

INSTITUT D'INTELLIGENCE ARTIFICIELLE APPLIQUÉE
Inteligência Artificial Aplicada aos Desafios Socioambientais da Amazônia
Projeto Final

**Modelagem Preditiva Baseada em IA para Classificação de Risco
Sanitário e Priorização de Investimentos em Saneamento na Região
Metropolitana de Belém (RMB)**

28/11/2025

Grupo 21

André Vinícius Neves Alves
Cícero Manoel dos Santos
Eric Narciso Pimentel dos Santos
Melissa Costa Freitas
Rafael da Silva Lopes
Verena Franco de Sá Farias
Yuri Cezar de Lima Ferreira

1. Introdução:

A Região Metropolitana de Belém (RMB) enfrenta vulnerabilidades estruturais em abastecimento de água, esgotamento sanitário, perdas no sistema e degradação ambiental, agravadas pela alta precipitação amazônica e por desigualdades territoriais. Essas condições contribuem para a elevada incidência de doenças de veiculação hídrica, especialmente as classificadas no CID-10 A00–A09, que afetam de forma desproporcional áreas de menor acesso ao saneamento.

Diante desse contexto, torna-se essencial integrar bases do sistema Nacional de Informações sobre Saneamento (SNIS)/ Sistema Nacional de Informações em Saneamento Básico (SINISA), sistema de Informação de Vigilância da Qualidade da Água para Consumo Humano (SISAGUA), Sistema de Informações Hospitalares do SUS (SIH)/ Departamento de Informática do SUS (DATASUS), Instituto Nacional de Meteorologia (INMET), Instituto Brasileiro de Geografia e Estatística (IBGE) e Sistema de Informações sobre Orçamentos Públicos em Saúde (SIOPS) para compreender a relação entre infraestrutura, qualidade da água, clima e saúde. Com o apoio de técnicas de Inteligência Artificial e modelagem estatística, é possível prever riscos sanitários e orientar decisões baseadas em evidências. Este trabalho

propõe um modelo preditivo para estimar o risco de internações por doenças hídricas na RMB e um dashboard interativo para monitoramento contínuo e simulação de cenários.

2. Justificativa

A Região Metropolitana de Belém enfrenta desafios relacionados à qualidade da água, saneamento básico e saúde pública. A baixa cobertura de abastecimento e esgotamento sanitário, aliada às elevadas perdas no sistema, contribui para o aumento de custos hospitalares e doenças de veiculação hídrica (especialmente diarreias infecciosas A00–A09: Doenças infecciosas intestinais, Classificação epidemiológica oficial CID-10).

Diante desse cenário, o projeto propõe integrar bases de dados públicos do SNIS, DATASUS e INMET, como Série Histórica (Água/Esgoto), dados meteorológicos e internações por doenças de veiculação hídrica, para identificar e classificar áreas de maior risco sanitário, permitindo a priorização de investimentos em saneamento.

Uma análise integrada desses dados, apoiada por inteligência artificial e modelagem estatística, possibilitará a construção de um painel de monitoramento e tomada de decisão, voltado à redução das desigualdades em infraestrutura hídrica e à melhoria da saúde das populações vulneráveis na região metropolitana de Belém.

A iniciativa se alinha diretamente aos Objetivos de Desenvolvimento Sustentável: ODS 3 (Saúde) e ODS 6 (Água e Saneamento), fortalecendo a tomada de decisão baseada em evidências para a região Amazônica.

A realização da COP-30 em Belém trouxe uma oportunidade estratégica para acelerar investimentos estruturantes em saneamento na Região Metropolitana, incluindo as obras de macrodrenagem dos canais urbanos e a inauguração da nova Estação de Tratamento de Esgoto (ETE) Una em 2025. Esses avanços recentes, somados à visibilidade internacional do evento, criam uma janela única para utilizar evidências ambientais, sanitárias e climáticas na priorização de investimentos, garantindo que as intervenções sejam direcionadas às áreas de maior vulnerabilidade e tenham impacto real na saúde pública e na qualidade da água na Amazônia.

