USO DE DADOS PÚBLICOS E MODELAGEM DE INCERTEZA NA GESTÃO SUSTENTÁVEL DA EXPOSIÇÃO AMBIENTAL A FÁRMACOS NO NORDESTE BRASILEIRO

USE OF PUBLIC DATA AND UNCERTAINTY MODELING IN THE SUSTAINABLE MANAGEMENT OF ENVIRONMENTAL EXPOSURE TO PHARMACEUTICALS IN NORTHEASTERN BRAZIL

SIMAO PEDRO DE OLIVEIRA URPIA
UNIVERSIDADE FEDERAL DA BAHIA-UFBA

KARLA PATRICIA SANTOS OLIVEIRA RODRIGUEZ ESQUERRE
UNIVERSIDADE FEDERAL DA BAHIA-UFBA

LUCIANO MATOS QUEIROZ
UNIVERSIDADE FEDERAL DA BAHIA-UFBA

Comunicação:
O XIV SINGEP foi realizado em conjunto com a 14th Conferência Internacional do CIK (CYRUS Institute of Knowledge) e com a SPARK (Society for Promoting Academic Research and Knowledge), em formato híbrido, com sede presencial na UNINOVE - Universidade Nove de Julho, no Brasil.

USO DE DADOS PÚBLICOS E MODELAGEM DE INCERTEZA NA GESTÃO SUSTENTÁVEL DA EXPOSIÇÃO AMBIENTAL A FÁRMACOS NO NORDESTE BRASILEIRO

Objetivo do estudo
Estimar e analisar as Concentrações Ambientais Previstas de dez princípios ativos nas nove capitais nordestinas, entre 2014 e 2019, incorporando fontes de incerteza por simulação de Monte Carlo, para identificar substâncias e localidades prioritárias ao monitoramento ambiental.

Relevância/originalidade
Apresenta uma abordagem replicável e de baixo custo que integra dados públicos regulatórios, parâmetros de excreção e indicadores de saneamento, propagando incertezas até as estimativas de exposição e relacionando tecnologia de dados, gestão ambiental e sustentabilidade urbana.

Metodologia/abordagem
Foram processados registros do SNGPC e indicadores do SNIS referentes a 2014–2019. O modelo adaptado da EMA calculou CAP para 6.480 cenários, com 100.000 iterações de Monte Carlo e distribuições associadas às variabilidades metrológica, normativa e metabólica.

Principais resultados
Amoxicilina e cefalexina apresentaram as maiores CAP. Dos 6.480 cenários, 89,38% permaneceram acima de 0,01 µg/L mesmo no limite inferior da faixa simulada. Natal e João Pessoa foram identificadas como localidades prioritárias para monitoramento complementar.

Contribuições teóricas/metodológicas
A integração das CAP com a propagação de incertezas demonstra que registros de comercialização podem sustentar triagens comparativas, desde que interpretados como estimativas conservadoras de exposição. A abordagem também evidencia a influência conjunta da carga farmacêutica e do volume disponível para diluição.

Contribuições sociais/para a gestão
Os resultados apoiam a seleção de fármacos, capitais e períodos para campanhas analíticas, auxiliando organizações públicas na alocação de recursos. A aplicação contribui para o planejamento do saneamento e do monitoramento ambiental associado à sustentabilidade urbana e aos ODS 6 e 11.

Palavras-chave: Contaminantes emergentes, Fármacos, Concentração Ambiental Prevista, Gestão ambiental, Monte Carlo.

USE OF PUBLIC DATA AND UNCERTAINTY MODELING IN THE SUSTAINABLE MANAGEMENT OF ENVIRONMENTAL EXPOSURE TO PHARMACEUTICALS IN NORTHEASTERN BRAZIL

Study purpose
To estimate and analyze the Predicted Environmental Concentrations of ten active pharmaceutical ingredients in nine Northeastern Brazilian capitals from 2014 to 2019, incorporating uncertainty through Monte Carlo simulation to identify substances and locations requiring priority environmental monitoring.

Relevance / originality
Provides a replicable, low-cost approach integrating public regulatory data, excretion parameters, and sanitation indicators, propagating uncertainty into exposure estimates and connecting data technology, environmental management, and urban sustainability.

Methodology / approach
SNGPC records and SNIS indicators for 2014–2019 were processed. The EMA-adapted model calculated PECs for 6,480 scenarios, using 100,000 Monte Carlo iterations and probability distributions representing metrological, regulatory, and metabolic variability.

Main results
Amoxicillin and cephalexin produced the highest PECs. Of the 6,480 scenarios, 89.38% remained above 0.01 µg/L even at the lower limit of the simulated range. Natal and João Pessoa were identified as priority locations for complementary monitoring.

Theoretical / methodological contributions
Integrating PECs with uncertainty propagation demonstrates that sales records can support comparative screening when interpreted as conservative exposure estimates. The approach also reveals the combined influence of pharmaceutical load and the volume available for dilution.

Social / management contributions
Results support the selection of pharmaceuticals, capitals, and periods for analytical campaigns, helping public organizations allocate resources. The application contributes to sanitation and environmental monitoring planning connected to urban sustainability and SDGs 6 and 11.

