

RESSALVA

Atendendo a solicitação do(a) autor(a), Daniel Hiroshi Rufino Toda, o texto completo deste documento será disponibilizado somente a partir de 03/04/2028.

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
Faculdade de Engenharia - Campus de Bauru

DANIEL HIROSHI RUFINO TODA

**INTEGRAÇÃO DE MODELOS FÍSICOS E DE APRENDIZADO DE MÁQUINA
PARA PREDIÇÃO DA VELOCIDADE DE SEDIMENTAÇÃO EM TRATAMENTO DE
ÁGUA**

Bauru
2026

DANIEL HIROSHI RUFINO TODA

**INTEGRAÇÃO DE MODELOS FÍSICOS E DE APRENDIZADO DE MÁQUINA
PARA PREDIÇÃO DA VELOCIDADE DE SEDIMENTAÇÃO EM TRATAMENTO DE
ÁGUA**

Dissertação apresentada à Universidade Estadual Paulista (UNESP), Faculdade de Engenharia, Bauru, para obtenção do título de Mestre em Engenharia Civil e Sanitária.

Área de Concentração: Saneamento

Orientador(a): Prof. Dr. Adriano Bressane

Bauru

2026

T633i

Toda, Daniel Hiroshi Rufino.

Integração de modelos físicos e de aprendizado de máquina para predição da velocidade de sedimentação em tratamento de água / Daniel Hiroshi Rufino Toda. - Bauru, 2026

135 p. : il., tabs

Dissertação (Mestrado)-Universidade Estadual Paulista (Unesp), Faculdade de Engenharia, Bauru
Orientador: Adriano Bressane

1. Machine learning. 2. Água Estações de tratamento. 3. Sedimentation analysis. I. Título.

ATA DA DEFESA PÚBLICA DA DISSERTAÇÃO DE MESTRADO DE DANIEL HIROSHI RUFINO TODA, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL E AMBIENTAL, DA FACULDADE DE ENGENHARIA - CÂMPUS DE BAURU.

Aos 02 de abril de 2026, às 10h30min, por meio de Videoconferência, realizou-se a defesa de DISSERTAÇÃO DE MESTRADO de DANIEL HIROSHI RUFINO TODA, intitulada "INTEGRAÇÃO DE MODELOS FÍSICOS E DE APRENDIZADO DE MÁQUINA PARA PREDIÇÃO DA VELOCIDADE DE SEDIMENTAÇÃO EM TRATAMENTO DE ÁGUA ", sob orientação do Prof. Dr. null. A Comissão Examinadora foi constituída pelos seguintes membros: Prof. Dr. ADRIANO BRESSANE (Participação Virtual) do(a) Departamento de Engenharia Ambiental / UNESP / Câmpus de São José dos Campos - ICT, Prof. Dr. HENRIQUE EWBANK DE MIRANDA VIEIRA (Participação Virtual) do(a) University of North Texas, Profª. Drª. LIANE YURI KONDO NAKADA (Participação Virtual) do(a) Departamento de Engenharia Ambiental / UNESP / Câmpus de São José dos Campos - ICT, Após a exposição pelo mestrando e arguição pelos membros da Comissão Examinadora que participaram do ato, de forma presencial e/ou virtual, o discente recebeu o conceito final: APROVADO. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelo(a) Presidente(a) da Comissão Examinadora.

Documento assinado digitalmente.
 **ADRIANO BRESSANE**
Data: 02/04/2026 14:23:31 -0300
Verifique em <https://validar.dfi.gov.br>

Prof. Dr. ADRIANO BRESSANE

RESUMO

A estimativa acurada da velocidade de sedimentação, uma característica crítica de escoamentos multifásicos, para a otimização de unidades de sedimentação no tratamento de água e efluentes, bem como para a compreensão e modelagem de processos em corpos hídricos naturais. Modelos tradicionais, como a Lei de Stokes, frequentemente falham em capturar o comportamento complexo de agregados fractais, cujas estruturas irregulares influenciam os coeficientes de arrasto e a dinâmica de sedimentação. Estudos têm demonstrado que esses agregados podem sedimentar mais rapidamente ou mais lentamente do que esferas perfeitas, enfatizando o papel crítico da morfologia do agregado para além de parâmetros simples de tamanho.

Embora modelos de aprendizado de máquina sejam capazes de tratar não linearidades, eles frequentemente carecem de embasamento físico, o que é crucial para a estimativa acurada da velocidade de sedimentação de agregados fractais influenciada por interações fluido-partícula e pela dinâmica do arrasto. Para contornar essa limitação, foram desenvolvidos e comparados dois modelos preditivos: um modelo tradicional de *Gradient Boosting Regressor* (GBR) e um modelo GBR estendido por meio de engenharia de atributos baseada em princípios físicos (PFE-GBR). Este último incorpora atributos derivados da Lei de Stokes, do número de Reynolds e da força de arrasto, aumentando sua aderência à teoria da sedimentação. Este trabalho estabelece um *baseline* crítico, demonstrando a eficácia da integração de princípios físicos em um algoritmo de aprendizado de máquina amplamente utilizado e reconhecido por sua robustez, como o GBR.

