

Análise de séries temporais para detecção de anomalias em aerogeradores

Gabriel Zapelini Barcellos¹ | gabriel.zb07@aluno.ifsc.edu.br
Sergio luciano avila² | sergio.avila@ifsc.edu.br
Marina Geremia³ | marina.geremia@ifsc.edu.br
João Victor Rochinski Vieira⁴ | joao.vrv2000@aluno.ifsc.edu.br
Eduardo Raimundo Rodycz⁵ | eduardo.r2003@aluno.ifsc.edu.br
Bruno Panerai Velloso⁶ | bruno.velloso@ifsc.edu.br
Glauco Cardozo⁷ | glauco.cardozo@ifsc.edu.br
Raiza Silveira Medina⁸ | raiza.sm1991@aluno.ifsc.edu.br

RESUMO

Este resumo descreve a validação de um método de detecção de anomalias em aerogeradores, utilizando um script Python aplicado a uma base de dados SCADA real e pública de 36 turbinas. O objetivo foi testar a eficácia do algoritmo na identificação de falhas de alta temperatura no Parque Eólico B (IDs 7 e 19). O algoritmo funciona com processos de ETL e decomposição da sazonalidade. A metodologia central envolve a aplicação de regressão linear diária para determinar a inclinação (slope), que é convertida em um ângulo geométrico. O método rastreia a variância acumulada desses ângulos para gerar um threshold adaptativo (com Window 14 e K 2.9) que define o limiar de anomalia. Os resultados demonstram que o método identificou anomalias consistentes e precoces em múltiplos sensores de temperatura (transformadores e mancais) durante os períodos críticos. A validação bem-sucedida em dados anonimizados confirma a aplicabilidade da metodologia para monitoramento da manutenção em ambientes operacionais reais.

Palavras-chave: detecção de anomalias; manutenção; aerogeradores;

1 INTRODUÇÃO

Este resumo tem como finalidade descrever e apresentar o processo de validação dos métodos de detecção de anomalias em aerogeradores. O método central proposto baseia-se na extração de características de tendência a partir da série temporal de sensores, convertendo a inclinação diária (slope) em um ângulo para rastrear a variância acumulada, permitindo a identificação precoce de desvios operacionais. Para a validação, foram utilizados dados SCADA públicos de aerogeradores, disponíveis nas plataformas Zenodo e Kaggle e referenciados no artigo 'CARE to Compare: A Real-World Benchmark Dataset for Early Fault Detection in Wind Turbine Data'. O conjunto de dados é composto por informações operacionais reais de turbinas eólicas coletadas a cada 10 minutos do Parque Eólico B. É importante notar que o dataset passou por um processo de anonimização para proteger informações confidenciais, o que incluiu mascarar IDs, anonimizar nomes de sensores e deslocar intencionalmente os carimbos



de data e hora. O arquivo `event_info.csv` serve como gabarito para classificação dos eventos (anomaly ou normal) e descrição da causa raiz

2 METODOLOGIA

Foi desenvolvido um notebook de validação no Google Colab com o objetivo de testar o método de detecção de anomalias. A análise se concentrou em eventos do Parque Eólico B (IDs 7 e 19), pois as falhas estavam relacionadas à alta temperatura, e os dados de temperatura apresentavam maior afinidade com os algoritmos propostos. No script foram utilizados os parâmetros Window 14 e K (grau de sensibilidade) 2.9.

2.1 O Algoritmo de Detecção de Anomalias

O algoritmo é estruturado em etapas que transformam a série temporal do sensor para extrair características de tendência e variabilidade:

2.2 Processos de ETL: Primeiramente, são executados processos de ETL, incluindo leitura dos dados, conversão de colunas, pivotagem e ordenação, preparando a base para a análise.

2.3 Decomposição da Sazonalidade: Calcula-se a média dos 144 pontos diários (um sinal a cada 10 minutos). Essa média é então subtraída do valor original da série temporal para remover a sazonalidade, separando a série em tendência, sazonalidade e ruído.