O enfrentamento aos desafios de saneamento e saúde na RMB deve ser orientado pela lente da Justiça Climática, visto que os impactos adversos das mudanças do clima acentuam as diferenças de raça, classe, gênero e território. O

racismo ambiental manifesta-se diretamente na saúde coletiva: em Belém, a taxa de internações por doenças de veiculação hídrica (CID-10 A00–A09), o principal indicador de desfecho sanitário deste trabalho, é até 30 vezes maior para pessoas negras do que para pessoas brancas. Essa disparidade alarmante ocorre porque pessoas negras e famílias de baixa renda estão desproporcionalmente mais expostas a desastres climáticos e habitam áreas de risco desprovidas de saneamento básico e drenagem. Assim, a utilização da inteligência artificial para identificar e classificar áreas de maior risco sanitário e priorizar investimentos em infraestrutura hídrica na RMB auxilia ativamente a correção dessas desigualdades territoriais e raciais.

3. Objetivo

3.1 objetivo geral:

Desenvolver e validar o modelo de IA para identificar e classificar as áreas de maior risco sanitário na RMB, priorizando a redução das internações por A00–A09. Analisar a contribuição relativa de fatores ambientais, climáticos e de infraestrutura para a predição de risco; e fornecer um dashboard de monitoramento para apoio à decisão.

3.2 Objetivos específicos:

1. Integrar e padronizar bases de dados oficiais relacionadas a saneamento, qualidade da água, clima, saúde e finanças públicas, consolidando um dataset para os municípios da RMB;
2. Realizar limpeza, tratamento e engenharia de atributos (gold features) que permitam análises consistentes e comparáveis entre municípios e períodos;
3. Explorar estatisticamente os indicadores de saneamento, qualidade da água e clima, identificando padrões, correlações e potenciais variáveis explicativas da ocorrência de internações por doenças de veiculação hídrica;.
4. Desenvolver, ajustar e validar um modelo preditivo, capaz de estimar o risco sanitário municipal associado a indicadores ambientais e de infraestrutura;
5. Construir um dashboard interativo com simulador de cenários, permitindo avaliar o impacto de melhorias em cobertura de água, cobertura de esgoto, redução de perdas e variação climática sobre as internações previstas.

4. Hipótese de trabalho

A combinação de indicadores de saneamento (cobertura de água, cobertura de esgoto e perdas na distribuição), qualidade da água (turbidez, cloro residual e conformidade microbiológica) e variáveis climáticas (chuva e extremos térmicos) apresenta relação estatisticamente significativa com a incidência de internações por doenças infecciosas intestinais (CID-10 A00–A09) na Região Metropolitana de Belém. Pressupõe-se que a melhoria desses indicadores, especialmente a ampliação da cobertura de esgotamento sanitário e a redução das perdas de água, resultaria em diminuição das internações. Espera-se que um modelo preditivo baseado em IA seja capaz de identificar padrões territoriais de risco sanitário, permitindo priorizar municípios e direcionar investimentos públicos de maneira mais eficiente e equitativa, contribuindo para a redução das desigualdades socioambientais e para o cumprimento das metas dos ODS 3 e 6.

5. Metodologia

A metodologia adotada neste estudo segue um pipeline de dados estruturado em etapas sequenciais e reproduzíveis, abrangendo engenharia de dados, análise exploratória, modelagem estatística/IA e desenvolvimento de produtos analíticos. O fluxo Bronze, Silver e Gold organizou a ingestão, padronização e integração das bases provenientes de SISAGUA, SNIS/SINISA, DATASUS, INMET, IBGE e SIOPS, garantindo rastreabilidade desde os arquivos brutos até os indicadores consolidados por município e ano. A partir da camada Gold, foram conduzidas a Análise Exploratória de Dados (EDA), a derivação de indicadores compostos, a definição das features e a construção do analysis dataset, que fundamentou a etapa preditiva e a geração dos artefatos destinados ao dashboard.