Ambos os modelos foram avaliados de forma rigorosa por meio de validação cruzada com cinco partições. O modelo PFE-GBR superou o modelo tradicional, alcançando um coeficiente de determinação R^2 de 0,951 no conjunto de teste, em comparação a 0,887, além de reduzir o erro absoluto médio de 234,31 $\mu\text{m/s}$ para 136,05 $\mu\text{m/s}$, juntamente com uma redução de 30,1 % no erro quadrático médio (RMSE). Essas melhorias demonstram que a incorporação de princípios físicos por meio da engenharia de atributos em estruturas de aprendizado de máquina aumenta a acurácia preditiva, ao mesmo tempo em que orienta implicitamente a generalização e a interpretabilidade do modelo. Os resultados obtidos ressaltam o

valor de abordagens informadas por física na modelagem de processos de sedimentação, fornecendo subsídios aplicáveis ao aprimoramento do projeto, do dimensionamento e da eficiência operacional de sistemas de tratamento de água, bem como à modelagem do transporte e deposição de partículas em sistemas aquáticos naturais e escoamentos multifásicos.

Palavras-chave: velocidade de sedimentação, agregados fractais, tratamento de água, sedimentação, aprendizado de máquina.

ABSTRACT

The accurate estimation of sedimentation velocity, a critical characteristic of multiphase flows, is essential for the optimization of sedimentation units in water and wastewater treatment, as well as for the understanding and modeling of processes in natural water bodies. Traditional models, such as Stokes' law, often fail to capture the complex behavior of fractal aggregates, whose irregular structures influence drag coefficients and sedimentation dynamics. Studies have shown that such aggregates may settle faster or slower than perfect spheres, highlighting the critical role of aggregate morphology beyond simple size parameters.

Although machine learning models are capable of handling nonlinearities, they often lack physical grounding, which is crucial for the accurate estimation of sedimentation velocity in fractal aggregates governed by fluid–particle interactions and drag dynamics. To address this limitation, two predictive models were developed and compared: a conventional Gradient Boosting Regressor (GBR) and a GBR extended through physics-based feature engineering (PFE-GBR). The latter incorporates attributes derived from Stokes' law, the Reynolds number, and drag force, thereby increasing its adherence to sedimentation theory. This work establishes a critical baseline by demonstrating the effectiveness of integrating physical principles into a widely used and robust machine learning algorithm such as GBR.

Both models were rigorously evaluated using five-fold cross-validation. The PFE-GBR outperformed the traditional model, achieving a coefficient of determination (R^2) of 0.951 on the test set, compared to 0.887, while reducing the mean absolute

error from 234.31 $\mu\text{m/s}$ to 136.05 $\mu\text{m/s}$, along with a 30.1% reduction in the root mean squared error (RMSE). These improvements demonstrate that incorporating physical principles through feature engineering within machine learning frameworks enhances predictive accuracy, while implicitly guiding model generalization and interpretability. The results underscore the value of physics-informed approaches in modeling sedimentation processes, providing actionable insights for improving the design, sizing, and operational efficiency of water treatment systems, as well as for modeling particle transport and deposition in natural aquatic systems and multiphase flows.

Keywords: Settling velocity, fractal aggregates, water treatment, sedimentation, machine learning.

LISTA DE ILUSTRAÇÕES

Artigo publicado I

Figure 1. Schematic representation of the experimental set up. Source: Modified from Moruzzi et al. (2020b).	39
Figure 2. Modelling process based on fuzzy machine learning. Source: modified from Bressane et al. (2017).	43
Figure 3. Process of the FIS: a) fuzzification, b) conjunction, c) disjunction, d) defuzzification. Source: modified from Bressane et al. (2022).	45
Figure 4. ANN based on MLP deep learning architecture. Source: modified from Bressane et al. (2024b).	47
Figure 5. Relationship between settling velocity and the key predictors.	50
Figure 6. Descriptive analysis of fractal aggregate features. Variables in red indicate those that were excluded due to high correlation.	53
Figure 7. Performance of predicting fuzzy models in the cross-validation tests.	54

Artigo publicado II

Figure. 1. Framework for accurate prediction of settling velocity and generalisation tests: (a) Fractal aggregate tracking using PIV (Bressane et al., 2024), (b) PFE feature engineering procedure.	84
Figure. 2. Comparison of Traditional and PIML GBR models: (a) Predicted vs. actual settling velocities, (b) Residuals vs. predicted velocities.	104
Figure. 3. Feature importance ranked by mean SHAP values from: (a) traditional, (b) PIML GBR.	104
Figure. 4. Model robustness under increasing noise.	106
Figure. 5. Error comparison across folds: Traditional vs. PIML GBR.	111

LISTA DE TABELAS

Artigo publicado I

Table 1. Raw water characteristics	38
Table 2. Summary of image analysis features for suspended particles.	40
Table 3. Relative importance of predictor variables.	48
Table 4. Settling velocity and key predictors prioritised for fuzzy ML- based models.	49
Table 5. Performance of predicting models in the training and testing sets.	54

