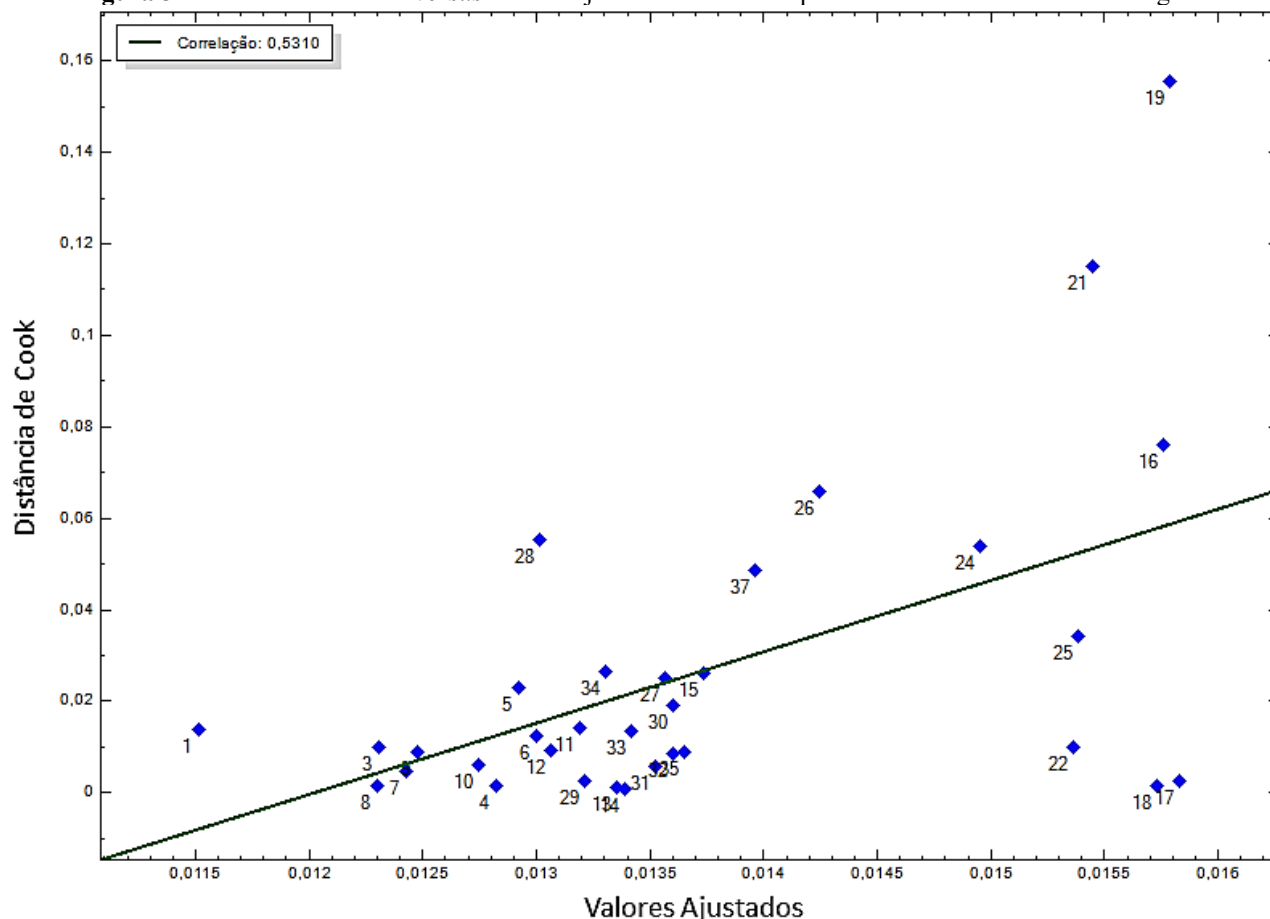
Tabela 8 - indicadores individuais da qualidade de ajuste das variáveis independentes do modelo preditor da fase II do reator acidogênico

Variável	Escala	Coefficientes	Desvio padrão	Estatística t	Nível de significância
Interseção		0,0268	0,0010	27,2868	0,0000
PH	1/(x) ²	-0,4557	0,0302	-15,0951	0,0000
ALCATOT	1/(x) ²	2137,2901	1066,9111	2,0033	0,0543
AGV	1/Raiz(x)				

Fonte: o autor (2023).

O modelo atendeu aos critérios de pontos influentes, como pode ser visto no gráfico das distâncias de Cook *versus* valores ajustados, Figura 3; de normalidade, como pode ser visto no histograma dos resíduos, Figura 4; da linearidade, como pode ser visto no gráfico dos valores observados *versus* valores ajustados, Figura 5;

Figura 3 – Distâncias de Cook *versus* valores ajustados do modelo preditor da fase II do reator acidogênico



O estudo do gráfico de resíduos padronizados *versus* valores ajustados, Figura 6, indicou a formação de *clusters*. Os pontos que compõem o *cluster* da direita são os pontos que vão do dia 127 ao 150. A localização desses pontos no gráfico de produção de AGV da Figura 1 demonstra que nesse período houve uma baixa na produção de AGV com recuperação em seguida, que foi explicada por Pinheiro (2023) pela instabilidade no valor de pH vivida pelo RA nesse período e que de fato são valores que destoam dos outros pontos. Desta forma, essa ocorrência era esperada e a fácil verificação, através dos recursos de modelagem empregados, de períodos com características atípicas demonstra o poder do método de revelar instabilidades, falhas ou eventos extraordinários no experimento estudado.

Figura 4 – Histograma dos resíduos do modelo preditor da fase II do reator acidogênico

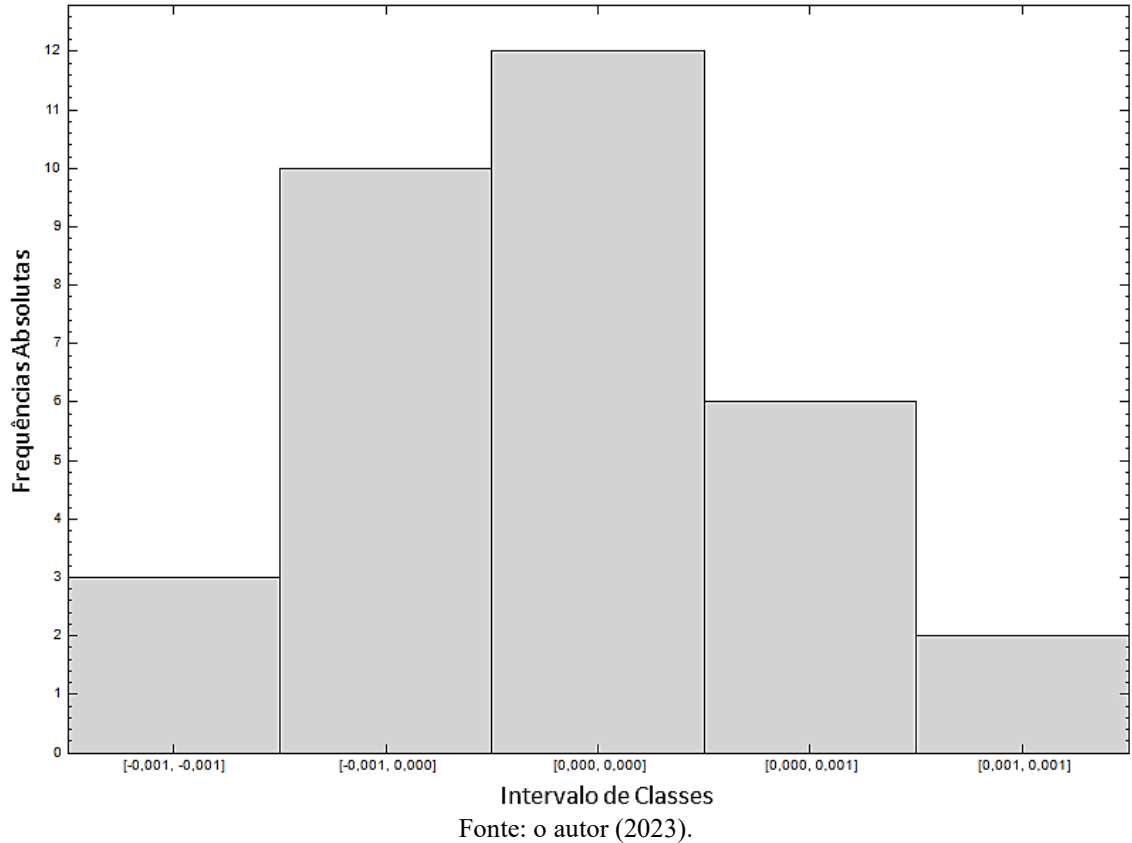
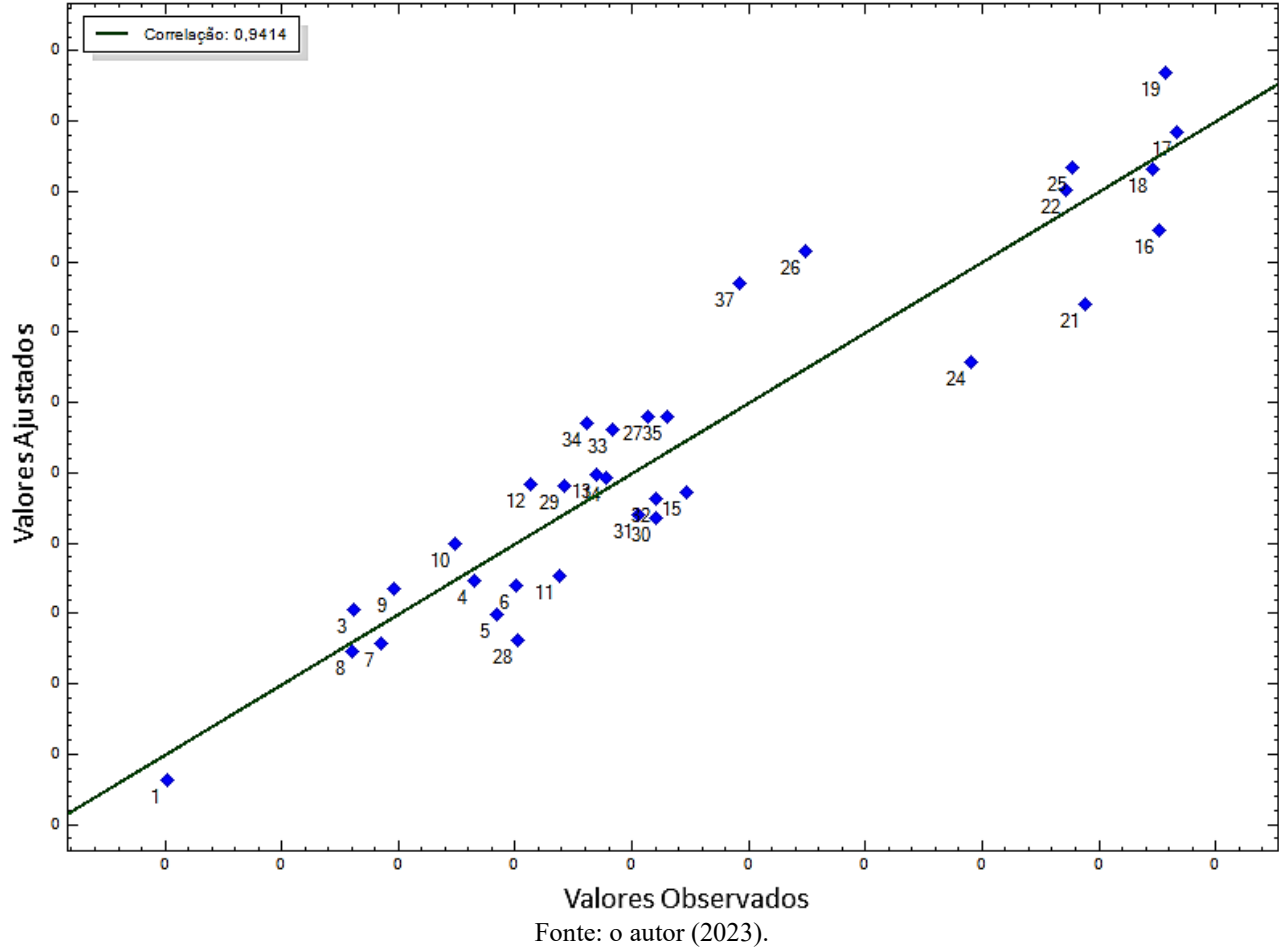


Figura 5 – Valores ajustados *versus* valores observados do modelo preditor da fase II do reator acidogênico



A heterocedasticidade no *cluster* da esquerda do gráfico de resíduos padronizados *versus* valores ajustados, Figura 6, com variância crescente à medida que os dias aumentam, é explicada pelo decréscimo contínuo na produção de AGV no fim da fase experimental II.

Percebe-se ainda a presença de dois *outliers* no gráfico de resíduos padronizados *versus* valores ajustados, o dado do dia 178 (dado 37) e o dado do dia 150 (dado 26). Contudo, o dado do dia 178 está seguindo o comportamento heterocedástico já explicado no parágrafo anterior e estaria acompanhado do ponto do dia 177 (dado 36), que já foi eliminado devido a ser um *outlier* mais severo, assim a eliminação também do dia 178 causaria a perda de informações que de fato fazem parte do experimento modelado, justificando a sua permanência no modelo. Em relação ao ponto do dia 150, esse é o ponto de transição entre os dois *clusters*, já explicados, e o ponto de transição entre o período de baixa produção de AGV e o da recuperação temporária da produção, sendo um ponto diferenciado em comparação com os demais por conta dessa peculiaridade, ao invés de ser um erro experimental ou analítico.

Além disso, quatro pontos atípicos já foram retirados do modelo, esses pontos não eram *outliers* no modelo do teste inicial, os indicadores da qualidade de ajuste do modelo estão ótimos apesar da presença destes *outliers* e a retirada de pontos do modelo não pode ser indiscriminada. Soma-se a isso que o gráfico de produção de AGV do RA, Figura 1, expõe que os dados possuem grande dispersão e a retirada de novos *outliers* levaria ao aparecimento de outros *outliers* continuamente, reduzindo demais a quantidade de valores observados no modelo, prejudicando a representatividade, normalidade e numerosidade do modelo.

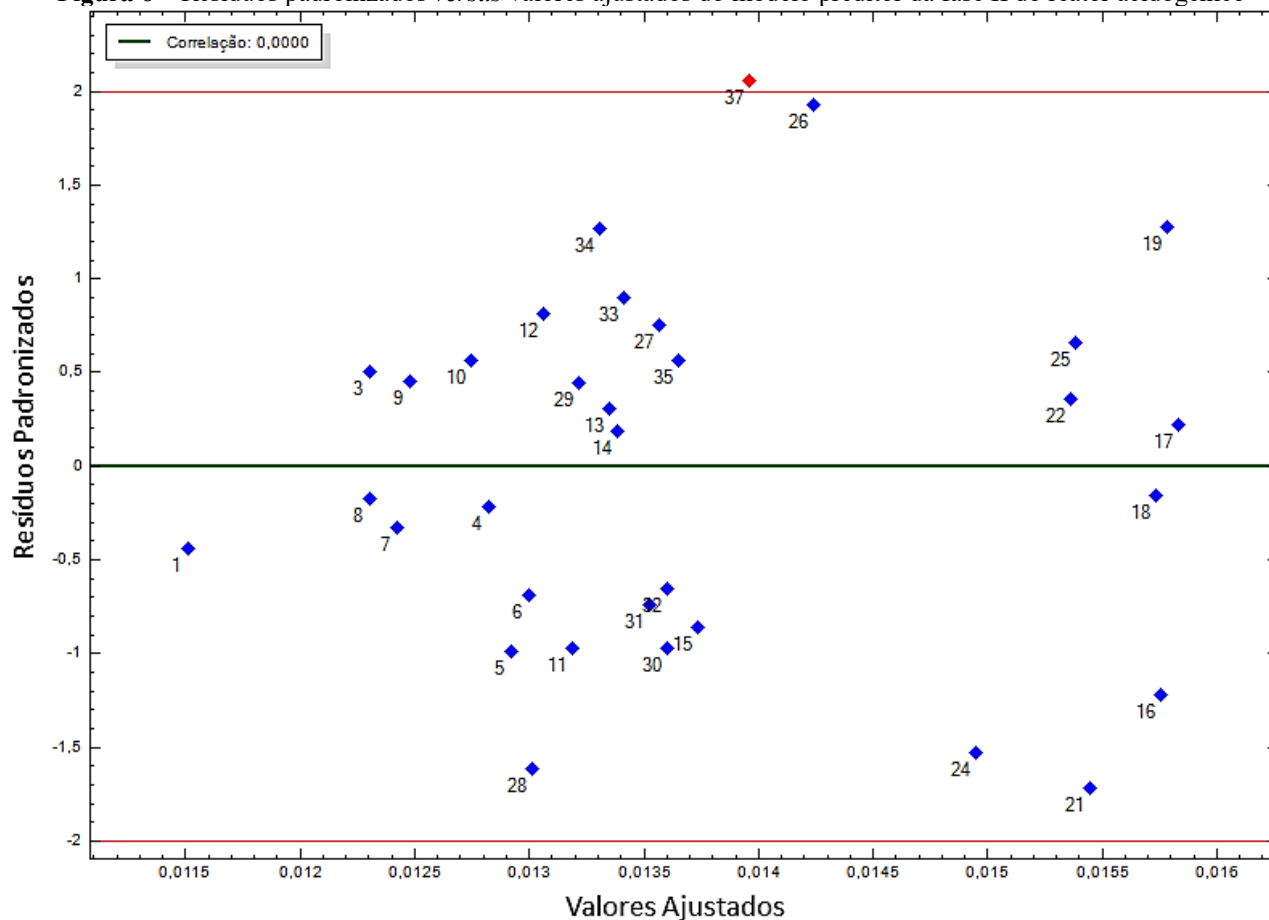
Tabela 9 – Matriz de correlação do modelo preditor da fase II do reator acidogênico

	PH	ALCATOT	AGV
PH	1,00		
ALCATOT	-0,03	1,00	
AGV	-0,93	0,15	1,00

Fonte: o autor (2023).