Keywords: Emerging contaminants, Pharmaceuticals, Predicted Environmental Concentration, Environmental management, Monte Carlo

USO DE DADOS PÚBLICOS E MODELAGEM DE INCERTEZA NA GESTÃO SUSTENTÁVEL DA EXPOSIÇÃO AMBIENTAL A FÁRMACOS NO NORDESTE BRASILEIRO

1 INTRODUÇÃO

Os poluentes emergentes englobam uma vasta gama de substâncias químicas sintéticas, naturais ou microrganismos que apresentam potencial toxicidade à ecologia e à saúde humana, mas cujos efeitos adversos e presença no ambiente ainda são pouco conhecidos (UNESCO, 2023). Trata-se de centenas de compostos frequentemente detectados nos diferentes compartimentos ambientais (ar, água e solo), sendo oriundos tanto de fontes naturais, quanto de atividades antrópicas, a exemplo dos efluentes domésticos, industriais, hospitalares e das práticas agropecuárias (Montagner et al., 2017). Apesar dos riscos que representam aos ecossistemas, essas substâncias caracterizam-se por não integrarem as normativas e legislações de controle ambiental vigentes, o que as exclui dos programas de monitoramento de rotina dos órgãos de meio ambiente e saúde (Moreira & Gonçalves, 2013). Por conta disso, essa categoria de contaminantes passou a ser importante objeto de estudo, principalmente em zonas urbanas com elevada densidade ocupacional (Montagner et al., 2017).

Além da contribuição à avaliação ambiental, a integração de registros regulatórios e indicadores de saneamento constitui uma abordagem de apoio à decisão baseada em dados públicos. Ao direcionar substâncias e localidades para campanhas analíticas, essa estratégia se relaciona aos Objetivos de Desenvolvimento Sustentável estabelecidos pela Organização das Nações Unidas (2015), com destaque para o ODS 6, voltado à água potável e ao saneamento, e o ODS 11, relacionado a cidades e comunidades sustentáveis. Nesse contexto, o uso de rotinas computacionais replicáveis amplia a capacidade de planejamento do monitoramento em cenários de restrição financeira e analítica.

Portanto, o presente trabalho se alinha a essa abordagem, focando na identificação de fármacos e localidades prioritárias e na avaliação da incerteza dos dados de massa vendida dos dez princípios ativos mais comercializados nas nove capitais brasileiras da região Nordeste entre 2014 e 2019. Utilizam-se dados de vendas do SNGPC e indicadores de saneamento do Sistema Nacional de Informações sobre Saneamento (SNIS), propondo-se um panorama analítico replicável para outras bases de dados.

REFERENCIAL TEÓRICO

Tratando-se especificamente da contaminação ambiental por fármacos, esta pode ocorrer de forma direta e indireta, o que influencia significativamente a abordagem do problema. A contaminação direta é causada, por exemplo, por efluentes da indústria farmacêutica gerados nas operações de limpeza ou no descarte inadequado de produtos fora de padrões de qualidade (Sales et al., 2012). Já a contaminação indireta, considerada a principal via de contaminação ambiental, pode ocorrer nos esgotos e provém, principalmente, da excreção e descarte inadequado de medicamentos – com prazo de validade vencido (Sales et al., 2012).

A maioria das Estações de Tratamento de Esgoto (ETEs) existentes não foi projetada para remover micropoluentes como os fármacos. Patel et al. (2019) explicam que a eficiência de remoção de um fármaco depende de suas propriedades físico-químicas (como solubilidade e hidrofobicidade) e das condições operacionais da ETE. Para diversos medicamentos, a diferença entre as concentrações detectadas no esgoto bruto e no esgoto tratado não é estatisticamente significativa, indicando que, mesmo após o tratamento, a maioria dos contaminantes desse grupo não é efetivamente eliminada (Valdez-Carrillo et al., 2020; Couto et al., 2019).

Nesse cenário, Welch et al. (2022) apontam que o cálculo de Concentrações Ambientais Previstas (CAP), a partir de dados de vendas, pode ser usado como uma ferramenta estratégica e de baixo custo para uma triagem inicial. O estudo demonstra que essa abordagem, quando refinada com as taxas de excreção, produz estimativas razoáveis da exposição ambiental, permitindo tanto a priorização de substâncias quanto de regiões para um monitoramento mais aprofundado.

Conforme observado por Welch et al. (2022), na Noruega, a utilização de dados de atacado permitiu capturar o volume total de fármacos destinados tanto a farmácias quanto a hospitais e instituições de saúde, revelando que as vendas de analgésicos comuns, como paracetamol e ibuprofeno, chegam a ser 1,5 a 2,3 vezes superiores ao que é registrado apenas em prescrições. No contexto brasileiro, essa distinção é fundamental, pois reforça que o monitoramento da movimentação física de medicamentos pelo SNGPC oferece uma visão mais fidedigna da carga potencial de poluentes do que sistemas baseados exclusivamente em registros médicos.

A utilização de dados oficiais para fins de pesquisa exige uma reflexão sobre o conceito de "erro". Nos bancos de dados como o SNGPC, o dado não é gerado por um experimento controlado, mas por um ato de escrituração (Caetano et al. 2025). A literatura aponta que a qualidade desses dados é frequentemente comprometida por falhas de preenchimento, ausência de padronização na descrição de apresentações e lacunas demográficas. Contudo, quando o objeto de estudo é a massa do fármaco, a incerteza passa também a incorporar a dúvida sobre o valor verdadeiro da massa contida em cada registro de venda.