Artigo publicado II

Table 1. Performance of the traditional and PFE GBR models.	99
Table 2. Performance of the PIML GBR under levels of Gaussian noise.	105
Table 3. Fold-wise errors comparison between models.	108
Table 4. Error metrics by morphological group (MAE in $\mu\text{m/s}$).	111

LISTA DE ABREVIATURAS E SIGLAS

ANN	Artificial Neural Network
ETA	Estação de Tratamento de Água
Fd	Força de arrasto
FIS	Fuzzy Inference System
GBR	Gradient Boosting Regressor
MAE	Mean Absolute Error
ML	Machine Learning
MLP	Multilayer Perceptron
PFE	Physics-Based Feature Engineering
PFE-GBR	Gradient Boosting Regressor com engenharia de atributos baseada em princípios físicos
PIML	Physics-Informed Machine Learning
PIV	Particle Image Velocimetry
R ²	Coeficiente de determinação
Re	Número de Reynolds
RMSE	Root Mean Squared Error
SHAP	Shapley Additive Explanations

LISTA DE SÍMBOLOS

v	Velocidade terminal de sedimentação dos agregados medida experimentalmente por PIV ($\mu\text{m/s}$)
v_s	Velocidade de sedimentação teórica estimada pela Lei de Stokes ($\mu\text{m/s}$)
d	Diâmetro esférico equivalente do agregado fractal, derivado da área projetada (μm)
ρ_p	Densidade da partícula (kg/m^3)
ρ_f	Densidade do fluido (kg/m^3)
μ	Viscosidade dinâmica do fluido ($\text{Pa}\cdot\text{s}$)
g	Aceleração gravitacional (m/s^2)
Re	Número de Reynolds, definido como $Re = \rho_f \cdot v \cdot d / \mu$
Cd	Coeficiente de arrasto (adimensional)
Fd	Força de arrasto atuando sobre o agregado (N)
A	Área projetada da seção transversal do agregado (m^2)
$Area/Box$	Razão entre a área do agregado e a área da caixa delimitadora (adimensional)
$Box\ X/Y$	Razão entre as dimensões da caixa delimitadora (adimensional)
$Clumpiness$	Medida da heterogeneidade interna e da irregularidade textural dos agregados (adimensional)
$Cluster$	Medida da densidade de agregação (adimensional)
$Compactness$	Grau de densidade e solidez da estrutura do floco (adimensional)

SUMÁRIO

1	INTRODUÇÃO.....	13
2	FUNDAMENTAÇÃO TEÓRICA.....	19
2.1	ENGENHARIA DE ATRIBUTOS BASEADA EM PRINCÍPIOS FÍSICOS PARA APRENDIZADO DE MÁQUINA (PHYSICS-BASED FEATURE ENGINEERING – PFE).....	19
2.2	CARACTERÍSTICAS FUNDAMENTAIS DO PROBLEMA INVESTIGADO.....	21
2.3	PROJETO DE UM MODELO PFE PARA PREDIÇÃO DA VELOCIDADE DE SEDIMENTAÇÃO.....	22
3	OBJETIVOS.....	26
3.1	OBJETIVO GERAL.....	26
3.2	OBJETIVOS ESPECÍFICOS.....	26
4	ESTRUTURA DA DISSERTAÇÃO.....	27
5	ARTIGO PUBLICADO I.....	29
6	ARTIGO PUBLICADO II.....	71
7	DISCUSSÃO.....	120
8	CONCLUSÕES E RECOMENDAÇÕES.....	127
8.1	RECOMENDAÇÕES DE ESTUDOS FUTUROS.....	128
	REFERÊNCIAS BIBLIOGRÁFICAS.....	130
	APÊNDICES.....	136

1 INTRODUÇÃO

A separação de material particulado de fluidos por meio da sedimentação constitui uma operação unitária fundamental em sistemas multifásicos, com aplicações que se estendem além do tratamento de água, abrangendo áreas como tratamento de efluentes, processamento mineral, indústria farmacêutica e remediação ambiental. A compreensão e a estimativa acurada da velocidade de sedimentação de partículas em suspensão têm sido objeto de investigação na comunidade científica por décadas, com contribuições significativas para o entendimento fundamental dos escoamentos multifásicos e dos processos de separação.

Nesse contexto, a investigação conduzida neste trabalho sobre a sedimentação de agregados fractais no tratamento de água oferece um estudo de caso relevante para o avanço do conhecimento fundamental na engenharia sanitária e ambiental, particularmente no que se refere à dinâmica do arrasto, aos efeitos da morfologia dos agregados e à sua influência sobre o comportamento de sedimentação em sistemas multifásicos.

