



APLICAÇÃO DO SISTEMA DE INFERÊNCIA FUZZY PARA ESTIMATIVA DO RESULTADO DA FERMENTAÇÃO ALCOÓLICA EM SISTEMAS DE BATELADA E BATELADA ALIMENTADA

Marcus Vinícius Costa Righetti ¹, Fernando Henrique da Silva ², José Eduardo Olivo ³

¹Mestrando em Engenharia Química, Universidade Estadual de Maringá - UEM, Maringá, Paraná. pg404719@uem.br

²Doutor em Engenharia Química, Universidade Estadual de Maringá - UEM, Maringá, Paraná. fhsilva@uem.br

³Orientador, Doutor, Departamento de Engenharia Química, Universidade Estadual de Maringá - UEM, Maringá, Paraná. jeolivo@uem.br

RESUMO

A fermentação alcoólica utilizando *Saccharomyces cerevisiae* é amplamente empregada no Brasil para a produção de etanol, sendo majoritariamente conduzida em batelada alimentada. Trata-se de um processo sujeito a incertezas associadas a parâmetros físicos, químicos e microbianos, dificultando o monitoramento em tempo real. Nesse contexto, este trabalho teve como objetivo desenvolver e avaliar um sistema de inferência fuzzy Mamdani-Assilian como estimador da massa de etanol no reator, visando aplicações como sensor de software. O modelo utilizou como variáveis de entrada o tempo de processo, a concentração inicial de açúcares redutores totais (S_0), a concentração inicial de células (X_0) e a razão entre o tempo de enchimento e o tempo de batelada (τ). A saída foi a massa de etanol. Foram empregadas funções de pertinência trapezoidais (3 e 5 termos) e a base de regras foi extraída pelo método Wang-Mendel, com poda por peso e Top-K. A avaliação considerou dados experimentais segmentados em treino (80%) e validação (20%), implementando a busca sistemática para seleção da melhor configuração. O arranjo com funções trapezoidais e cinco termos para entradas e saída apresentou o melhor desempenho. Apesar de acompanhar a tendência geral, o modelo apresentou subestimação de picos, simplificação de transições e resíduos com dispersão ampla e autocorrelação em curtos intervalos. Conclui-se que, embora os resultados mostrem desempenho moderado, a abordagem tem potencial como ferramenta de monitoramento, desde que sejam aplicadas melhorias na granularidade das funções de pertinência, na diversidade de dados e na etapa de geração e poda de regras.

PALAVRAS-CHAVE: Batelada; Batelada alimentada; Etanol; Fermentação alcoólica; Predição fuzzy.

1 INTRODUÇÃO

Os biocombustíveis oferecem uma alternativa sustentável e renovável aos combustíveis fósseis e são produzidos com a utilização de biomassa vegetal ou animal, ou seja, não são de origem mineral (Padder *et al.*, 2024). O etanol é um exemplo de biocombustível, podendo ser proveniente de processos químicos ou microbiológicos. O método mais utilizado mundialmente para a produção de etanol utiliza o microrganismo *Saccharomyces cerevisiae* em que, por meio do processo de fermentação dos açúcares provenientes de vegetais, obtém-se como produto o etanol (Lopes *et al.*, 2016).

No Brasil, a matéria-prima utilizada como fonte de açúcares para o microrganismo é proveniente da cana-de-açúcar que, quando comparada aos outros vegetais, é a que apresenta maiores vantagens energéticas e econômicas para a produção. Isso se deve ao reaproveitamento do CO_2 da atmosfera para o plantio da cana e ao menor consumo de combustível fóssil em sua produção (Dias *et al.*, 2009; Belincanta *et al.*, 2016).

Diversos estudos científicos foram cruciais para o avanço da produção de etanol. Essas pesquisas buscaram otimizar e descobrir novos processos de fermentação, selecionar e criar microrganismos mais eficazes, realizar o controle de bactérias contaminantes e melhorias nos controles químico e microbiológico nas destilarias (Amorim *et al.*, 2010; Zhang *et al.*, 2015; Liu *et al.*, 2013; Naghshbandi *et al.*, 2019).

Ainda existem incertezas associadas ao processo de fermentação que resultam em um modelo matemático complexo para sua descrição. A maioria dos dados é obtida em



laboratório e necessitam de equipamentos custosos para medições realizadas em tempo real e que são de difícil escalonamento para processos industriais de grande escala (Feng *et al.*, 2021).

Estudos recentes buscam o desenvolvimento de “sensores de software”, ou *soft-sensors*, que utilizam algoritmos inteligentes para estimar os dados futuros do processo. Esses sensores mitigam a necessidade de interferências constantes no processo para realizar medições e ensaios, além de dispensarem a utilização de sensores específicos para a medição de todos os parâmetros do procedimento. Portanto, podem ser utilizados como uma estratégia de controle em tempo real da fermentação (Brunner *et al.*, 2021; Meleiro; Maciel Filho, 2000).

A lógica *fuzzy* é utilizada como uma possível estratégia para o início da criação desses “sensores de software”, pois utilizam um sistema de inferência para prever o estado futuro de um processo com base em conhecimentos previamente obtidos sobre ele. É uma alternativa promissora para atuar em processos de fermentação, pois sistemas *fuzzy* são baseados nos dados de entrada e de saída do processo, ou performance do sistema, são robustos em relação às incertezas do processo e não necessitam de um modelo matemático para sua operação (Patnaik, 1997; Vaidehi *et al.*, 2008; Wen *et al.*, 2019).