Alguns estudos analisaram a qualidade dos dados registrados no SNGPC. A primeira avaliação sistemática da base foi realizada por Caetano et al. (2025) e identificou vulnerabilidades importantes na qualidade e na consistência dos registros, que exigem uma etapa de pré-processamento. Entre os problemas encontrados, destaca-se a falta de padronização do campo "DESCRICAO_APRESENTACAO". Por ser um campo de texto aberto, ele agrega dose, forma farmacêutica e quantidade de maneira desestruturada, gerando grande variabilidade de registros para um mesmo produto. Para a amoxicilina, por exemplo, foram encontradas 227 descrições diferentes, embora existissem apenas 39 apresentações comerciais distintas (CAETANO et al., 2025).

Em uma análise comparativa de dados de risco ambiental publicamente disponíveis, Ahkola et al. (2024) identificaram inconsistências significativas nos valores de PNEC para um mesmo fármaco reportados por diferentes fontes, como o banco de dados sueco e os próprios fabricantes. Para a amoxicilina, por exemplo, a diferença no PNEC alcançou sete vezes, uma variação atribuída ao uso de diferentes espécies de algas e tempos de exposição nos ensaios ecotoxicológicos.

O estudo de Ahkola et al. (2024) assim como o de Giebułtowicz e Nałęcz-Jawecki (2014) validaram também os PECs calculados a partir de dados de vendas com as Concentrações Ambientais Medidas (MECs) em águas superficiais. Os resultados demonstraram que, de modo geral, os valores de PEC foram 10 vezes ou mais, superiores aos valores de MEC, o que indica que os modelos baseados no Cálculo de PEC, de modo geral, apresentam previsões conservadoras. Os estudos apontam que a exclusão dos metabólitos oriundos da degradação dos fármacos nos cálculos de PEC é possivelmente uma das principais razões para a superestimação das concentrações ambientais.

2 METODOLOGIA

Para viabilizar a análise do cenário de exposição ambiental, avaliaram-se registros do SNGPC (2014–2019) nas nove capitais do Nordeste brasileiro. O estudo selecionou os dez princípios ativos com maior massa comercializada: amoxicilina, azitromicina, cefalexina, clonazepam, cloridrato de ciprofloxacino, dexametasona, fosfato, levofloxacino, metronidazol

e sulfametoxazol. O critério de seleção baseou-se na premissa de que maiores volumes de vendas representam maior carga potencial nos efluentes. Os cerca de 90 milhões de registros originais foram processados no ambiente R, passando por etapas de padronização textual das apresentações comerciais e agregação dos dados validados por município, mês e ano.

O fosfato integra este recorte por sua relevância para a análise comparativa das capitais nordestinas e para a interpretação conjunta dos dez princípios ativos apresentados nos resultados gráficos.

2.1 Tratamento e consolidação dos dados do SNGPC

Para obter projeções de exposição e critérios de priorização confiáveis, os dados de referência devem ser consistentes. Portanto, inicialmente realizou-se um extenso tratamento dos dados extraídos do SNGPC.

Os dados do SNGPC são disponibilizados no portal do governo federal e armazenados em 82 arquivos no formato .csv. Cada arquivo reúne os dados escriturados referentes a um mês dos sete anos em que o registro ocorreu de forma obrigatória. Esses bancos de dados estão originalmente divididos em 15 campos, com os seguintes títulos: “ANO_VENDA”, “MES_VENDA”, “UF_VENDA”, “MUNICIPIO_VENDA”, “PRINCIPIO_ATIVO”, “DESCRICAO_APRESENTACAO”, “QTD_VENDIDA”, “UNIDADE_MEDIDA”, “CONSELHO_PRESCRITOR”, “TIPO_RECEITUARIO”, “UF_CONSELHO_PRESCRITOR”, “CID10”, “SEXO”, “IDADE” e “UNIDADE_IDADE” (Agência Nacional de Vigilância Sanitária, 2020).

Neste estudo, um grande desafio foi manipular e visualizar grandes volumes de dados. Cada arquivo .csv continha inicialmente cerca de 6.000.000 (seis milhões) de linhas, totalizando aproximadamente 90.000.000 (noventa milhões) de registros. No entanto, o Microsoft Excel® apresentou limitações devido ao número de linhas suportadas. Para otimizar a manipulação e a visualização, foi realizada uma série de operações nos bancos de dados, reduzindo seu tamanho sem comprometer informações relevantes. Para isso, todos os arquivos foram convertidos de .csv para .RData; estes são processados pelo R de forma mais rápida e ocupam cerca de 23 vezes menos espaço na memória do dispositivo que os mesmos bancos em .csv.

O processo de importação foi automatizado por meio de uma função de leitura que iterou sobre cada arquivo, carregou seu conteúdo em um ambiente isolado e extraiu o primeiro objeto do tipo data.frame encontrado. Para assegurar a compatibilidade entre arquivos com esquemas de colunas distintos, todas as variáveis foram convertidas ao tipo caractere antes da consolidação, evitando conflitos de tipos durante a união. Os objetos resultantes foram concatenados por meio da função bind_rows, que preenche automaticamente com NA as colunas ausentes em determinados arquivos. Ao final, o banco unificado foi salvo em disco como banco_completo.RData para preservar o estado consolidado.

Devido à complexidade das informações presentes no banco de dados, foi necessário fazer algumas considerações e manipulações no banco original. Nesse sentido, os medicamentos com mais de um princípio ativo foram tratados como possuindo apenas o princípio ativo principal. Assim, a coluna “PRINCIPIO_ATIVO” foi usada como base para extrair a informação referente ao primeiro princípio ativo apresentado que foi armazenada na coluna “PRINCIPIO_ATIVO_PRINCIPAL”.