Portanto, a sedimentação constitui um elemento central das operações de tratamento de água, nas quais a velocidade de sedimentação dos flocos determina a eficiência da separação sólido-líquido. Uma sedimentação eficiente é essencial não apenas para otimizar a remoção de material particulado e poluentes, mas também para aumentar a eficiência global das estações de tratamento de água (ETAs), reduzir custos operacionais e mitigar impactos ambientais (Bürger et al., 2023; Toledo-Zucco et al., 2305). Essa otimização contribui para a mitigação do risco de doenças de veiculação hídrica ao melhorar a claridade e a segurança da água tratada (Chen e Fu, 2023). Além disso, permite a redução do consumo de produtos químicos e da demanda energética, favorecendo processos de tratamento mais sustentáveis (Adam e Lukashevich, 2023).

Sob a perspectiva da engenharia sanitária e ambiental, a otimização dos processos de tratamento de água por meio do aprimoramento da sedimentação não apenas fortalece a oferta de água limpa e segura, um desafio global crítico, como também contribui diretamente para os princípios da sustentabilidade. Ao possibilitar a remoção mais eficiente de poluentes e material particulado, esta pesquisa apoia o

desenvolvimento de tecnologias de tratamento de água mais eficientes no uso de recursos e ambientalmente responsáveis. A capacidade de estimar com acurácia as velocidades de sedimentação, por meio de uma modelagem que incorpore atributos informados por princípios físicos e características morfológicas fractais dos agregados floculentos, representa um avanço tanto em relação aos métodos clássicos (como lei de Stokes) como aos métodos empíricos tradicionais. Assim, oferece uma ferramenta estratégica para engenheiros no projeto e na operação mais eficaz desses sistemas essenciais.

A abordagem empregada neste trabalho para estimar as velocidades de sedimentação sustenta-se em princípios fundamentais da física e da química, como a mecânica de fluidos, que rege o movimento dos agregados no meio aquoso e os efeitos associados ao arrasto; os mecanismos de transferência de massa, particularmente no que se refere à influência da morfologia dos agregados em sua interação com o fluido; e a caracterização do material particulado, cujas propriedades, tamanho, forma, densidade e características superficiais, frequentemente descritas por meio de dimensões fractais, determinam o comportamento macroscópico do sistema.

Sendo assim, este trabalho busca superar as limitações dos modelos tradicionais, sejam eles mecanicistas ou puramente empíricos, ao incorporar explicitamente esses princípios fundamentais em um arcabouço preditivo avançado.

Compreender em profundidade a velocidade de sedimentação é essencial não apenas para melhorar a claridade da água, mas também para reduzir custos operacionais e aumentar a vida útil dos sistemas de filtração subsequentes. Os flocos apresentam ampla variedade de estruturas morfológicas, frequentemente com características não esféricas e porosas, cujas dimensões fractais afetam significativamente a densidade efetiva e comportamento de sedimentação (Moruzzi et al., 2020; Bankole et al., 2024).

De modo geral, flocos maiores e mais porosos tendem a sedimentar mais lentamente em razão do aumento da flutuabilidade e da força de arrasto, o que demanda o uso de modelos avançados para a estimação acurada de sua dinâmica (Asensi e Alemany, 2022). Uma exceção a essa tendência ocorre em casos específicos descritos na literatura, como o de Johnson e Logan (JOHNSON et al., 1996), que observaram a possibilidade de escoamento de água através de

macroporos nos flocos, o que pode alterar significativamente o padrão de sedimentação.

Modelos clássicos, como a Lei de Stokes, simplificam excessivamente essa dinâmica, frequentemente falhando em capturar a natureza multifacetada das propriedades dos flocos, incluindo tamanho, textura e morfologia (Vahedi e Gorczyca, 2014; Chakraborti e Kaur, 2014).

Nesse contexto, Bressane et al. (Bressane et al., 2024) enfatizaram o papel crítico dos agregados fractais durante a sedimentação, demonstrando que fatores como a compactação (clumpiness), marginação e a dimensão fractal influenciam significativamente a velocidade de sedimentação. Os autores ressaltaram a importância de integrar essas características morfológicas aos modelos preditivos, de modo a aumentar sua acurácia e utilidade prática em cenários de tratamento de água.

Avanços recentes em aprendizado de máquina (ML) oferecem ferramentas poderosas para a modelagem de sistemas não lineares; contudo, muitos modelos de ML carecem da capacidade de incorporar conhecimento específico do domínio. Para superar essas limitações, abordagens computacionais como a engenharia de atributos baseada em princípios físicos (physics-based feature engineering – PFE) e modelos fuzzy têm emergido, capazes de integrar leis físicas ao processo de modelagem e lidar com incertezas nos dados, aumentando a acurácia preditiva (Bressane et al., 2024; Ferdowsi et al., 2022). Essa abordagem é especialmente relevante no contexto da operação unitária de sedimentação, em que a dinâmica do fluido e as interações partícula-fluido são influenciadas por fatores como o número de Reynolds (Re), força de arrasto (F_d) e porosidade dos agregados (DELCEY et al., 2024).

