ReGeo ISSN: 2177-3246

RANDOM FOREST E XAI NA MODELAGEM MULTIESCALAR DA ATIVIDADE DO FOGO E FATORES METEOROLÓGICOS

RANDOM FOREST AND CAI IN MULTI-SCALE MODELING OF FIRE ACTIVITY AND METEOROLOGICAL FACTORS

BOSQUE ALEATORIO Y CAI EN LA MODELACIÓN MULTIESCALA DE LA ACTIVIDAD DE INCENDIOS Y LOS FACTORES METEOROLÓGICOS

https://doi.org/10.56238/revgeov17n10-130

Data de submissão: 19/09/2026 **Data de publicação:** 19/10/2026

Regivaldo Ferreira
Mestre
Instituição: Universidade Estadual de Campinas (Unicamp)
E-mail: r225153@dac.unicamp.br

Rogério Gomes Matias
Mestre
Instituição: Universidade Estadual de Feira de Santana (UEFS)
E-mail: rgmatias@uefs.br

Luiz Roberto Albano Junior
Especialista
Instituição: Universidade Estadual de Campinas (Unicamp)
E-mail: l272746@dac.unicamp.br

Takaaki Ohishi
Doutor
Instituição: Universidade Estadual de Campinas (Unicamp)
E-mail: takaaki@unicamp.br

Gilney Figueira Zebende
Doutor
Instituição: Universidade Estadual de Feira de Santana (UEFS)
E-mail: gfzebende@uefs.br

RESUMO

A resolução temporal pode modificar tanto os padrões aprendidos por modelos de aprendizado de máquina quanto a interpretação das contribuições ambientais atribuídas por esses modelos. Neste estudo, utilizamos Random Forest como estrutura fixa de modelagem e SHAP como método de inteligência artificial explicável para investigar como a agregação anual, mensal e semanal modifica a representação da relação entre condições meteorológicas e atividade do fogo. A análise combinou detecções de fogo ativo e sete variáveis meteorológicas provenientes de 28 estações em Mato Grosso do Sul, Brasil, no período de 2005–2024, e os modelos foram avaliados por validação cruzada de séries temporais com janela expansiva. O R^2 médio na validação cruzada foi de −0.2743, 0.1646 e 0.3245 para os modelos anual, mensal e semanal, respectivamente; a agregação semanal apresentou o melhor desempenho. A umidade relativa foi o preditor dominante em todas as resoluções, enquanto as contribuições da precipitação, da velocidade do vento, da temperatura do ponto de orvalho e da temperatura do ar variaram entre as escalas. Os padrões SHAP semanais revelaram associações não

ISSN: 2177-3246

lineares e heterogêneas. A agregação temporal constitui, portanto, uma escolha estrutural de modelagem que modifica tanto a capacidade de generalização quanto a estrutura de atribuição aprendida pelo modelo, devendo ser explicitamente avaliada em modelos ambientais de incêndios.

Palavras-chave: Detecção de Fogo Ativo. Monitoramento do Fogo. Fatores Meteorológicos. Agregação Temporal. Random Forest. Inteligência Artificial Explicável SHAP.

ABSTRACT
Temporal resolution can modify both the patterns learned by machine learning models and the interpretation of environmental contributions attributed by these models. In this study, we used Random Forest as a fixed modeling structure and SHAP as an explainable artificial intelligence method to investigate how annual, monthly, and weekly aggregation modifies the representation of the relationship between meteorological conditions and fire activity. The analysis combined active fire detections and seven meteorological variables from 28 stations in Mato Grosso do Sul, Brazil, from 2005 to 2024, and the models were evaluated by cross-validation of time series with an expanding window. The average R^2 in the cross-validation was -0.2743, 0.1646, and 0.3245 for the annual, monthly, and weekly models, respectively; weekly aggregation showed the best performance. Relative humidity was the dominant predictor at all resolutions, while the contributions of precipitation, wind speed, dew point temperature, and air temperature varied across scales. Weekly SHAP patterns revealed non-linear and heterogeneous associations. Temporal aggregation, therefore, constitutes a structural modeling choice that modifies both the generalization capacity and the learned attribution structure of the model, and should be explicitly evaluated in environmental fire models.

Keywords: Active Fire Detection. Fire Monitoring. Meteorological Factors. Temporal Aggregation. Random Forest. SHAP Explainable Artificial Intelligence.

RESUMEN
La resolución temporal puede modificar tanto los patrones aprendidos por los modelos de aprendizaje automático como la interpretación de las contribuciones ambientales atribuidas por estos modelos. En este estudio, utilizamos Random Forest como una estructura de modelado fija y SHAP como un método de inteligencia artificial explicable para investigar cómo la agregación anual, mensual y semanal modifica la representación de la relación entre las condiciones meteorológicas y la actividad de incendios. El análisis combinó detecciones de incendios activos y siete variables meteorológicas de 28 estaciones en Mato Grosso do Sul, Brasil, de 2005 a 2024, y los modelos se evaluaron mediante validación cruzada de series temporales con una ventana expansiva. El R^2 promedio en la validación cruzada fue de -0,2743, 0,1646 y 0,3245 para los modelos anual, mensual y semanal, respectivamente; la agregación semanal mostró el mejor rendimiento. La humedad relativa fue el predictor dominante en todas las resoluciones, mientras que las contribuciones de precipitación, velocidad del viento, temperatura del punto de rocío y temperatura del aire variaron a través de las escalas. Los patrones semanales de SHAP revelaron asociaciones no lineales y heterogéneas. La agregación temporal, por lo tanto, constituye una opción de modelado estructural que modifica tanto la capacidad de generalización como la estructura de atribución aprendida del modelo, y debe evaluarse explícitamente en los modelos ambientales de incendios.

ISSN: 2177-3246

Palabras clave: Detección Activa de Incendios. Monitoreo de Incendios. Factores Meteorológicos. Agregación Temporal. Bosque Aleatorio. SHAP (Inteligencia Artificial Explicable).

ISSN: 2177-3246

1 INTRODUÇÃO

Os incêndios em vegetação constituem uma perturbação ecológica com consequências para a estrutura e o funcionamento dos ecossistemas, a saúde pública, a produção agrícola e as emissões de carbono (SAYEDI et al., 2024; WALKER et al., 2025; CHEN et al., 2024; SOBREIRA et al., 2025). A atividade do fogo emerge em um sistema acoplado no qual as condições atmosféricas interagem com a vegetação e as características dos combustíveis, as práticas de manejo da terra, as fontes de ignição e outras atividades humanas. No Brasil, períodos de baixa umidade e altas temperaturas ocorrem simultaneamente ao uso do fogo no manejo da terra e estão associados à maior suscetibilidade a incêndios nos diferentes biomas (ARRUDA et al., 2024; OLIVEIRA et al., 2022; SILVEIRA et al., 2022). Este estudo concentra-se especificamente no componente meteorológico desse sistema mais amplo, examinando como as condições atmosféricas contemporâneas contribuem para a representação, baseada em dados, das contagens de detecções de fogo ativo. As sete variáveis meteorológicas consideradas são, portanto, tratadas como descritores preditivos parciais do sistema, e não como uma representação exaustiva dos processos de fogo ou como evidências de efeitos causais.

A agregação temporal é uma decisão de modelagem relevante nesse contexto, pois representar as mesmas observações subjacentes em intervalos semanais, mensais ou anuais pode alterar a variabilidade e a distribuição dos dados, assim como o número de observações disponíveis para o ajuste e a avaliação dos modelos. Intervalos mais curtos podem preservar episódios de secura atmosférica, calor, vento e atividade do fogo que são progressivamente suavizados em agregações mais amplas, enquanto intervalos mais longos enfatizam padrões sazonais ou interanuais mais abrangentes. Essas mudanças podem influenciar não apenas os padrões preditivos que um modelo baseado em dados consegue aprender, mas também a distribuição das contribuições preditivas entre as variáveis meteorológicas.

A resolução temporal deve, portanto, ser tratada como um componente estrutural da especificação do modelo, e não apenas como uma escolha de pré-processamento. Ao definir a unidade de observação, ela determina a amostra disponível para treinamento e validação e a variação apresentada ao modelo. Consequentemente, modelos ajustados em diferentes resoluções temporais representam formulações distintas do sistema meteorologia–fogo, mesmo quando a região de estudo, os preditores e o algoritmo de modelagem permanecem inalterados. Abordagens de aprendizado de máquina são cada vez mais utilizadas na avaliação de incêndios e no mapeamento de suscetibilidade (MOGHIM; MEHRABI, 2024; MASOUDIAN; MIRZAEI; BAGHERI, 2025), enquanto pesquisas correlatas baseadas em dados incluem recursos globais para modelagem de incêndios (KARASANTE et al., 2025) e análises da recuperação da vegetação após o fogo (ZAHURA et al., 2024).

Comparativamente, pouca atenção tem sido dedicada a verificar se diferentes representações temporais das observações meteorológicas e de fogo conduzem a desempenho preditivo e interpretação dos modelos consistentes entre escalas. Assim, ainda não está claro quais padrões de atribuição meteorológica permanecem consistentes entre resoluções temporais e quais dependem da escala.