Para assegurar transparência, reprodutibilidade e versionamento das etapas, foi criado um repositório estruturado no GitHub, contendo todos os notebooks, scripts, documentação técnica e logs do pipeline. Em complemento, um repositório no Google Drive armazena as bases intermediárias (Bronze, Silver e Gold), os CSVs finais utilizados no Looker Studio e os relatórios gerados automaticamente pelo projeto. Essa arquitetura integrada permite que qualquer parte do processo, da coleta de dados à geração das previsões e simulações, seja auditável, atualizável e reutilizável, reforçando a consistência metodológica de todo o estudo.

5.1. Limpeza e preparação de dados

Os dados utilizados foram coletados de fontes oficiais brasileiras, cada uma relacionada a um aspecto do estudo: serviços de água/esgoto, qualidade da água, clima, saúde e finanças públicas. Para organizar o fluxo de tratamento, o projeto utilizou três camadas:

Bronze (dados brutos) → Silver (padronizados) → Gold (prontos para análise).

A) De onde vieram os dados

- SNIS(Sistema Nacional de Informações sobre Saneamento) /SINISA (Sistema Nacional de Informações em Saneamento Básico) – indicadores de abastecimento de água, coleta e tratamento de esgoto.
- SISAGUA (Sistema de Informação de Vigilância da Qualidade da Água para Consumo Humano)– resultados de análises de qualidade da água (turbidez, cloro, pH).
- INMET (Instituto Nacional de Meteorologia) – informações climáticas (chuva e calor extremo).
- SIH (Sistema de Informações Hospitalares do SUS)/ DATASUS (Departamento de Informática do SUS) – internações por doenças relacionadas à água (CID-10 A00–A09).
- IBGE – SIDRA (Sistema IBGE de Recuperação Automática) – estimativas de população dos municípios.
- SIOPS (Sistema de Informações sobre Orçamentos Públicos em Saúde) – dados de despesas e receitas da saúde.

Todos foram coletados para os municípios da Região Metropolitana de Belém (Ananindeua, Marituba, Benevides, Santa Bárbara do Pará, Santa Izabel do Pará), entre 2018 e 2025.

B) Como os dados foram preparados

Etapa 1 — Organização em “Bronze”: dados brutos

- Todos os arquivos foram baixados no formato original (CSV, ZIP, XLSX ou DBC) sem qualquer modificação.

Objetivo: manter exatamente o que veio da fonte.

Etapa 2 — Conversão e padronização (“Silver”)

Nessa etapa, os dados foram organizados para ficarem consistentes e fáceis de combinar.

As principais ações foram:

- Padronizar nomes de colunas, como “cod_mun”, “município” e “ano”.
- Converter arquivos especiais, como os arquivos .dbc do SIH para .csv.
- Ajustar tipos de dados (datas, números, categorias).
- Organizar por município e ano, garantindo a mesma estrutura para todas as bases.
- Remover duplicações e valores impossíveis (ex.: números negativos onde não deveriam existir).

(Nenhum cálculo analítico foi feito nesta etapa, apenas limpeza e padronização.)

Etapa 3 — Construção dos indicadores (“Gold”)

Após a padronização, cada fonte gerou um conjunto de indicadores prontos para análise.

Principais passos:

- **SISAGUA**: cálculo de médias anuais de turbidez, cloro e pH, além do percentual de amostras conformes.
- **INMET**: agregação diária → mensal → anual (chuva total, dias de calor extremo).
- **SIH**: contagem de internações filtradas pelos CIDs relevantes, normalização por população (internações por 10 mil habitantes).
- **SIOPS**: cálculo de despesas em saúde por habitante e proporções de receitas.
- **SNIS**: seleção e correção de índices oficiais (água, esgoto, perdas, hidrômetros).

Etapa 4 — Junção final das bases

Todos os indicadores foram combinados em um único dataset final (**gold_features_ano.csv**), unificando:

- Serviço (SNIS)
- Qualidade da água (SISAGUA)
- Clima (INMET)
- Saúde (SIH + IBGE)
- Finanças (SIOPS)

A junção foi feita usando os campos `cod_mun`, `município` e `ano`, garantindo alinhamento entre municípios e anos sem perda de registros.