Portanto, este trabalho tem como objetivo desenvolver um Sistema de Inferência *Fuzzy* (FIS) como possível estimador de resultados para um processo de fermentação alcoólica em sistemas de batelada e batelada alimentada a partir de dados experimentais previamente identificados.

2 MATERIAIS E MÉTODOS

A metodologia adotada seguiu um fluxo estruturado (Figura 1), cobrindo da organização e normalização dos dados à construção, teste e avaliação do FIS do tipo Mamdani–Assilian. Todos os dados experimentais foram obtidos a partir de Silva (2021), no qual o autor adotou um Delineamento Composto Central Rotacional (DCCR) buscando avaliar a influência que a concentração inicial de açúcar redutores totais (ART) (S_0), concentração de células (X_0) e variações na taxa do tempo de enchimento da batelada alimentada e o tempo de término da fermentação batelada nas mesmas condições (τ), exercem nas variáveis de resposta de interesse, com foco na velocidade geral de produção e no rendimento em etanol (Silva, 2021). No presente trabalho, a variável-alvo para predição pelo FIS é a massa de etanol no reator. As demais respostas podem ser tratadas de forma análoga.

O escopo de atuação deste trabalho se limita pelos valores máximos e mínimos utilizados e encontrados por Silva (2021). A Tabela 1 apresenta o valor máximo e mínimo de cada variável utilizada e observada durante o processo (Silva, 2021). Esse conhecimento sobre as variáveis do processo, permite a “fuzzificação” dos dados seguindo a metodologia proposta por Zadeh (1973) e Mamdani e Assilian (1975) e explicada em mais detalhes por Sivanandam *et al.* (2007).

A partir desses resultados, estruturou-se uma base de dados para modelagem. No contexto do FIS, as entradas consideradas foram *Tempo*, S_0 , X_0 e τ , e a saída foi a massa de etanol no reator. O FIS foi implementado e validado em *MATLAB* sem uso de toolboxes proprietárias, o que favorece reprodutibilidade e portabilidade. Estudos anteriores indicam a adequação dessa abordagem para problemas reais de engenharia (Wen *et al.*, 2019; Dholawala e Christian, 2020; Zhang e Zhou, 2021).

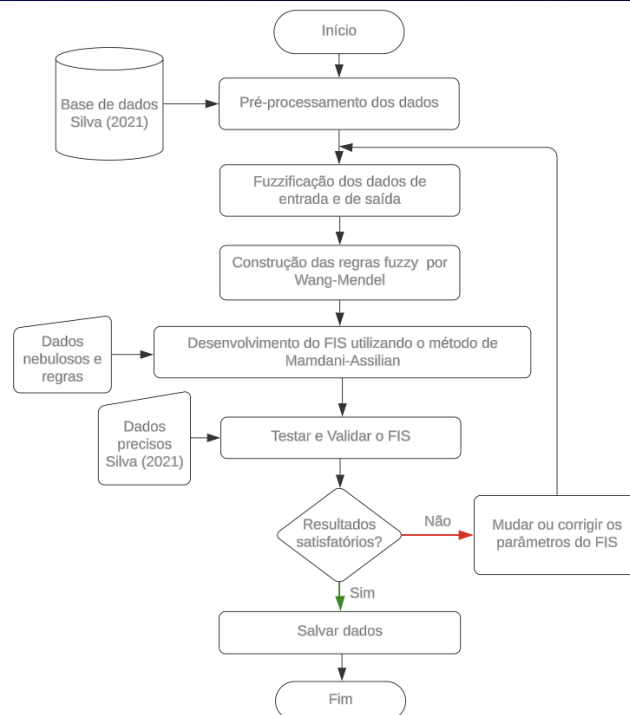


Figura 1: Fluxo de desenvolvimento e validação do FIS Mamdani–Assilian

Fonte: Os autores (2025)

O pré-processamento teve início com a segmentação dos dados, onde eles foram divididos em duas matrizes de dados, referentes às entradas e saídas do FIS, respectivamente. Em seguida, foi desenvolvida uma lógica de programação para realizar automaticamente a separação dos ensaios. Essa rotina permite identificar a nomenclatura específica de cada ensaio na base de dados e associar corretamente as variáveis de entrada e saída correspondentes. Dessa forma, garante-se a integridade dos dados utilizados, assegurando que todos os 28 ensaios únicos sejam reconhecidos e processados individualmente, sem perdas, duplicações ou sobreposição de informações.

Tabela 1: Intervalos observados para as variáveis de entrada e saída do processo fermentativo

Variável	Valor mínimo	Valor máximo
Concentração inicial de ART (So) (g/L)	150	290
Concentração celular massa seca inicial (Xo) (g/L)	15	55
Razão do tempo de enchimento pelo tempo de fermentação descontínua equivalente (τ)	0,15	0,65
Massa ART Consumida (g)	0	1051
Massa de etanol gerado (g)	0	377
Massa Frutose Consumida (g)	0	608
Massa de Células no Reator (g)	58	233
Massa Glicose Consumida (g)	0	442
Tempo (h)	0	16

Fonte: Adaptado de Silva (2021)

Com os ensaios identificados, os dados foram divididos em dois conjuntos, treino e validação, respeitando uma proporção de 80% para 20%, respectivamente. Essa separação foi realizada mantendo a informação completa dos ensaios, ou seja, foram utilizados os dados de 22 ensaios para o treinamento e 6 ensaios para validação. Para garantir a aleatoriedade e evitar viés na seleção, os ensaios foram embaralhados antes da divisão, assegurando que todas as amostras de um mesmo ensaio permanecessem



agrupadas em apenas um dos conjuntos. Essa divisão apresentou resultados satisfatórios e foi validada por autores como Butt *et al.* (2024) e Olatunji *et al.* (2022).