No contexto da sedimentação, a incorporação de princípios físicos clássicos, como a Lei de Stokes, o número de Reynolds (Re) e a força de arrasto (F_d), permite que os modelos representem de forma mais realista as interações complexas entre tamanho, densidade e morfologia das partículas. Essa fundamentação baseada em leis físicas é especialmente relevante para assegurar que os modelos preditivos se mantenham consistentes com o comportamento observado em sistemas reais, mesmo diante de desafios como a não linearidade das relações envolvidas e a escassez de dados (MORUZZI et al., 2020; BAKIRI & NACEF, 2022). Estudos recentes também indicam que a incorporação de

características morfológicas, em especial dimensões fractais, aumenta significativamente a confiabilidade dos modelos, reduzindo a lacuna entre o entendimento teórico e a aplicação prática em contextos reais (Asensi e Alemany, 2022; Jin et al., 2019).

Apesar dos avanços e das contribuições relevantes dos modelos de sedimentação existentes, muitos ainda apresentam dificuldades em representar a complexidade da morfologia de agregados fractais, especialmente no ambiente dinâmico dos processos de tratamento de água. Modelos empíricos frequentemente carecem de generalização quando aplicados a diferentes tipos de agregados e condições operacionais, enquanto abordagens puramente teóricas tendem a simplificar excessivamente as estruturas intrincadas dos flocos e suas interações com o fluido (Gregory e Barany, 2011). Esse cenário evidencia uma lacuna importante na literatura: a necessidade de modelos preditivos que integrem de forma eficaz descritores físicos detalhados de agregados fractais às capacidades adaptativas de técnicas modernas de aprendizado de máquina.

O presente trabalho aborda essa lacuna ao introduzir uma estrutura inovadora de engenharia de atributos informada por princípios físicos (physics-based feature engineering – PFE). Especificamente, este estudo tem como objetivo quantificar os ganhos de desempenho obtidos ao aprimorar um modelo de aprendizado de máquina consolidado e robusto, como o Gradient Boosting Regressor (GBR), por meio da incorporação de atributos baseados em leis físicas. Ao focar em uma comparação direta e controlada com sua contraparte puramente orientada por dados (modelo empírico), este trabalho fornece uma demonstração clara e fundamental do valor da abordagem PFE. Essa proposta distingue-se de outras abordagens informadas por física que buscam derivar novas formulações matemáticas explícitas a partir dos dados.

A inovação da abordagem proposta reside na integração criteriosa de leis físicas bem estabelecidas, como a Lei de Stokes, bem como de considerações relacionadas à força de arrasto (F_d) e ao número de Reynolds (Re), diretamente na arquitetura de um modelo de aprendizado de máquina. Essa metodologia PFE representa um afastamento significativo em relação à modelagem empírica tradicional comumente empregada, oferecendo uma estrutura capaz de explorar as capacidades adaptativas dos algoritmos de aprendizado de máquina, enquanto assegura consistência física e maior interpretabilidade, aspectos essenciais para o

avanço da compreensão e do projeto de processos. O impacto dessa abordagem estende-se à disponibilização de ferramentas mais confiáveis e precisas para a otimização das etapas de sedimentação no tratamento de água, atendendo diretamente à demanda por tecnologias eficientes e sustentáveis nessa área crítica de aplicação da sanitária e ambiental.

Métodos tradicionais para a estimativa da velocidade de sedimentação frequentemente não conseguem representar a estrutura complexa e irregular de agregados fractais, o que resulta em imprecisões no projeto e na operação de unidades de sedimentação (Kim et al., 2023; Alalm et al., 2016). Para superar essas limitações, este estudo propõe uma abordagem de engenharia de atributos informada por princípios físicos (physics-based feature engineering – PFE) que incorpora características morfológicas extraídas por meio de análise de imagens em um modelo Gradient Boosting Regressor (GBR). Essa abordagem oferece uma predição da velocidade de sedimentação mais robusta e fisicamente consistente, avançando além das limitações de modelos puramente empíricos ou teóricos comumente empregados no tratamento de água.

Este trabalho compara duas estratégias de modelagem preditiva para a estimativa da velocidade de sedimentação de agregados fractais. A primeira consiste em um modelo tradicional de aprendizado de máquina, que utiliza um GBR treinado exclusivamente com descritores morfológicos e estatísticos, como área, razão de aspecto, porosidade e circularidade, obtidos por meio de análises de imagens e medições experimentais, sem a incorporação explícita de leis físicas, explorando sua reconhecida capacidade de capturar relações não lineares complexas. A segunda estratégia corresponde a um modelo GBR aprimorado por PFE, que estende a abordagem tradicional ao incorporar variáveis baseadas em princípios físicos ao processo de aprendizagem, aumentando assim sua consistência física e robustez preditiva.

A utilização do mesmo algoritmo de regressão em ambas as abordagens assegura uma comparação justa, uma vez que controla a influência da arquitetura do modelo e isola o efeito da incorporação de atributos informados por física. Por meio dessa análise comparativa, a pesquisa busca demonstrar como a incorporação de conhecimento físico em modelos de aprendizado de máquina pode melhorar a acurácia e a aplicabilidade em processos reais de sedimentação.