C) Resultado final

Ao fim da preparação dos dados, obteve-se:

- Bases limpas, padronizadas e comparáveis entre si.
- Indicadores consolidados por município e ano.
- Um dataset final que permite análise estatística, modelagem e construção do dashboard para gestores.

5.2. Exploração de Dados (EDA)

A etapa de Análise Exploratória de Dados foi conduzida a partir do notebook `analise_exploratoria_IA2A.ipynb`, cujo objetivo foi estruturar, validar e analisar os indicadores essenciais para compreender os padrões de risco sanitário na Região Metropolitana de Belém (RMB). Essa etapa estabeleceu a base analítica utilizada posteriormente na modelagem preditiva e na geração das saídas destinadas ao dashboard no Looker Studio.

O escopo adotado incluiu os oito municípios da RMB — Belém, Ananindeua, Marituba, Benevides, Santa Bárbara do Pará, Santa Izabel do Pará, Castanhal e Barcarena — cobrindo o período de 2018 a 2025. A EDA integrou múltiplas fontes consolidadas na camada *gold*, principalmente `gold_features_ano.parquet`, complementada pelos indicadores do SNIS/SINISA (`snis_rmb_indicadores_v2.parquet`) e pelas estimativas populacionais do IBGE (`populacao.parquet`), todas harmonizadas segundo a configuração oficial em `rmb_municipios.csv`.

O notebook foi estruturado em etapas reprodutíveis, iniciando pela configuração do ambiente analítico, definição dos diretórios raiz utilizados pelo projeto (Bronze, Silver e Gold) e criação da pasta `dashboard/material_para_dashboard`, onde são armazenadas as saídas finais. Em seguida, foi realizado um diagnóstico inicial da camada *gold*, verificando número de registros, consistência das colunas e integridade temporal das séries 2018–2025.

A partir da leitura dos dados populacionais do IBGE, procedeu-se à normalização de códigos municipais e à verificação da completude dos oito municípios da RMB. Em seguida, foi executada a preparação analítica central, que consistiu na fusão da camada *gold* com dados complementares do SNIS para preenchimento de lacunas em indicadores de serviço. Essa etapa incluiu validações de cobertura anual e geração de alertas automáticos em caso de ausência de séries de qualidade da água (*gold_qualidade_agua*).

Com o conjunto consolidado, foram derivadas as principais métricas analíticas utilizadas no estudo, incluindo:

- imputação de valores ausentes por mediana municipal;
- criação de variáveis de ação, como `deficit_atendimento`, `deficit_tratamento`, `perdas_excesso` e `alerta_qualidade`;
- derivação de `chuva_total_mm_lag1`, incorporando defasagem temporal para análise de efeitos climáticos;
- cálculo do `score_priorizacao`, composto por pesos definidos para taxa hídrica (40%), déficits de tratamento (20%) e atendimento (15%), alertas de qualidade (15%) e perdas no sistema (10%);
- classificação dos municípios em categorias *Estável*, *Atenção* e *Crítico*;
- construção dos *drivers* de correlação, ranking de risco para 2025, estimativas de retorno por município e impactos elásticos de investimentos.

Além disso, foram produzidos indicadores agregados para o ODS Tracker, incluindo verificação automática das metas associadas às ODS 3 e 6. Entre os destaques obtidos na EDA, observou-se correlação positiva entre chuva total anual e taxa de internações ($p = 0,41$), relação negativa com investimento per capita em saúde ($p = -0,34$), desempenho crítico de Ananindeua em taxa média de internações e baixa capacidade de tratamento em Belém, além de gargalos persistentes em qualidade da água e tratamento de esgoto que impedem a aproximação das metas ODS em todos os anos avaliados.