Do ponto de vista da gestão ambiental, a resolução temporal não é apenas uma escolha estatística. Informações anuais podem apoiar o planejamento estratégico e a alocação de recursos em longo prazo, enquanto informações mensais e semanais podem ser mais úteis para a preparação sazonal, a priorização do monitoramento e a organização das atividades de prevenção de incêndios. Determinar como a agregação temporal afeta tanto o desempenho preditivo quanto a interpretação dos fatores meteorológicos é, portanto, importante para selecionar escalas temporais compatíveis com diferentes decisões de gestão. Embora a análise se concentre em Mato Grosso do Sul, a comparação também é relevante para outras regiões propensas a incêndios, caracterizadas por forte sazonalidade climática e ecossistemas heterogêneos.

1.1 RANDOM FOREST

O algoritmo Random Forest (RF) combina árvores de decisão e representa relações não lineares sem impor uma forma funcional aos dados (BREIMAN, 2001). O método tem sido aplicado à previsão de incêndios, de áreas queimadas e da suscetibilidade ao fogo, utilizando variáveis climáticas, ambientais e de sensoriamento remoto (FREITAS et al., 2025; MIRANDA et al., 2025; MUHAMMAD et al., 2025). Estudos recentes incorporam informações históricas e índices de seca para identificar detecções de fogo ativo ou antecipar anomalias sazonais na área queimada (CHEERALA et al., 2025; PÁDUA; LEVADA; VALENTE, 2026; TORRES-VÁZQUEZ et al., 2025). Essas propriedades tornam o RF adequado para examinar relações entre meteorologia e fogo sob uma estrutura comum de modelagem em diferentes resoluções temporais, permitindo avaliar mudanças no desempenho preditivo sem impor formas funcionais distintas em cada escala.

1.2 SHAP

Embora seja capaz de representar padrões complexos, o RF apresenta interpretabilidade limitada, o que motiva o uso de métodos de inteligência artificial explicável (XAI). O método SHapley Additive exPlanations (SHAP) utiliza valores de Shapley para quantificar a contribuição individual das variáveis para cada predição (LUNDBERG; LEE, 2017). Aplicações a incêndios mostram que essa abordagem permite identificar a importância dos preditores, a direção de suas contribuições e os padrões não lineares aprendidos pelo modelo, além de comparar fatores de influência entre diferentes

ISSN: 2177-3246

regiões (ANEES et al., 2025; LI et al., 2024; SENGUPTA; WOODFORD, 2025; TEKE; KAVZOGLU, 2025; TONBUL; VERAVERBEKE, 2026). Em uma estrutura com múltiplas resoluções, o SHAP pode, portanto, ser empregado não apenas para ordenar os preditores por importância, mas também para examinar se a estrutura de atribuição aprendida pelo modelo permanece robusta ou se modifica com a agregação temporal.

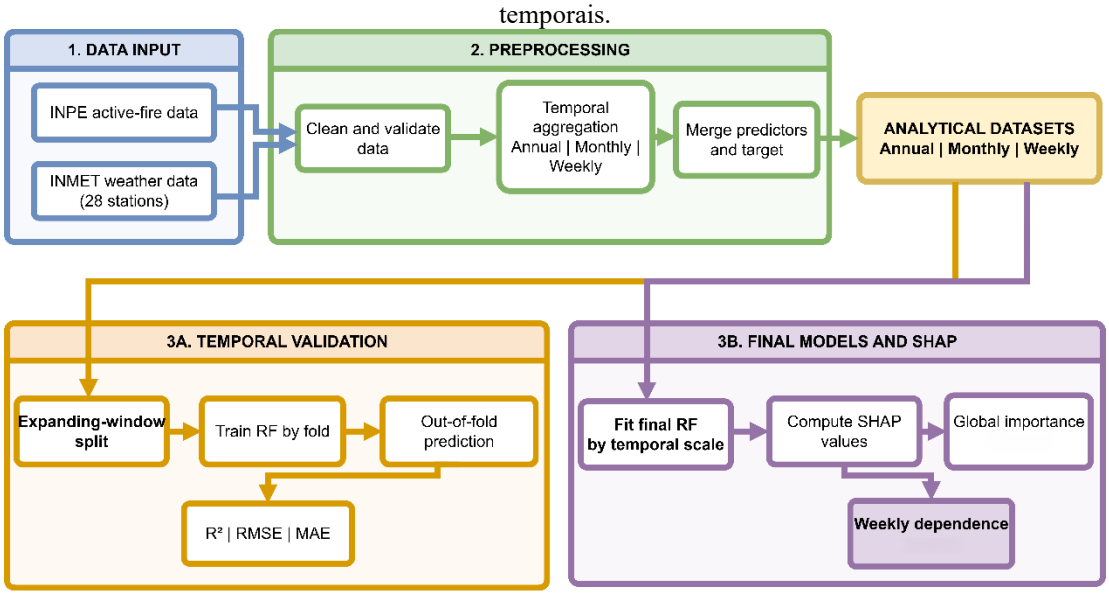
Este estudo utiliza observações de fogo ativo de Mato Grosso do Sul entre 2005 e 2024 como caso de aplicação de um procedimento metodológico desenvolvido para avaliar a sensibilidade do desempenho preditivo e da interpretação dos modelos à escala temporal. O objetivo deste estudo é investigar como diferentes resoluções temporais modificam o comportamento do Random Forest e a estrutura de atribuição dos preditores meteorológicos revelada pelo SHAP. São abordadas três questões de pesquisa: (1) como o desempenho preditivo fora da amostra se modifica quando as contagens de detecções de fogo ativo e as variáveis meteorológicas são agregadas nas escalas anual, mensal e semanal; (2) se a importância relativa e a direção das contribuições dos preditores meteorológicos permanecem consistentes entre essas escalas temporais; e (3) quais relações não lineares caracterizam os preditores na configuração temporal com melhor desempenho preditivo. Em conjunto, essas questões abordam um problema mais amplo de modelagem: se as conclusões obtidas a partir de modelos ambientais baseados em dados permanecem estáveis quando se altera a representação temporal de um mesmo sistema subjacente.

Além da aplicação regional, a contribuição metodológica deste estudo consiste em uma estrutura de avaliação transferível, com múltiplas resoluções, para examinar como a agregação temporal afeta o desempenho preditivo e a interpretação de modelos ambientais baseados em dados. Essa estrutura combina agregação temporal consistente, validação cronológica fora da amostra e análise de atribuição sob uma configuração comum de modelagem. Embora se espere que a ordenação dos preditores e as resoluções temporais ótimas permaneçam específicas de cada sistema, a estratégia de avaliação proposta pode ser aplicada a outros processos ambientais nos quais a escala de observação possa influenciar o comportamento e a interpretação dos modelos. Neste trabalho, predição designa a estimação fora da amostra com condições meteorológicas contemporâneas, e não a previsão operacional com antecedência predefinida. Assim, a comparação não isola o efeito da resolução temporal nem quantifica diretamente a perda de informação, e as relações identificadas são interpretadas como associações preditivas, e não como efeitos causais ou limiares operacionais de risco de incêndio.

2 MATERIAIS E MÉTODOS

A análise foi organizada como uma estrutura de modelagem e interpretação em múltiplas resoluções, composta por cinco etapas gerais: (1) harmonização dos preditores ambientais e dos dados da variável resposta; (2) construção de representações temporais alternativas; (3) aplicação de uma especificação comum do modelo preditivo entre resoluções; (4) avaliação cronológica fora da amostra; e (5) comparação das estruturas de atribuição dos modelos entre escalas temporais. A Figura 1 resume essa estrutura e sua aplicação aos dados meteorológicos e de fogo.

Figura 1: Estrutura de modelagem e interpretação em múltiplas resoluções aplicada a dados meteorológicos e de fogo. A estrutura compreende a harmonização dos dados, a construção de representações temporais alternativas, a especificação comum dos modelos, a avaliação cronológica fora da amostra e a comparação das estruturas de atribuição entre escalas temporais.