2.4 Cálculo da Tendência (Slope): É aplicada uma regressão linear calculada para cada bloco de um dia, gerando uma reta que representa a **tendência** dos dados, já sem a sazonalidade. O resultado desse processo é a **inclinação (slope)**, que representa o dia específico e pode ser neutra, positiva ou negativa.

2.5 Conversão para Ângulo: O *slope* da regressão linear é convertido em um **ângulo geométrico** utilizando a função **arcotangente**.

2.6 Cálculo da Variância Acumulada: O ângulo de cada dia é guardado em uma lista. A cada novo dia (N), calcula-se a **variância acumulada** dos ângulos, considerando todos os ângulos desde o início até o dia N. Isso é feito calculando a média do período e aplicando a fórmula da variância amostral.

2.7 Geração do Threshold Adaptativo: Utiliza-se uma **janela deslizante** de dias (definida como 14 no *notebook*) para calcular a média e o desvio padrão sobre a variância



acumulada daquele período (*window*). O grau de sensibilidade ($K = 2.9$) é usado como fator de corte para definir o limiar (*threshold*):

$$\text{Limite do threshold} = \text{média} + (K \times \text{desvio padrão})$$

O período de análise utilizado foi de aproximadamente um ano para cada série temporal. Os resultados gerados pelo notebook incluíram gráficos temporais e uma tabela que filtra as anomalias detectadas no período de interesse. O foco da análise permaneceu nos casos da data da anomalia designada.

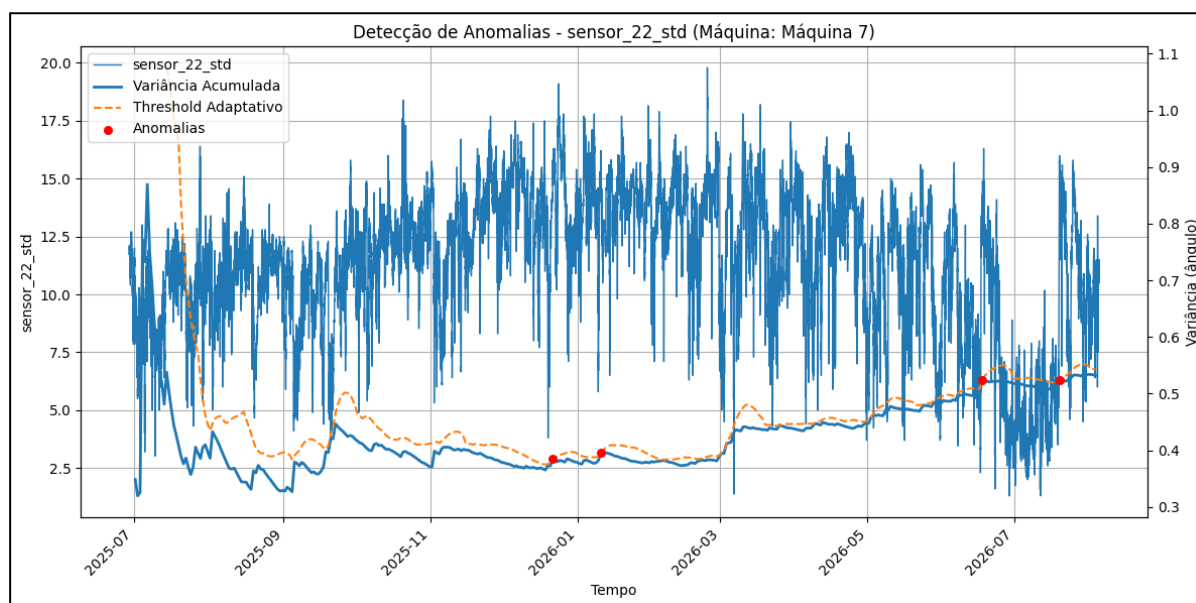
3 RESULTADOS

O objetivo da validação era encontrar os registros de anomalia relacionados a problemas de temperatura na Fazenda B, conforme consta no arquivo `event_info`. Abaixo, são apresentadas as descobertas para os dois eventos analisados, focando nos sensores alvo de temperatura.