Por fim, todas as métricas analíticas e indicadores derivados foram exportados para os arquivos `painel_prioridade.csv` e `ods_tracker.csv`, ambos com *timestamp* em UTC, constituindo as fontes oficiais do dashboard final. Essa etapa consolidou a EDA como componente essencial para a interpretação dos padrões territoriais, identificação de vulnerabilidades e fundamentação da modelagem preditiva subsequente.

5.3. Modelagem (Estatística/IA)

A etapa de modelagem preditiva foi desenvolvida com foco em estimar a taxa de internações por agravos hídricos (CID-10 A00–A09) por 10 mil habitantes, utilizando como preditores os indicadores consolidados de saneamento, qualidade da água, clima e finanças municipais. O escopo temporal compreende os anos de 2018 a 2025 para os oito municípios da Região Metropolitana de Belém (RMB), com o ano de 2025 reservado como *holdout* para avaliação temporal. Após a validação, o ano é reintegrado ao conjunto de treino para gerar previsões históricas completas.

A base utilizada para modelagem foi derivada do arquivo `gold_features_ano.parquet`, complementada por colunas provenientes do SNIS em casos de lacunas pontuais e harmonizada com a configuração municipal contida em `rmb_municipios.csv`, de modo a reproduzir integralmente o *analysis_df* gerado na etapa exploratória. Em seguida, selecionou-se a variável-alvo `internacoes_hidricas_10k`, enquanto 21 variáveis acionáveis — abrangendo cobertura de serviços, perdas, indicadores climáticos e despesas em saúde — foram definidas como *features* do modelo.

Três algoritmos foram testados em um pipeline padronizado composto por imputação mediana e padronização dos preditores: Regressão Linear, Lasso Regression e Random Forest Regressor. A validação foi realizada com *GroupKFold* ($k=4$), onde cada município constitui um grupo, evitando vazamento de informação entre cidades e garantindo robustez temporal e espacial. Entre os modelos avaliados, o Random Forest apresentou o melhor desempenho, com Mean Absolute Error (MAE) médio de aproximadamente 4,8 e Root Mean Squared Error (RMSE) de 6,3 nas validações cruzadas, além de estabilidade entre os *folds*.

A avaliação no *holdout* de 2025 indicou MAE de 4,18 e RMSE de 5,46, embora com R^2 negativo — comportamento esperado em séries curtas e com forte oscilação anual. Assim, o modelo é recomendado para interpretação de tendências relativas, e não para previsões pontuais absolutas. A análise de explicabilidade apontou população, despesa em saúde por habitante, percentual de amostras de água conformes e variáveis climáticas como os principais determinantes da taxa de internações. Adicionalmente, foram simuladas *elasticidades* para estimar o impacto médio de melhorias nos indicadores, como +5 p.p. em cobertura/tratamento ou +0,5 p.p. em investimentos.

Por fim, os resultados modelados deram origem aos arquivos `modelagem_metricas.csv`, `modelagem_importancias.csv`, `modelagem_previsoes.csv` e `modelagem_cenarios.csv`, que compõem a base do dashboard, abastecendo principalmente as páginas *Priorizar & Simular* e *Elasticidades*. Esses artefatos garantem a integração entre o processo analítico e a camada de visualização, permitindo que gestores explorem cenários e priorizações baseadas em dados e em inferência estatística.

5.4. Construção do Dashboard

A construção do dashboard no Google Looker Studio foi estruturada para integrar, de forma consistente, as bases derivadas das etapas de engenharia de dados e modelagem. As fontes utilizadas incluem tanto o dataset consolidado (Gold), responsável pelos indicadores históricos de saneamento, saúde, clima e finanças, quanto os artefatos gerados pelos modelos preditivos (previsões, importâncias e cenários). Para garantir desempenho e estabilidade, os arquivos CSV foram carregados em Google Sheets, permitindo consultas rápidas e maior integridade nas relações entre município, ano e indicadores.