A fim de viabilizar o cálculo da massa de princípio ativo por transação, o banco de dados foi enriquecido com parâmetros de dosagem extraídos de uma planilha auxiliar (BANCO_AUX.xlsx). Este dicionário, elaborado manualmente a partir do processamento da coluna DESCRICAO_APRESENTACAO, permitiu converter as descrições textuais em três vetores de informação fundamentais: a dosagem por unidade (DI), a quantidade de unidades por embalagem (NUI) e o volume total de diluição para soluções líquidas (VI).

A integração dos dados foi realizada por meio de um left join, utilizando a coluna de

descrição original como chave de ligação para garantir a preservação de todos os registros do banco principal. Previamente à junção, a tabela auxiliar foi tratada para manter apenas a primeira ocorrência de cada apresentação, eliminando o risco de duplicações redundantes. O banco resultante, denominado "dados_combinados", constituiu a base analítica consolidada para as etapas subsequentes.

Após a consolidação dos registros de venda, foram incorporadas as bases auxiliares necessárias à estimativa da exposição.

2.2 Bases de dados auxiliares

A caracterização da diluição nas capitais baseou-se em dados do Sistema Nacional de Informações sobre Saneamento (SNIS). Para estimar o volume de diluição na rede de esgotamento, utilizou-se o indicador AG010 (Volume de água consumida). O Índice de Hidrometração foi incorporado como parâmetro para o cálculo da incerteza sobre esse volume, integrando a variabilidade da micromedição à simulação de Monte Carlo.

Devido à inexistência de repositórios centralizados para os fármacos analisados, elaborou-se uma base de dados via levantamento bibliográfico. Este banco consolida a fração de excreção do princípio ativo inalterado e os limites de variação percentual desta fração necessários para o modelo, permitindo conversão de apresentações comerciais em massa de princípio ativo excretado. Como os registros de venda não equivalem perfeitamente ao consumo local e o descarte inadequado não foi modelado, adotou-se o cenário conservador de maior impacto ambiental, que considera 100% de consumo para o que foi vendido.

A fração de excreção humana foi parametrizada individualmente para cada princípio ativo com base nos valores centrais e nas amplitudes extraídos da literatura farmacocinética. Em cada iteração, amostrou-se uma distribuição uniforme limitada ao intervalo específico do princípio ativo e ao domínio físico entre zero e um. A escolha uniforme representa ausência de informação suficiente sobre a forma da distribuição dentro dos limites publicados; portanto, esses intervalos devem ser entendidos como premissas de entrada, e não como distribuições populacionais estimadas para as capitais.

2.3 Cálculo da CAP

Com base em uma adaptação do modelo para cálculo da CAP estabelecido pela Agência Europeia de Medicamentos (European Medicines Agency, 2006), realizou-se o cálculo do valor de CAP de referência em conjunto com a respectiva incerteza associada.

As bases de dados relativas às vendas de medicamentos (SNGPC), aos indicadores de saneamento (SNIS) e aos parâmetros farmacocinéticos (Excreção) foram carregadas e submetidas a diagnósticos de integridade para a identificação de valores não conversíveis ou inconsistentes. Para a base SNGPC, foram validadas as colunas de massa total, quantidade vendida, ano e mês; no SNIS, inspecionaram-se o volume de esgoto, a população atendida e o índice de hidrometração; e, na base de Excreção, verificaram-se a fração de excreção e a coluna referente a variabilidade. Após a validação, todos os registros foram convertidos ao formato numérico, assegurando a consistência para os cálculos de massa, a integração ao modelo computacional e a posterior propagação de incertezas via simulação de Monte Carlo.

Para o cálculo das variáveis finais, foi criada uma rotina computacional para computar a massa total vendida. Utilizou-se a lógica de propagação para multiplicar a quantidade vendida pela massa de princípio ativo de cada unidade, conforme definido nos modelos apresentados pelas equações 3 e 4. Esse procedimento inicial foi essencial para converter os registros de venda em dados de massa total por registro de venda, permitindo a subsequente avaliação da incerteza segundo as diretrizes do GUM (Associação Brasileira de Normas Técnicas [ABNT] & Instituto Nacional de Metrologia, Qualidade e Tecnologia [Inmetro], 2008).

A determinação da $Massa\,Vendida$ baseou-se nas equações 2, 3 e 4 inter-relacionadas, aplicadas a cada um dos "n" registros da base de dados, e estruturadas para permitir o cálculo

final da CAP pela equação 1. As grandezas foram classificadas em termos de sua dependência metrológica e fluxo de informação.

$$CAP = \frac{Massa\,Vendida\ \times F_{Excreção} \times F_{consumo} \times F_{conversão}}{Volume \times Coeficiente\ de\ retorno} \quad \textbf{(eq. 1)}$$

$$Massa\,Vendida = \sum_{i=1}^{n} MVTI \quad \text{(eq. 2)}$$

$$MVTI = QTDI \times MPUI \quad \text{(eq. 3)}$$

$$MPUI = DI\ \times NUI\ \times VI \quad \text{(eq. 4)}$$

Sendo: CAP: Concentração Ambiental Prevista; $Massa\,Vendida$: Massa Total Vendida por cidade e recorte temporal; $F_{Excreção}$: Porcentagem de princípio ativo que é excretada inalterada após o consumo humano ; $F_{consumo}$: Porcentagem de medicamentos comprados que de fato foram consumidos; $F_{conversão}$: Fator de conversão de unidades kg para µg; $Volume$: Volume de água consumida por cidade e recorte temporal ; $Coeficiente\ de\ retorno$: Fator padrão de 0,8 (ABNT, 1986) ; MTVI: Massa Total de princípio ativo Vendida por registro; MPUI: Massa princípio ativo por unidade vendida; QTDI: Quantidade de caixas ou frascos vendida por registro; n: Quantidade total de registros; DI: Dosagem por caixa ou frasco; NUI: Unidades por caixa (em caso de comprimido ou cápsula); VI: Volume total em que a dosagem está diluída (em caso de solução líquida);