A matriz de correlação, Tabela 9, expõe dependência linear fortíssima entre AGV e pH pelo coeficiente de correlação de -0,93 (93%). Essa correlação inversa é explicada pelo fato de o pH médio do período de maior produção de AGV da fase II (dias 72 a 122) ter sido menor, $5,59 \pm 0,10$ (coeficiente de variação de 2%), enquanto que o pH médio do intervalo de menor produção de AGV (dias 127 a 150) ter sido maior, $6,14 \pm 0,14$ (coeficiente de variação de 2%), refletindo a dependência inversa nessa fase experimental. Destaca-se que os valores médios apresentados foram calculados eliminando os *outliers* e pontos influentes indicados pelo modelo.

Não há autocorrelação, como também pode ser visto no gráfico de resíduos padronizados *versus* valores ajustados na Figura 6. Também não há problema de micronumerosidade dos dados, uma vez que a regra de referência $n \geq 3(k+1)$ leva ao mínimo de 9 dados e foram utilizados 33 dados.

Figura 6 – Resíduos padronizados *versus* valores ajustados do modelo preditor da fase II do reator acidogênico

Fonte: o autor (2023).

A equação do modelo preditor da fase II do reator acidogênico é a equação (19).

$$AGV = \left(0,0268 - \frac{0,4557}{PH^2} + \frac{2137,2901}{ALCATOT^2} \right)^2 \quad (19)$$

O modelo de regressão linear sem a transformação das variáveis alcançou R^2 de 0,85 e R' de 0,84, Tabela 10, contra R^2 de 0,89 e R' de 0,88 do modelo com transformação, um aumento de 4% devido à transformação. Todos os pressupostos foram atendidos pelo modelo sem transformação e os critérios de aceitação foram semelhantes ao do modelo com transformação. A equação (20) é a do modelo sem transformação das escalas.

Tabela 10 - Indicadores globais do ajuste do modelo de regressão sem transformação das escalas fase II do reator acidogênico

Modelo	Clássico de Regressão
Desvio Padrão	354,9686
Estatística Fc	87,9874
Nível de Significância do Modelo	0,0001
Coeficiente de determinação	0,8544
Coeficiente de determinação ajustado	0,8446
Observações	33

Fonte: o autor (2023).

$$AGV = 22379,6552 - (3213,1394 \cdot PH) + (0,8450 \cdot ALCATOT) \quad (20)$$

5.2 MODELO PREDITOR DO REATOR ACIDOGÊNICO – FASE III

O modelo preditor da fase III (dias 204 a 270) do RA foi gerado com 16 dados (56 a 71), que podem ser vistos no apêndice A. Salienta-se que, por motivos operacionais, houve um maior intervalo entre as quantificações dos AGV na fase III, o que levou à menor disponibilidade e uma certa descontinuidade nos dados.

Chegou-se a um R^2 de 95%, que pode ser visualizado com demais indicadores globais na Tabela 11. O desvio padrão do modelo foi aceitável, 180,27, que comparado com a produção média de AGV da fase III, 2339,67 mgAGV/L, leva a coeficiente de variação de 7,71%, abaixo do valor máximo usual de 10%. No caso mais extremo, comparado à menor produção de AGV do período no dia 226, 1466,98 mgAGV/L, levaria a um coeficiente de variação de 12,28%, próximo de 10%. O único ponto influenciante é indicado na Tabela 12, e não houve *outliers*.

Tabela 11 - Indicadores globais do ajuste do modelo preditor da fase III do reator acidogênico

Modelo	Clássico de Regressão
Desvio Padrão	180,2715
Estatística Fc	43,7926
Nível de Significância do Modelo	0,0001
Coeficiente de determinação	0,9460
Coeficiente de determinação ajustado	0,9244
Observações	15

Fonte: o autor (2023).

Tabela 12 - Ponto influenciante retirado do modelo preditor da fase III do reator acidogênico

Dado	Dias	PH	REDOX	ALCATOT	AMON	AGV
15	73	5,72	-182,4	1558,3645	86,94	4031,95

Fonte: o autor (2023).

Todas as variáveis permaneceram no modelo e seus indicadores podem ser verificados na Tabela 13.

Tabela 13 - Indicadores individuais da qualidade de ajuste das variáveis independentes do modelo preditor da fase III do reator acidogênico

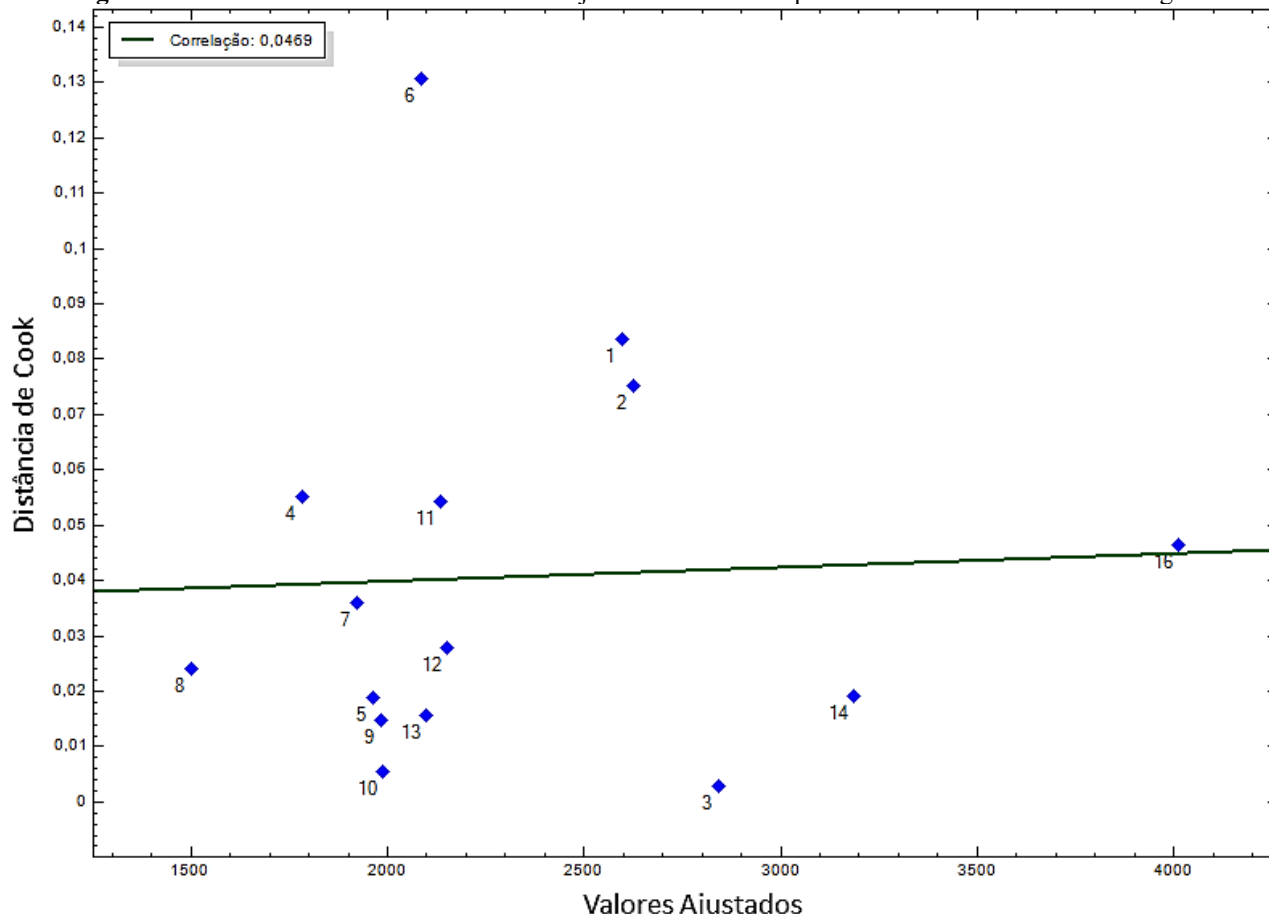
Variável	Escala	Coeficientes	Desvio Padrão	Estatística t	Nível de Significância
PH	1/(x) ²	242970,7308	28075,6913	8,6541	0,0000
REDOX	(x) ²	0,0153	0,0044	3,5078	0,0057
ALCATOT	1/(x) ²	-652429923,4562	159441286,7077	-4,0920	0,0022
AMON	1/(x)	-661,0759	261,9522	-2,5237	0,0302
AGV	X				

Fonte: o autor (2023).

O modelo atendeu aos critérios de pontos influentes, Figura 7; dos *outliers*, da formação de *clusters* e da autocorrelação, Figura 8; da linearidade, Figura 9;

O estudo do gráfico de resíduos padronizados *versus* valores ajustados, Figura 8, indicou heterocedasticidade dos dados, com decrescimento da variância. Contudo, não foi possível associar essa heterocedasticidade a um evento no experimento, uma vez que não há concentração de maior variância nos pontos iniciais ou finais do período avaliado, tendo pontos do começo e do fim do período tanto com menores como com maiores variâncias.

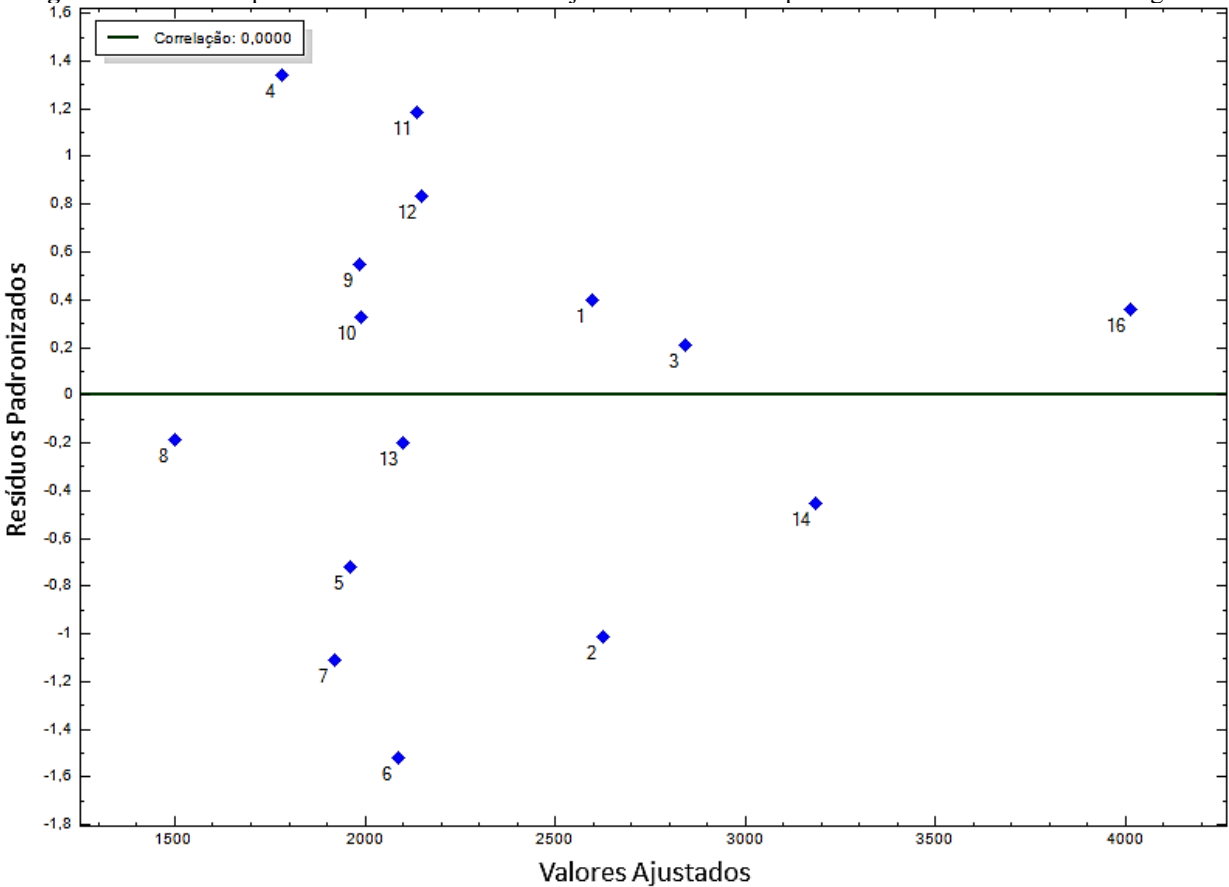
Figura 7 – Distâncias de Cook *versus* valores ajustados do modelo preditor da fase III do reator acidogênico



Fonte: o autor (2023).

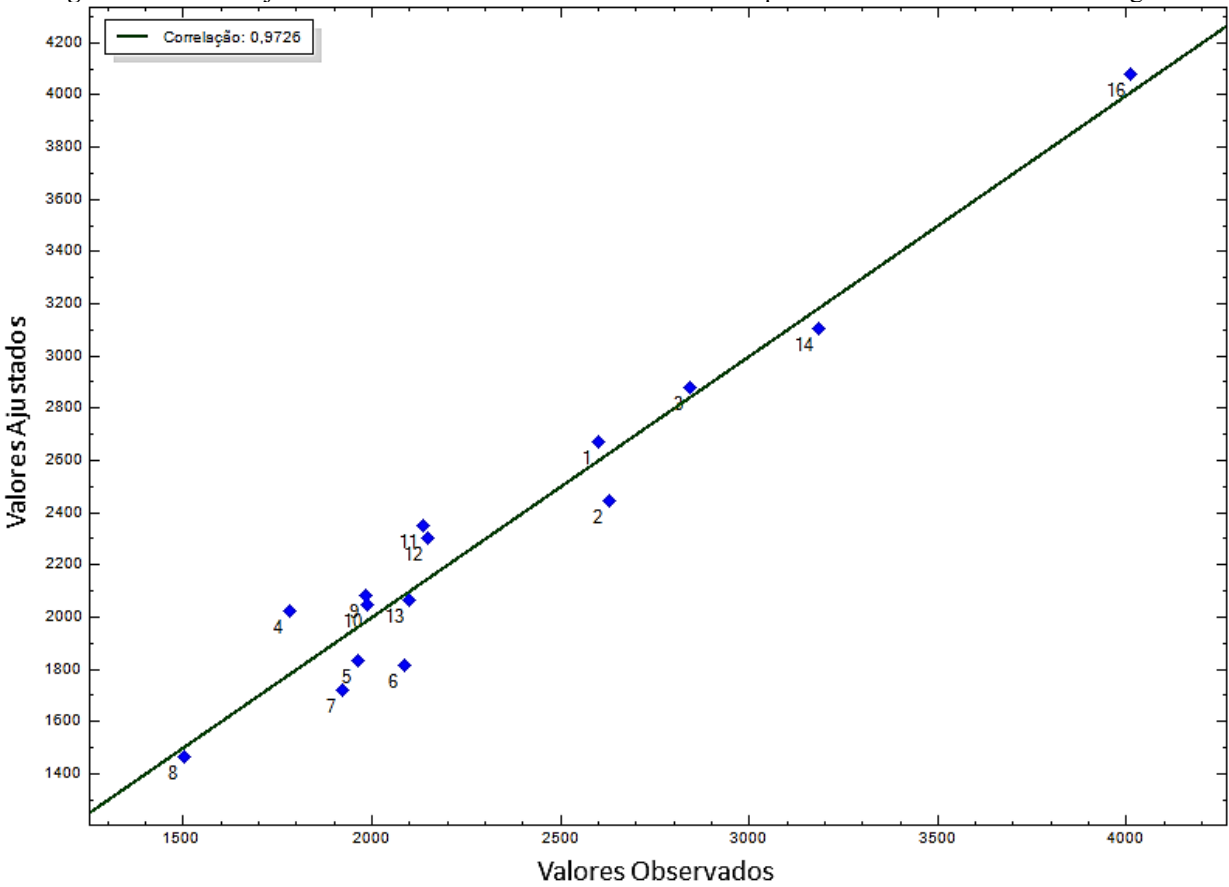
O histograma dos resíduos, Figura 10, sugere a falta de normalidade dos dados devido a não haver uma concentração de 68% dos resíduos no intervalo $[-1;1]$, havendo uma diferenciação do histograma em relação à forma de normalidade. Isso é explicado pelos maiores intervalos e descontinuidade entre as quantificações da produção de AGV na fase III do experimento, conforme explicação no primeiro parágrafo da corrente seção.