A base principal utilizada na construção dos gráficos foi o arquivo `gold_features_ano.csv`, que reúne todos os indicadores padronizados no nível municipal (2018–2025). Essa estrutura alimentou séries temporais, mapas temáticos e resumos agregados da primeira e segunda páginas do dashboard. Em paralelo, o arquivo `modelagem_previsoes.csv` sustentou a página de Priorização, contendo as taxas previstas para cada município, os erros do modelo e a categorização de risco sanitário. Esses elementos permitiram a criação de rankings, mapas de calor e matrizes comparativas entre risco observado e risco estimado.

A funcionalidade de simulação (“what-if”) foi construída com base no arquivo `modelagem_cenarios.csv`, que contém as elasticidades estimadas pelo modelo Random Forest. Esse dataset possibilita projetar a variação na taxa de internações a partir de alterações hipotéticas nos indicadores de saneamento (ex.: aumento da cobertura de esgoto ou redução de perdas). A implementação do simulador foi realizada por meio de parâmetros ajustáveis no Looker Studio, vinculados a expressões que aplicam as elasticidades às previsões do modelo.

Por fim, a página de explicabilidade e ODS utilizou o arquivo `modelagem_importancias.csv` para apresentar os fatores que mais influenciaram o modelo, reforçando transparência e reprodutibilidade. Esse recurso foi combinado ao `ods_tracker.csv` (gerado na etapa de EDA), que acompanha o avanço dos municípios em relação às metas dos ODS 3 e 6. O resultado final é um painel interativo composto por quatro páginas — Visão Geral da RMB, Visão Municipal, Priorização e Simulador, e ODS Tracker — que consolida análises históricas, preditivas e exploratórias, fornecendo uma ferramenta robusta para suporte à decisão em políticas de saneamento e saúde ambiental.

6. Resultados e discussão

6.1. Resultados

A integração das bases SNIS/SINISA, SISAGUA, DATASUS, INMET, IBGE e SIOPS resultou em um dataset completo para os oito municípios da RMB no período 2018–2025, permitindo análises comparáveis e a construção de indicadores consolidados de saneamento, clima, saúde e finanças públicas. A análise exploratória mostrou que déficits de atendimento e tratamento, baixa conformidade da qualidade da água e perdas elevadas são fatores diretamente associados ao aumento das internações por doenças de veiculação hídrica (CID-10 A00–A09). O score de priorização sintetizou esses elementos, classificando os municípios entre Estável, Atenção e Crítico.

Os resultados evidenciaram padrões territoriais relevantes: Ananindeua, Belém e Santa Izabel do Pará apresentaram risco sanitário elevado em 2025, enquanto Castanhal e Barcarena exibiram condições mais favoráveis. A influência climática também se destacou, com chuva total anual (lag de 1 ano) apresentando correlação positiva com internações, indicando que períodos de alta precipitação antecedem picos de agravos hídricos. Investimentos em saúde, por sua vez, mostraram correlação negativa moderada, sugerindo efeito protetivo.

Na modelagem preditiva, três algoritmos foram avaliados: Regressão Linear, Lasso e Random Forest, utilizando pipelines padronizados e validação com GroupKFold para evitar vazamento entre municípios. O Random Forest apresentou o melhor desempenho, com MAE médio de 4,8 (cross-validation) e MAE de 4,18 no holdout de 2025. Apesar do R^2 negativo no holdout, as previsões se mantiveram dentro de erro absoluto reduzido, sendo adequadas para análise de tendências relativas.

As elasticidades derivadas do modelo mostraram que melhorias de +5 p.p. em qualidade da água, cobertura de abastecimento e tratamento de esgoto reduzem a taxa projetada de internações de forma consistente. O ODS Tracker evidenciou que cobertura de água é o indicador mais próximo das metas globais, enquanto tratamento de esgoto e conformidade da água permanecem críticos. Todos os resultados foram consolidados em arquivos .csv destinados ao dashboard para apoiar tomadas de decisão.