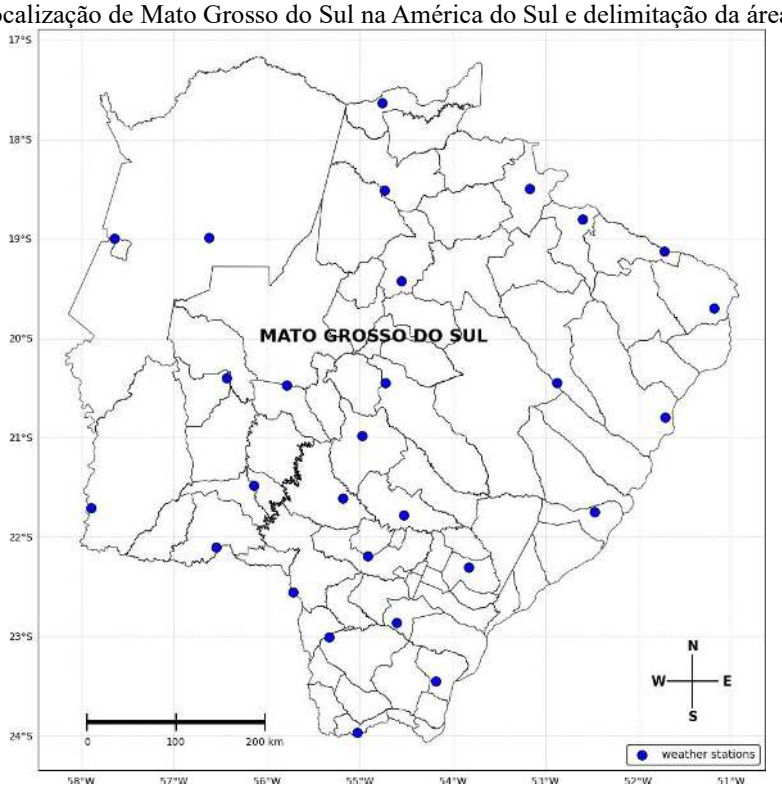


Fonte: Autores (2026).

2.1 ÁREA DE ESTUDO

Mato Grosso do Sul está localizado na região Centro-Oeste do Brasil e abrange porções de três importantes biomas brasileiros: Pantanal, Cerrado e Mata Atlântica. O clima regional é caracterizado por uma estação chuvosa concentrada principalmente entre outubro e março e uma estação seca que, em geral, se estende de abril a setembro. Essa sazonalidade acentuada afeta a umidade atmosférica, as condições da vegetação e a distribuição temporal da atividade do fogo (OLIVEIRA-JÚNIOR et al., 2021; SOUZA et al., 2022). A Figura 2 apresenta a localização e os limites da área de estudo.

Figura 2: Localização de Mato Grosso do Sul na América do Sul e delimitação da área de estudo.



Nota: As linhas do mapa delimitam áreas de estudo e não representam necessariamente fronteiras nacionais reconhecidas.
Fonte: Autores (2026).

2.2 FONTES DE DADOS E ABRANGÊNCIA TEMPORAL

As detecções georreferenciadas de fogo ativo de 2005–2024 foram obtidas no Programa Queimadas/BDQueimadas do Instituto Nacional de Pesquisas Espaciais (INPE) (Instituto Nacional de Pesquisas Espaciais, 2026a). Os recursos disponíveis on-line incluem o portal do Programa Queimadas do INPE e a interface de consulta e exportação de dados do BDQueimadas. Foi utilizada a série temporal do satélite de referência (ref) disponibilizada pelo INPE, destinada a apoiar comparações temporais das detecções de fogo ativo. Informações metodológicas sobre os produtos de fogo ativo e o conceito de satélite de referência estão disponíveis nas perguntas frequentes sobre a metodologia do INPE (Instituto Nacional de Pesquisas Espaciais, 2026b).

As observações meteorológicas foram obtidas em 28 estações meteorológicas automáticas operadas pelo Instituto Nacional de Meteorologia (INMET). Os dados históricos das estações automáticas foram obtidos no portal oficial do INMET (Instituto Nacional de Meteorologia, 2026).

O período geral da análise foi 2005–2024. A disponibilidade das estações variou ao longo do tempo; portanto, esse período não implica cobertura contínua de todas as 28 estações. Para cada intervalo e preditor, a média regional foi calculada com base nas observações válidas disponíveis.

2.3 VARIÁVEL RESPOSTA E VARIÁVEIS PREDITORAS

A variável resposta foi o número de detecções de fogo ativo registradas em cada intervalo de agregação temporal. As detecções de fogo ativo são observações de anomalias térmicas por satélite e não foram interpretadas como o número de incêndios individuais nem como uma medida direta da área queimada.

Foram considerados sete preditores meteorológicos: umidade relativa (%), velocidade do vento (m s−1), precipita-ção (mm), pressão atmosférica ao nível da estação (hPa), temperatura do ar de bulbo seco (◦C), radiação solar global (kJ m−2) e temperatura do ponto de orvalho (◦C). Foram preservadas as unidades de medida originais fornecidas pelo INMET.

2.4 PRÉ-PROCESSAMENTO DOS DADOS

Os registros de detecções de fogo ativo foram examinados quanto à presença de observações duplicadas. Não foram encontradas entradas duplicadas para os identificadores foco_id ou id_bdq. Um registro potencialmente duplicado foi identificado com base na combinação exata de latitude, longitude e data_pas no conjunto completo de dados de 2004–2024. Como a análise utilizou apenas a série temporal do satélite de referência, não foi necessário consolidar detecções de diferentes satélites como observações de um mesmo evento.

Os arquivos meteorológicos horários das 28 estações do INMET foram combinados em uma única base de dados. As vírgulas decimais foram convertidas em pontos decimais antes do processamento numérico. Valores registrados como null, campos vazios e valores que não puderam ser convertidos para formato numérico foram tratados como ausentes.

Os valores negativos de radiação solar global também foram tratados como ausentes, pois não representam radiação solar incidente fisicamente plausível para os propósitos desta análise. As demais variáveis foram verificadas com base em intervalos predefinidos de plausibilidade física. Não foram identificadas observações fora desses intervalos. A duplicidade de observações meteorológicas foi avaliada pela combinação de estação, data e hora, e não foram encontradas duplicatas.

2.5 AGREGAÇÃO TEMPORAL E INTEGRAÇÃO DOS DADOS

Os preditores meteorológicos e as detecções de fogo ativo foram agregados nas resoluções temporais anual, mensal e semanal. Os intervalos semanais foram definidos como semanas do calendário encerradas aos domingos. Os intervalos mensais e anuais seguiram, respectivamente, o mês e o ano do calendário.

Para cada intervalo temporal, calculou-se, independentemente para cada preditor

meteorológico, a média aritmética de todas as observações horárias válidas disponíveis nas 28 estações. Consequentemente, os preditores agregados representam as condições médias de todas as observações disponíveis das estações, e não medições de uma estação individual.

Como a agregação foi realizada diretamente sobre todos os registros horários válidos, as estações com maior número de observações válidas contribuíram proporcionalmente mais para as médias resultantes. O procedimento produziu, portanto, uma média regional ponderada pelas observações, e não uma média espacial com pesos iguais entre as estações.

A precipitação foi agregada pelo mesmo procedimento de média aritmética aplicado às demais variáveis mete-orológicas. Assim, o preditor de precipitação representa a média das observações de precipitação disponíveis em cada intervalo temporal, e não o aporte hídrico total acumulado nesse intervalo. Essa distinção foi considerada na interpretação dos resultados dos modelos.

As detecções de fogo ativo foram somadas em cada intervalo semanal, mensal ou anual. Em seguida, os preditores meteorológicos agregados foram associados às contagens de detecções de fogo na resolução temporal correspondente. Cada observação dos modelos continha, portanto, sete preditores meteorológicos e o número de detecções de fogo ativo registradas no mesmo intervalo.

Os intervalos sem informações suficientes para construir o conjunto completo de preditores foram excluídos após a integração temporal. Os conjuntos finais de dados analíticos continham 20 observações anuais, 240 mensais e 1.044 semanais.

2.6 MODELAGEM COM RANDOM FOREST

O Random Forest foi selecionado como instrumento fixo de modelagem, e não como candidato em um exercício de seleção de algoritmos. O objetivo deste estudo não foi determinar qual algoritmo de aprendizado de máquina oferece a maior acurácia preditiva, mas examinar se a alteração da representação temporal de um mesmo sistema ambiental modifica o desempenho preditivo e a atribuição dos modelos. Manter o algoritmo de aprendizado constante reduziu uma fonte adicional de variação metodológica e permitiu concentrar a análise na resolução temporal como principal fator experimental. O RF foi particularmente adequado para esse propósito por representar relações não lineares e interações sem exigir uma forma funcional predefinida, permanecendo compatível com a interpretação baseada em SHAP em todas as resoluções temporais avaliadas (BREIMAN, 2001; LUNDBERG; LEE, 2017).

Modelos de regressão RF separados foram ajustados aos conjuntos anuais, mensais e semanais, utilizando uma configuração deliberadamente comum. Todos os modelos utilizaram 200 árvores; o número máximo de preditores considerados em cada divisão foi definido como $\sqrt{p}$, em que $p = 7$ era

ISSN: 2177-3246

o número total de preditores; e o número mínimo de observações por folha terminal foi definido como um. Não foi realizada otimização de hiperparâmetros específica para cada resolução, pois o objetivo era comparar o comportamento dos modelos sob uma especificação comum, e não otimizar o desempenho preditivo de forma independente em cada escala temporal. Caso contrário, o ajuste específico por resolução poderia confundir diferenças associadas à representação temporal com diferenças introduzidas pela otimização dos modelos.