Figura 8 – Resíduos padronizados *versus* valores ajustados do modelo preditor da fase III do reator acidogênico

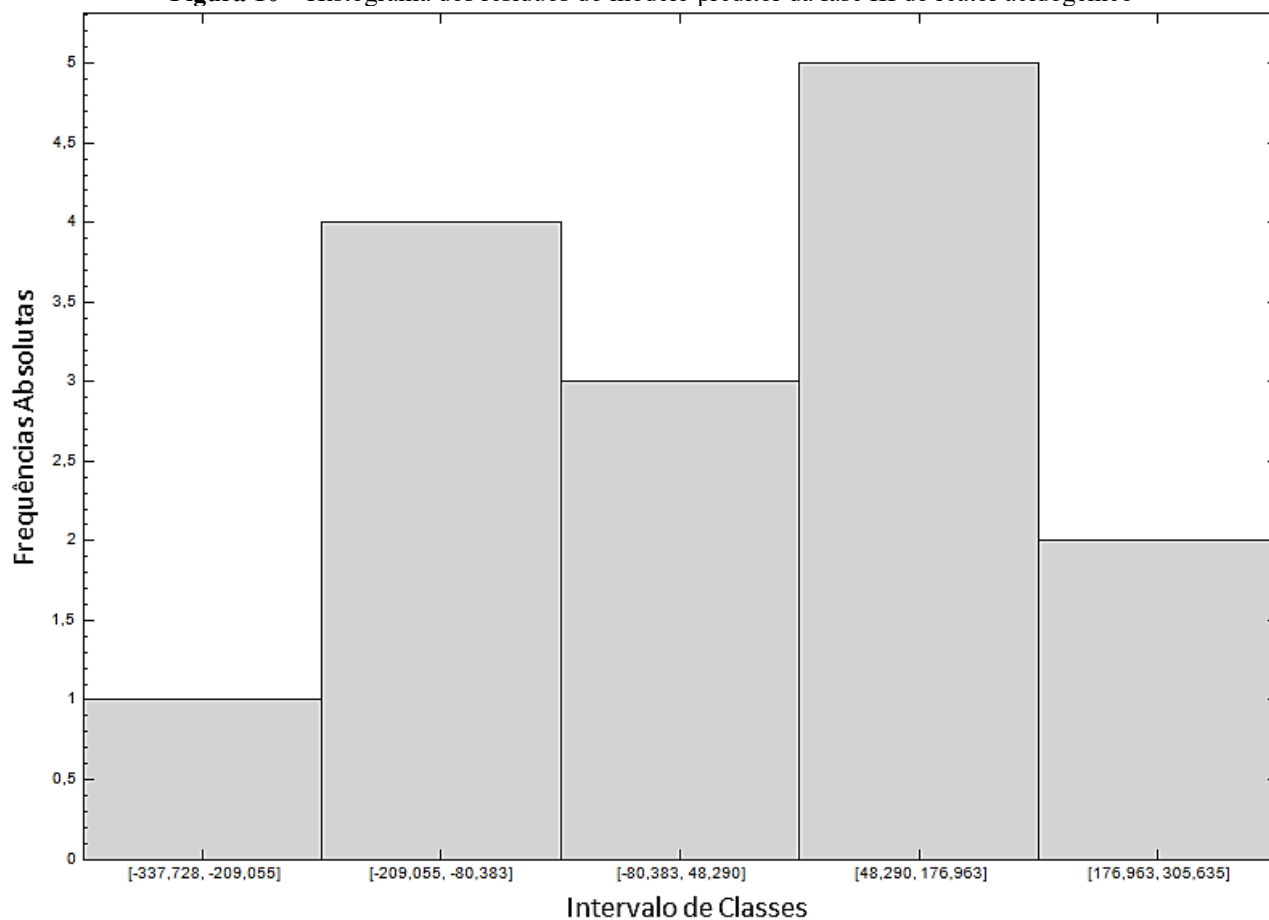


Fonte: o autor (2023).

Figura 9 – Valores ajustados *versus* valores observados do modelo preditor da fase III do reator acidogênico



Fonte: o autor (2023).

Figura 10 – Histograma dos resíduos do modelo preditor da fase III do reator acidogênico

Fonte: o autor (2023).

Tabela 14 – Matriz de correlação do modelo preditor da fase III do reator acidogênico

	PH	REDOX	ALCATOT	AMON	AGV
PH	1,00				
REDOX	-0,67	1,00			
ALCATOT	0,80	-0,38	1,00		
AMON	0,27	-0,13	-0,10	1,00	
AGV	0,91	-0,50	0,62	0,25	1,00

Fonte: o autor (2023).

A matriz de correlação, Tabela 14, expõe colinearidade forte de mesmo sentido entre PH e ALCATOT, com coeficiente de correlação de 0,80 (80%), sendo um restritor de uso do modelo devido a ser maior do que o limite de 70%. Por conta disto, o modelo só pode ser utilizado para prever produções de AGV em reatores que tenha como condição operacional altos valores de pH acompanhados de altos valores de alcalinidade total, ou baixos valores de pH acompanhados de baixos valores de alcalinidade total. Combinações de valores de pH e alcalinidade total diferentes dessas não podem ser explicadas pelo modelo gerado, sendo a proporção direta entre as duas variáveis a restrição desse modelo. Essa colinearidade já era esperada devido à alcalinidade influenciar forte e diretamente o pH pelo efeito químico de tamponamento, não sendo um problema para o modelo e as combinações permitidas são as possíveis na prática. O sinal negativo na correlação com REDOX era

esperado devido à maior anaerobiose favorecer a hidrólise e acidogênese. Devido à fraca dependência linear entre AMON e AGV, 25%, estudou-se a retirada da variável independente, no entanto o resultado foi perda no poder de explicação com redução do R^2 e piora da normalidade dos dados, assim AMON foi mantida.

A equação do modelo preditor a fase III do reator acidogênico é a equação (21).

$$AGV = -5425,6001 + \frac{2,4297 \cdot 10^5}{PH^2} + (0,0153 \cdot REDOX^2) - \frac{652429923,4562}{ALCATOT^2} - \frac{661,0759}{AMON} \quad (21)$$

O modelo de regressão sem a transformação das variáveis alcançou R^2 de 0,81 e R' de 0,75, Tabela 15, contra R^2 de 0,95 e R' de 0,92 do modelo com transformação, significando um aumento de 14% graças à transformação. O modelo sem transformações de escala ficou com uma variável a menos, AMON, finalizando com as variáveis PH, REDOX e ALCATOT, não havendo inclusive problema de micronumerosidade mesmo com a retirada de um ponto a mais. A equação (22) é a do modelo sem transformação das escalas.

Tabela 15 - Indicadores globais do ajuste do modelo de regressão sem transformação das escalas da fase III do reator acidogênico

Modelo	Clássico de Regressão
Desvio Padrão	212,3571
Estatística Fc	14,3075
Nível de Significância do Modelo	0,0006
Coeficiente de determinação	0,8110
Coeficiente de determinação ajustado	0,7544
Observações	14

Fonte: o autor (2023).

$$AGV = 6993,8850 - (1453,9347 \cdot PH) - (9,9495 \cdot REDOX) + (0,5848 \cdot ALCATOT) \quad (22)$$

5.3 MODELO PREDITOR DO REATOR METANOGENICO – FASE II-1

A fase experimental II durou dos dias 72 a 203. Por motivos detalhados na seção 4.13

Geração do modelo estatístico para o reator metanogênico, a fase II foi dividida em fase II-1, dos dias 72 a 154, e fase II-2, dos dias 155 a 203, permitindo modelagens mais ajustadas e trabalhando com faixas de rendimento de metano compatíveis em cada modelo. A fase II-1 35 dados (31 a 65) disponíveis que podem ser conferidos no apêndice A. A menor disponibilidade de dados de AGV,

explicada na seção anterior, permanece e o modelo que apresentou melhor resultado foi o que conteve variáveis de AGV, reduzindo para 26 dados, sem apresentar micronumerosidade.

Para a fase experimental II-1 chegou-se a um R^2 de 0,89, Tabela 16. O modelo teve três pontos retirados, Tabela 17.

Tabela 16 - Indicadores globais do ajuste do modelo preditor da fase II-1 do reator metanogênico

Modelo	Clássico de Regressão
Desvio Padrão	1,0123
Estatística Fc	28,5805
Nível de Significância do Modelo	0,0001
Coefficiente de determinação	0,8937
Coefficiente de determinação ajustado	0,8624
Observações	23

Fonte: o autor (2023).

Tabela 17 - *Outliers* e ponto influenciante retirados do modelo preditor da fase II-1 do reator metanogênico

Dado	DIAS	REDOX	ALCATOT	ALCAPAR	AIAT	ACE	PROP	VALE
16	127	-340	1216,25	982,5	0,80781	852,4	510	604,8
19	134	-319,5	1156,25	935	0,80865	1484	407	567,509
23	142	-305,7	1063,75	860	0,80846	662,904	450,904	613,348

Fonte: o autor (2023).

Conforme a metodologia de seleção das variáveis independentes já explicada (seção 4.5), dentre as 14 variáveis testadas inicialmente, 5 permaneceram no modelo definitivo e podem ser vistas com seus respectivos níveis de significância na Tabela 18.

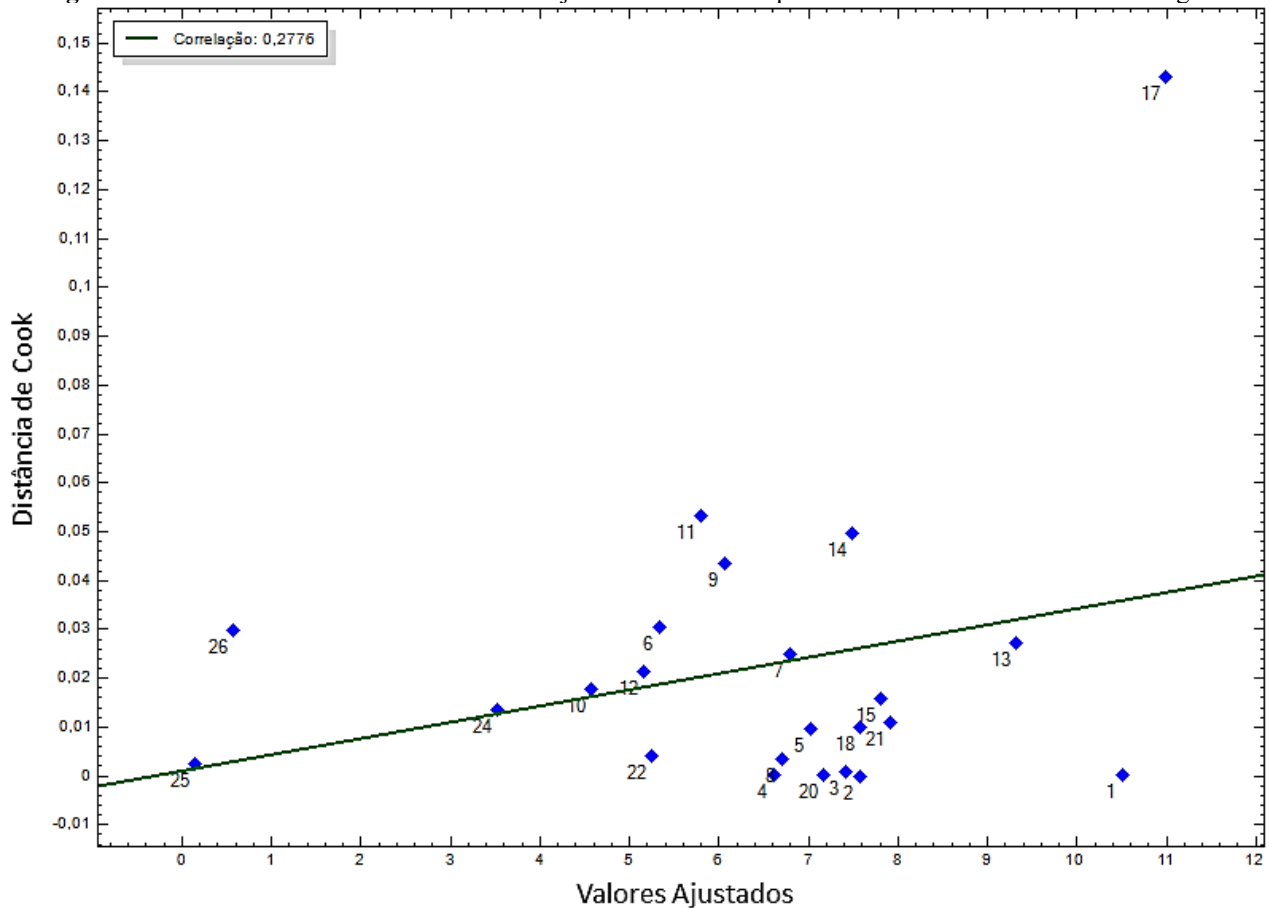
Tabela 18 - Indicadores individuais do ajuste das variáveis do modelo preditor da fase II-1 do reator metanogênico

Variável	Escala	Coefficientes	Desvio Padrão	Estatística t	Nível de Significância
Interseção		68,8193	8,2084	8,3840	0,0000
REDOX	$1/(x)^2$	95326,3153	22071,3075	4,3190	0,0005
ALCAPAR	$1/(x)^2$	-4318930,7940	655059,2440	-6,5932	0,0000
ACE	$\ln(x)$	-19,1512	2,1758	-8,8018	0,0000
PROP	$1/(x)$	3328,5825	542,9488	6,1306	0,0000
AGV	Raiz(x)	1,0588	0,1355	7,8156	0,0000
RENDCH4	Raiz(x)				

Fonte: o autor (2023).

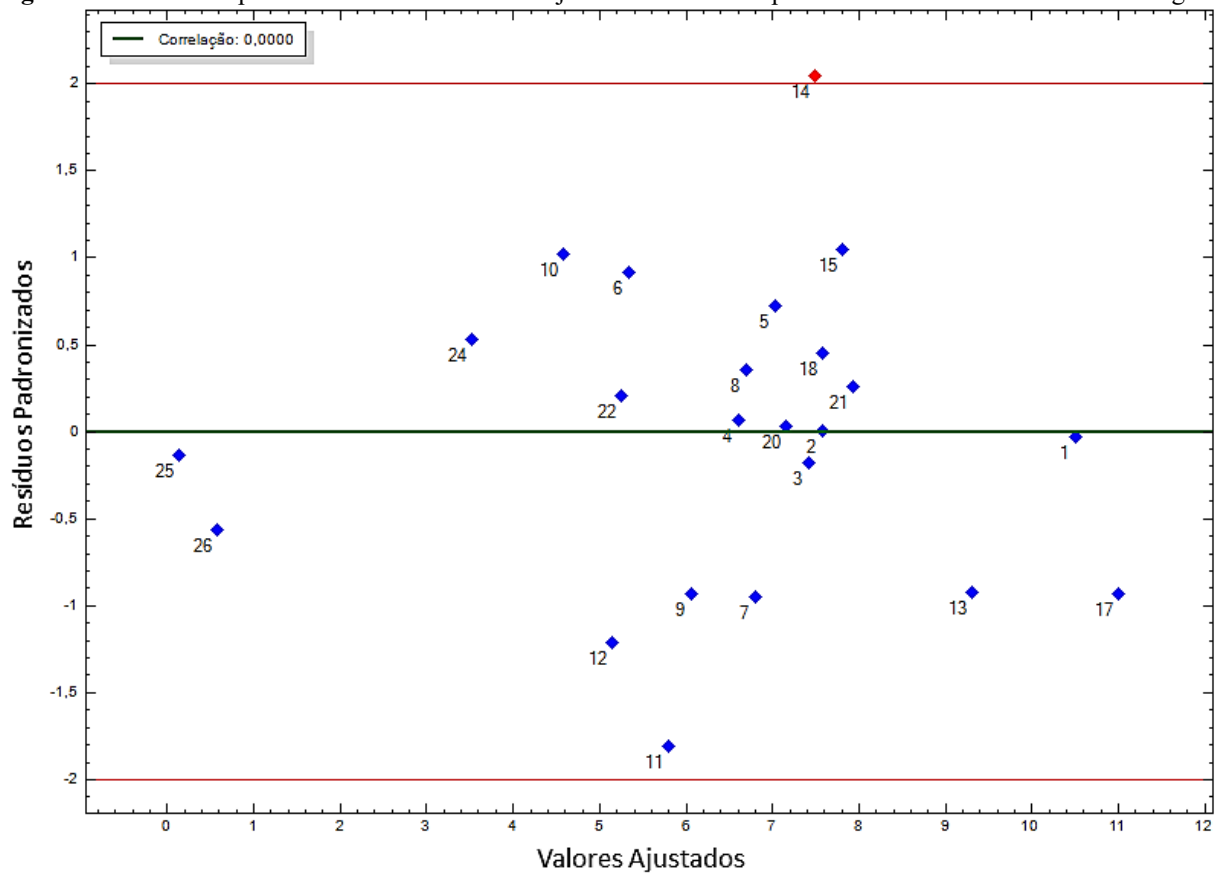
O modelo não apresentou pontos influentes, Figura 11; permaneceu um *outlier*, e não houve formação de *clusters* ou autocorrelação, Figura 12; o modelo foi linear, Figura 13; por fim, o modelo atendeu ao critério da normalidade, Figura 14.