O desempenho dos modelos é avaliado utilizando métricas padrão de regressão: erro absoluto médio (MAE), raiz do erro quadrático médio (RMSE) e coeficiente de determinação (R^2). Os resultados ressaltam o valor agregado da abordagem PFE na modelagem da sedimentação, oferecendo subsídios práticos para a otimização do projeto e da eficiência operacional de estações de tratamento de água (ETAs).

8 CONCLUSÕES E RECOMENDAÇÕES

A pesquisa analisou comparativamente estratégias baseadas em aprendizado de máquina para a predição da velocidade terminal de sedimentação de agregados fractais em processos de tratamento de água, evidenciando a relevância de integrar princípios físicos à modelagem preditiva. Os resultados, obtidos a partir dos artigos publicados demonstram de forma consistente que abordagens híbridas, como a engenharia de atributos baseada em princípios físicos para aprendizado de máquina (*physics-based feature engineering – PFE*) e a modelagem fuzzy, superam limitações de modelos puramente empíricos ou heurísticos, oferecendo maior robustez, coerência física e interpretabilidade.

Assim, os resultados alcançados contribuem para o avanço de técnicas voltadas à otimização da etapa de sedimentação, particularmente em Estações de Tratamento de Água (ETAs), alinhando-se às diretrizes de eficiência operacional e sustentabilidade.

Em particular, destaca-se que:

(i) O primeiro estudo verificou que atributos morfológicos de agregados fractais, tais como marginação, raio médio e clumpiness, exercem influência determinante na velocidade de sedimentação, reforçando a necessidade de descritores detalhados para a adequada representação da complexidade dos flocos.

(ii) A modelagem fuzzy mostrou-se eficaz na representação de incertezas e variabilidades experimentais, atingindo elevado desempenho preditivo ($R^2 = 0.923$) mesmo em condições de dados ruidosos, o que respalda sua aplicabilidade em cenários reais de ETAs.

(iii) O segundo estudo demonstrou que a integração explícita de variáveis físicas, como número de Reynolds, força de arrasto e expressões derivadas da Lei de Stokes, ao modelo Gradient Boosting Regressor (GBR) resulta em ganhos expressivos de acurácia e redução de erro absoluto médio (MAE) e erro quadrático médio (RMSE).

(iv) A comparação entre as abordagens GBR Tradicional vs PFE - GBR evidencia que o segundo apresenta maior capacidade de generalização frente a ruídos artificiais nos dados, destacando-se como uma estratégia promissora para aplicação em contextos operacionais dinâmicos e incertos.

(v) A análise por meio de valores SHAP permitiu interpretar as contribuições individuais das variáveis físicas e morfológicas para o desempenho dos modelos, validando sua relevância no processo de previsão e oferecendo suporte para decisões de projeto e operação de unidades de sedimentação.

(vi) De forma geral, os resultados indicam que abordagens híbridas, que combinam conhecimento físico consolidado com técnicas modernas de aprendizado de máquina, representam uma direção robusta e sustentável para o aprimoramento contínuo dos sistemas de tratamento de água.

8.1 RECOMENDAÇÕES DE ESTUDOS FUTUROS

A continuidade deste estudo pode seguir diferentes direções promissoras. Primeiramente, recomenda-se realizar uma comparação sistemática entre os modelos baseados em aprendizado de máquina e os resultados obtidos por meio da Lei de Stokes, tanto em sua forma clássica quanto em versões adaptadas que incorporam propriedades fractais dos agregados.

Outra possibilidade consiste em expandir o conjunto de variáveis explicativas do modelo, integrando parâmetros físico-químicos, como o tipo e a dosagem de coagulantes, pH e condutividade. A inclusão dessas variáveis pode oferecer uma compreensão mais ampla dos fatores que influenciam a sedimentação, superando as limitações do enfoque puramente geométrico e físico.

Adicionalmente, recomenda-se aplicar a metodologia proposta em diferentes contextos operacionais, contemplando variações nos mecanismos de coagulação e floculação. A análise da performance dos modelos em cenários diversos permitirá avaliar com mais rigor sua capacidade de generalização, considerando que o presente estudo foi conduzido em um arranjo experimental específico que, embora represente a dinâmica fundamental do fenômeno de sedimentação, pode não contemplar integralmente as complexidades hidrodinâmicas e físico-químicas presentes em sistemas reais de tratamento de água.

Visando ao desenvolvimento de soluções ainda mais interpretáveis e robustas, sugere-se a investigação de outras arquiteturas de aprendizado informado por física, como redes neurais com restrições físicas explícitas (Physics-Informed Neural Networks – PINNs) e modelos simbólicos baseados em equações

diferenciais. Essas abordagens podem contribuir para o avanço de modelos capazes de aliar desempenho preditivo e consistência física.

Considerando a natureza dinâmica do fenômeno da sedimentação, destaca-se ainda o potencial de abordagens baseadas em séries temporais e modelos de ordem reduzida, especialmente diante da possibilidade de utilização do arranjo experimental e das imagens já coletadas para a extração de informações temporais relacionadas à morfologia dos agregados e ao coeficiente de arrasto.