Grandezas	Unidade	Dependência Metrológica				Fluxo de Informação			
		eq 1	eq 2	eq 3	eq 4	eq 1	eq 2	eq 3	eq 4
$Concentração$	Ug/l	DPD	NA	NA	NA	Saída	NA	NA	NA
Volume	m³	IND	NA	NA	NA	Entrada	NA	NA	NA
$Coeficiente\ de\ retorno$	NA	IND	NA	NA	NA	Entrada	NA	NA	NA
$F_{excreção}$	NA	IND	NA	NA	NA	Entrada	NA	NA	NA
$F_{consumo}$	NA	IND	NA	NA	NA	Entrada	NA	NA	NA
$F_{conversão}$	NA	IND	NA	NA	NA	Entrada	NA	NA	NA
$Massa\ Vendida$	kg	IND	DPD	NA	NA	Entrada	Saída	NA	NA
MTVI	kg	NA	IND	DPD	NA	NA	Entrada	Saída	NA
QTDI	NA	NA	NA	IND	NA	NA	NA	Entrada	NA
MPUI	kg	NA	NA	IND	DPD	NA	NA	Entrada	Saída
DI	kg/m³ ou kg	NA	NA	NA	IND	NA	NA	NA	Entrada
NUI	NA	NA	NA	NA	IND	NA	NA	NA	Entrada
VI	m³	NA	NA	NA	IND	NA	NA	NA	Entrada

Quadro 1. *Classificação das grandezas quanto a dependência metrológica e ao fluxo de informação.*

Nota: NA – Não se aplica; DPD – Dependente; IND – Independente.

2.4 Identificação e quantificação das fontes de incerteza

As fontes de incerteza foram identificadas em 3 grupos principais, conforme apresentado na Quadro 2. Cada grandeza do modelo foi avaliada quanto à sua origem (normativa, metrológica ou metabólica) e ao tipo de variabilidade associada, permitindo a parametrização apropriada de suas distribuições de probabilidade. A seguir descreve-se a justificativa técnica para cada componente de incerteza incorporada ao modelo.

Variável	Incerteza	Função de Densidade de Probabilidade (PDF)	Justificativa
DI	±15,0 %	Distribuição Triangular	Limite de variabilidade de produção (Agência Nacional de Vigilância Sanitária, 2019)
Volume	Variável[i]	Distribuição Retangular	Índice de Hidrometração (SNIS) por capital e ano
$F_{excreção}$	Específica[ii]	Distribuição Retangular	Variabilidade metabólica reportada nas bulas dos medicamentos

Quadro 2. *Fontes de incerteza incorporadas ao modelo*

A propagação das incertezas foi realizada por meio do Método de Monte Carlo (MMC), seguindo as recomendações do Suplemento 1 do Joint Committee for Guides in Metrology (2008). As concentrações foram simuladas via iterações com parâmetros perturbados e, posteriormente, agregadas por grupo (ano, mês, município e fármaco). A distribuição de saída foi caracterizada pela mediana, desvio padrão, incerteza relativa e Intervalo de Confiança de 95 % (quantis 2,5 % e 97,5 %).

2.5 Critérios de classificação

Os resultados estatísticos foram confrontados com o valor de referência de 0,01 µg/L adotado internacionalmente para recursos hídricos (European Medicines Agency, 2006). Com base na abrangência do intervalo de confiança (IC) de 95%, o algoritmo classificou automaticamente os resultados segundo a posição do intervalo em relação ao referencial. A categoria “Não ultrapassa” corresponde aos casos em que o limite superior não excede 0,01 µg/L; “Risco por precaução”, aos casos em que o intervalo cruza esse valor; e “Ultrapassa confirmado”, aos casos em que até o limite inferior o excede.

3 ANÁLISE DOS RESULTADOS E DISCUSSÕES

3.1 Qualidade e consistência da base analítica

A completude das variáveis essenciais foi elevada. QTD_VENDIDA, Massa_Total_MG_UI, ANO_VENDA, MES_VENDA, MUNICIPIO_VENDA e PRINCIPIO_ATIVO não apresentaram valores ausentes na saída final. A variável Vol_ou_Qtd_X, que de fato não é necessária para algumas apresentações específicas, apresentou 6.499 ausências, equivalentes a 0,02% dos registros, mantendo 99,98% de valores válidos.

A elevada completude observada nas variáveis de identificação e quantificação física usadas no estudo deve-se ao caráter compulsório do preenchimento desses campos pelas farmácias privadas para a homologação das transmissões eletrônicas. Esse padrão contrasta com a alta incidência de dados ausentes usualmente relatada para variáveis de natureza clínica e demográfica em sistemas de informação em saúde. Conforme apontam Ferreira et al. (2023), o SNGPC apresenta taxas críticas de omissão em campos não obrigatórios, como a classificação CID-10 e a idade do paciente, que chegam a superar 98% de valores nulos em determinadas análises. A coexistência desses padrões evidencia que a qualidade da base não é homogênea, mas condicionada à função regulatória e administrativa de cada variável.