2.7 VALIDAÇÃO CRONOLÓGICA E MÉTRICAS DE DESEMPENHO

O desempenho preditivo foi avaliado por validação cruzada de séries temporais com janela expansiva, implemen-tada com a rotina de divisão de séries temporais do scikit-learn. Foram utilizadas três partições para o conjunto anual e cinco para os conjuntos mensal e semanal. Em cada partição, o modelo foi treinado exclusivamente com observações anteriores ao período de teste correspondente. Esse procedimento preservou a ordem cronológica dos dados e impediu que observações futuras fossem utilizadas para predizer períodos anteriores.

As predições geradas para a parte de teste de cada partição foram mantidas como predições fora da amostra de treinamento. O desempenho dos modelos foi avaliado pelo coeficiente de determinação ($R2$), pela raiz do erro quadrático médio (RMSE) e pelo erro absoluto médio (MAE). A média e o desvio-padrão de cada métrica foram calculados entre as partições de validação.

Como a variável resposta consistia em contagens agregadas em intervalos de durações distintas, RMSE e MAE foram interpretados apenas dentro de cada resolução temporal. Seus valores absolutos não foram utilizados para comparações diretas entre os conjuntos anual, mensal e semanal. As comparações entre resoluções basearam-se principalmente no $R2$, considerando também a variabilidade de todas as métricas entre as partições.

Neste estudo, predição refere-se à estimação fora da amostra sob validação cronológica. Os modelos não utilizaram preditores meteorológicos temporalmente defasados e, portanto, não foram avaliados como previsões operacionais com antecedência predefinida.

2.8 INTERPRETAÇÃO DOS MODELOS BASEADA EM SHAP

Após a conclusão da validação cronológica, um modelo RF final foi ajustado ao conjunto completo de dados disponível em cada resolução temporal. Esses modelos finais foram utilizados exclusivamente para interpretação. A capacidade de generalização preditiva foi avaliada pelos resultados fora das amostras de treinamento, enquanto a interpretação foi realizada com os modelos ajustados a todas as observações disponíveis.

Para cada escala temporal, as explicações SHAP foram calculadas para o modelo Random Forest final ajustado ao conjunto completo de dados. O valor de referência do SHAP, $E[f\ (X)]$, representa a saída esperada do modelo antes de considerar as características de uma observação específica e equivale aproximadamente à predição média na distribuição de referência. No Random Forest, essa saída esperada é derivada dos valores armazenados nas folhas terminais das árvores, ponderados pela distribuição das observações de treinamento ao longo de seus caminhos de decisão e, posteriormente, promediados entre as 200 árvores. A predição para a observação i pode, portanto, ser decomposta como:

$$\hat{y}_i = E[f(X)] + \sum_{j=1}^{7} \phi_{ij}, \qquad (1)$$

em que $\hat{y}i$ é o número predito de detecções de fogo ativo e ϕij é a contribuição SHAP do preditor meteorológico j para a observação i. Valores SHAP positivos indicam contribuições que aumentam a predição em relação ao valor de referência, enquanto valores negativos indicam contribuições que a reduzem. Como a variável alvo é uma contagem, as contribuições SHAP são expressas em detecções de fogo ativo na escala anual, mensal ou semanal correspondente.

Foram gerados gráficos de resumo global SHAP para os modelos anual, mensal e semanal, com o objetivo de comparar a magnitude, a direção e a ordenação das contribuições dos preditores entre resoluções temporais.

A interpretação entre resoluções concentrou-se em dois aspectos complementares da estrutura de atribuição SHAP: a magnitude relativa das contribuições dos preditores e a direção dessas contribuições ao longo dos valores dos preditores. Preditores com padrões de atribuição amplamente consistentes entre resoluções temporais foram considerados consistentes entre resoluções, enquanto mudanças acentuadas na importância relativa ou na direção das atribuições foram interpretadas como padrões dependentes da escala. Essas comparações foram descritivas e baseadas nos modelos, e não testes estatísticos de equivalência entre escalas, especialmente porque o modelo anual foi ajustado a um número substancialmente menor de observações e apresentou capacidade limitada de generalização fora da amostra. Por ter alcançado o melhor desempenho na validação cruzada, o modelo semanal final foi selecionado para a análise dos padrões não lineares de atribuição por meio de gráficos de dependência. Os valores SHAP foram interpretados

como descrições do comportamento preditivo do modelo, e não como estimativas de efeitos

causais.

Os valores de referência SHAP da Figura 5 foram estimados separadamente para cada preditor no modelo semanal final. Para o preditor j, os pares de observações (xij, ϕij) foram ordenados de forma crescente pelo valor do preditor xij . Uma observação com ϕij = 0 foi tratada como um cruzamento exato; quando ocorreram múltiplos zeros exatos, foi mantido o menor valor correspondente do preditor. Caso contrário, foram identificados todos os pares adjacentes que satisfaziam $\phi kj\phi(k{+}1)j < 0$. Quando ocorreram múltiplas mudanças de sinal, selecionou-se o par com o menor valor do preditor, e o valor de referência foi estimado por interpolação linear:

$$x_{0j} = x_{kj} - \phi_{kj}\frac{x_{(k+1)j} - x_{kj}}{\phi_{(k+1)j} - \phi_{kj}}. \quad (2)$$

Na aplicação deste estudo, a Equação 2 foi utilizada para localizar, para cada preditor meteorológico, o valor de $x0j$ em que a contribuição SHAP estimada cruza zero. Esse ponto representa a transição entre contribuições positivas e negativas do preditor para a saída do modelo, considerando os valores SHAP ordenados segundo os valores observados do próprio preditor. Quando não havia zero exato nem mudança de sinal, o valor de referência foi definido como o valor observado do preditor associado a min$i$ ϕij . Assim, o procedimento selecionou de forma consistente a menor transição detectada no padrão ordenado de dependência e forneceu uma alternativa para os preditores sem cruzamento. Esses valores de referência foram utilizados apenas como indicadores exploratórios de transições no comportamento aprendido pelo modelo e não foram interpretados como limiares causais ou operacionais de ocorrência de fogo.

3 RESULTADOS

3.1 DESEMPENHO PREDITIVO ENTRE ESCALAS TEMPORAIS

Após o pré-processamento, a agregação temporal e a remoção dos intervalos sem o conjunto completo de preditores, os conjuntos de dados analíticos continham 20 observações anuais, 240 mensais e 1.044 semanais.

A Tabela 1 apresenta as divisões temporais com janela expansiva e as métricas preditivas de cada partição. O desempenho dos modelos variou tanto ao longo do tempo quanto entre resoluções temporais. A primeira partição nas três escalas apresentou valores negativos de $R2$, indicando desempenho inferior ao referencial correspondente baseado na média. As partições subsequentes

apresentaram valores positivos, embora o desempenho tenha diminuído no período de teste mais recente nas escalas mensal e semanal.

A Tabela 2 resume o desempenho entre as partições. O $R2$ médio na validação cruzada aumentou entre as resoluções temporais avaliadas, de −0.2743 na escala anual para 0.1646 na mensal e 0.3245 na semanal. O modelo semanal também apresentou a menor dispersão de $R2$ (0.2843), em comparação com 0.4582 no modelo mensal e 0.9311 no anual. RMSE e MAE não foram comparados entre escalas, pois representam contagens agregadas em intervalos de durações diferentes. A Figura 3 mostra maior aderência às flutuações temporais observadas nas resoluções mais finas, embora tenham persistido discrepâncias nos maiores picos.

Tabela 1 - Divisões da validação cruzada temporal com janela expansiva e desempenho preditivo dos modelos Random Forest nas escalas anual, mensal e semanal.

Escala	Partição	Intervalo de treinamento	Intervalo de teste	R^2	RMSE	MAE
Anual	1	2005-12-31–2009-12-31	2010-12-31–2014-12-31	-1.5872	3423.50	3099.18
Anual	2	2005-12-31–2014-12-31	2015-12-31–2019-12-31	0.4692	2235.26	1549.64
Anual	3	2005-12-31–2019-12-31	2020-12-31–2024-12-31	0.2952	3508.66	2723.69
Mensal	1	2005-01-31–2008-04-30	2008-05-31–2011-08-31	-0.6628	647.51	362.23
Mensal	2	2005-01-31–2011-08-31	2011-09-30–2014-12-31	0.5682	356.59	208.55
Mensal	3	2005-01-31–2014-12-31	2015-01-31–2018-04-30	0.0292	538.29	251.26
Mensal	4	2005-01-31–2018-04-30	2018-05-31–2021-08-31	0.5615	602.71	347.38
Mensal	5	2005-01-31–2021-08-31	2021-09-30–2024-12-31	0.3271	778.80	510.17
Semanal	1	2005-01-02–2008-04-27	2008-05-04–2011-08-28	-0.1295	151.60	78.35
Semanal	2	2005-01-02–2011-08-28	2011-09-04–2014-12-28	0.1493	131.28	67.92
Semanal	3	2005-01-02–2014-12-28	2015-01-04–2018-04-29	0.6663	85.35	46.31
Semanal	4	2005-01-02–2018-04-29	2018-05-06–2021-08-29	0.5355	194.30	101.83
Semanal	5	2005-01-02–2021-08-29	2021-09-05–2024-12-29	0.4009	204.73	120.31

Nota: RMSE e MAE são expressos como contagens de detecções de fogo ativo no intervalo de agregação correspondente. Portanto, seus valores absolutos não devem ser comparados diretamente entre escalas temporais. Valores negativos de $R2$ indicam desempenho inferior ao da predição pela média do conjunto de teste correspondente.
Fonte: Autores.