Figura 11 – Distâncias de Cook *versus* valores ajustados do modelo preditor da fase II-1 do reator metanogênico



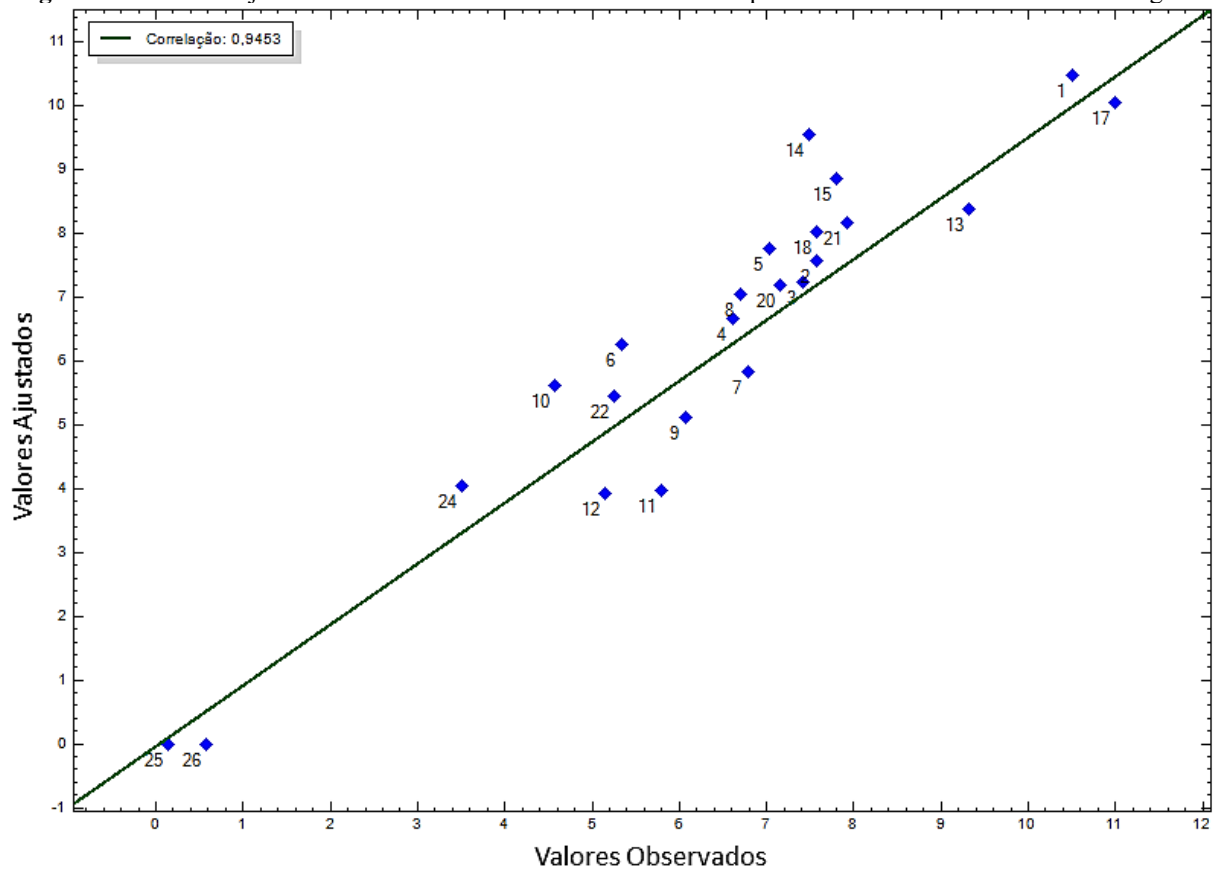
Fonte: o autor (2023).

Figura 12 – Resíduos padronizados *versus* valores ajustados do modelo preditor da fase II-1 do reator metanogênico



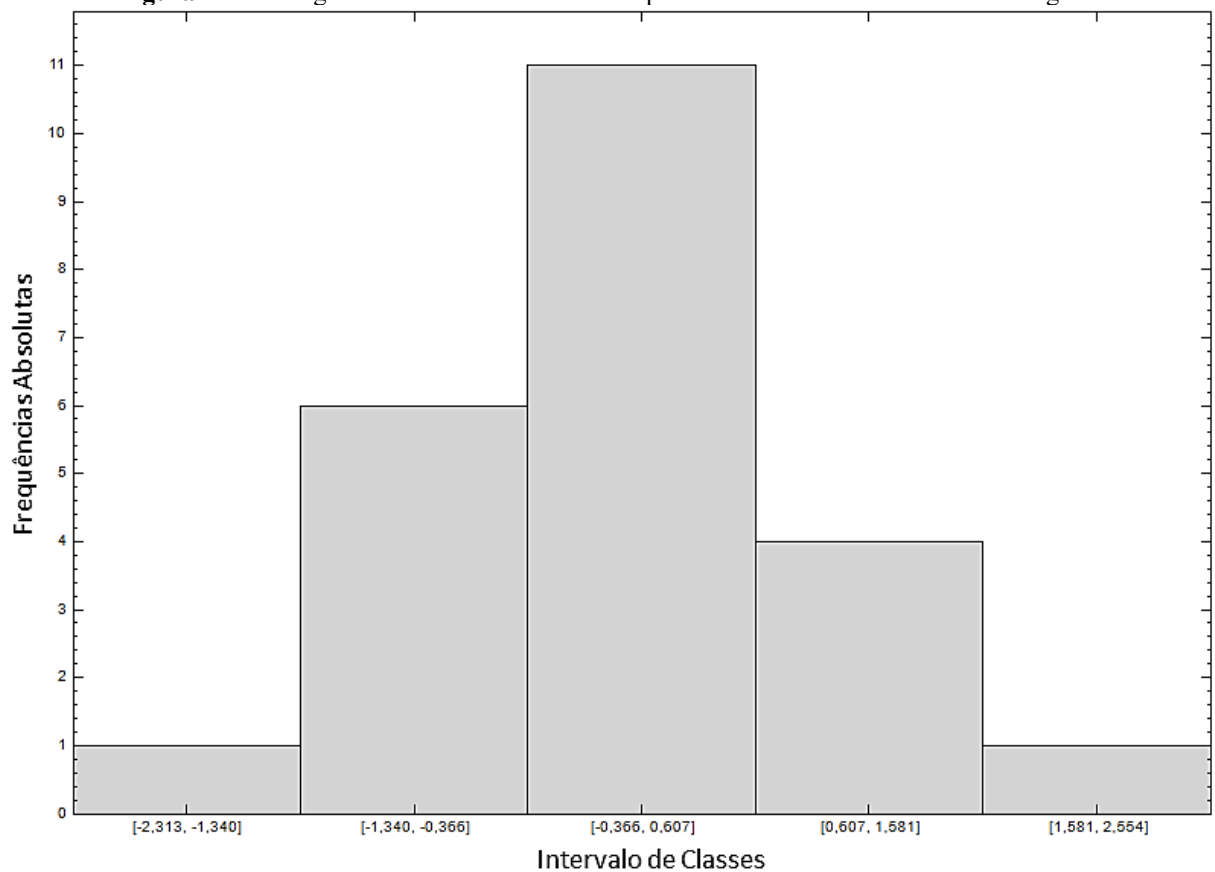
Fonte: o autor (2023).

Figura 13 – Valores ajustados *versus* valores observados do modelo predictor da fase II-1 do reator metanogênico



Fonte: o autor (2023).

Figura 14 – Histograma dos resíduos do modelo predictor da fase II-1 do reator metanogênico



Fonte: o autor (2023).

A colinearidade entre as variáveis ACE e AGV, Tabela 19, era esperada com conta de o ácido acético ter apresentado o seu período de maior concentração de todo os 270 dias de experimento na fase II-1, como pode ser visto no gráfico de produção de AGV na Figura 1, inclusive sendo o AGV mais abundante nesse período. Apesar da colinearidade já comentada e de haver correlação média apenas com ALCAPAR, os indicadores de qualidade de ajuste individual das variáveis foram satisfatórios, concluindo-se que nenhuma das variáveis deve ser eliminada do modelo.

Tabela 19 – Matriz de correlação do modelo preditor da fase II-1 do reator metanogênico

	REDOX	ALCAPAR	ACE	PROP	AGV	RENDCH4
REDOX	1,00					
ALCAPAR	0,25	1,00				
ACE	-0,36	-0,71	1,00			
PROP	-0,09	-0,11	-0,19	1,00		
AGV	-0,36	-0,62	0,93	-0,46	1,00	
RENDCH4	0,07	-0,58	0,14	0,17	0,22	1,00

Fonte: o autor (2023).

A equação do modelo preditor da fase II-1 do reator metanogênico é a (23).

$$RENDCH4 = \left[68,8193 + \frac{95326}{REDOX^2} - \frac{4318930,7940}{ALCAPAR^2} - (19,1512 \cdot \ln ACE) + \frac{3328,5825}{PROP} + (1,0588 \cdot AGV) \right]^2 \quad (23)$$

O modelo de regressão linear sem a transformação das variáveis alcançou R^2 de 0,88 e R' de 0,84, Tabela 20, contra R^2 de 0,89 e R' de 0,86 do modelo com transformação, aumento de apenas 1 ponto percentual atingido com a transformação. O modelo sem transformações de escala ficou sem intersecção e com uma variável a mais, VALE, finalizando com as variáveis REDOX, ALCAPAR, ACE, PROP, VALE e AGV. A equação (24) é a do modelo sem transformação das escalas.

Tabela 20 - Indicadores globais da qualidade de ajuste do modelo preditor da fase II-1 do reator metanogênico

Modelo	Clássico de Regressão
Desvio Padrão	11,9845
Estatística Fc	20,2871
Nível de Significância do Modelo	0,0001
Coefficiente de determinação	0,8775
Coefficiente de determinação ajustado	0,8414
Observações	23

Fonte: o autor (2023).

$$RENDCH4 = (0,1928 \cdot REDOX) + (0,0290 \cdot ALCAPAR) - (0,2070 \cdot ACE) - (0,0921 \cdot PROP) - (0,1778 \cdot VALE) + (0,1101 \cdot AGV) \quad (24)$$

5.4 MODELO PREDITOR DO REATOR METANOGÊNICO – FASE II-2

A fase II-2 durou dos dias 155 a 203, sendo um período de rendimento de metano muito acima dos 83 dias anteriores da fase II. Os dados utilizados vão de 66 a 80 e podem ser vistos no apêndice A. Por causa de o período em questão ser de poucos dias, apenas 15 dados foram referentes a esse período, o que levou a um problema de micronumerosidade, por conta de o modelo ter sido significativo com 4 variáveis independentes e um ponto influenciante ter sido descartado. Caso esse ponto não fosse retirado, o modelo ficaria exatamente com o número mínimo aceitável de dados [$3(k+1)=15$], no entanto avaliou-se que o prejuízo pela manutenção do ponto influenciante seria maior do que o problema de micronumerosidade, uma vez que o ponto influenciante alterava demais a direção da reta de regressão.

Para a fase experimental II-2 chegou-se a um R^2 de 0,90, Tabela 21. O modelo teve um único ponto influenciante, Tabela 22.

Tabela 21 - Indicadores globais do ajuste do modelo preditor da fase II-2 do reator metanogênico

Modelo	Clássico de Regressão
Desvio Padrão	22,7806
Estatística Fc	21,3340
Nível de Significância do Modelo	0,0001
Coeficiente de determinação	0,9046
Coeficiente de determinação ajustado	0,8622
Observações	14

Fonte: o autor (2023).

Tabela 22 - Ponto influenciante retirado do modelo preditor da fase II-2 do reator metanogênico

Dado	DIAS	PH	REDOX	ALCATOT	ALCAPAR	AIAT	DQO	RENDCH4
1	155	6,84	-349,2	960	756,25	0,78776	0,70368	221,17

Fonte: o autor (2023).

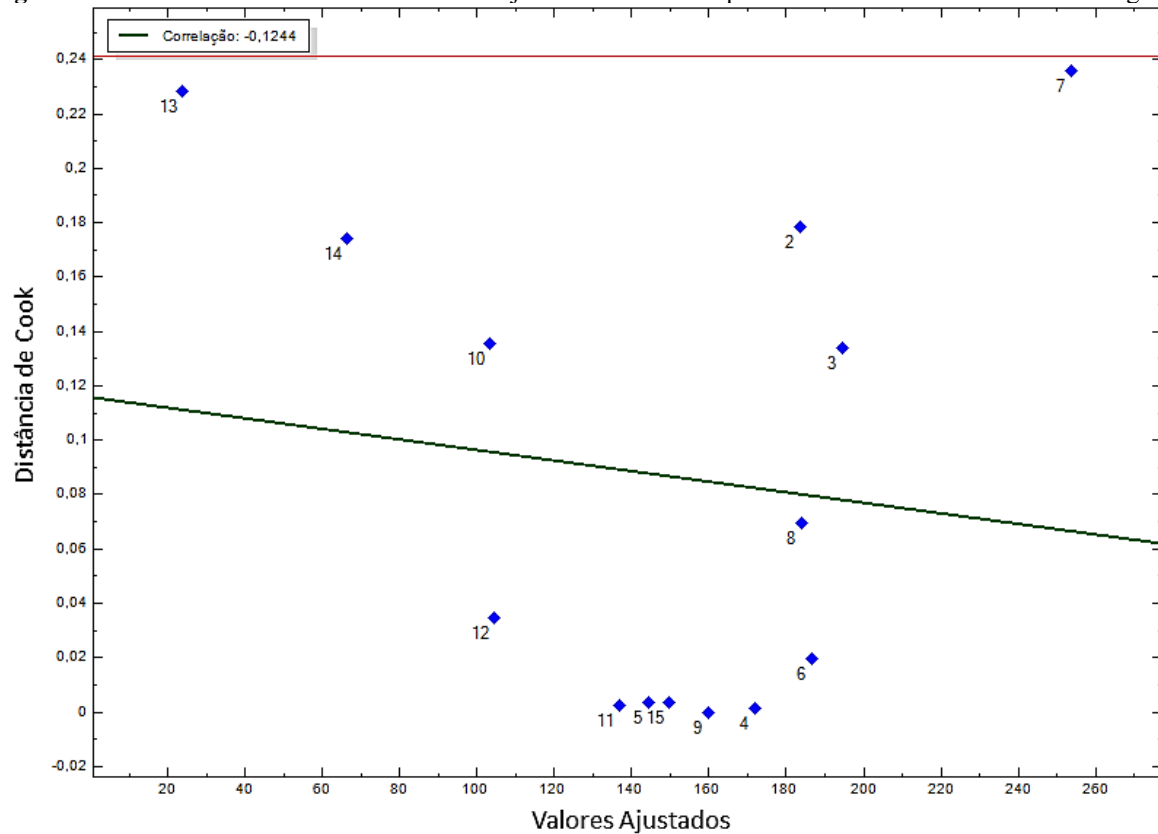
Dentre as variáveis testadas inicialmente, 4 permaneceram no modelo definitivo e foram REDOX, ALCATOT, ALCAPAR, AIAT, que podem ser vistas com seus respectivos níveis de significância na Tabela 23. A variável com menor significância foi AMON, 99,73%.

Tabela 23 - Indicadores individuais do ajuste das variáveis do modelo preditor da fase II-2 do reator metanogênico

Variável	Escala	Coefficientes	Desvio Padrão	Estatística t	Nível de Significância
Interseção		14911,7725	3457,5056	4,3129	0,0020
REDOX	(x) ²	0,0010	0,0002	4,4203	0,0017
ALCATOT	1/(x)	-8323382,4695	2024374,2627	-4,1116	0,0026
ALCAPAR	1/(x) ²	2437526807,4043	595424010,0607	4,0938	0,0027
AIAT	1/(x)	-8267,9880	1888,8304	-4,3773	0,0018
RENDCH4	x				

Fonte: o autor (2023).