Por fim, a aplicação em escala real pode demandar ajustes metodológicos adicionais, uma vez que a velocidade necessária à geração de atributos físicos não é diretamente observável a priori, configurando-se como uma grandeza implícita no processo de modelagem. Nesse contexto, o uso de previsões intermediárias como entrada para etapas subsequentes de aprendizado, abordagem alinhada aos métodos de stacked generalization (WOLPERT, 1992; BREIMAN, 1996), desponta como uma linha promissora para investigações futuras, ao viabilizar a estimação de grandezas necessárias à construção de descritores fisicamente fundamentados.

REFERÊNCIAS BIBLIOGRÁFICAS

BÜRGER, R.; CAREAGA, J.; DIEHL, S.; PINEDA, R. A model of reactive settling of activated sludge: Comparison with experimental data. *Chemical Engineering Science*, v.267, p.118244, 2023. <https://doi.org/10.1016/j.ces.2023.118244>

TOLEDO-ZUCCO, J.P.; SBARBARO, D.; SILVA JR, J.M.G.D. Turbidity Control in Sedimentation Columns by Direction Dependent Models. *arXiv preprint arXiv:2305.18605*, 2023. <https://arxiv.org/abs/2305.18605>

CHEN, Q.; FU, S. Quantitative analysis and management of sustainable development of ecological water resources and digital financial system based on an intelligent algorithm. *Water Supply*, v.23, n.7, p.2881–2898, 2023. <https://doi.org/10.2166/ws.2023.159>

ADAM, A.M.; LUKASHEVICH, O. Efficiency of the best available techniques in water utilization on water management site of the river Tom. *Bulletin of the Tomsk State University of Architecture and Civil Engineering*, v.25, n.5, p.120–131, 2023.

MORUZZI, R.B.; BRIDGEMAN, J.; SILVA, P.A.G. A combined experimental and numerical approach to the assessment of floc settling velocity using fractal geometry. *Water Science & Technology*, v.81, n.5, p.915–924, 2020. <https://doi.org/10.2166/wst.2020.055>

ASENSI, E.; ALEMANY, E. A hindered settling velocity model related to the fractal dimension and activated sludge flocs characteristics: Application to a sludge with a previous fragmentation and flocculation process. *Separation and Purification Technology*, v.300, p.121812, 2022. <https://doi.org/10.1016/j.seppur.2022.121812>

VAHEDI, A.; GORCZYCA, B. Settling velocities of multifractal flocs formed in chemical coagulation process. *Water Research*, v.53, p.322–328, 2014. <https://doi.org/10.1016/j.watres.2014.01.012>

CHAKRABORTI, R.K.; KAUR, J. Noninvasive measurement of particle-settling velocity and comparison with Stokes' Law. *Journal of Environmental Engineering*, v.140, n.2, p.04013008, 2014. [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0000819](https://doi.org/10.1061/(ASCE)EE.1943-7870.0000819)

BRESSANE, A.; MELO, C.P.; SHARIFI, S.; SILVA, P.G.; RUFINO TODA, D.H.; MORUZZI, R. Fuzzy machine learning predictions of settling velocity based on fractal aggregate physical features in water treatment. *Journal of Water Process Engineering*, v.67, p.106138, 2024. <https://doi.org/10.1016/j.jwpe.2024.106138>

FERDOWSI, A.; VALIKHAN-ANARAKI, M.; FARZIN, S.; MOUSAVI, S.-F. A new combination approach for optimal design of sedimentation tanks based on hydrodynamic simulation model and machine learning algorithms. *Physics and Chemistry of the Earth*, v.127, p.103201, 2022. <https://doi.org/10.1016/j.pce.2022.103201>

DELCEY, M.; CHENY, Y.; KECK, J.B.; GANS, A.; DE RICHTER, S.K. Identification of Settling Velocity with Physics-Informed Neural Networks for Sediment-Laden Flows. *arXiv preprint arXiv:2404.06802*, 2024. <https://arxiv.org/abs/2404.06802>

BAKIRI, Z.; NACEF, S. Estimation of solids distribution and settling velocity of solid particles in secondary clarifiers: Large-scale measurements and numerical modeling. *Desalination and Water Treatment*, v.278, p.273–282, 2022. <https://doi.org/10.5004/dwt.2022.28666>

JIN, Y.; ZABARAS, N.; KOUTSOURELAKIS, P.-S.; PERDIKARIS, P. Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data. *Journal of Computational Physics*, v.394, p.56–81, 2019. <https://doi.org/10.1016/j.jcp.2019.05.016>

GREGORY, J.; BARANY, S. Adsorption and flocculation by polymers and polymer mixtures. *Advances in Colloid and Interface Science*, v.169, n.1, p.1–12, 2011. <https://doi.org/10.1016/j.cis.2011.08.004>

KIM, D.H.; YOO, H.J.; BANG, Y.J.; LEE, S.O. Quantification of floc growth for sediment with mixing intensity. *Sustainability*, v.15, n.5, p.4073, 2023. <https://doi.org/10.3390/su15054073>