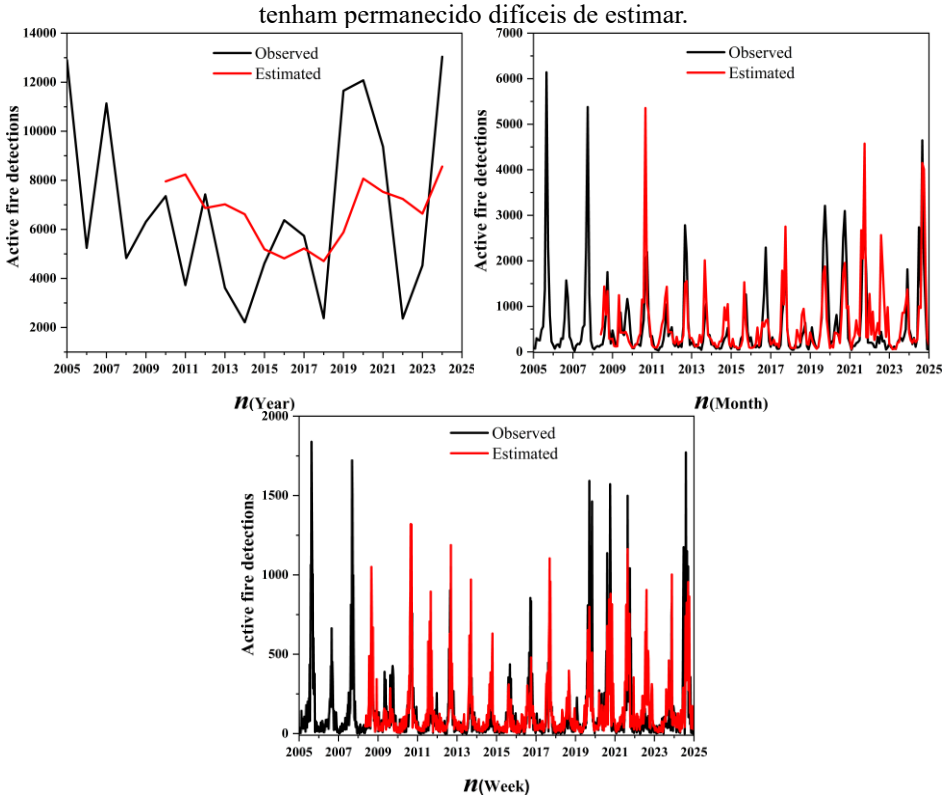
Tabela 2 - Desempenho preditivo médio entre as partições da validação cruzada temporal.

Escala	R^2	RMSE	MAE
Anual	−0.2743 ± 0.9311	3055.81 ± 581.25	2457.50 ± 660.01
Mensal	0.1646 ± 0.4582	584.78 ± 138.66	335.92 ± 104.48
Semanal	0.3245 ± 0.2843	153.45 ± 43.43	82.94 ± 25.85

Nota: Os desvios-padrão foram calculados entre as partições com divisor n. RMSE e MAE são expressos como contagens de detecções de fogo ativo na escala de agregação correspondente.
Fonte: Autores.

ISSN: 2177-3246

Figura 3: Contagens observadas e estimadas de detecções de fogo ativo nas escalas temporais anual, mensal e semanal. Os valores estimados correspondem às predições produzidas para os períodos de validação cronológica. As resoluções temporais mais finas reproduziram uma parcela maior da variabilidade temporal observada, embora os maiores picos tenham permanecido difíceis de estimar.



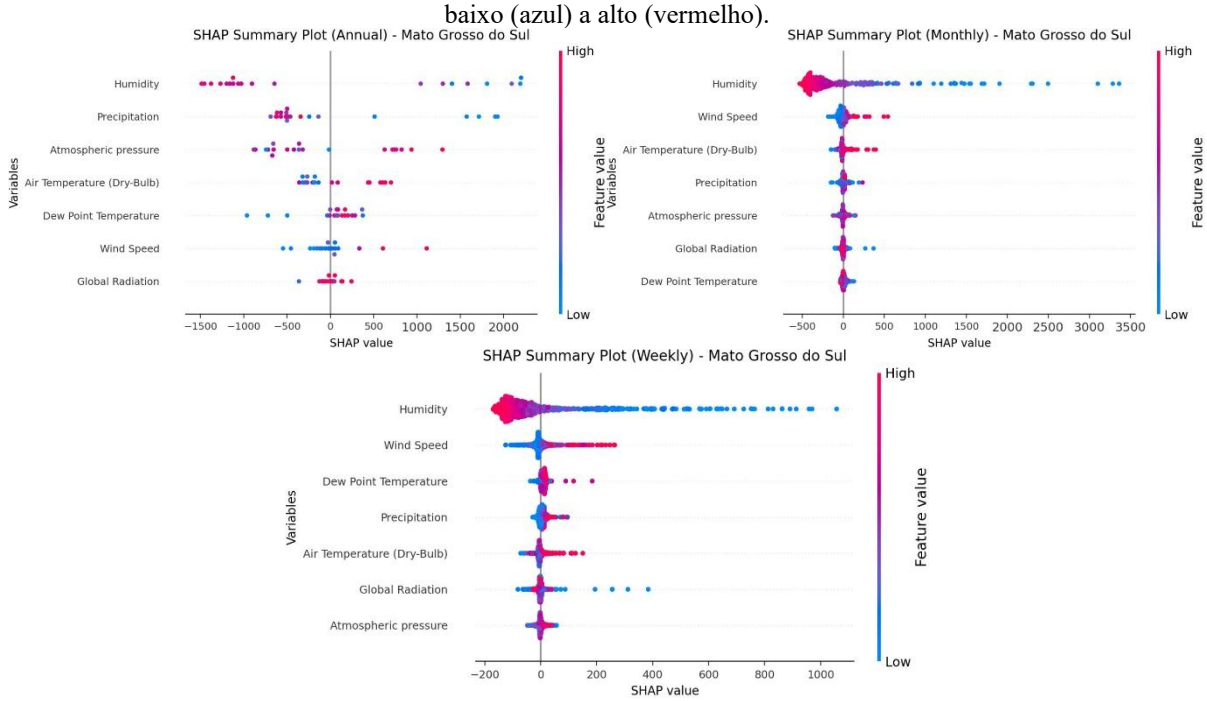
Fonte: Autores (2026).

3.2 CONTRIBUIÇÕES DOS PREDITORES BASEADAS EM SHAP

3.2.1 Importância global dos preditores entre escalas temporais

A Figura 4 apresenta os gráficos de resumo global SHAP dos modelos finais anual, mensal e semanal. Os preditores foram ordenados de acordo com a média dos valores SHAP absolutos, que quantifica a magnitude média de suas contribuições para as saídas dos respectivos modelos.

Figura 4: Gráficos de resumo global SHAP para as escalas temporais anual, mensal e semanal. A posição horizontal representa a contribuição de cada variável para a predição, enquanto a cor indica o valor relativo do preditor, variando de baixo (azul) a alto (vermelho).



Fonte: Autores (2026).

A ordenação dos preditores diferiu entre resoluções temporais. A umidade relativa apresentou a maior média dos valores SHAP absolutos nos três modelos. A precipitação ocupou a segunda posição na escala anual, enquanto a velocidade do vento ocupou a segunda posição nas escalas mensal e semanal. A temperatura do ar de bulbo seco ocupou a terceira posição no modelo mensal. Na escala semanal, a temperatura do ponto de orvalho ocupou a terceira posição, com contribuição SHAP absoluta média inferior à da umidade relativa e à da velocidade do vento, mas superior à da precipitação, da temperatura do ar de bulbo seco, da radiação global e da pressão atmosférica.

Assim, a estrutura de atribuição apresentou tanto consistência entre resoluções quanto dependência da escala. A umidade relativa foi o único preditor que manteve a maior importância global nas três resoluções temporais, enquanto as posições relativas dos demais preditores meteorológicos se modificaram com a agregação temporal.

A ordenação anual deve ser interpretada com cautela, pois o modelo foi ajustado a apenas 20 observações e apresentou generalização temporal limitada. Sua hierarquia de preditores é, portanto, exploratória, e não uma evidência robusta de importância na escala anual.