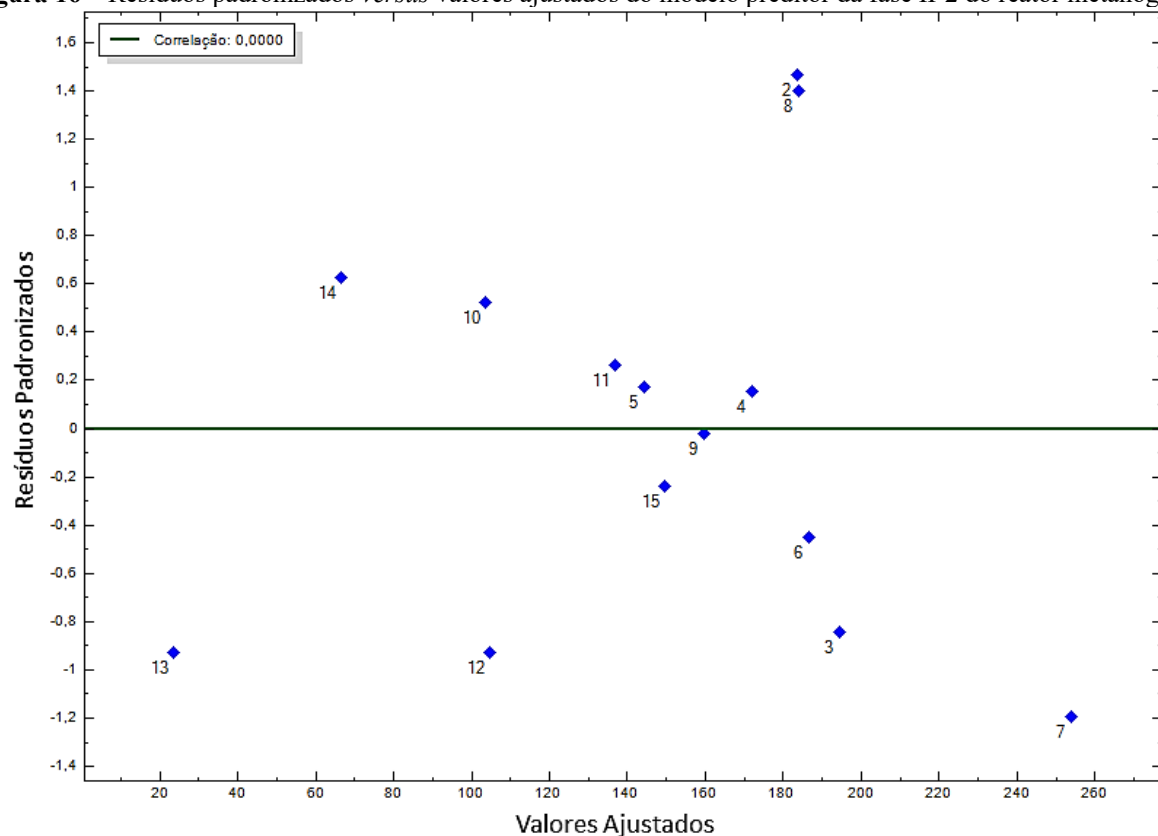
Figura 15 – Distâncias de Cook *versus* valores ajustados do modelo preditor da fase II-2 do reator metanogênico



Fonte: o autor (2023).

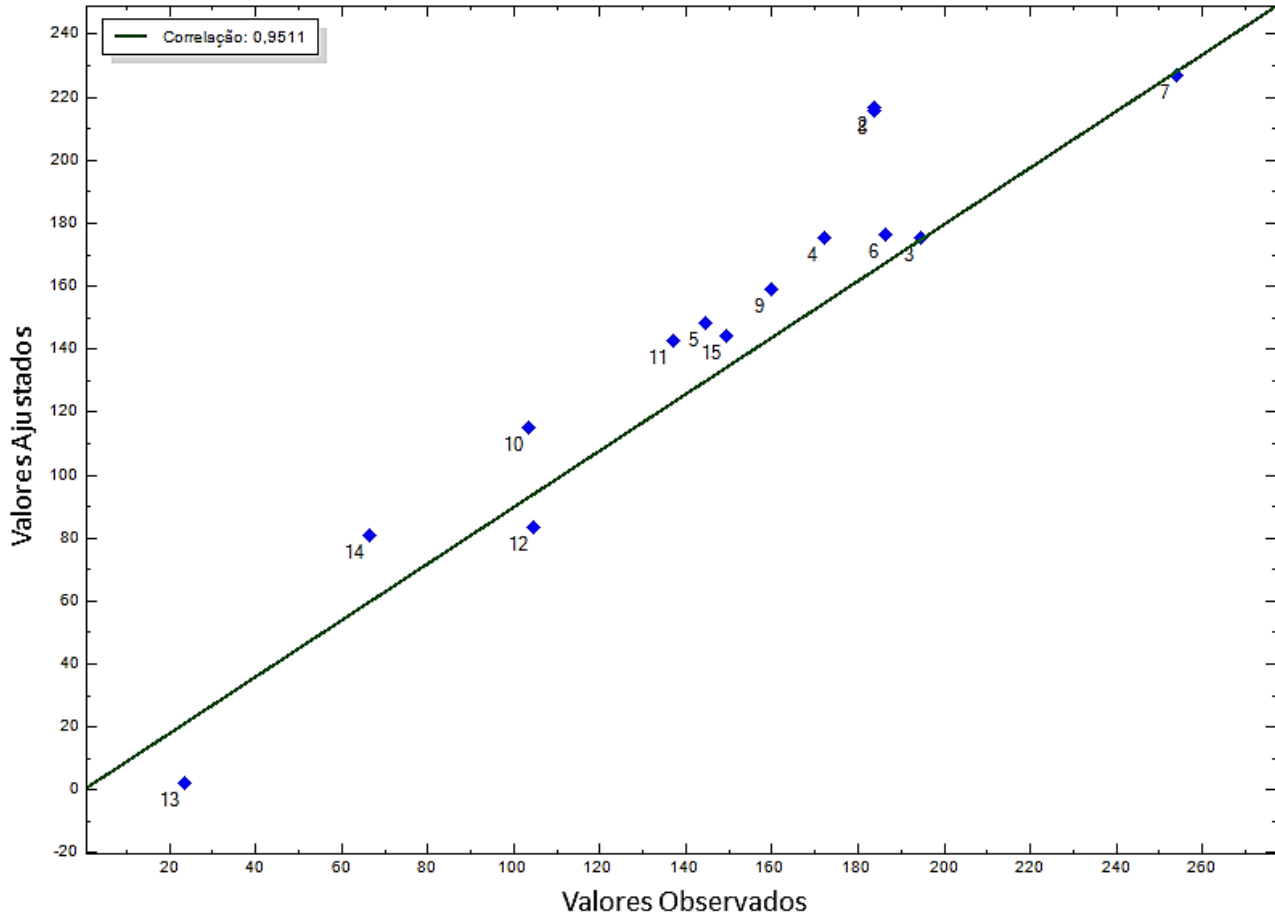
O modelo final não apresentou pontos influentes, Figura 15; nenhum *outlier* foi identificado, não houve formação de *clusters* ou autocorrelação, Figura 16; o modelo foi linear, Figura 17.

Figura 16 – Resíduos padronizados *versus* valores ajustados do modelo preditor da fase II-2 do reator metanogênico



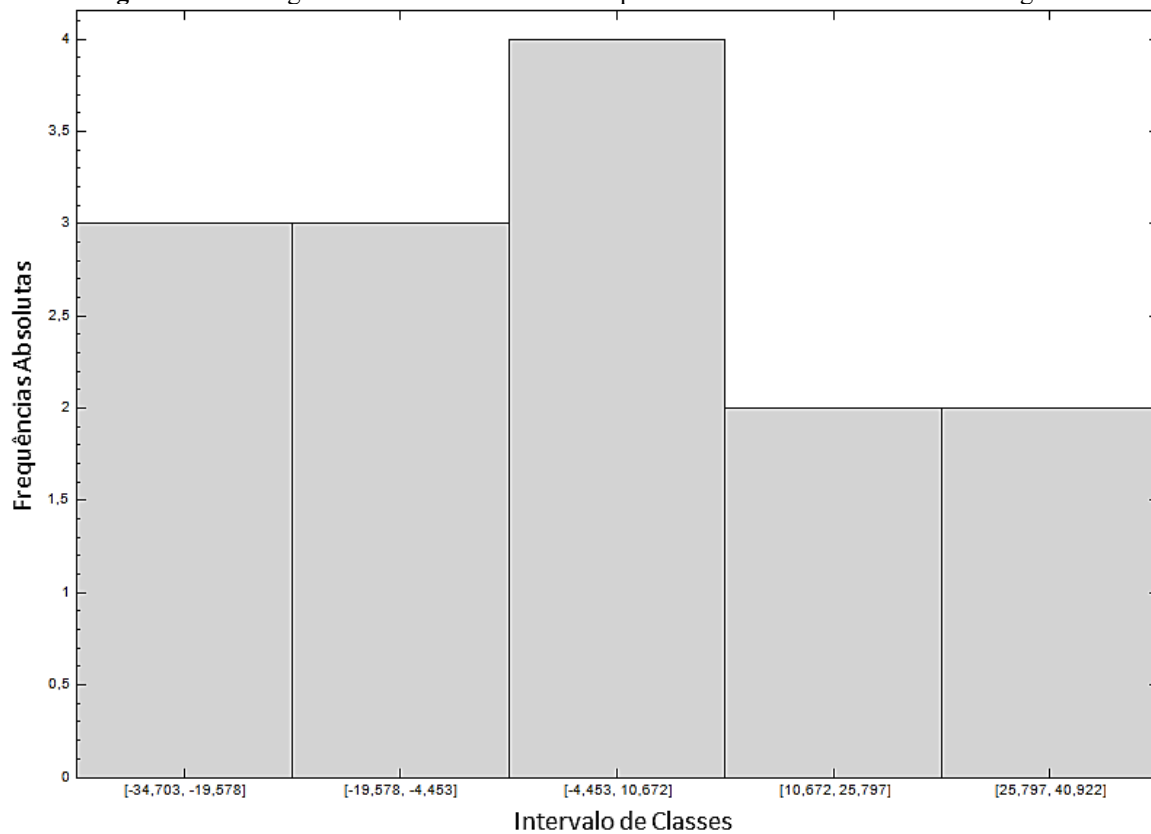
Fonte: o autor (2023).

Figura 17 – Valores ajustados *versus* valores observados do modelo preditor da fase II-2 do reator metanogênico



Fonte: o autor (2023).

Figura 18 – Histograma dos resíduos do modelo preditor da fase II-2 do reator metanogênico



Fonte: o autor (2023).

O modelo apresentou problema de normalidade, como pode ser visto no histograma da Figura 18. Essa falta de normalidade é associada à micronumerosidade e à repentina alteração no valor médio de rendimento de metano na fase II-2, que pode ter levado o reator a uma fase semelhante a novo estado transiente até nova estabilização.

A colinearidade entre as variáveis ALCATOT e ALCAPAR, Tabela 24, é normal, uma vez que são variáveis de mesma natureza.

Tabela 24 – Matriz de correlação do modelo preditor da fase II-2 do reator metanogênico

	REDOX	ALCATOT	ALCAPAR	AIAT	RENDCH4
REDOX	1,00				
ALCATOT	-0,38	1,00			
ALCAPAR	-0,28	0,83	1,00		
AIAT	-0,02	0,18	0,69	1,00	
RENDCH4	0,65	-0,39	-0,58	-0,56	1,00

Fonte: o autor (2023).

A equação do modelo preditor da fase experimental II-2 do reator metanogênico é a (25).

$$RENDCH4 = 14911,7725 + (0,0010 \cdot REDOX^2) - \frac{8323382,4695}{ALCATOT} + \frac{2,4375 \cdot 10^9}{ALCAPAR^2} - \frac{8267,9880}{AIAT} \quad (25)$$

Não foi possível gerar um modelo com alto poder de explicação para a fase II-2 sem a transformação das escalas das variáveis, chegando ao R² de 0,34 com a única variável explicativa significativa REDOX e sem interseção. Os modelos gerados com maiores poderes de explicação não tiveram nível de significância aceitável para as variáveis e/ou não atenderam os pressupostos básicos do método clássico de regressão linear. Os indicadores globais do ajuste e sua equação são mostrados a seguir na Tabela 25 e equação ($RENDCH4 = (-0,5229 \cdot REDOX)$) (26), respectivamente. Associa-se esse ocorrido à pouca disponibilidade de dados devido ao curto período analisado.

Tabela 25 - Indicadores globais do ajuste do modelo de regressão sem transformação das escalas fase II-2 do reator metanogênico

Modelo	Clássico de Regressão
Desvio Padrão	35,2816
Estatística Fc	6,2408
Nível de Significância do Modelo	0,0280
Coeficiente de determinação	0,3421
Coeficiente de determinação ajustado	0,3421
Observações	13

Fonte: o autor (2023).

$$RENDCH4 = (-0,5229 \cdot REDOX) \quad (26)$$

5.5 MODELO PREDITOR DO REATOR METANOGÊNICO – FASE III

A fase III durou dos dias 203 a 270. Os dados utilizados vão de 81 a 106 e podem ser vistos no apêndice A. Salienta-se que, por motivos operacionais, houve um maior intervalo entre as quantificações dos AGV na fase III, o que levou à menor disponibilidade e uma certa descontinuidade nos dados.

Para a fase experimental III chegou-se a um R^2 de 0,96, Tabela 26. O modelo teve um único *outlier*, Tabela 27.

Tabela 26 - Indicadores globais do ajuste do modelo predictor da fase III do reator metanogênico

Modelo	Clássico de Regressão
Desvio Padrão	7046,8685
Estatística Fc	55,4249
Nível de Significância do Modelo	0,0001
Coefficiente de determinação	0,9568
Coefficiente de determinação ajustado	0,9396
Observações	15

Fonte: o autor (2023).

Tabela 27 - *Outlier* retirado do modelo predictor da fase III do reator metanogênico

Dado	Dia	ph	Redox	alcatot	alcapar	aiat	amon	ace	prop	vale	Agv	dqo	ch ₄
2	205	6,79	-258,2	6,79	655	2,811	5,83	321,6	311,6	642,5	2444,2	0,604	162,74

Fonte: o autor (2023).

Dentre as 14 variáveis testadas inicialmente, 4 foram significantes, sendo ALCAPAR, AIAT, PROP, DQO que podem ser vistas com seus respectivos níveis de significância na Tabela 28. Todas as variáveis tiveram significância de 99,99%.

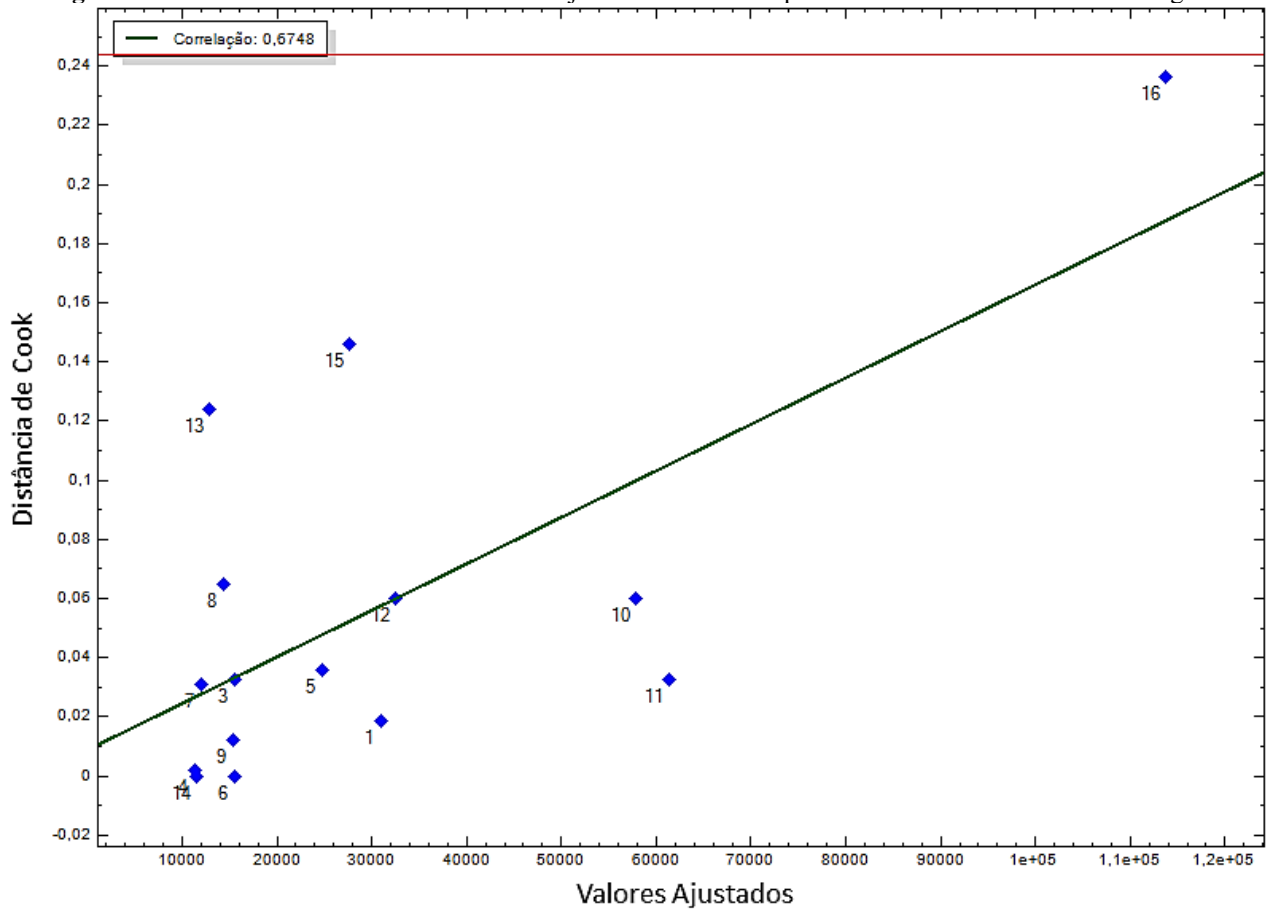
Tabela 28 - Indicadores individuais do ajuste das variáveis do modelo predictor da fase III do reator metanogênico

Variável	Escala	Coefficientes	Desvio Padrão	Estatística t	Nível de Significância
Interseção		-349885,4255	37777,4067	-9,2618	0,0000
ALCAPAR	1/(x) ²	37740633609,5095	2786007193,3094	13,5465	0,0000
AIAT	1/(x)	160129,4615	20573,4034	7,7833	0,0000
PROP	(x) ²	0,0730	0,0069	10,5895	0,0000
DQO	Raiz(x)	257645,4138	38346,7599	6,7188	0,0001
RENDCH4	(x) ²				

Fonte: o autor (2023).