Essa divisão foi feita aleatoriamente, para manter a reprodutibilidade e consistência entre testes, foi adicionado uma semente fixa do gerador de números aleatórios, garantindo que os mesmos ensaios serão atribuídos ao treino e à validação em cada execução. Ao final do processo, obteve-se as matrizes de dados de entrada e de saída separadas por conjunto, resultando em dados específicos para treinamento e validação, com integridade preservada entre os pares de entrada e saída de cada ensaio.

Após o término da segmentação, foi realizada a normalização das variáveis de entrada e de saída do processo, visando padronizar as escalas numéricas e prevenir que variáveis com magnitudes muito elevada ou baixa interfiram negativamente no processo de inferência criado. Utilizou-se a padronização via *Z-score* (padronização estatística). Essa normalização transforma os dados para que eles tenham média zero e desvio padrão unitário. Essa escolha se justifica pela necessidade de estabilidade numérica durante o treinamento, melhoria da eficiência na convergência dos algoritmos de otimização e prevenção de saturação em etapas de regressão local ponderada (Pujaru *et al.*, 2024; Polatgil, 2022).

A “fuzzificação” das entradas foi realizada considerando cardinalidades linguísticas de 3 e 5 termos e três formatos de funções de pertinência: triangular, trapezoidal e gaussiana. Os parâmetros das funções de entrada foram obtidos a partir dos dados normalizados do conjunto de treinamento, definindo-se centros uniformemente distribuídos no domínio observado e larguras ajustadas de forma proporcional ao espaçamento entre centros, evitando extrapolações. Para a saída normalizada em $[0,1]$, adotaram-se centros fixos: 0,10; 0,50; 0,90 (três termos) e 0,05; 0,25; 0,50; 0,75; 0,95 (cinco termos).

A base de regras foi extraída automaticamente pelo método de Wang–Mendel (WM), aplicado exclusivamente ao conjunto de treinamento. Esse método segue um processo sistemático:

- 1) Para cada amostra, identifica-se, em cada variável de entrada, o termo linguístico cujo grau de pertinência é máximo;
- 2) Na variável de saída, determina-se também o termo de maior pertinência;
- 3) Forma-se então uma regra do tipo “Se (Entrada 1 é Termo X) e (Entrada 2 é Termo Y) ... então (Saída é Termo Z)”;
- 4) Atribui-se a essa regra um peso calculado como o produto dos graus de pertinência de todos os antecedentes pelo grau do consequente;
- 5) Em caso de conflito, quando diferentes amostras geram o mesmo conjunto de antecedentes, mas com consequentes diferentes, mantém-se apenas a regra com maior peso, pois ela é considerada mais representativa para aquele padrão de entrada.

Após a geração inicial, foram aplicadas estratégias de poda para reduzir redundâncias e priorizar regras mais relevantes. Utilizou-se poda por *Top-K*, selecionando as *K* regras mais representativas de acordo com seu peso, e por limiar mínimo de peso, eliminando regras com baixa ativação no conjunto de treinamento. As regras resultantes foram organizadas em matrizes específicas para cada combinação de cardinalidade e tipo de função de pertinência, bem como em versões concatenadas por variável de saída, permitindo seu carregamento direto no modelo Mamdani–Assilian sem a necessidade de regeneração (Gou, *et al.*, 2015).

A inferência foi conduzida segundo a arquitetura clássica de Mamdani–Assilian, utilizando operador *AND* por mínimo, implicação por mínimo, agregação por máximo e “desfuzzificação” pelo centroide calculado numericamente sobre uma malha uniforme no intervalo $[0,1]$. As previsões foram posteriormente desnormalizadas para o domínio original, obtendo-se a massa de etanol em gramas (Zadeh, 1973).



Para a variável de saída considerada neste estudo (etanol), foi empregada uma busca sistemática (*grid search*) para avaliar todas as combinações possíveis entre cardinalidades linguísticas (3 e 5 termos) e formatos de funções de pertinência (triangular, trapezoidal e gaussiana), sempre utilizando a base de regras gerada pelo método WM correspondente. O *grid search* consistiu em definir uma grade de parâmetros discretos, testar exaustivamente cada configuração e registrar as métricas de desempenho obtidas. Esse procedimento garantiu que a comparação fosse justa, pois todas as combinações foram avaliadas sob as mesmas condições de dados, normalização e conjunto de regras, permitindo identificar o arranjo de funções de pertinência mais adequado ao problema preditivo (Zhang *et al.*, 2023; Duranoglu, Altin, Kuçuk, 2024).

A seleção do arranjo final foi realizada no conjunto de validação, priorizando o coeficiente de determinação (R^2) como critério principal e reportando também o Erro Quadrático Médio (Root Mean Square Error – RMSE) e o Erro Médio Absoluto (Mean Absolute Error – MAE). Essas métricas foram calculadas tanto de forma global, considerando todos os dados, quanto por ensaio individual, de modo a identificar variações de desempenho entre diferentes experimentos.