ALALM, A.; TAWFIK, A.; OOKAWARA, A. Enhancing wastewater treatment and reuse with forward osmosis system: Feasibility and performance assessment. *Desalination and Water Treatment*, v.57, n.24, p.11363–11374, 2016. <https://doi.org/10.1080/19443994.2015.1042063>

RAISSI, M.; PERDIKARIS, P.; KARNIADAKIS, G.E. Machine learning of linear differential equations using Gaussian processes. *Journal of Computational Physics*, v.348, p.683–693, 2017. <https://doi.org/10.1016/j.jcp.2017.07.050>

KARNIADAKIS, G.E.; KEVREKIDIS, I.G.; LU, L.; PERDIKARIS, P.; WANG, S.; YANG, L. Physics-informed machine learning. *Nature Reviews Physics*, v.3, n.6, p.422–440, 2021. <https://doi.org/10.1038/s42254-021-00314-5>

LU, L.; MENG, X.; MAO, Z.; KARNIADAKIS, G.E. DeepXDE: A deep learning library for solving differential equations. *SIAM Review*, v.63, n.1, p.208–228, 2021. <https://doi.org/10.1137/19M1274067>

RAISSI, M.; PERDIKARIS, P.; KARNIADAKIS, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, v.378, p.686–707, 2019. <https://doi.org/10.1016/j.jcp.2018.10.045>

ZHU, Y.; ZABARAS, N.; KOUTSOURELAKIS, P.-S.; PERDIKARIS, P. Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data. *Journal of Computational Physics*, v.394, p.56–81, 2019. <https://doi.org/10.1016/j.jcp.2019.05.016>

WANG, S.; YU, H.; PERDIKARIS, P. When and why PINNs fail to train: A neural tangent kernel perspective. *Journal of Computational Physics*, v.449, p.110768, 2021. <https://doi.org/10.1016/j.jcp.2021.110768>

JOHNSON, C.P.; LI, X.; LOGAN, B.E. Settling velocities of fractal aggregates. *Environmental Science & Technology*, v.30, n.6, p.1911–1918, 1996. <https://doi.org/10.1021/es950731a>

RICHARDSON, J.F.; ZAKI, W.N. Sedimentation and fluidisation: Part I. *Chemical Engineering Science*, v.3, n.2, p.65–73, 1954. [https://doi.org/10.1016/0009-2509\(54\)85015-9](https://doi.org/10.1016/0009-2509(54)85015-9)

YUKSELEN, Z.; GREGORY, J. The effect of pH and ionic strength on the zeta potential of kaolinite. *Journal of Colloid and Interface Science*, v.278, n.2, p.497–502, 2004. <https://doi.org/10.1016/j.jcis.2004.06.042>

LUO, X.; ZHANG, Y.; LI, W. Impact of shipping activities on river water quality: A case study in the Yangtze River. *Environmental Pollution*, v.306, p.119402, 2022. <https://doi.org/10.1016/j.envpol.2022.119402>

MORUZZI, R.B.; CAMPOS, L.C.; SHARIFI, S.; GREGORY, J. Nonintrusive investigation of large Al-kaolin fractal aggregates with slow settling velocities. *Water Research*, v.185, p.116287, 2017. <https://doi.org/10.1016/j.watres.2020.116287>

BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, v.13, n.2, p.281–305, 2012. <http://jmlr.org/papers/v13/bergstra12a.html>

LUNDBERG, S.M.; LEE, S.I. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, v.30, 2017. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

NARAYANAN, A.; KAPOOR, S. Why an overreliance on AI-driven modelling is bad for science. *Nature*, v.640, n.8058, p.312–314, 2025. <https://doi.org/10.1038/d41586-025-00915-3>

KARNIADAKIS, G.E.; KEVREKIDIS, I.G.; LU, L.; PERDIKARIS, P.; WANG, S.; YANG, L. *Physics-informed machine learning*. Nature Reviews Physics, v.3, n.6, p.422–440, 2021. <https://doi.org/10.1038/s42254-021-00314-5>

RAISSI, M.; PERDIKARIS, P.; KARNIADAKIS, G.E. *Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*. Journal of Computational Physics, v.378, p.686–707, 2019. <https://doi.org/10.1016/j.jcp.2018.10.045>

WOLPERT, D. H. Stacked generalization. *Neural Networks*, v. 5, n. 2, p. 241–259, 1992. [https://doi.org/10.1016/S0893-6080\(05\)80023-1](https://doi.org/10.1016/S0893-6080(05)80023-1)

BREIMAN, L. *Stacked regressions*. Machine Learning, v. 24, n. 1, p. 49–64, 1996.
DOI: <https://doi.org/10.1023/A:1018046112532>

APÊNDICES

Os códigos-fonte, conjuntos de dados e instruções necessárias para a reprodução dos resultados desta dissertação estão disponíveis no repositório público: <https://github.com/daniel todaDS/terminal-velocity-prediction>.