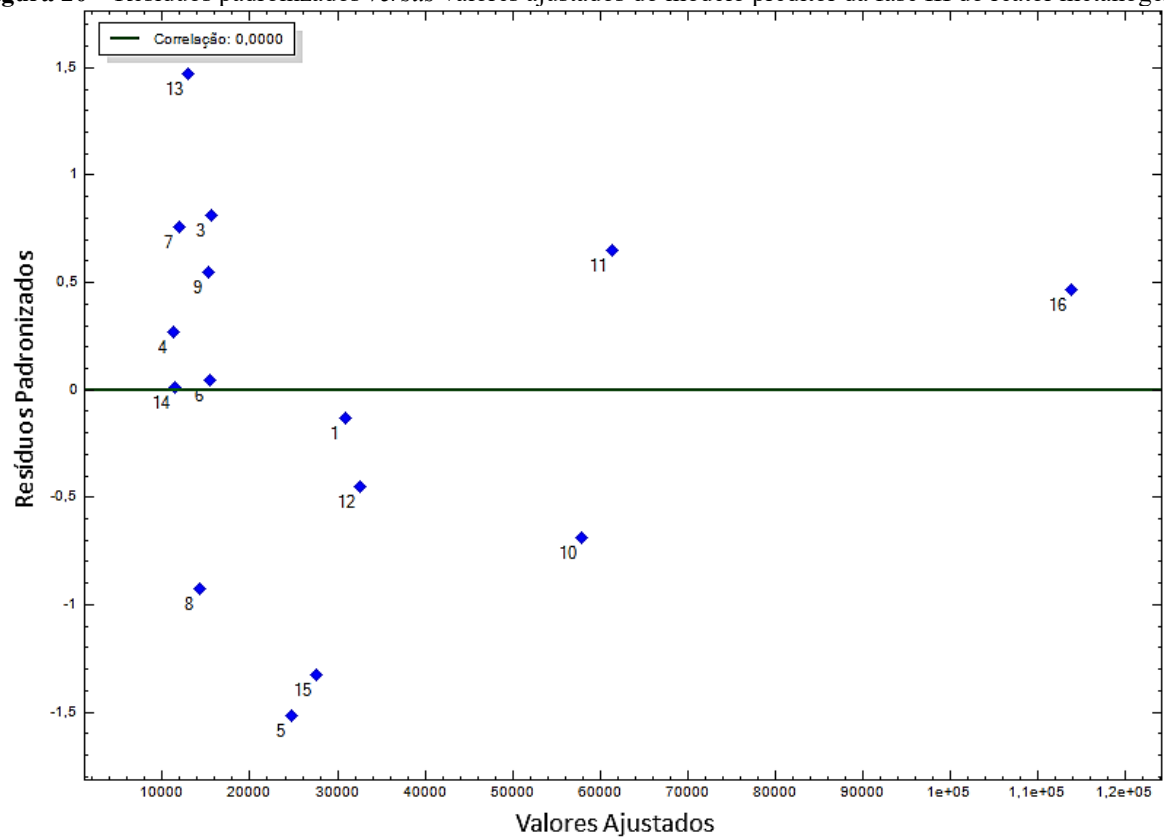
O modelo final não apresentou pontos influentes, Figura 19; nenhum *outlier* foi identificado, não houve formação de *clusters* ou autocorrelação, Figura 20; o modelo foi linear, Figura 21.

Figura 19 – Distâncias de Cook *versus* valores ajustados do modelo preditor da fase III do reator metanogênico



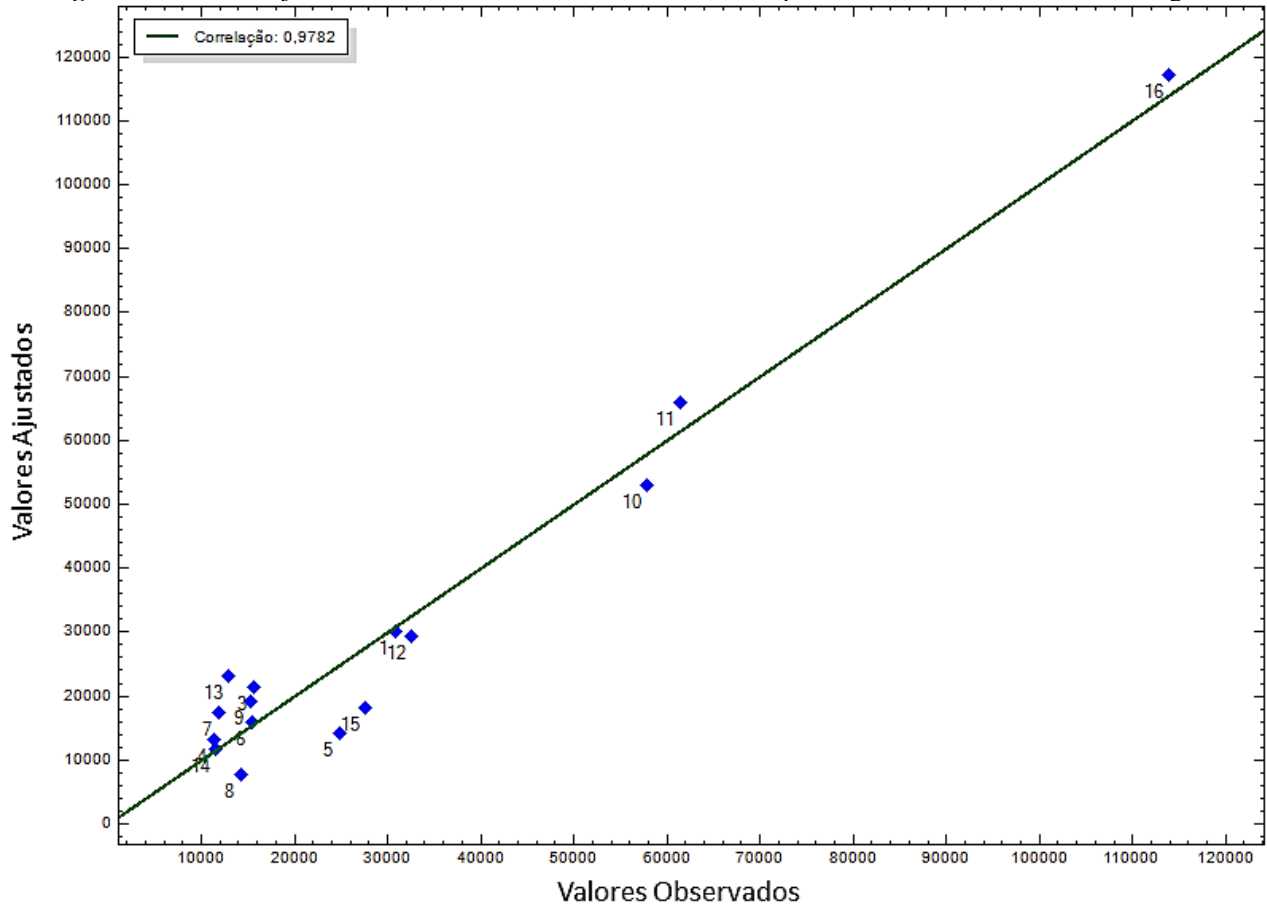
Fonte: o autor (2023).

Figura 20 – Resíduos padronizados *versus* valores ajustados do modelo preditor da fase III do reator metanogênico



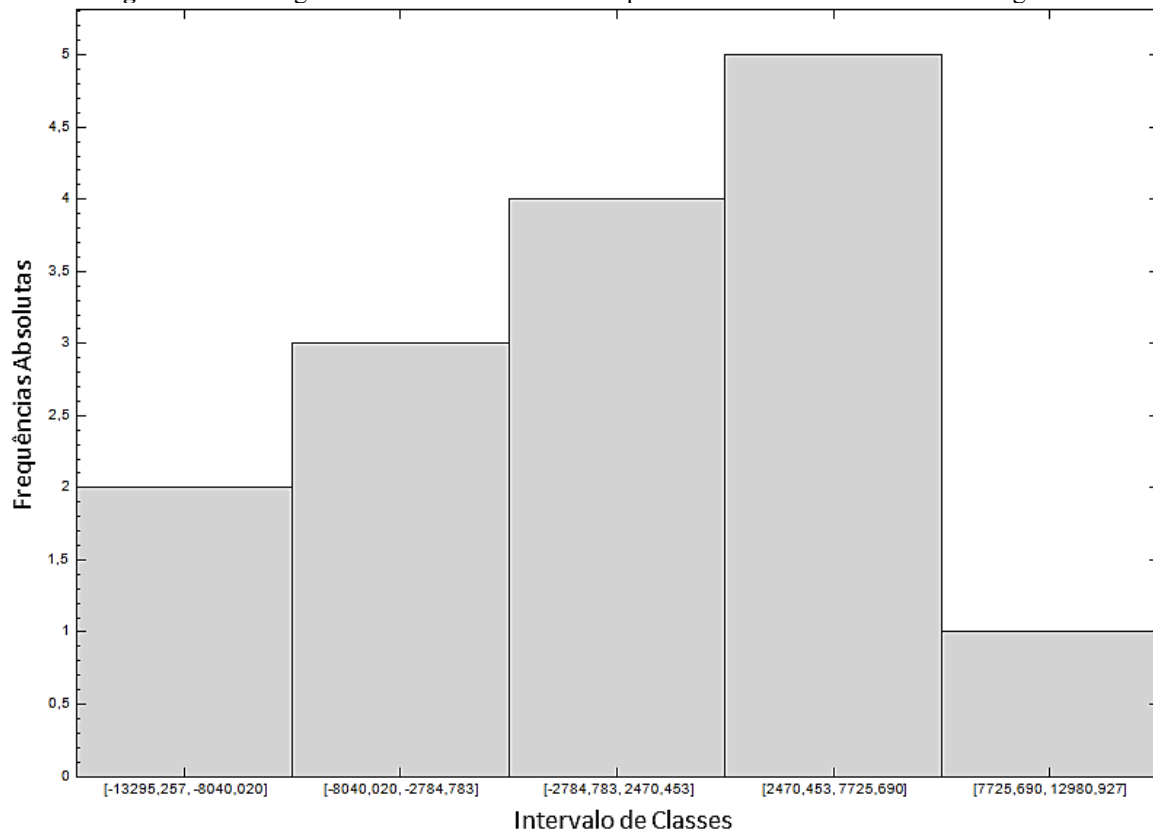
Fonte: o autor (2023).

Figura 21 – Valores ajustados *versus* valores observados do modelo preditor da fase III do reator metanogênico



Fonte: o autor (2023).

Figura 22 – Histograma dos resíduos do modelo preditor da fase III do reator metanogênico



Fonte: o autor (2023).

A normalidade apresentada pelo histograma dos resíduos não foi ideal, porém, esteve no limite da aceitação, como pode ser visto na Figura 22. Isso se deve aos maiores intervalos entre as quantificações da produção de AGV, refletidos na presença da variável independente PROP.

Não houve multicolinearidade, como pode ser visto na matriz de correlação, Tabela 29. Os valores de correlação com a variável dependente foram aceitáveis e não conflitantes com os níveis de significância.

Tabela 29 – Matriz de correlação do modelo preditor da fase III do reator metanogênico

	ALCAPAR	AIAT	PROP	DQO	RENDCH4
ALCAPAR	1,00				
AIAT	-0,51	1,00			
PROP	0,03	-0,46	1,00		
DQO	-0,55	0,37	-0,46	1,00	
RENDCH4	0,66	-0,18	0,39	-0,28	1,00

Fonte: o autor (2023).

A equação do modelo preditor da fase III do reator metanogênico é a (27).

$$RENDCH4 = \sqrt{\left[-349885,4255 + \frac{3,7741 \cdot 10^{10}}{ALCAPAR^2} + \frac{1,6013 \cdot 10^5}{AIAT} + (0,0730 \cdot PROP^2) + (2,5765 \cdot 10^5 \cdot \sqrt{DQO}) \right]} \quad (27)$$

Não foi possível gerar um modelo com poder de explicação aceitável e que atendesse ao pressuposto da normalidade para a fase III sem a transformação das escalas das variáveis, chegando ao R² de 0,39 com a única variável explicativa significativa AGV e sem intercepto. Os modelos gerados com maiores poderes de explicação não tiveram nível de significância aceitável para as variáveis e/ou não atenderam os pressupostos básicos do método clássico de regressão linear. Os indicadores globais do ajuste e sua equação são mostrados a seguir na Tabela 30 e equação (28), respectivamente. Associa-se esse ocorrido à pouca disponibilidade de dados de ácidos graxos voláteis, que foram as variáveis mais significantes para a fase experimental III.

Tabela 30 - Indicadores globais do ajuste do modelo de regressão sem transformação das escalas da fase III do reator metanogênico

Modelo	Clássico de Regressão
Desvio Padrão	19,3633
Estatística Fc	7,1519
Nível de Significância do Modelo	0,0216
Coeficiente de determinação	0,3940
Coeficiente de determinação ajustado	0,3940
Observações	12

Fonte: o autor (2023).

$$RENDCH4 = (0,0634 \cdot AGV) \quad (28)$$

5.6 RESUMO DAS EQUAÇÕES E COMPARAÇÃO COM OUTROS TRABALHOS COM DIFERENTES MÉTODOS

Um resumo dos valores de R^2 alcançados no presente trabalho é mostrado no Quadro 1.

Quadro 1 - Valores de R^2 alcançados de acordo com as fases experimentais com e sem transformações de linearização das variáveis

reator	Fase experimental II				Fase experimental III	
	Com transformação		Sem transformação		Com transformação	Sem transformação
RA	0,89		0,85		0,95	0,81
RM	II-1 = 0,89	II-2 = 0,90	II-1 = 0,88	II-2 = 0,34	0,96	0,39

Fonte: o autor (2023).

Podemos ver no Quadro 1 que o método de transformação das variáveis para linearização e otimização do R^2 permitiu gerar modelos com poder de explicação maior em todos os casos, chegando a 54 pontos percentuais a mais no R^2 no caso mais crítico (fase II-2 do RM). Além disso, conseguiu desenvolver modelo satisfatório em um caso que não seria possível sem esse artifício (III do RM). Ressalva-se que Niquini et al. (2019), Xu et al. (2014), Gunaseelan (2007) também tiveram modelos de regressão linear simples com poder de explicação muito baixo para a digestão anaeróbia, com melhores resultados na RLM mas ainda inferiores aos que incluíram termos não lineares e interação entre variáveis. A lista das equações dos modelos de cada fase dos reatores é mostrada no Quadro 2:

Rossi, Pecorini e Iannelli (2022), chegaram a R^2 de 0,91 no desenvolvimento de um modelo de predição pelo método da regressão linear múltipla da produção específica de metano (SMP, na sigla em inglês) na digestão anaeróbia seca de FORSU em um reator de fluxo em pistão. As variáveis utilizadas por eles foram sólidos voláteis totais, COV, TDH, relação C/N teor de lignina e AGV.

Rossi, Pecorini e Iannelli (2022) também resumiram o poder de explicação alcançado em diferentes métodos de geração de modelos de predição em três diferentes trabalhos. Os dados de digestão anaeróbia usados nesses trabalhos foram coletados de revisões bibliográficas e de ETE. Seguem as informações resumidas no Quadro 3:

Os valores de R^2 encontrados para os modelos gerados no presente trabalho estiveram compatíveis os valores de R^2 dos 4 trabalhos acima citados. Assim sendo, os modelos são válidos e aceitos.