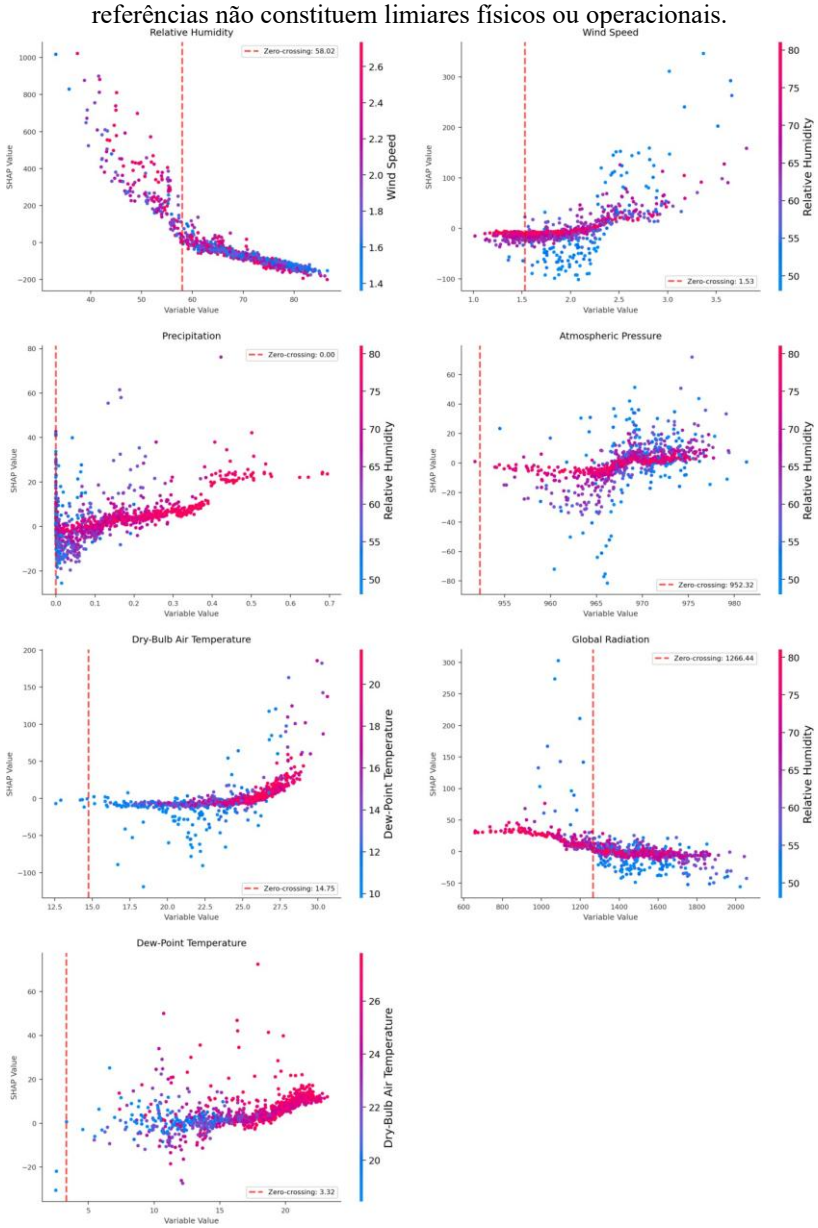
3.2.2 Padrões de dependência SHAP no modelo semanal

A Figura 5 apresenta os gráficos de dependência SHAP dos sete preditores meteorológicos no modelo semanal final. A configuração semanal foi selecionada por ter alcançado o maior *R*2 médio na

ISSN: 2177-3246

validação cruzada entre as três resoluções temporais.

Figura 5: Gráficos de dependência SHAP dos sete preditores meteorológicos no modelo semanal final. As cores representam o preditor utilizado na codificação por cores e apoiam a inspeção exploratória de possíveis padrões de interação. As linhas vermelhas tracejadas indicam os valores de referência algorítmicos definidos na Seção 2.8. Essas referências não constituem limiares físicos ou operacionais.



Fonte: Autores.

Os gráficos de dependência SHAP semanais revelaram padrões não lineares distintos na forma como o Random Forest utilizou os preditores meteorológicos. A umidade relativa exibiu a relação mais clara, com valores baixos produzindo grandes contribuições positivas para as contagens preditas de detecções de fogo ativo e valores progres-sivamente mais altos produzindo contribuições neutras ou negativas. A velocidade do vento apresentou contribuições comparativamente pequenas na faixa de valores mais baixos, mas contribuições cada vez mais positivas e dispersas nos valores mais altos. O

ISSN: 2177-3246

modelo sugere que condições secas constituem o principal contexto associado ao aumento das detecções de fogo, enquanto velocidades maiores do vento parecem reforçar essa contribuição sobretudo quando a atmosfera já apresenta baixa umidade. A temperatura do ar de bulbo seco contribuiu pouco na maior parte de sua distribuição, enquanto as observações na faixa superior de temperatura estiveram, em geral, associadas a contribuições positivas; a codificação por cores mostrou variação simultânea com a temperatura do ponto de orvalho. A precipitação apresentou contribuições heterogêneas próximas de zero e uma tendência positiva modesta em parte de sua distribuição, enquanto o aumento da radiação global esteve, em geral, associado a contribuições neutras ou negativas. A contribuição da radiação global parece depender fortemente do estado atmosférico no qual ocorre. Valores elevados de radiação não produziram, isoladamente, contribuições positivas para a atividade do fogo, indicando que o modelo pode estar utilizando essa variável em combinação com a umidade, a sazonalidade e outras condições meteorológicas. A pressão atmosférica apresentou contribuições locais comparativamente pequenas e heterogêneas. A temperatura do ponto de orvalho, apesar de ocupar a terceira posição na importância global medida pela média dos valores SHAP absolutos, também apresentou heterogeneidade local substancial. De modo geral, a dispersão e os gradientes de cores mostraram que as contribuições dos preditores variaram entre observações e combinações de valores meteorológicos. Como as cores fornecem indicações exploratórias, e não estimativas formais de interação SHAP, elas não foram interpretadas como evidência de interações ou de efeitos causais.

Os cruzamentos de zero do SHAP também foram examinados de forma exploratória como pontos de referência baseados no modelo. Para a umidade relativa, a referência próxima de 58 % coincidiu, de modo geral, com a transição de contribuições predominantemente positivas para negativas. Para vários outros preditores, entretanto, a primeira mudança de sinal detectada ocorreu fora da faixa em que uma transição mais ampla se tornou visualmente aparente. Essas referências não foram, portanto, interpretadas como limiares físicos, ecológicos ou operacionais.

Os gráficos de dependência foram gerados a partir do modelo final ajustado ao conjunto completo de dados semanais, e não das partições individuais da validação cruzada. Assim, eles caracterizam padrões aprendidos pelo modelo final, mas não constituem uma avaliação independente da generalização preditiva.

4 DISCUSSÃO

O aumento do $R2$ médio da escala anual (−0.2743) para a semanal (0.3245) mostra que o desempenho preditivo diferiu substancialmente entre as representações temporais avaliadas. As médias

ISSN: 2177-3246

anuais suavizam episódios de secura atmosférica, calor e vento, enquanto a agregação semanal preserva parte dessas flutuações. O baixo desempenho anual também deve ser interpretado à luz da redução drástica do tamanho amostral causada pela agregação anual: a série anual completa continha apenas 20 observações, e a primeira partição de validação treinou o Random Forest com apenas cinco observações anuais, aumentando a instabilidade das estimativas de desempenho fora da amostra. Como a agregação temporal também modifica o tamanho amostral, a distribuição da variável resposta e os períodos de avaliação, essa comparação não isola o efeito da resolução temporal. Ainda assim, o modelo semanal apresentou desempenho apenas moderado, com variações entre as partições de validação; portanto, sua superioridade relativa não implica acurácia preditiva suficiente para uso operacional sem validação prospectiva. O modelo também tendeu a suavizar temporadas extremas de fogo, indicando capacidade limitada para reproduzir os maiores picos observados. Comparações diretas dos erros absolutos entre escalas temporais devem ser evitadas, pois cada resolução temporal representa contagens acumuladas em intervalos de durações diferentes.

Os resultados SHAP demonstram que a ordenação dos preditores foi sensível à resolução temporal, apesar de a umidade relativa ocupar a primeira posição nos três modelos. A precipitação ocupou a segunda posição na escala anual, enquanto a velocidade do vento ocupou a segunda posição nas resoluções mensal e semanal. A velocidade do vento tornou-se mais proeminente no ranking nas escalas mensal e semanal, enquanto a precipitação passou a ocupar posições inferiores à medida que a resolução temporal se tornou mais fina. A temperatura do ar de bulbo seco ocupou a terceira posição no modelo mensal. A temperatura do ponto de orvalho ocupou a terceira posição no modelo semanal, em comparação com a quinta posição no anual e a sétima no mensal, enquanto a precipitação ocupou a quarta posição na escala semanal. Essa reordenação mostra que a agregação temporal esteve associada a mudanças tanto no desempenho dos modelos quanto nos padrões de atribuição aos preditores.