3.1 Análise 1 – Evento ID 7

O período de anomalia designado foi de 30/06/2026 a 31/07/2026, com datas em 2026 devido ao deslocamento temporal da anonimização. Período da Anomalia: O sensor_22 (Hub temperature) registrou anomalias em 19/07/2026.

Figura 1 – Detecção de anomalia evento ID 7



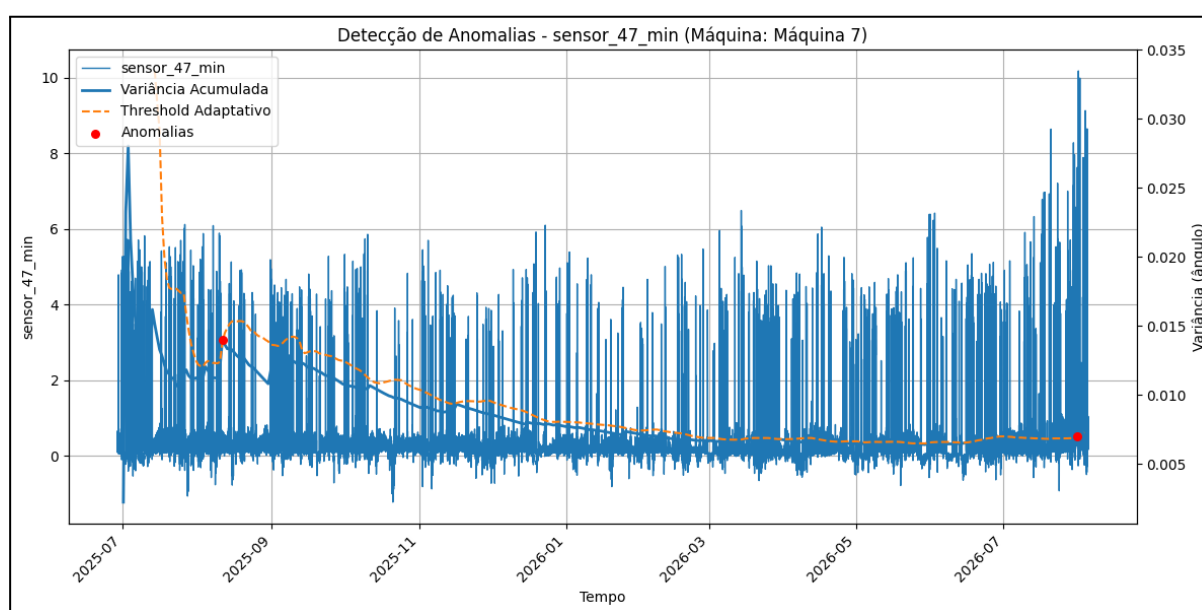


Fonte: Google colab do autor

3.2 Análise 2 - Evento ID 19

O período de anomalia designado foi de 02/02/2025 a 22/02/2025. Período da Anomalia: O sensor_47_min continuou a apresentar anomalias em 11/02/2025 e 12/02/2025.

Figura 2 – Detecção de anomalia evento ID 19



Fonte: Google colab do autor

4 CONCLUSÃO

O processo de validação dos métodos de detecção de anomalias, aplicado aos dados reais de temperatura do Parque Eólico B, demonstrou a eficácia do script proposto. O método conseguiu identificar precocemente picos de temperatura em diversos componentes críticos, como transformadores, mancais e hubs, nos períodos que antecederam as datas designadas das anomalias nos eventos 7 e 19. A validação bem-sucedida reforça a utilidade da metodologia para a implementação de estratégias de manutenção preditiva em parques eólicos.

REFERÊNCIAS

GÜCK, Christian; ROELOFS, Cyriana M.A.; FAULSTICH, Stefan. CARE to Compare: A real-world dataset for anomaly detection in wind turbine data. *arXiv:2404.10320v2*, 2024.