Quadro 2 - Equações dos modelos de acordo com o reator e com a fase experimental

Reator	Fase	Equação
Reator acidogênico	II	$AGV = \left(0,0268 - \frac{0,4557}{PH^2} + \frac{2137,2901}{ALCATOT^2}\right)^2$
	III	$AGV = -5425,6001 + \frac{2,4297 \cdot 10^5}{PH^2} + (0,0153 \cdot REDOX^2) - \frac{652429923,4562}{ALCATOT^2} - \frac{661,0759}{AMON}$
Reator metanogênico	II	II-1 $RENDCH4 = \left[68,8193 + \frac{95326}{REDOX^2} - \frac{4318930,7940}{ALCAPAR^2} - (19,1512 \cdot \ln ACE) + \frac{3328,5825}{PROP} + (1,0588 \cdot AGV)\right]^2$
		II-2 $RENDCH4 = 14911,7725 + (0,0010 \cdot REDOX^2) - \frac{8323382,4695}{ALCATOT} + \frac{2,4375 \cdot 10^9}{ALCAPAR^2} - \frac{8267,9880}{AIAT}$
	III	$RENDCH4 = \sqrt{[-349885,4255 + \frac{3,7741 \cdot 10^{10}}{ALCAPAR^2} + \frac{1,6013 \cdot 10^5}{AIAT} + (0,0730 \cdot PROP^2) + (2,5765 \cdot 10^5 \cdot \sqrt{DQO})]}$

Fonte: o autor (2023).

Quadro 3 - Valores de R² alcançados através de diferentes métodos de geração de modelos preditores

Método de geração do modelo de predição	Fonte de dados	Variáveis independentes	Poder de explicação
Rede neural artificial	50 pontos de dados coletados de dez publicações	Extrativos, lignina, celulose, relação substrato/inóculo	R = 0,91
Regressão linear múltipla	86 pontos de dados de 13 publicações	Extrativos, lignina, celulose, relação substrato/inóculo, sólidos voláteis, hemicelulose	R = 0,83
Rede adaptativa baseada em sistemas de inferência Fuzzy	Estação de tratamento de esgoto	AGV, sólido totais, sólido fixos, pH, taxas de alimentação	R = 0,93

Fonte: adaptado de Rossi, Pecorini e Iannelli (2022).

5.7 AVALIAÇÃO DAS VARIÁVEIS MAIS IMPORTANTES PARA CADA REATOR

As variáveis pH e alcalinidade total estiveram em todos os modelos das fases experimentais do reator acidogênico e apresentaram maior nível de significância, sendo, portanto, as mais importantes na geração de modelos preditores da produção de AGV, ambas na escala (1/x²). Em segundo lugar, a variável potencial de oxi-redução, que apareceu em um modelo a menos, também na escala (1/x²). A variável nitrogênio amoniacal (1/x) foi significativa apenas na fase III, com

significância inferior que as demais. A escala que melhor descreveu a produção de AGV foi a transformação para $[1/\text{raiz}(x)]$, com maior aparição nos modelos, dentre todas as escalas, sendo a escala direta (x) a segunda melhor opção.

A variável alcalinidade parcial esteve em todos os modelos das fases e sub-fases experimentais do reator metanogênico, com significância entre as maiores, senão a maior, em cada modelo, sendo, portanto, a mais importante na geração de modelos preditores do rendimento de metano, na escala $1/x^2$. Em segundo lugar, as variáveis potencial de oxi-redução e relação AI/AT, que estiveram em um modelo a menos, nas escalas ($1/x^2$ ou secundariamente x^2) e ($1/x$ ou secundariamente x), respectivamente, também com significância entre as maiores. Em terceiro lugar as variáveis alcalinidade total ($1/\text{raiz}(x)$ e $1/x$) e ácido propiônico ($1/x$ e x^2), apareceram em dois modelos, porém com segunda maior significância em alguns casos. PH ($1/x^2$) e DQO adicionada ($\text{raiz}(x)$) apareceram uma única vez em modelos e com significâncias inferiores. Ácido acético ($\text{Ln}(x)$) e AGV ($\text{raiz}(x)$) apareceram apenas no modelo da fase II-1, contudo, foram as duas maiores significâncias desse modelo. Ressalta-se que a fase II-2 foi a fase de maior concentração de AGV no afluente do RM. Não houve uma escala de transformação da variável dependente comum entre os modelos, não sendo possível indicar a transformação de melhor ajuste do rendimento de metano. As escalas do rendimento de metano nos 3 modelos para os quais se apresentaram resultados foram: II-1= $\text{raiz}(x)$, II-2= x , III= x^2 .

A maioria dos estudos de modelos de regressão linear preditores da digestão anaeróbia têm trabalhado com biomassa lignocelulósica, dando ênfase a variáveis independentes, como teor de celulose, lignina, hemicelulose e extrativos e a variáveis de processo, como as relações matéria-prima efluente e carbono/nitrogênio (Rossi *et al.* 2022; Raposo *et al.* 2020; Bayard *et al.* 2018; Dandikas *et al.* 2015; Thomsen *et al.* 2014; Triolo *et al.* 2011). No entanto, Niquini *et al.* (2019) estudou a regressão linear simples e múltipla com termos de ordens superiores até sete, inversas das variáveis e interações entre 11 variáveis, tendo algumas variáveis em comum com o presente trabalho e encontrando significância para as variáveis alcalinidade e nitrogênio amoniacal. Asadi, Guo e McPhedran (2020) tiveram resultados satisfatórios em modelos que incluíram pH e AGV gerados através de rede neural artificial (ANN) de percepção de múltiplas camadas e sistema de inferência fuzzy baseado em rede adaptativa (ANFIS). Por fim, Tufaner e Demirci (2020), gerando modelos de predição da taxa de produção de biogás de reator híbrido anaeróbio por redes neurais artificiais e modelos de regressões não lineares, usaram as variáveis independentes taxa de enchimento do reator, pH afluente e efluente, alcalinidade afluente e efluente, COV, DQO do efluente, série de sólidos suspensos do efluente, também com algumas em comum ao presente trabalho.

6 CONCLUSÕES

Os modelos preditores da produção de AGV no RA, gerados pela RLM através do MMQ com transformações de linearização e otimização do R^2 , foram satisfatórios. Os modelos tiveram significância global acima de 99,99%; valores de R^2 acima de 0,89, máximo de 0,95; e bons resultados em todos os outros indicadores da qualidade de ajuste. Chama-se a atenção para a necessidade de manter a regularidade dos intervalos entre as análises de AGV e a continuidade dos dados.

Os modelos preditores do rendimento de metano no reator metanogênico foram satisfatórios. Os modelos tiveram nível de significância global acima de 99,99%; valores de R^2 acima de 0,89, máximo de 0,96; e bons resultados em todos os outros indicadores da qualidade de ajuste.

Nas variáveis e suas escalas de transformação significativas para o RA, pH ($1/x^2$) e alcalinidade total ($1/x^2$) estiveram no primeiro nível de importância. Potencial de oxi-redução ($1/x^2$) esteve no segundo nível de importância. Nitrogênio amoniacal ($1/x$) esteve no terceiro nível de importância, com significância inferior às demais. A escala que melhor descreveu a variável dependente produção de AGV foi $1/\text{raiz}(x)$, seguida da escala direta (x).

Nas variáveis e suas escalas de transformação significativas para o RM, alcalinidade parcial foi a mais importante, na escala $1/x^2$ (exceção da fase I, $1/\text{raiz}(x)$). Potencial de oxi-redução ($1/x^2$ ou, secundariamente, x^2) e relação AI/AT ($1/x$ ou, secundariamente, x), estiveram no segundo nível de importância. Alcalinidade total ($1/\text{raiz}(x)$ e $1/x$) e ácido propiônico ($1/x$ e x^2) estiveram no terceiro nível de importância. Ácido acético ($\text{Ln}(x)$) e AGV ($\text{raiz}(x)$) se mostraram particularmente importantes no momento de mais alta concentração de AGV no afluente do RM. A variável dependente rendimento de metano teve as seguintes escalas de acordo com cada modelo: I= $1/x$, II-1= $\text{raiz}(x)$, II-2= x , III= x^2 .

O método de transformações para linearização das variáveis e otimização do poder de explicação do modelo representado pelo R^2 se mostrou uma ferramenta eficaz e valiosa na geração dos modelos, permitindo a geração de modelos com altos valores de R^2 em caso em que a regressão linear múltipla falhou e o aumento de até 54 pontos percentuais em casos em que a regressão linear múltipla foi capaz de gerar modelos satisfatórios.

7 *SUGESTÕES DE TRABALHOS FUTUROS*

Seleção das variáveis para geração dos modelos pelo método de análise de componentes principais (PCA, na sigla em inglês), o que é uma alternativa com potencial de economia de tempo de cálculo em relação ao método de retirada das variáveis não significativas após a verificação das combinações de transformações de linearização e otimização do valor do R^2 com todas as variáveis, empregado no presente trabalho. E, em seguida, a modelagem combinando termos de ordens superiores, acima da ordem 2 já testada no presente trabalho, e interações entre as variáveis juntamente com as transformações nas escalas verificadas no presente trabalho na otimização do valor do R.

Realização de um planejamento fatorial do experimento de codigestão da biomassa de microalgas que permita a análise dos resultados pela metodologia de superfície de resposta (RSM) usando o método de design composto central (CCD) para identificar os valores ótimos das variáveis identificadas como mais relevantes no presente trabalho e na análise PCA sugerida acima.

Geração de modelos preditores da codigestão anaeróbia de biomassa de microalgas utilizando aprendizado de máquina pelos métodos de rede neural artificial (ANN) e/ou sistema de inferência fuzzy baseado em rede adaptativa (ANFIS).

Análise da elasticidade dos modelos, através de incrementos percentuais nos valores de entrada de uma variável independente, mantendo as demais variáveis constantes, para saber a porcentagem de impacto de cada variável sobre os resultados do modelo. Inclusive, utilizar essa análise da elasticidade para fundamentar as variáveis mais importantes, através das maiores alterações causadas na variável dependente. Além da medição percentual da influência de cada variável, observar o sentido da influência e os casos em que a mudança na variável dependente é melhor expressa em múltiplos de um valor fixo, ao invés de em porcentagem.

REFERÊNCIAS

- APHA, AWWA, and WPCF. **Standard methods for the examination of water and wastewater**, 25th ed. Baltimore: Port City Press, 2017.
- AQUINO, S. F. et al. Metodologias para determinação da atividade metanogênica específica (AME) em lodos anaeróbios. **Engenharia Sanitaria e Ambiental**, v. 12, n. 2, p. 192–201, 2007.
- ASADI, M.; GUO, H.; MCPHEDRAN, K. Biogas production estimation using data-driven approaches for cold region municipal wastewater anaerobic digestion. **Journal of environmental management**, v. 253, p. 109708, 2020.
- Atkinson, A., C., (1985), **Plots, Transformations and Regression**, Oxford University Press, Oxford.
- BARROSO JÚNIOR, J. C. A. **Produção de biomassa de algas em lagoas de alta taxa alimentadas com esgoto sanitário com posterior separação por flotação por ar dissolvido**. 2015. 114 f. Dissertação (Mestrado em Recursos Hídricos E Saneamento Ambiental)-Instituto De Pesquisas Hidráulicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2015.
- BLIGH, E. G.; DYER, W. J. A rapid method of total lipid extraction and purification. **Canadian Journal of Biochemistry and Physiology**, v. 37, p. 911–917, 1959.
- BRASIL. **Lei nº 14.026**, de 15 de julho de 2020. Atualiza o marco legal do saneamento básico, e dá outras providências. Diário Oficial da União: Brasília, 16 de julho de 2020. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/lei/114026.htm#:~:text=%E2%80%9CEstabelece%20as%20diretrizes%20nacionais%20para,Art. Acesso em: maio 2022.
- CHEN, W.; WANG, J.; LIU, W. A View of Anaerobic Digestion: Microbiology, Advantages and Optimization. **Academic Journal of Environment & Earth Science**, v. 5, n. 1, p. 1–8, 2023.
- CHERNICHARO, C. A. DE L. Reatores anaeróbios. **Belo Horizonte: Departamento de Engenharia Sanitária e Ambiental–UFMG**, v. 5, n. 2, p. 380, 1997.
- Cook, R. D. e Weisberg, S., (1999), **Applied Regression Including Computing and Graphics**, John Wiley&Sons, New York, NY.
- COSTA, L. G. da. **Gestão de lodos de esgoto sanitário: avaliação comparativa de viabilidade econômica entre processo convencional e com aproveitamento de biogás para geração de energia**. 2022. Trabalho de Conclusão de Curso.
- DANDIKAS, V. et al. Correlation between biogas yield and chemical composition of energy crops. **Bioresource technology**, v. 174, p. 316–320, 2014.
- DANDIKAS, V. et al. Correlation between biogas yield and chemical composition of grassland plant species. **Energy & Fuels**, v. 29, n. 11, p. 7221–7229, 2015.
- DANTAS, R. A. **Engenharia de Avaliações: uma introdução à metodologia científica**. [s.l.] Pini, 2005.
- GOULARTE, P. G. **Produção de biogás a partir da co-digestão de resíduos da geração de energia**. 2014. 86 f. Dissertação (Mestrado em Engenharia e Ciência de Alimentos)-Escola Química e Alimentos, Universidade Federal do Rio Grande, Rio Grande, 2014.
- GUNASEELAN, V. N. Regression models of ultimate methane yields of fruits and vegetable solid wastes, sorghum and napiergrass on chemical composition. **Bioresource technology**, v. 98, n. 6, p. 1270–1277, 2007.

- MACHADO, F. L. O. **Co-digestão anaeróbia de microalgas e de glicerol residual do biodiesel**. 2012. 108 f. Dissertação (Mestrado em Engenharia Civil: Saneamento Ambiental)-Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, 2012.
- MYERS, R. H.; MONTGOMERY, D. C. A tutorial on generalized linear models. **Journal of Quality Technology**, v. 29, n. 3, p. 274-291, 1997.
- MYERS, R. H., MONTGOMERY, D. C., VINING, G. G., (2002) **Generalized Linear Models: with Applications in Engineering and the Sciences**, John Wiley&Sons, New York, NY.
- NAZARPOUR, M. et al. Optimization of biohydrogen production from microalgae by response surface methodology (RSM). **Energy**, v. 253, p. 124059, 2022.
- NIQUINI, G. R. et al. Feedstock and inoculum characteristics and process parameters as predictors for methane yield in mesophilic solid-state anaerobic digestion. **Anais da Academia Brasileira de Ciências**, v. 91, 2019.
- PINHEIRO, A. A. D. **Codigestão anaeróbia de biomassa de algas e vinhaça para produção de metano**. 2023. 86 f. Dissertação (Mestrado em Engenharia Civil: Tecnologia Ambiental e Recursos Hídricos)-Programa De Pós-Graduação Em Engenharia Civil, Universidade Federal de Pernambuco, Recife, 2023.
- Portal tratamento de água. Lagoas de estabilização, 17 de Abril de 2019. **Tratamento de efluentes**. Disponível em: <<https://tratamentodeagua.com.br/artigo/lagoas-estabilizacao/>>. Acesso em: 27 de Jun. de 2022.
- RAPOSO, F.; BORJA, R.; IBELLI-BIANCO, C. Predictive regression models for biochemical methane potential tests of biomass samples: Pitfalls and challenges of laboratory measurements. **Renewable and Sustainable Energy Reviews**, v. 127, p. 109890, 2020.
- ROSSI, E.; PECORINI, I.; IANNELLI, R. **Multilinear Regression Model for Biogas Production Prediction from Dry Anaerobic Digestion of OFMSW. Sustainability 2022, 14, 4393.** , [s.d.].
- SHOBANA, S. et al. Fermentative hydrogen production from mixed and pure microalgae biomass: key challenges and possible opportunities. **International Journal of Hydrogen Energy**, v. 42, n. 42, p. 26440–26453, 2017.
- SIALVE, B.; BERNET, N.; BERNARD, O. Anaerobic digestion of microalgae as a necessary step to make microalgal biodiesel sustainable. **Biotechnology advances**, v. 27, n. 4, p. 409–416, 2009.
- SILVA, E. M. da. **Co-digestão anaeróbia de biomassa algas com diferentes resíduos orgânicos para otimização da produção de biogás**. 2022. 82 f. TCC (Graduação) - Curso de Engenharia Civil, Universidade Federal de Pernambuco, Recife, 2022.
- TEPARI, E. A. et al. Co-fermentation of carbohydrates and proteins for biohydrogen production: Statistical optimization using Response Surface Methodology. **International Journal of Hydrogen Energy**, v. 45, n. 4, p. 2640–2654, 2020.
- THOMSEN, S. T.; SPLIID, H.; ØSTERGÅRD, H. Statistical prediction of biomethane potentials based on the composition of lignocellulosic biomass. **Bioresource Technology**, v. 154, p. 80–86, 2014.
- TRIOLO, J. M. et al. A new algorithm to characterize biodegradability of biomass during anaerobic digestion: Influence of lignin concentration on methane production potential. **Bioresource technology**, v. 102, n. 20, p. 9395–9402, 2011.

TUFANER, F.; DEMIRCI, Y. Prediction of biogas production rate from anaerobic hybrid reactor by artificial neural network and nonlinear regressions models. **Clean Technologies and Environmental Policy**, v. 22, p. 713–724, 2020.

Universidade Federal do Rio de Janeiro: Centro Experimental de Saneamento. **Lagoa de estabilização.** Infraestrutura. Disponível em: <<http://www.saneamento.poli.ufrj.br/index.php/br/infraestrutura/lagoa-estabilizacao>>. Acesso em: 27 de Jun. de 2022.