6.2. Discussão

Os resultados obtidos demonstram que a dinâmica das internações por doenças hídricas na RMB é fortemente condicionada por desigualdades estruturais de saneamento e pela interação entre clima e infraestrutura urbana. A correlação

entre chuva total, especialmente com defasagem temporal, e aumento das interações revela a vulnerabilidade das áreas urbanas a eventos extremos, reforçando a importância de políticas integradas de drenagem, controle de perdas e monitoramento da qualidade da água. Essa evidência dialoga com estudos que apontam o papel central dos ciclos hidrológicos amazônicos no agravamento dos riscos sanitários.

A modelagem preditiva reforça que, embora o cenário de curto horizonte (2018–2025) e o número reduzido de observações imponham limitações estatísticas, o modelo Random Forest oferece boa capacidade de identificar tendências e drivers estruturais. Seu uso é recomendado como instrumento de priorização, simulação e planejamento, evitando interpretação direta de valores absolutos de previsão. A forte influência de variáveis como população, despesas em saúde e conformidade da água evidencia que a capacidade municipal de gestão e os investimentos recorrentes exercem maior impacto do que aumentos marginais em outras categorias orçamentárias.

Outro ponto relevante é a coerência entre o score de priorização e os resultados da modelagem. Municípios classificados como críticos também apresentaram maiores taxas observadas e previstas de interações, além de déficits persistentes de saneamento. Isso valida o uso combinado da EDA e da modelagem para orientar tomadas de decisão, principalmente no contexto da COP-30, onde investimentos estruturantes ganharam visibilidade e urgência. A presença de gargalos em indicadores-chave reforça que intervenções incrementais podem ser insuficientes para reduzir o risco sanitário de forma significativa.

Por fim, a avaliação frente às metas dos ODS 3 e 6 evidencia avanços limitados e desigualdades internas entre os municípios da RMB. A persistência de status “Alerta” para tratamento de esgoto e qualidade da água indica necessidade de políticas de maior escala, com foco em equidade territorial e justiça socioambiental. A abordagem integrada, combinando dados; análises estatísticas; modelagem de IA e visualização interativa, oferece uma base técnica sólida para apoiar gestores públicos, pesquisadores e agentes comunitários na construção de estratégias coerentes para redução das desigualdades sanitárias na Amazônia.

6.3. Impacto Socioambiental do Projeto

A RMB concentra mais de 2 milhões de habitantes e enfrenta desafios históricos ligados à urbanização acelerada, precariedade de infraestrutura básica e

forte influência de fatores climáticos típicos da Amazônia, como chuvas intensas e redes drenantes frágeis. O projeto tem potencial de gerar impacto socioambiental direto ao identificar áreas mais vulneráveis, orientar investimentos públicos com maior retorno sanitário e apoiar intervenções capazes de reduzir internações evitáveis.

Além disso, o dashboard proposto cria uma ferramenta contínua de monitoramento, útil para gestores, pesquisadores, consórcios intermunicipais e comunidades urbanas periféricas. A solução é escalável para outras regiões que enfrentam desafios semelhantes, fortalecendo políticas públicas integradas de saneamento, vigilância da qualidade da água e saúde ambiental. Ao transformar dados brutos em inteligência aplicada, o projeto contribui para decisões mais eficazes, transparentes e alinhadas à sustentabilidade regional.

7. Referências

AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (ANA). Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH) – Painéis Atlas Águas e Atlas Esgotos. Brasília, DF: ANA/SNIRH, [s. d.]. Disponível em: <<https://www.snirh.gov.br/portal/snirh-1/paineis-de-indicadores>>. Acesso em: [3 nov. 2025].

BASE DOS DADOS. INMET/BDMEP – Espelho da base BDMEP no BigQuery. São Paulo: Base dos Dados, [s. d.]. Disponível em: <https://basedosdados.github.io/mais/access_data_bq/>. Acesso em: [3 nov. 2025].

BRASIL. Ministério da Saúde. DATASUS. SISAGUA – Controle Mensal – Demais Parâmetros. Brasília, DF: OpenDataSUS/CKAN, [s. d.]. Disponível em: <<https://opendatasus.saude.gov.br/dataset/sisagua-controle-mensal-demais-parametros>>. Acesso em: [3 nov. 2025].