3 RESULTADOS E DISCUSSÕES

As regras obtidas são interpretáveis e coerentes com a lógica físico-química do processo fermentativo. O Tempo aparece em todas elas, reforçando seu papel central na determinação do nível de produção de etanol. Combinações com S_0 e X_0 médios ou altos tendem a gerar consequentes Altos ou Muito Altos, especialmente quando associadas a valores elevados de τ . Por outro lado, configurações com Tempo Muito Baixo e/ou X_0 Muito Baixo resultam predominantemente em saídas Muito Baixo ou Baixo. Há ainda regras que indicam possibilidade de produção elevada mesmo com S_0 Baixo, quando as demais condições são favoráveis, comportamento compatível com a cinética adaptativa da fermentação. Observa-se a predominância de consequentes Muito Baixo e Baixo, que juntos representam cerca de 73% das regras, enquanto Alto e Muito Alto correspondem a 27%. Esse padrão reflete o conjunto experimental, no qual condições de baixa carga inicial de substrato e/ou baixa concentração celular são mais frequentes e resultam em menor produção de etanol.

A Tabela 3 apresenta os parâmetros e valores avaliados no processo de *grid search* aplicado ao modelo Mamdani–Assilian com regras geradas pelo método WM. Observa-se que a variação do número de termos linguísticos nas entradas e na saída (3 e 5 termos) permite avaliar o impacto da granularidade da representação *fuzzy* sobre o desempenho do modelo. Configurações com maior número de termos tendem a capturar nuances mais finas do comportamento do processo, mas podem aumentar a complexidade e o risco de sobreajuste. A inclusão de diferentes formatos de funções de pertinência (triangular, trapezoidal e gaussiana) possibilita verificar qual forma de distribuição oferece melhor ajuste aos dados observados.

A Tabela 4 apresenta o ranking das cinco melhores configurações obtidas, ordenadas de forma decrescente pelo coeficiente de determinação R^2 . Cada configuração é identificada pelo formato da função de pertinência e número de termos linguísticos das entradas e da saída. Também são reportadas as métricas de desempenho RMSE e MAE, permitindo comparar simultaneamente a precisão e o erro médio de cada arranjo. Essa análise evidencia como diferentes combinações de formatos e cardinalidades influenciam o desempenho do sistema, facilitando a seleção da configuração final mais adequada. Observa-se que as funções de pertinência trapezoidais com cinco termos para entradas e saídas obtiveram os melhores resultados, atingindo R^2 de 0,54569, acompanhadas de valores de RMSE e MAE relativamente menores em comparação com outras combinações.



Tabela 3: Parâmetros e valores testados no *grid search* para o modelo Mamdani–Assilian com regras Wang–Mendel

Parâmetro	Valores testados	Descrição
Número de termos linguísticos nas entradas	3 e 5	Quantidade de rótulos linguísticos atribuídos a cada variável de entrada.
Formato das funções de pertinência (MF)	Triangular, trapezoidal e gaussiana	Forma matemática utilizada para representar cada termo linguístico.
Número de termos linguísticos na saída	3 e 5	Quantidade de rótulos linguísticos atribuídos à variável de saída.

Fonte: Os autores (2025)

Embora as funções triangulares e gaussianas apresentem desempenhos próximos em alguns casos, a predominância das funções trapezoidais na parte superior do ranking indica maior capacidade de modelar as variações do processo fermentativo para a metodologia selecionada. O melhor desempenho observado para funções trapezoidais pode estar relacionado à sua maior flexibilidade na definição de regiões de pertinência com transições mais suaves e intervalos centrais mais amplos, favorecendo a robustez em regiões de sobreposição entre termos linguísticos. Em contraste, as funções triangulares, embora simples, apresentam transições mais abruptas, o que pode limitar a capacidade de capturar comportamentos intermediários. Já as gaussianas, apesar de oferecerem suavidade máxima, podem ser menos adaptáveis quando a distribuição dos dados não segue um padrão simétrico, resultando em menor ajuste em algumas regiões do domínio.

Tabela 4: Ranking das cinco melhores configurações do Mamdani–Assilian com regras Wang–Mendel para a saída Massa de Etanol

Tipo de MF	Nº de Termos Entrada	Nº de Termos Saída	R ²	RMSE	MAE
Trapezoidal	5	5	0,54569	55,857	46,874
Trapezoidal	5	3	0,49766	61,348	50,398
Triangular	5	5	0,49654	61,553	50,653
Triangular	5	3	0,49239	62,326	51,168
Gaussiana	5	3	0,48884	69,975	60,503

Fonte: Os autores (2025)

Ainda assim, observa-se que a diferença de R² entre as melhores configurações não é extremamente elevada, indicando que todas as combinações listadas possuem potencial preditivo similar. A predominância das configurações com cinco termos linguísticos tanto para entradas quanto para saídas pode ser explicada pela maior capacidade de discriminação das variáveis do processo, permitindo que o método WM capture relações mais sutis entre os antecedentes e o consequente. Esse aumento na granularidade tende a gerar regras mais específicas, o que, aliado à etapa de poda por peso e/ou *Top-K*, contribui para a eliminação de redundâncias e a manutenção das regras mais representativas.