No presente estudo, os indexadores de princípio ativo e apresentação comercial, por estarem vinculados à escrituração sanitária e ao banco de registros da Anvisa, apresentaram consistência suficiente para sustentar o cálculo da carga de massa comercializada (Caetano et al., 2025). Assim, embora o SNGPC possua limitações relevantes para análises clínicas ou epidemiológicas, sua utilização para a estimativa da exposição ambiental potencial mostra-se metodologicamente adequada nas variáveis diretamente empregadas nos cálculos. A elevada completude reduz a necessidade de imputação de dados e limita uma fonte importante de incerteza no modelo; desse modo, a etapa seguinte concentra-se na avaliação e no tratamento de valores extremos.

3.2 Panorama global do risco e triagem de contaminantes

A viabilidade da modelagem matemática utilizando dados de comercialização para a estimativa de concentrações ambientais, conforme proposto neste trabalho, encontra suporte nos resultados de Welch et al. (2022), que confirmam a robustez do uso de registros de vendas em nível nacional como ferramenta estratégica de triagem (screening).

As distribuições das cinco macrorregiões apresentaram assimetria à direita e caudas longas. Por essa razão, a comparação foi baseada na mediana, no intervalo entre Q1 e Q3 e nos percentis 90 e 95, sempre acompanhados pelo número de capitais e de grupos. A Região Norte apresentou a maior mediana, 0,1936 µg/L, seguida pelo Nordeste, com 0,1758 µg/L. Esses valores são sínteses descritivas das combinações incluídas e não demonstram diferenças causais entre regiões.

4 Tabela 3. *Síntese robusta das CAP por macrorregião*

Macrorregião	Capitais	n de grupos	Mediana (Q1–Q3), µg/L	P95 (µg/L)	Máximo (µg/L)	Avaliação complementar
Norte	7	4.536	0,1936 (0,0475–0,8140)	2,5787	78,3563	82,16%
Nordeste	9	5.832	0,1758 (0,0386–0,6119)	2,2788	213,3030	88,32%
Centro-Oeste	4	2.592	0,1075 (0,0264–0,4346)	1,5879	19,4563	87,38%
Sudeste	4	2.592	0,1018 (0,0265–0,3986)	2,2486	9,8768	88,00%
Sul	3	1.944	0,1429 (0,0359–0,4220)	1,9315	135,3430	87,81%

Fonte: elaboração própria a partir das simulações de Monte Carlo já concluídas.

No nordeste a análise evidenciou que a Amoxicilina e a Cefalexina são os dois medicamentos com maior CAP. Geograficamente, os maiores níveis ocorreram em Natal, com medianas de CAP de 1,168 µg/L (IC 95%: 0,835 – 1,705 µg/L) para Amoxicilina e 0,778 µg/L (IC 95%: 0,556 – 1,135 µg/L) para Cefalexina. João Pessoa apresentou o segundo pior cenário, com 0,825 µg/L (IC 95%: 0,589 – 1,200 µg/L) e 0,540 µg/L (IC 95%: 0,386 – 0,785 µg/L), respectivamente. A análise dos intervalos de confiança confirmou a magnitude da exposição ambiental: mesmo na capital com o menor impacto estimado (Salvador, com limite inferior do IC de 0,143 µg/L para Amoxicilina), as concentrações superam em mais de 10 vezes a referência ecotoxicológica de 0,01 µg/L.

A distribuição de concentração não mostra relação com os dados demográficos. Os valores máximos (picos) ocorrem em João Pessoa (833.932 habitantes) e Natal (751.300 habitantes), enquanto Salvador (2.417.678 habitantes) e Fortaleza (2.428.708 habitantes) apresentam medianas diferentes entre si. Isso indica, para esse conjunto de dados, que o tamanho populacional, determinante no comportamento das vendas de fármacos (Couto et al., 2019), não se traduz automaticamente em concentrações ambientais proporcionais.

4.1 Comparação com modelos e evidências da literatura

Em uma análise comparativa de dados de risco ambiental publicamente disponíveis, Ahkola et al. (2024) identificaram inconsistências significativas nos valores de PNEC para um mesmo fármaco reportados por diferentes fontes, como o banco de dados sueco e os próprios fabricantes. Para a amoxicilina, por exemplo, a diferença entre os valores de PNEC chegou a sete vezes, variação atribuída ao uso de diferentes espécies de algas e tempos de exposição nos ensaios ecotoxicológicos.

Os estudos de Ahkola et al. (2024) e de Giebułtowicz e Nałęcz-Jawecki (2014) também compararam as PEC calculadas a partir de dados de vendas com as concentrações ambientais medidas (MEC) em águas superficiais. De modo geral, os valores de PEC foram dez vezes ou mais superiores aos valores de MEC, indicando que os modelos baseados no cálculo de PEC tendem a produzir previsões conservadoras. Os estudos apontam que a exclusão dos metabólitos oriundos da degradação dos fármacos nos cálculos de PEC é possivelmente uma das principais razões para a superestimação das concentrações ambientais.