VAVILIN, V. A.; RYTOW, S. V; LOKSHINA, L. Y. Modelling hydrogen partial pressure change as a result of competition between the butyric and propionic groups of acidogenic bacteria. **Bioresource technology**, v. 54, n. 2, p. 171–177, 1995.

VISHTAL, A.; KRASLAWSKI, A. Challenges in industrial applications of technical lignins. **BioResources**, v. 6, n. 3, 2011.

WANG, Y. et al. Effects of volatile fatty acid concentrations on methane yield and methanogenic bacteria. **Biomass and bioenergy**, v. 33, n. 5, p. 848–853, 2009.

XU, F.; WANG, Z.-W.; LI, Y. Predicting the methane yield of lignocellulosic biomass in mesophilic solid-state anaerobic digestion based on feedstock characteristics and process parameters. **Bioresource technology**, v. 173, p. 168–176, 2014.

APÊNDICE A – DADOS OBSERVADOS NO EXPERIMENTO

Dados observados no reator acidogênico

DADOS	DIAS	PH	REDOX	ALCATOT	AMON	AGV
1	23	4,94	-43	425	6,16	467,41
2	24	4,9	-71,3	430	6,72	537,88
3	29	4,98	-103,5	485	5,04	559,90
4	30	4,93	-19,5	506	10,08	667,42
5	31	4,9	-88	435	6,72	803,63
6	36	4,91	-156,5	576	8,53	1494,12
7	43	5,11	-92,7	730	2,67	2006,20
8	44	5,04	-192,7	776	3,20	2035,68
9	45	5,03	-86,9	607	6,40	1865,29
10	50	4,98	-96,1	724	3,20	2645,05
11	51	4,94	-165	732	3,20	2546,38
12	52	5,06	-157	1071	2,13	2651,54
13	57	5,08	-197,1	994	6,40	2775,12
14	58	5,05	-117,7	885	2,67	3210,03
15	59	4,99	-224,9	870	4,80	3196,53
16	64	5,29	-122,8	1201	2,67	3516,06
17	65	5,28	-138,7	1371	1,60	3717,84
18	66	5,2	-132,2	1241	14,40	3573,02
19	72	5,32	-149,5	1645	5,51	7802,17
20	73	5,72	-182,4	1558	86,94	4031,95
21	79	5,51	-125	2082	50,67	6372,36
22	80	5,61	-143,65	2105	61,87	6168,67
23	85	5,62	-162,3	2017	59,73	6409,73
24	87	5,64	-199,8	2061	78,93	6199,83
25	92	5,54	-275,3	2177	82,67	6626,28
26	93	5,52	-258,4	2218	96,00	6688,21
27	94	5,55	-180,6	2180	87,48	6222,73
28	113	5,57	-257,5	1871	62,40	5924,85
29	114	5,66	-242,1	1905	73,07	6135,52
30	115	5,62	-238,8	1789	44,27	5556,40
31	120	5,68	-246,8	1810	34,67	5499,47
32	121	5,65	-251,7	1596	43,77	5513,00
33	122	5,71	-243,5	1550	34,77	5602,40
34	127	6,2	-322	1645	33,13	4312,80
35	128	6,25	-321,4	1780	17,18	3941,80
36	129	6,24	-339,4	1869	30,47	4076,64
37	134	6,3	-377,6	2193	10,23	3745,01
38	135	6,12	-345,4	1899	16,67	3791,16
39	136	6,16	-357,1	1842	25,46	4628,75
40	141	6,19	-366,6	2229	52,56	4154,48
41	142	6,08	-363,3	1798	6,03	3375,59
42	143	6,02	-334,8	1746	10,33	4901,27
43	149	6,06	-334,4	1485	6,44	4071,20
44	150	5,84	-327,6	1654	21,47	4396,80
45	155	5,67	-323,2	1525	5,93	5179,20
46	156	5,59	-322,7	1664	73,11	6599,00
47	157	5,65	-320,6	1793	85,38	5564,92

48	162	5,7	-331,6	1627	58,29	5754,18
49	163	5,71	-323,3	1773	117,59	5733,96
50	164	5,7	-325,2	1627	119,13	5634,94
51	169	5,67	-324,9	1669	91,72	5245,53
52	170	5,65	-322,2	1680	62,58	5207,08
53	171	5,7	-322,9	1582	43,46	5178,22
54	177	5,81	-327	1812	72,60	3815,07
55	178	5,81	-329,8	1832	88,35	4532,32
56	204	5,78	-328,3	1461	1,12	2670,22
57	205	5,75	-326,8	1090	1,74	2444,20
58	206	5,69	-305,5	1073	6,85	2880,66
59	218	6,51	-348,5	1378	16,83	2022,38
60	219	6,42	-356,1	1262	18,51	1832,24
61	220	6,4	-362,3	1284	22,09	1813,80
62	225	6,38	-339	1360	25,87	1721,77
63	226	6,34	-284,7	1432	16,77	1466,98
64	227	6,37	-330,1	1761	20,35	2084,47
65	233	6,45	-362,7	1287	15,13	2047,65
66	240	6,25	-358	1075	12,58	2350,06
67	247	5,98	-341,7	863	5,11	2300,01
68	248	5,75	-333,4	674	7,57	2062,19
69	254	5,3	-308,5	676	10,12	3102,70
70	260	5,16	-285,7	627	16,87	2558,29
71	267	5,02	-287,1	688	7,77	4077,08

Fonte: o autor (2023).

Dados observados no reator metanogênico

dado	dia	ph	redox	alcatot	alcapar	alcaint	aiat	Aiap	amon	dqo	ace	propan	Vale	agv	ch ₄
1	1	7,96	-422	4350	3855	494	0,89	0,13	991,5						0,00
2	2	7,74	-431	4737	2079	2658	0,44	1,28	925,8						113,26
3	3	7,71	-428	6652	2352	4299	0,35	1,83	786,7						19,29
4	8	7,41	-401	2938	2185	753	0,74	0,34	647,6						12,47
5	9	7,11	-391	2950	2200	750	0,75	0,34	701,8						0,00
6	10	6,69	-348	2803	2190	613	0,78	0,28	705,5		1118	32	0,6	1278	2,49
7	15	6,63	-381	2790	1845	945	0,66	0,51	912,8		790	72	1,2	874	42,41
8	16	6,66	-361	2993	2070	923	0,69	0,45	560,0		1320	33	0,2	1363	14,97
9	17	6,68	-371	2378	1580	798	0,66	0,50	523,6		1169	21	0,3	1200	7,48
10	22	7,18	-391	2283	1625	658	0,71	0,40	476,0		202	25	0,7	240	72,34
11	23	7,05	-394	2363	1668	695	0,71	0,42	390,6		430	20	7,0	467	119,74
12	24	7,43	-376	3348	2493	855	0,74	0,34	410,5		500	27	0,9	538	154,66
13	29	7,44	-399	2325	1828	498	0,79	0,27	317,8		513	34	0,8	560	52,38
14	30	7,2	-403	3390	2848	543	0,84	0,19	387,8		631	24	0,7	667	84,81
15	31	7,1	-380	2225	1833	393	0,82	0,21	375,2		759	32	1,1	804	107,26
16	36	7,18	-380	1968	1570	398	0,80	0,25	304,6		1216	36	0,7	1494	87,31
17	37	7,13	-385	2050	1693	358	0,83	0,21	285,6						77,33
18	38	7,44	-380	1905	1563	343	0,82	0,22	454,7						44,90
19	43	7,11	-376	1788	1475	313	0,83	0,21	265,4		822	8	1,1	2006	69,85
20	44	7,08	-366	1908	1628	280	0,85	0,17	280,0		893	9	1,1	2036	84,81
21	45	7,08	-375	1938	1625	313	0,84	0,19	297,9		917	36	0,8	1865	94,79

22	50	7	-360	2046	1725	321	0,84	0,19	229,6		1236	64	1,6	2645	92,30
23	51	7,08	-359	1996	1653	344	0,83	0,21	157,9		1217	144	2,0	2546	42,41
24	52	7,06	-347	1918	1555	363	0,81	0,23	215,0		1235	259	3,6	2652	52,38
25	57	7,06	-92	1924	1600	324	0,83	0,20	201,6		1261	243	31,3	2775	4,99
26	58	7,08	-340	1964	1690	274	0,86	0,16	164,6		1459	299	32,7	3210	54,88
27	59	7	-356	1808	1478	330	0,82	0,22	170,2		1500	292	29,8	3197	109,76
28	64	6,99	-336	1855	1519	336	0,82	0,22	125,4		1439	274	50,0	3516	37,42
29	65	7,05	-360	1799	1501	298	0,83	0,20	154,6		1466	338	46,5	3718	104,77
30	66	7,02	-352	1765	1471	294	0,83	0,20	136,6		1614	276	37,0	3573	52,38
31	72	7	-362	1334	1104	230	0,83	0,21	107,1	0,71	2875	952	158,8	7802	109,76
32	73	7,07	-350	1700	1450	250	0,85	0,17	117,3	0,71	1565	244	62,6	4032	57,37
33	78	7,11	-346	1700	1443	257	0,85	0,18	117,3	0,68					62,92
34	79	7,1	-347	1657	1329	329	0,80	0,25	117,3	0,68	2669	487	150,2	6372	52,44
35	80	7,05	-343	1657	1329	329	0,80	0,25	117,3	0,68	2533	523	173,3	6169	44,57
36	85	7,07	-333	1688	1375	313	0,81	0,23	58,8	0,68	2449	772	212,8	6410	60,30
37	86	7,12	-339	1669	1328	341	0,80	0,26	33,6	0,68					39,41
38	87	7,28	-348	1630	1343	288	0,82	0,21	82,9	0,68	2370	931	200,1	6200	39,25
39	92	7,31	-340	1579	1271	308	0,81	0,24	81,2	0,68	2352	1441	217,6	6626	34,08
40	93	7,15	-342	1573	1286	287	0,82	0,22	76,2	0,68	2427	1414	282,4	6688	49,81
41	94	7,07	-203	1576	1270	306	0,81	0,24	70,8	0,68	2343	1239	276,7	6223	26,22
42	99	7,12	-337	1594	1288	306	0,81	0,24	73,1	0,68					31,46
43	100	6,91	-353	1537	1217	320	0,79	0,26	67,2	0,68					31,46
44	101	7,14	-373	1544	1244	300	0,81	0,24	59,7	0,68					26,22
45	106	7,04	-376	1500	1194	306	0,80	0,26	66,1	0,68					34,08
46	107	7,18	-356	1475	1196	279	0,81	0,23	68,3	0,68					34,08
47	108	7,17	-364	1473	1198	275	0,81	0,23	66,2	0,68					13,11
48	113	7	-357	1250	1011	239	0,81	0,24	34,7	0,68	2148	714	492,0	5925	31,46
49	114	7,1	-355	1420	1143	278	0,80	0,24	52,3	0,68	2172	892	508,0	6136	15,73
50	115	7,03	-366	1245	1050	195	0,84	0,19	34,7	0,69	1950	584	492,4	5556	15,39
51	120	7,05	-362	1281	1051	230	0,82	0,22	29,2	0,70	1596	520	535,2	5499	70,22
52	121	7,05	-355	1261	1036	225	0,82	0,22	33,7	0,82	1679	601	510,4	5513	91,24
53	122	7,03	-348	1193	1055	138	0,88	0,13	29,9	0,90	1766	554	503,2	5602	78,65
54	127	7,07	-340	1216	983	234	0,81	0,24	29,0	0,70	852	510	604,8	4313	43,52
55	128	7,1	-325	1136	944	193	0,83	0,20	19,6	0,64	763	492	593,6	3942	101,13
56	129	7,03	-337	1161	941	220	0,81	0,23	51,2	0,84	949	514	622,0	4077	64,55
57	134	7,08	-320	1156	935	221	0,81	0,24	3,8	0,50	1484	407	567,5	3745	23,08
58	135	7,06	-145	1080	863	218	0,80	0,25	3,8	0,66	1031	456	632,1	3791	51,76
59	136	7	-145	1051	829	223	0,79	0,27	27,8	0,68	1112	1060	631,0	4629	66,98
60	141	6,95	-255	1173	898	275	0,77	0,31	26,9	0,73	926	1154	569,9	4154	29,75
61	142	6,94	-306	1064	860	204	0,81	0,24	4,8	0,72	663	451	613,3	3376	1,88
62	143	6,93	-282	910	738	173	0,81	0,23	3,9	0,65	1139	1594	583,5	4901	16,40
63	148	6,98	-322	1075	875	200	0,81	0,23	3,9	0,72					4,03
64	149	6,98	-305	800	650	150	0,81	0,23	6,1	0,70	1010	612	736,0	4071	0,00
65	150	6,97	-335	835	691	144	0,83	0,21	8,3	0,66	1208	590	741,0	4397	0,00
66	155	6,84	-349	960	756	204	0,79	0,27	1,3	0,70	1575	573	690,0	5179	221,17
67	156	6,85	-350	1025	814	211	0,79	0,26	5,2	0,70	1816	1463	692,0	6599	217,00
68	157	6,83	-349	1075	846	229	0,79	0,27	8,0	0,72	1503	793	671,2	5565	175,37
69	169	6,89	-336	856	666	190	0,78	0,29	28,8	0,73	1207	478	679,4	5246	175,48

70	170	6,76	-357	899	704	195	0,78	0,28	23,2	0,94	1153	479	672,2	5207	148,31
71	171	6,78	-356	856	666	190	0,78	0,29	9,9	0,87	1152	461	627,9	5178	176,23
72	177	6,83	-331	955	803	153	0,84	0,19	26,1	0,86	900	360	524,3	3815	226,75
73	178	6,78	-330	935	766	169	0,82	0,22	3,5	0,62	975	398	642,1	4532	215,69
74	183	6,83	-285	936	769	168	0,82	0,22		0,82			712,5		159,24
75	184	6,72	-302	894	631	263	0,71	0,42		0,80			705,2		115,29
76	185	6,66	-309	870	667	203	0,77	0,31		0,85			575,1		142,85
77	190	6,84	-266	882	700	182	0,79	0,26		0,63	653	417	714,9	3465	83,48
78	191	6,8	-276	918	670	248	0,73	0,37		0,78	433	464	639,45	3320	2,43
79	192	6,84	-117	853	688	165	0,81	0,24		0,71	480	325	701,9	3263	80,74
80	197	6,78	-319	873	708	165	0,81	0,23		0,72	409	304	680,9	2855	143,98
81	204	6,75	-257	810	655	155	0,81	0,24	1,5	0,57	368	353	718,7	2670	173,21
82	205	6,79	-258	785	655	130	0,83	0,20	5,8	0,60	322	312	642,5	2444	162,74
83	206	6,7	-288	660	577	83	0,87	0,14	4,8	0,61	431	337	699,0	2881	146,11
84	212	6,68	-302	758	567	192	0,75	0,34	7,3	0,54					178,55
85	213	6,67	19	802	623	178	0,78	0,29	5,6	0,58					145,37
86	218	6,765	-302	737	602	135	0,82	0,22	7,7	0,51	128	746	806,1	2022	114,75
87	219	6,73	-309	720	562	158	0,78	0,28	2,0	0,53	85	735	696,2	1832	118,97
88	220	6,75	-304	741	593	147	0,80	0,25	6,3	0,58	80	721	700,1	1814	125,73
89	225	6,72	-279	755	621	133	0,82	0,21	8,2	0,64	85	717	561,4	1722	131,51
90	226	6,74	-359	735	595	140	0,81	0,24	6,4	0,67	60	624	490,8	1467	88,32
91	227	6,76	-290	756	602	154	0,80	0,26	7,7	0,60	59	790	731,2	2084	138,32
92	233	6,71	-328	720	574	146	0,80	0,25	9,0	0,51	4	1142	637,7	2048	230,20
93	234	6,74	-329	727	574	153	0,79	0,27	3,9	0,50					129,62
94	240	6,68	-338	674	535	139	0,79	0,26	8,0	0,50	287	1069	561,1	2350	256,80
95	241	6,7	-333	706	554	152	0,79	0,27	5,5	0,49					197
96	246	6,72	-351	737	572	165	0,78	0,29	3,1	0,33					258,32
97	247	6,67	-353	693	530	163	0,76	0,31	3,9	0,31	375	1128	524,7	2300	171,30
98	248	6,71	-355	611	518	93	0,85	0,18	1,4	0,41	459	782	471,8	2062	152,43
99	253	6,65	-338	504	448	56	0,89	0,13	8,4	0,35					207,05
100	254	6,68	-329	605	481	125	0,79	0,26	3,2	0,41	1207	547	416,6	3103	107,55
101	255	6,65	-335	600	470	130	0,78	0,28	8,0	0,42					137,23
102	260	6,59	-327	598	465	133	0,78	0,29	5,9	0,48	854	445	418,8	2558	135,16
103	261	6,59	-333	519	465	54	0,90	0,12	6,9	0,45					183,12
104	262	6,57	-336	549	398	103	0,73	0,25	7,4	0,42					184
105	267	6,55	-350	549	398	151	0,73	0,38	7,9	0,44	1575	783	541,8	4077	342,16
106	268	6,54	-339	560	451	109	0,81	0,24	6,9	0,43					310,83

Fonte: o autor (2023).