As funções de pertinência das variáveis de entrada, foram definidas no espaço real utilizando cinco termos trapezoidais: Muito Baixo, Baixo, Médio, Alto e Muito Alto. Conforme ilustrado na Figura 2, essa configuração garante cobertura completa do domínio experimental, com sobreposição suave entre termos adjacentes e platôs centrais mais largos, favorecendo a robustez da inferência em regiões de transição. Entre as variáveis, destaca-se que τ apresenta faixas mais estreitas, refletindo a menor variação observada



experimentalmente, enquanto Tempo cobre de forma quase linear todo o ciclo fermentativo, permitindo ativação diferenciada de regras ao longo do processo.

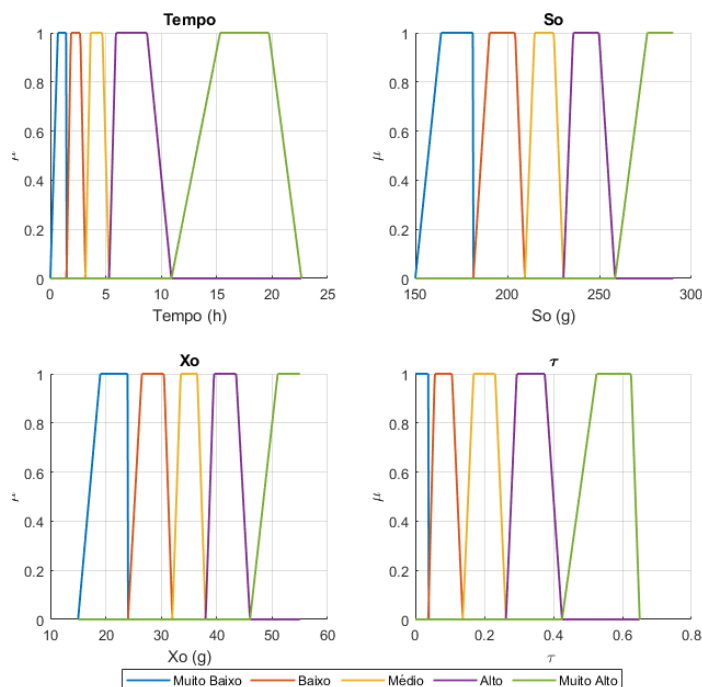


Figura 2: Funções de pertinência das entradas (Tempo, S_0 , X_0 e τ) no espaço real
Fonte: Os autores (2025)

A variável de saída, massa de etanol no reator, foi igualmente modelada com cinco funções de pertinência trapezoidais distribuídas de forma equidistante ao longo do domínio real, conforme apresentado na Figura 3. Essa abordagem facilita a “desfuzzificação” por centroide e mantém coerência com o padrão adotado nas entradas, embora a suavização nas regiões extremas possa contribuir para a leve subestimação de picos observada posteriormente.

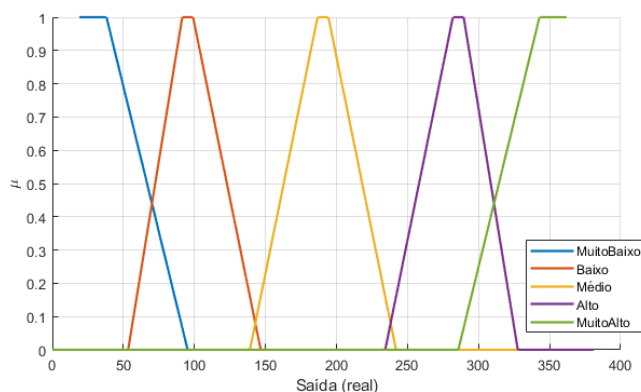


Figura 3: Funções de pertinência da saída (massa de etanol no reator) no espaço real
Fonte: Os autores (2025)

A Figura 4 apresenta a comparação entre os valores reais (linha azul contínua com marcadores circulares) e os valores preditos pelo modelo *fuzzy* Mamdani–Assilian (linha tracejada vermelha com marcadores em “x”) para a variável de saída massa de etanol no reator para os 6 ensaios de validação. No eixo horizontal estão as amostras temporais ordenadas por ensaio, e no eixo vertical, os valores em gramas da saída observada e estimada. Visualmente, observa-se que o modelo é capaz de acompanhar moderadamente



a tendência de crescimento e estabilização da produção de etanol ao longo de cada ensaio. Em diversos trechos, especialmente nas fases iniciais de subida, há boa sobreposição entre as curvas real e predita, indicando que o sistema *fuzzy* capturou adequadamente o comportamento dinâmico inicial da fermentação.

No entanto, a análise crítica revela que, nos platôs e picos máximos, o modelo apresenta subestimação significativa, com valores preditos sistematicamente abaixo dos reais. Além disso, há trechos em que as transições entre níveis de produção são representadas de forma abrupta nas previsões, enquanto os valores experimentais mostram variações mais graduais. Isso indica que, apesar de o modelo ter captado os padrões gerais, ele tende a simplificar a curva de resposta, possivelmente devido ao número limitado de regras ativas e à forma das funções de pertinência escolhidas.