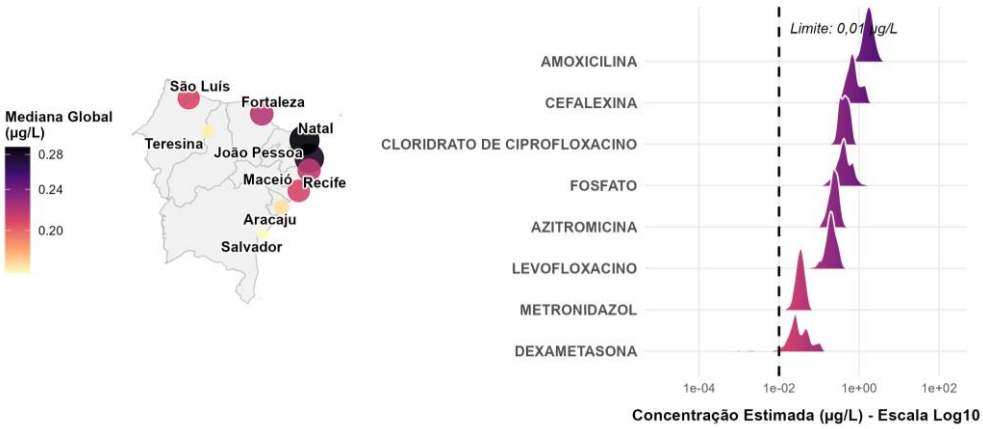


Figura 1. *Panorama de Risco Ambiental no Nordeste Brasileiro*

4.2 A influência da incerteza na modelagem ambiental

A modelagem estocástica gerou 6.480 cenários, combinando cada um dos meses dos seis anos de dados, nove cidades e dez princípios ativos, categorizados conforme o comportamento do intervalo de confiança (IC 95%) frente ao valor de referência de 0,01 µg/L (Quadro 3). Os resultados mostram uma carga elevada: 89,38% dos cenários (5.792

ocorrências) foram classificados na categoria "Ultrapassa confirmado", caso em que a concentração estimada supera o valor de referência mesmo no limite inferior do IC.

A frequência muito reduzida da categoria "Risco por precaução" (0,08%) evidencia que os valores estimados são suficientemente elevados ou reduzidos para que a incerteza do modelo, concentrada em cerca de 20% para a maioria dos dados, não altere o diagnóstico. A incerteza simulada a partir das lacunas nos dados originais, portanto, não invalida as conclusões.

Relação com limite de 0,01 ug/l	Frequência Absoluta	Frequência Percentual
Ultrapassa Confirmado	5792	89,38 %
Ultrapassa por Precaução	5	0,08 %
Não Ultrapassa Confirmado	683	10,54 %
Total	6480	100 %

Quadro 3. *Frequência de ocorrências para cada cenário relativo ao limite de 0,01 µg/L*

4.3 Implicações para a gestão ambiental e sociedades sustentáveis

A Figura 2 foi elaborada com foco na amoxicilina, destacada por seus altos valores de CAP. Ao observar o comportamento desse princípio ativo, é possível confirmar que a exposição ambiental estimada não é um reflexo isolado da massa de medicamentos comercializada (carga), mas depende também da capacidade da infraestrutura de saneamento para reduzir a concentração desses poluentes e da capacidade de diluição dos cursos d'água receptores de efluentes tratados (Reis et al., 2019), o que se reflete na abordagem para o enfrentamento da problemática.

Diante do panorama de emergência climática, a frequência de eventos de escassez hídrica atuará como catalisador do risco, agravando a baixa diluição e concentrando a carga de contaminantes emergentes nas redes (Rodrigues et al., 2025). Portanto, a identificação desses *hotspots* fornece subsídios práticos para que gestores ambientais direcionem investimentos em infraestrutura para a remoção de compostos que as tecnologias convencionais, baseadas predominantemente em lodo ativado, não são capazes de eliminar (Couto et al., 2019; Montagner et al., 2017).

Welch et al. (2022) demonstraram a utilidade de dados nacionais de comercialização para estimar concentrações previstas; He et al. (2020) observaram concordância variável entre estimativas baseadas em vendas e concentrações medidas; e Verlicchi e Zambello (2016) destacaram a influência da vazão, excreção e amostragem. Em conjunto, essas fontes apoiam o uso das CAP para triagem e desenho de campanhas, não como substitutas do monitoramento químico.

Em termos de gestão, as CAP podem apoiar a seleção de princípios ativos, capitais e períodos para monitoramento ambiental de maior custo, orientando a alocação de recursos e o planejamento de campanhas analíticas. Amoxicilina, cefalexina, cloridrato de ciprofloxacino e levofloxacino constituem prioridades de exposição no cenário analisado. A integração dos dados do SNGPC e do SNIS configura, assim, uma ferramenta de apoio à decisão para organizações públicas de saneamento, vigilância sanitária e controle ambiental, sem substituir a confirmação por monitoramento químico. Sua aplicação em outras regiões depende da disponibilidade e da compatibilização das bases, preservando o caráter replicável da abordagem.