Considerando conjuntamente as três resoluções, os preditores podem ser organizados descritivamente segundo a estabilidade de sua importância relativa entre escalas. A umidade relativa apresentou atribuição estável, permanecendo como o preditor dominante nos modelos anual, mensal e semanal. A velocidade do vento apresentou estabilidade moderada, pois se manteve entre os preditores mais relevantes e passou a ocupar uma posição mais proeminente nas resoluções mais finas. A temperatura do ponto de orvalho, a precipitação e a temperatura do ar foram mais sensíveis à escala, com mudanças acentuadas em suas posições relativas entre os modelos. A pressão atmosférica e a radiação global apresentaram, em geral, contribuições relativas menores, embora sua ordenação também tenha variado conforme a resolução. Essa classificação é descritiva e resume a estabilidade

das atribuições aprendidas pelo Random Forest, não constituindo um teste estatístico de invariância entre escalas.

Dadas as limitações da amostra anual e da generalização temporal, a ordenação SHAP anual deve ser interpretada com cautela. A liderança da umidade relativa e a segunda posição da precipitação nessa escala podem ser sensíveis a anos individuais e não devem ser interpretadas como evidência de controles anuais dominantes. Em contrapartida, a permanência da umidade relativa e da velocidade do vento entre os principais preditores nas escalas mensal e semanal fornece evidência mais consistente de que a secura atmosférica e as condições relacionadas ao vento contêm informações relevantes para estimar as contagens de detecções de fogo ativo. Essa interpretação é compatível com estudos anteriores que associam baixa umidade atmosférica ao aumento da inflamabilidade dos combustíveis e da atividade do fogo (ARRUDA et al., 2024; OLIVEIRA et al., 2022).

Os gráficos de dependência SHAP mostraram que a contribuição dos preditores meteorológicos não ocorreu de forma isolada, mas variou conforme o estado atmosférico representado pelas demais variáveis. A relação conjunta entre umidade relativa e velocidade do vento foi particularmente evidente. Valores baixos de umidade produziram as maiores contribuições positivas para as contagens preditas de fogo ativo, enquanto velocidades mais elevadas do vento tenderam a apresentar contribuições positivas sobretudo sob condições de menor umidade. Assim, o modelo não está apenas reconhecendo baixa umidade; ele parece reconhecer estados atmosféricos secos e ventosos como configurações particularmente associadas ao aumento das detecções de fogo.

Um comportamento conjunto também foi observado entre a temperatura do ar de bulbo seco e a temperatura do ponto de orvalho. Temperaturas mais elevadas do ar estiveram associadas a contribuições SHAP crescentemente positivas, especialmente na faixa superior da distribuição, enquanto a contribuição da temperatura do ponto de orvalho variou de acordo com a temperatura do ar. Esses padrões indicam que o modelo utiliza informações compartilhadas entre temperatura e umidade atmosférica para representar diferentes estados meteorológicos, de modo que suas atribuições individuais não devem ser interpretadas como efeitos independentes.

A radiação global apresentou comportamento distinto, com valores elevados predominantemente associados a contribuições neutras ou negativas. A distribuição das cores, entretanto, indica que esse padrão varia com a umidade relativa, sugerindo que a radiação pode estar atuando como marcador de condições atmosféricas ou sazonais mais amplas, e não como um controle isolado da atividade do fogo. De forma semelhante, o padrão observado para a precipitação deve ser interpretado com cautela, uma vez que o preditor utilizado representa a média das observações horárias no intervalo temporal e não a precipitação acumulada.

Essas relações conjuntas são exploratórias. A codificação por cores dos gráficos de dependência permite identificar padrões condicionais aprendidos pelo Random Forest, mas não constitui uma estimativa formal de interação entre variáveis nem evidência de relações causais.

A validação cruzada de séries temporais com janela expansiva impediu o treinamento dos modelos com ob-servações posteriores aos respectivos períodos de teste e, assim, proporcionou uma avaliação mais realista do que uma divisão aleatória. Entretanto, os valores SHAP explicam os modelos finais ajustados aos conjuntos completos de dados, enquanto as métricas de validação apresentadas descrevem o desempenho preditivo fora das amostras de treinamento. Consequentemente, o SHAP explica como os modelos finais distribuem suas predições entre as variáveis meteorológicas; ele não valida essas relações de forma independente. Além disso, as contribuições SHAP dos modelos anual, mensal e semanal são expressas em detecções de fogo ativo em escalas temporais diferentes. Suas magnitudes absolutas não devem ser comparadas diretamente entre resoluções; apenas a ordenação dos preditores e os padrões qualitativos dentro de cada modelo são comparáveis.

4.1 IMPLICAÇÕES METODOLÓGICAS E TRANSFERIBILIDADE

A implicação metodológica mais ampla desta análise é que a resolução temporal deve ser tratada como parte da especificação dos modelos ambientais, e não como uma decisão neutra de pré-processamento. O fluxo de trabalho proposto oferece uma forma estruturada de diagnosticar a sensibilidade à escala temporal, avaliando simultaneamente três componentes: generalização preditiva, atribuição aos preditores e padrões não lineares de resposta. Essa distinção é importante porque um modelo pode manter desempenho preditivo semelhante enquanto modifica sua estrutura de atribuição ou, inversamente, preservar os preditores dominantes enquanto apresenta diferentes capacidades preditivas entre representações temporais. A estrutura não se restringe à modelagem de incêndios e pode, em princípio, ser aplicada a outros sistemas ambientais agregados temporalmente, desde que as características dos dados e as restrições de validação específicas de cada escala sejam explicitamente consideradas.

Do ponto de vista da gestão ambiental, a comparação fornece evidências para selecionar escalas temporais e indicadores meteorológicos a serem considerados no desenvolvimento posterior do monitoramento em Mato Grosso do Sul. O maior $R2$ médio semanal apoia a priorização da agregação semanal para avaliações subsequentes, enquanto a importância recorrente da umidade relativa e o destaque da velocidade do vento nas escalas mensal e semanal identificam variáveis que merecem atenção especial. Entretanto, os modelos apresentados não estabelecem limiares de alerta nem demonstram melhorias na alocação de recursos, nas atividades de combate a incêndios ou em outros

ISSN: 2177-3246

resultados de gestão. Sua utilidade para essas decisões deve ser testada prospectivamente. O fluxo de trabalho comparativo também pode ser avaliado em outras regiões sazonalmente secas e propensas a incêndios, desde que a cobertura dos dados locais, os regimes de fogo e o desempenho preditivo sejam avaliados de forma independente; não se deve pressupor que os resultados semanais e a ordenação dos preditores apresentados aqui sejam generalizáveis para além de Mato Grosso do Sul.

Algumas limitações devem ser consideradas. Os preditores meteorológicos foram agregados espacialmente pelas médias das observações válidas das estações em cada período, e o número e a distribuição espacial das estações contribuintes variaram ao longo do tempo. Os modelos também não representaram explicitamente o uso da terra, as condições da vegetação, a disponibilidade de combustível, as fontes de ignição, as práticas de manejo do fogo ou outros fatores antrópicos. As detecções de fogo ativo são observações derivadas de satélites e não medem diretamente a área queimada, a duração ou a severidade do fogo. Além disso, os modelos utilizaram condições meteorológicas agregadas nos mesmos períodos das detecções de fogo. O desempenho moderado fora da amostra e o uso de preditores contemporâneos limitam a previsão operacional. O delineamento observacional e as atribuições SHAP baseadas nos modelos não estabelecem efeitos causais.

O uso de um único algoritmo de aprendizado também delimita o alcance dos resultados apresentados. O Random Forest foi intencionalmente mantido fixo entre resoluções temporais para reduzir a variabilidade relacionada ao algoritmo e permitir que a análise se concentrasse nas mudanças associadas à representação temporal. Portanto, os resultados não estabelecem que as diferenças observadas no desempenho preditivo e na atribuição dos modelos, dependentes da escala, sejam independentes da classe do modelo. Outras abordagens estatísticas e de aprendizado de máquina podem responder de modo diferente à agregação temporal devido a diferenças em sua estrutura, seus pressupostos e seus mecanismos de aprendizado. Estudos futuros devem avaliar a estrutura proposta em múltiplas resoluções utilizando algoritmos alternativos, para determinar quais padrões são robustos entre classes de modelos e quais são específicos do Random Forest.

Apesar dessas limitações, a agregação semanal alcançou o maior $R2$ médio na validação cruzada entre as escalas avaliadas. A umidade relativa permaneceu como o preditor semanal mais influente, enquanto a velocidade do vento, a temperatura do ponto de orvalho, a precipitação e a temperatura do ar forneceram contribuições preditivas complementares. Estudos futuros devem avaliar defasagens meteorológicas, déficit de pressão de vapor, preditores ambientais e antrópicos, intervalos de incerteza e validação prospectiva em períodos posteriores aos dados de treinamento. Essas extensões são necessárias antes que a estrutura possa ser considerada para aplicações operacionais de alerta antecipado ou de gestão do risco de incêndios.