Outro ponto relevante é que, embora o modelo apresente baixa dispersão de erros em alguns ciclos, em outros há divergências consistentes ao longo de todo o segmento, sugerindo que o conjunto de regras não representa com a mesma eficiência todas as combinações de condições de entrada. Esse comportamento pode estar relacionado à variabilidade intrínseca dos ensaios e à presença de regiões do espaço de entrada pouco exploradas no conjunto de treinamento. Portanto, confirma-se a capacidade do modelo de seguir a tendência global, mas evidencia a necessidade de ajustes finos, como o aumento da granularidade em regiões críticas, reavaliação da parametrização das funções de pertinência ou inclusão de mais regras relevantes, para reduzir os desvios observados nas fases de estabilização e nos valores de pico.

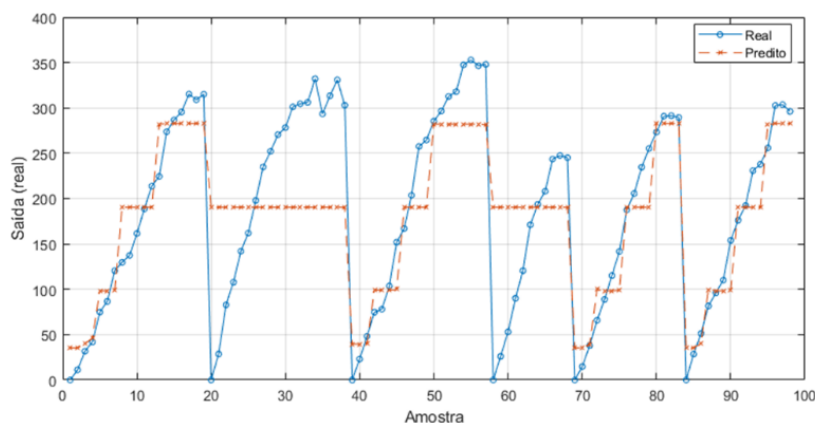


Figura 4: Comparação entre valores reais e preditos para a massa de etanol no reator utilizando o modelo Mamdani–Assilian com funções de pertinência trapezoidais (5 termos).

Fonte: Os autores (2025)

O gráfico Q-Q (Figura 5) compara os *quantis* dos resíduos do modelo com os *quantis* de uma distribuição normal padrão. A região central apresenta bom alinhamento com a linha de referência (vermelha tracejada), indicando que a maior parte dos resíduos segue aproximadamente a distribuição normal nessa faixa intermediária. Entretanto, nas extremidades inferiores e superiores, observa-se afastamento acentuado da linha, especialmente para valores negativos extremos, evidenciando a presença de *outliers* e comportamento não gaussiano nas caudas. Esses desvios sugerem que, em condições específicas do processo fermentativo, o modelo apresenta erros mais elevados e sistemáticos, possivelmente associados a combinações de variáveis pouco representadas no conjunto de treinamento.

O histograma de resíduos (Figura 6) apresenta a distribuição dos erros obtidos entre os valores reais e preditos pelo modelo Mamdani–Assilian. Observa-se que a maior parte dos resíduos está concentrada próxima de zero, indicando que o modelo tende a não gerar desvios sistemáticos significativos na média.

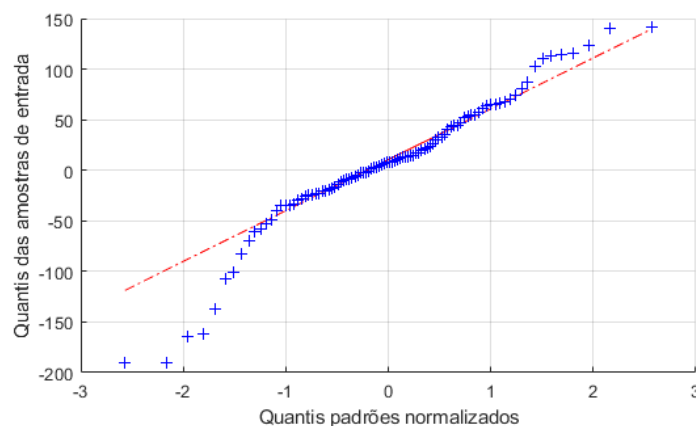


Figura 5: Gráfico Q-Q dos resíduos do modelo Mamdani–Assilian para a predição da massa de etanol no reator.

Fonte: Os autores (2025)

No entanto, a dispersão é relativamente ampla, com ocorrência de resíduos positivos e negativos de magnitude elevada, o que revela que em alguns pontos há diferenças substanciais entre a previsão e o valor observado. Essa distribuição apresenta caudas mais alongadas do que o esperado para uma distribuição normal, reforçando a presença de *outliers* já identificada na análise do gráfico Q-Q. O histograma evidencia que, apesar da centralização dos erros, existe assimetria e excesso de dispersão. Tal fato sugere que a modelagem pode ser otimizada por meio de ajustes na granularidade das funções de pertinência ou pelo aumento da representatividade de regiões específicas do espaço de entrada no conjunto de treinamento.

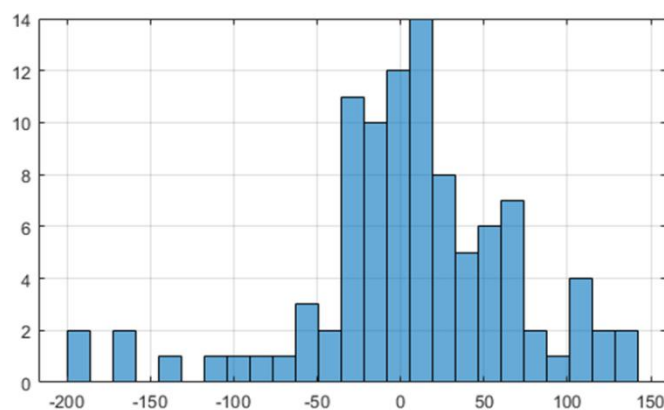


Figura 6: Histograma dos resíduos do modelo Mamdani–Assilian para a predição da massa de etanol no reator.