BRASIL. Ministério da Saúde. DATASUS. SISAGUA – Controle Mensal – Parâmetros Básicos. Brasília, DF: OpenDataSUS/CKAN, [s. d.]. Disponível em: <<https://opendatasus.saude.gov.br/dataset/sisagua-controle-mensal-parametros-basicos>>. Acesso em: [3 nov. 2025].

BRASIL. Ministério da Saúde. DATASUS. SISAGUA – Vigilância – Demais Parâmetros. Brasília, DF: OpenDataSUS/CKAN, [s. d.]. Disponível em: <<https://opendatasus.saude.gov.br/dataset/sisagua-vigilancia-demais-parametros>>. Acesso em: [3 nov. 2025].

BRASIL. Ministério da Saúde. DATASUS. SIH/SUS – Autorizações de Internação Hospitalar (AIH) – Base Reduzida (RD). Brasília, DF: DATASUS/FTP, [s. d.].

Disponível em: ftp://ftp.datasus.gov.br/dissemin/publicos/SIHSUS/200801_/Dados/. Acesso em: [3 nov. 2025].

BRASIL. Ministério da Saúde. OpenDataSUS. SISAGUA – Controle Mensal – Infraestrutura Operacional. Brasília, DF: OpenDataSUS/CKAN, [s. d.]. Disponível em:

<<https://opendatasus.saude.gov.br/dataset/sisagua-controle-mensal-infraestrutura-operacional>>. Acesso em: [3 nov. 2025].

BRASIL. Ministério da Saúde. SIOPS. Indicadores Municipais de Receitas e Despesas em Saúde. Brasília, DF: SIOPS/MS, [s. d.]. Disponível em: <<https://www.gov.br/saude/pt-br/acesso-a-informacao/siops/indicadores>>. Acesso em: [3 nov. 2025].

BRASIL. Ministério das Cidades. Sistema Nacional de Informações sobre Saneamento (SNIS) – Diagnósticos, Planilhas e Indicadores de Água e Esgoto. Brasília, DF: Ministério das Cidades, [s. d.]. Disponível em: <<https://www.gov.br/cidades/pt-br/acesso-a-informacao/acoes-e-programas/saneamento/snis>>. Acesso em: [3 nov. 2025].

CANAL RURAL. Água potável e saneamento básico são os maiores desafios da Amazônia, diz ministro. São Paulo: Canal Rural, [s. d.]. Disponível em: <<https://www.canalrural.com.br/cop30/agua-potavel-e-saneamento-basico-sao-os-maiores-desafios-da-amazonia-diz-ministro/>>. Acesso em: [20 nov. 2025].

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Cidades e Estados – Metadados Municipais e Códigos Oficiais do IBGE. Rio de Janeiro: IBGE, [s. d.]. Disponível em: <<https://www.ibge.gov.br/cidades-e-estados>>. Acesso em: [3 nov. 2025].

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). SIDRA – Tabela 6579: População residente – Estimativas. Rio de Janeiro: IBGE/SIDRA, [s. d.]. Disponível em: <<https://apisidra.ibge.gov.br/values/t/6579>>. Acesso em: [3 nov. 2025].

INSTITUTO NACIONAL DE METEOROLOGIA (INMET). BDMEP – Banco de Dados Meteorológicos para Ensino e Pesquisa. Brasília, DF: INMET, [s. d.]. Disponível em: <<https://bdmep.inmet.gov.br/>>. Acesso em: [3 nov. 2025].

INSTITUTO PÓLIS. Racismo ambiental e injustiça climática: o papel das cidades no enfrentamento às mudanças do clima e no combate às desigualdades socioterritoriais. São Paulo, 2025.

METRÓPOLES. Governo do Pará entrega a maior Estação de Tratamento de Esgoto do estado. Brasília, DF: Metrôpoles, [s. d.]. Disponível em:

<<https://www.metropoles.com/conteudo-especial/governo-do-para-entrega-a-maior-e-stacao-de-tratamento-de-esgoto-do-estado>>. Acesso em: [12 out. 2025].