ISSN: 2177-3246

5 CONCLUSÕES

Este estudo mostrou que a resolução temporal modifica tanto o comportamento do Random Forest quanto a estrutura de atribuição dos fatores meteorológicos associados à atividade do fogo. Entre as escalas avaliadas, a representação semanal apresentou o maior desempenho fora da amostra, com $R2$ médio de 0.3245, enquanto as representações mensal e anual apresentaram desempenho inferior. A umidade relativa apresentou a atribuição mais estável entre as três resoluções, permanecendo como o preditor dominante, enquanto a velocidade do vento, a precipitação, a temperatura do ponto de orvalho, a temperatura do ar, a pressão atmosférica e a radiação solar apresentaram diferentes graus de sensibilidade à escala temporal.

Esses resultados mostram que a agregação temporal não constitui apenas uma etapa de pré-processamento, mas uma escolha estrutural da modelagem meteorologia–fogo. Diferentes representações temporais estiveram associadas a mudanças tanto na capacidade de generalização dos modelos quanto na estrutura de atribuição aprendida pelo Random Forest. Essas diferenças, entretanto, não podem ser atribuídas exclusivamente à resolução temporal, uma vez que o tamanho amostral, a distribuição da variável resposta e os períodos de avaliação também variaram entre escalas. As atribuições SHAP devem, portanto, ser interpretadas como descrições do comportamento aprendido pelo modelo, e não como evidência de efeitos meteorológicos causais.

O principal resultado generalizável deste estudo é metodológico: a resolução temporal deve ser explicitamente avaliada como parte da especificação de modelos ambientais baseados em dados. A combinação de uma estrutura comum de modelagem, validação cronológica e inteligência artificial explicável oferece uma estratégia transferível para avaliar se o desempenho e a interpretação dos modelos permanecem estáveis ou dependem da escala temporal. Os resultados específicos observados para Mato Grosso do Sul, incluindo a superioridade relativa da escala semanal e a ordenação dos preditores, devem ser validados de forma independente antes de serem extrapolados para outros contextos.

FINANCIAMENTO

Esta pesquisa não recebeu financiamento específico de nenhuma agência de fomento dos setores público, comercial ou sem fins lucrativos.

ISSN: 2177-3246

DISPONIBILIDADE DOS DADOS

Os dados brutos de detecções de fogo ativo utilizados neste estudo estão disponíveis publicamente no Programa Queimadas do Instituto Nacional de Pesquisas Espaciais (INPE), por meio da plataforma BDQueimadas (Instituto Nacional de Pesquisas Espaciais, 2026a). Os dados meteorológicos brutos estão disponíveis publicamente no portal de dados históricos do Instituto Nacional de Meteorologia (INMET) (Instituto Nacional de Meteorologia, 2026). Os conjuntos de dados analisados neste estudo foram derivados dessas fontes públicas por meio dos procedimentos de limpeza, agregação temporal e integração descritos no manuscrito.

DECLARAÇÃO DE CONFLITOS DE INTERESSE

Os autores declaram não ter interesses financeiros concorrentes conhecidos nem relações pessoais que possam ter influenciado o trabalho relatado neste artigo.

ISSN: 2177-3246

REFERÊNCIAS

ANEES, S. A. et al. Advancing forest fire burn severity and vegetation recovery assessments using remote sensing and machine learning approaches. Ecological Informatics, v. 92, p. 103446, 2025.

ARRUDA, V. L. da S. et al. Assessing four decades of fire behavior dynamics in the Cerrado biome (1985 to 2022). Fire Ecology, v. 20, n. 1, p. 64, 2024.

BREIMAN, L. Random forests. Machine Learning, v. 45, n. 1, p. 5–32, 2001.

CHEERALA, U. B. et al. Probabilistic wildfire susceptibility from remote sensing using random forests and SHAP. In: 2025 IEEE Asia-Pacific Conference on Geoscience, Electronics and Remote Sensing Technology (AGERS). [S.l.: s.n.], 2025. p. 578–585.

CHEN, G. et al. Continuous wildfires threaten public and ecosystem health under climate change across continents. Frontiers of Environmental Science & Engineering, v. 18, n. 10, p. 130, 2024.

FREITAS, K. M. et al. Prediction of forest fire susceptibility using machine learning tools in the Triunfo do Xingu environmental protection area, Amazon, Brazil. Journal of South American Earth Sciences, v. 153, p. 105366, 2025.

Instituto Nacional de Meteorologia. Dados meteorológicos históricos. 2026. Accessed 23 September 2026. Disponível em: <https://portal.inmet.gov.br/dadoshistoricos>.

Instituto Nacional de Pesquisas Espaciais. BDQueimadas: Banco de Dados de Queimadas. 2026. Programa Queimadas. Accessed 23 September 2026. Disponível em: <https://terrabrasilis.dpi.inpe.br/queimadas/bdqueimadas/>.

Instituto Nacional de Pesquisas Espaciais. Programa Queimadas: Perguntas frequentes. 2026. Accessed 23 September 2026. Disponível em: <https://terrabrasilis.dpi.inpe.br/queimadas/portal/pages/secao_informacoes/faq/>.

KARASANTE, I. et al. SeasFire cube–a multivariate dataset for global wildfire modeling. Scientific Data, v. 12, n. 1, p. 368, 2025.

LI, H. et al. Spatio-temporal feature attribution of European summer wildfires with explainable artificial intelligence (XAI). Science of the Total Environment, v. 916, p. 170330, 2024.

LUNDBERG, S. M.; LEE, S.-I. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems. [S.l.]: Curran Associates, Inc., 2017. v. 30, p. 4765–4774.

MASOUDIAN, E.; MIRZAEI, A.; BAGHERI, H. Assessing wildfire susceptibility in Iran: Leveraging machine learning for geospatial analysis of climatic and anthropogenic factors. Trees, Forests and People, v. 19, p. 100774, 2025.

MIRANDA, J. da R. et al. Climate-vegetation intersection in determining the burn rate in an area of the Brazilian Cerrado. Earth Science Informatics, v. 18, n. 1, p. 56, 2025.

ISSN: 2177-3246

MOGHIM, S.; MEHRABI, M. Wildfire assessment using machine learning algorithms in different regions. Fire Ecology, v. 20, n. 1, p. 104, 2024.

MUHAMMAD, S. et al. Machine learning-based forest fire vulnerability assessment in subtropical chir pine forests of Pakistan. Fire Ecology, v. 21, n. 1, p. 72, 2025.

OLIVEIRA-JÚNIOR, J. F. de et al. Confronting CHIRPS dataset and in situ stations in the detection of wet and drought conditions in the Brazilian Midwest. International Journal of Climatology, v. 41, n. 9, p. 4478–4493, 2021.

OLIVEIRA, U. et al. Determinants of fire impact in the Brazilian biomes. Frontiers in Forests and Global Change, v. 5, p. 735017, 2022.

PÁDUA, J. E. S.; LEVADA, A. L. M.; VALENTE, F. J. A retrospective benchmark of spatiotemporal covariates for daily active-fire detection in Cerrado conservation units. 2026. ArXiv preprint. Accessed 4 August 2026. Disponível em: <https://arxiv.org/abs/2606.04170>.

SAYEDI, S. S. et al. Assessing changes in global fire regimes. Fire Ecology, v. 20, n. 1, p. 18, 2024.

SENGUPTA, A.; WOODFORD, B. J. Recent advances in explainable machine learning models for wildfire prediction. Applied Computing and Geosciences, v. 27, p. 100266, 2025.

SILVEIRA, M. V. F. et al. Amazon fires in the 21st century: The year of 2020 in evidence. Global Ecology and Biogeography, v. 31, n. 10, p. 2026–2040, 2022.

SOBREIRA, E. et al. Wildfires and their toll on Brazil: Who's counting the cost? Perspectives in Ecology and Conservation, v. 23, n. 3, p. 214–217, 2025.

SOUZA, A. de et al. Climate regionalization in Mato Grosso do Sul: A combination of hierarchical and non-hierarchical clustering analyses based on precipitation and temperature. Brazilian Archives of Biology and Technology, v. 65, p. e22210331, 2022.

TEKE, A.; KAVZOGLU, T. Explainable artificial intelligence to unveil intrinsic characteristics of conditioning factors governing forest fire susceptibility. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-M-6-2025, p. 281–287, 2025.

TONBUL, H.; VERAVERBEKE, S. Machine-learning wildfire susceptibility mapping with SHAP-based explainability in Türkiye's fire-prone regions. Stochastic Environmental Research and Risk Assessment, v. 40, n. 1, p. 2, 2026.

TORRES-VÁZQUEZ, M. Á. et al. Enhancing seasonal fire predictions with hybrid dynamical and random forest models. npj Natural Hazards, v. 2, n. 1, p. 20, 2025.

WALKER, X. J. et al. Increasing wildfire frequency decreases carbon storage and leads to regeneration failure in Alaskan boreal forests. Fire Ecology, v. 21, n. 1, p. 57, 2025.

ISSN: 2177-3246

ZAHURA, F. T. et al. Impact of topography and climate on post-fire vegetation recovery across different burn severity and land cover types through random forest. Ecological Informatics, v. 82, p. 102757, 2024.