Fonte: Os autores (2025)

O gráfico de autocorrelação (Figura 7) evidencia que, embora o modelo capture tendências gerais, ajustes na base de regras ou nas funções de pertinência podem reduzir a persistência de erros em intervalos curtos e melhorar a independência temporal dos resíduos. A autocorrelação dos resíduos avalia a presença de padrões temporais não capturados pelo modelo. Nota-se que, nos primeiros *lags*, os valores de autocorrelação são positivos e ultrapassam os limites de confiança (linhas azuis), indicando correlação significativa entre erros próximos no tempo. Esse comportamento sugere que o modelo não conseguiu representar totalmente variações de curto prazo do processo fermentativo.

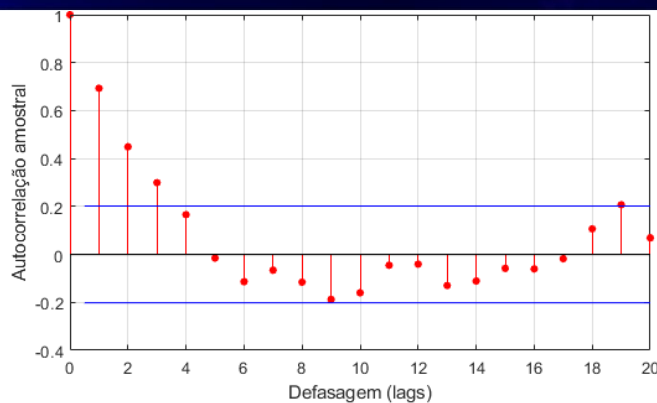


Figura 7: Autocorrelação dos resíduos do modelo Mamdani–Assilian para a predição da massa de etanol no reator ($NumLags = 20$).
Fonte: Os autores (2025)

A partir de *lag* 5, as autocorrelações tendem a se aproximar de zero e permanecem, em sua maioria, dentro do intervalo de confiança, caracterizando comportamento mais aleatório.

4 CONCLUSÃO

O sistema de inferência *fuzzy* Mamdani–Assilian, parametrizado com funções de pertinência trapezoidais e base de regras gerada pelo método Wang–Mendel, foi aplicado à predição da massa de etanol no reator. Embora tenha reproduzido parcialmente o comportamento do processo, os resultados apresentaram desempenho moderado, com coeficiente de determinação limitado e erros relativamente elevados. Observou-se subestimação nos picos de produção e simplificação excessiva das transições, sugerindo que a granularidade adotada e o conjunto de regras não foram suficientes para representar adequadamente as variações do processo.

As análises dos resíduos, via histograma, gráfico Q-Q e autocorrelação, confirmaram dispersão elevada, presença de *outliers* e padrões temporais não capturados, especialmente em curtos intervalos. A predominância de regras com consequentes Muito Baixo e Baixo indica desbalanceamento na cobertura da saída, possivelmente relacionado à distribuição não uniforme dos dados de treinamento. Esses fatores comprometem a precisão preditiva e limitam a aplicação prática do modelo em cenários que demandam estimativas confiáveis, como o controle em tempo real.

Para trabalhos futuros, recomenda-se ampliar e equilibrar a base de dados, ajustar o posicionamento e a forma das funções de pertinência, revisar a poda de regras para preservar combinações críticas e considerar abordagens híbridas ou adaptativas, como ANFIS. Apesar das limitações identificadas, a abordagem *fuzzy* mantém potencial como ferramenta de monitoramento, desde que seja otimizada para capturar de forma mais precisa e robusta a dinâmica do processo fermentativo.

REFERÊNCIAS

AMORIN, H. V.; GRYSCHKEK, M.; LOPES, M. L. The success and Sustainability of the Brazilian Sugarcane-Fuel Ethanol Industry. **Sustainability of the sugar and sugar-ethanol industries**. Volume 5, p. 73-82. 2010.



BELICANTA, J.; ALCHORNE, J. A.; SILVA, M. T. The Brazilian Experience with Ethanol Fuel: Aspects of Production, Use, Quality and Distribution Logistics. **Brazilian Journal of Chemical Engineering**. Volume 33. p. 1091-1102. 2016.

BRUNNER, V.; SIEGL, M.; GEIER, D.; BECKER, T. Challenges in the Development of Soft Sensors for Bioprocesses: A Critical Review. **Frontiers in Bioengineering and Biotechnology**. Volume 9. 2021.

BUTT, N. A.; AWAIS, M. M.; SHAHZADI, S.; KIM, T.H.; ASHRAF, I. Towards the development of believable agents: Adopting neural architectures and adaptive neuro-fuzzy inference system via playback of human traces. **Journal of King Saud University - Computer and Information Sciences**. Volume 36. p. 1319-1578. 2024.

DHOLAWALA, M. J.; CHRISTIAN, R. A. A Unique Variable Selection Approach in *Fuzzy* Modeling to Predict Biogas Production in Upflow Anaerobic Sludge Blanket Reactor (UASBR) Treating Distillery Wastewater. **Arabian Journal for Science and Engineering**. Volume 45. p. 7511-7527. 2020.

DIAS, M. O. S.; ENSINAS, A. V.; NEBRA, S. A.; MACIEL FILHO, R.; ROSSELL, C. E. V.; MACIEL, M. R. W. Production of bioethanol and other bio-based materials from sugarcane bagasse: Integration to conventional bioethanol production process. **Chemical Engineering Research and Design**. Volume 87. p. 1206-1216. 2009.

DURANOĞLU, D.; ALTIN, E. S.; KÜÇÜK, İ. Optimization of adaptive neuro-fuzzy inference system (ANFIS) parameters via Box-Behnken experimental design approach: The prediction of chromium adsorption. **Heliyon**, Volume. 10, e25813, 2024.

FENG, Y.; TIAN, X; CHEN, Y.; WANG, Z.; QIAN, J.; ZHUANG, Y.; CHU, J. Real-time and on-line monitoring of ethanol fermentation process by viable cell sensor and electronic nose. **Bioresources Bioprocess**. Volume 8. p. 1099-1113. 2021.

LIU, C. G.; XUE, C.; LIN, Y. H. BAI, F. W. Redox potential control and applications in microaerobic and anaerobic fermentations. **Biotechnology Advances**. Volume 31, p. 257-265. 2013.

LOPES, M. L.; PAULILLO, S. C. L; GODOY, A.; CHERUBIN, R. A.; LORENZI, M. S.; GIOMETTI, F. H. C.; BERNARDINO, C. D.; AMORIM NETO, H. B.; AMORIM, H. V. Ethanol production in Brazil: a bridge between science and industry. **Brazilian Journal of Microbiology**. Volume 47. p. 64-76. 2016.

MAMDANI, E. H.; ASSILIAN, S. An experiment in linguistic synthesis with a *fuzzy* logic controller. *Int. J. Human-Computer Studies*. Volume 51. p. 135-147. 1999.

MELEIRO, L. A. C.; MACIEL FILHO, R. State and parameter estimation based on a nonlinear filter applied to an industrial process control of ethanol production. **Brazilian Journal of Chemical Engineering**. Volume 17. p. 4-7. 2000.

NAGHSHBANDI, M. P; TABATABAEI, M.; AGHBASHLO, M.; GUPTA, V. K.; SULAIMAN, A.; KARIMI, K.; MOGHIMI, H.; MALEK, M. Progress toward improving ethanol production through decreased glycerol generation in *Saccharomyces cerevisiae* by metabolic and



genetic engineering approaches. **Renewable and Sustainable Energy Reviews**. Volume 115. 109353. 2019.

OLATUNJI, O. M.; HORSFALL, I. T.; ONUOHA, E. U.; ARIA, K. O. Application of hybrid ANFIS-based non-linear regression modeling to predict the %oil yield from grape peels: Effect of process parameters and FIS generation techniques. **Cleaner Engineering and Technology**. Volume 6. 2022.

PADDER, S. A.; KHAN, R.; RATHER, R. A. Biofuel generations: New insights into challenges and opportunities in their microbe-derived industrial production. **Biomass and Bioenergy**. Volume 185. 107220. 2024.

POLATGIL, M. Investigation of the Effect of Normalization Methods on ANFIS Success: Forestfire and Diabetes Datasets. I.J. **Information Technology and Computer Science**. Volume 14. p. 1–8. 2022.

PUJARU, K.; ADAK, S.; KAR, T. K.; PATRA, S.; JANA, S. Mamdani fuzzy inference system with trapezoidal membership functions for investigating fishery production. **Decision Analytics Journal**. Volume 11. 2024.

SILVA, F. H. '**Scale-up**' populacional de fermentação alcoólica de mel invertido em sistema de batelada alimentada. Tese de Doutorado. Universidade Estadual de Maringá – UEM. 119 p. Maringá, 2021.

SINAVANDAM, S. N.; SUMATHI, S.; DEEPA, S. N. **Introduction to Fuzzy Logic using MATLAB**. São Paulo: Springer. Berlin. 2007.

VAIDEHI, V.; MONICA S.; MOHAMED SHEIK SAFEER S.; DEEPIKA M.; SANGEETHA S. A prediction system based on *fuzzy* logic. **Proceedings of the World Congress on Engineering and Computer Science**. WCECS. San Francisco, USA. 2008.

WEN, H.; ZHAO, X.; CHEN, X. *Fuzzy* prediction of metro traffic flow. **International Conference on Fuzzy Theory and its Applications**. 2019.

ZADEH, L. A. Outline of a new approach to the analysis of complex systems and decision processes. **IEEE Transactions on Systems, Man and Cybernetics**. Volume SMC – 3, n° 1. 1973.

ZHANG, D.; CHEN, T. Scikit-ANFIS: A Scikit-Learn Compatible Python Implementation for Adaptive Neuro-Fuzzy Inference System. **International Journal of Fuzzy Systems**. Volume 26. p. 2039–2057. 2024.

ZHANG, Z.; ZHOU, Q. The study of multivariate *fuzzy* neural scheme in controlling temperature during beer fermentation. **Journal of Physics: Conference Series**. 1732. 012140. 2021.