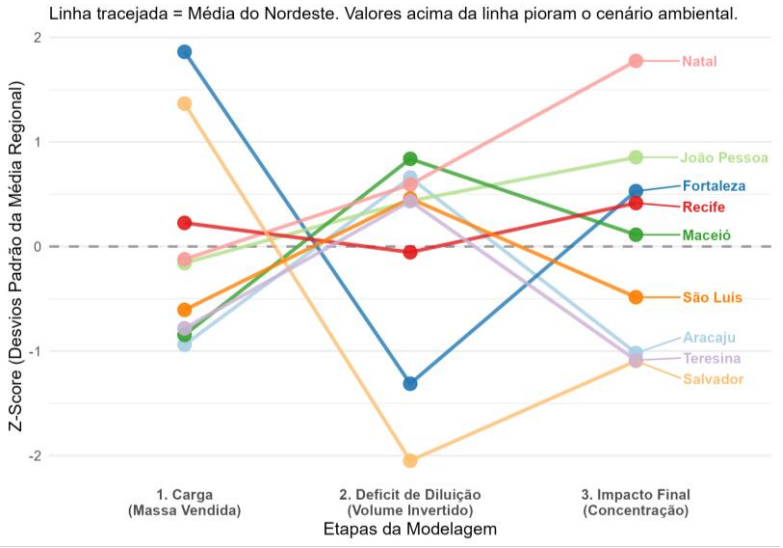


Figura 2. *Perfil Z-score da amoxicilina (princípio ativo com maiores valores de CAP)*

Outros recortes podem ampliar a capacidade de direcionamento da análise: capital por ano, para avaliar mudanças interanuais; capital por mês, para investigar sazonalidade; capital por princípio ativo, para reconhecer prioridades locais; e macrorregião por princípio ativo. Recortes por bacia hidrográfica, área de atendimento de estações de tratamento ou corpo receptor exigem ligação espacial adicional e não podem ser inferidos com os dados públicos disponíveis. Em todos os casos, o tamanho amostral nominal e o número de unidades efetivamente independentes devem ser informados separadamente.

5 CONCLUSÕES

O estudo, por limitações de disponibilidade de dados, assume uma equivalência direta entre as vendas e o consumo local. Como o modelo adota uma postura conservadora de proteção ambiental e a modelagem da remoção desses fármacos é uma tarefa de elevada complexidade, as estimativas refletem a exposição bruta nas redes de esgotamento, sem incorporar potenciais eficiências de remoção nas ETEs.

A propagação das incertezas associadas à modelagem que integra os dados do SNGPC e do SNIS demonstrou que as variabilidades dos registros originais não alteram a classificação final, com 89,38% das simulações resultando em concentrações acima do valor de referência. A decomposição das variáveis estabelece que a concentração dos fármacos nas redes de esgotamento é determinada essencialmente pela relação entre a massa total comercializada e o volume de água disponível para diluição, afetando principalmente as cidades com as piores relações entre esses fatores. Assim, embora os cenários de exposição não substituam o monitoramento analítico tradicional, a metodologia reforça a viabilidade dos registros de comercialização para o cálculo da CAP como ferramenta de triagem, fornecendo subsídios quantitativos preliminares para orientar a priorização de substâncias e localidades críticas para análises laboratoriais. Essa contribuição aproxima o uso de tecnologias de dados e da modelagem de incerteza das necessidades da gestão ambiental, fornecendo elementos para o planejamento de ações relacionadas aos ODS 6 e 11. A abordagem pode ser replicada em outras regiões, desde que as bases disponíveis sejam compatibilizadas e as particularidades locais de saneamento sejam consideradas.

6 REFERÊNCIAS

Agência Nacional de Vigilância Sanitária. (2019). *Farmacopeia brasileira* (6ª ed., Vol. 1). https://www.gov.br/anvisa/pt-br/assuntos/farmacopeia/farmacopeia-brasileira/volume-1-fb6.pdf

Agência Nacional de Vigilância Sanitária. (2020). *Dados abertos: Venda de medicamentos*

industrializados sujeitos à escrituração no SNGPC: Documentação e dicionário de dados.

Ahkola, H., Äystö, L., Sikanen, T., Riikonen, S., Pihlaja, T., & Kauppi, S. (2024). Current uncertainties and challenges of publicly available pharmaceutical environmental risk assessment data. *European Journal of Pharmaceutical Sciences, 197*, 106769. https://doi.org/10.1016/j.ejps.2024.106769

Associação Brasileira de Normas Técnicas. (1986). *ABNT NBR 9649: Projeto de redes coletoras de esgoto sanitário — Procedimento*. https://www.normas.com.br/visualizar/abnt-nbr-nm/5590/abnt-nbr9649-projeto-de-redes-coletoras-de-esgoto-sanitario-procedimento

Associação Brasileira de Normas Técnicas, & Instituto Nacional de Metrologia, Qualidade e Tecnologia. (2008). *Avaliação de dados de medição: Guia para a expressão de incerteza de medição (GUM 2008)*. https://www.gov.br/inmetro/pt-br/centrais-de-conteudo/publicacoes/documentos-tecnicos-em-metrologia/gum_final.pdf

Caetano, M. C., Campos, M. R., Emmerick, I. C. M., & Luiza, V. L. (2025). Análise da qualidade da base de dados do Sistema Nacional de Gerenciamento de Produtos Controlados (SNGPC). *Saúde em Debate, 49*(144), e9350. https://doi.org/10.1590/2358-289820251449350P

Couto, C. F., Lange, L. C., & Amaral, M. C. S. (2019). Occurrence, fate and removal of pharmaceutically active compounds (PhACs) in water and wastewater treatment plants: A review. *Journal of Water Process Engineering, 32*, 100927. https://doi.org/10.1016/j.jwpe.2019.100927

European Medicines Agency, Committee for Medicinal Products for Human Use. (2006). *Guideline on the environmental risk assessment of medicinal products for human use*. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-environmental-risk-assessment-medicinal-products-human-use-first-version_en.pdf

Ferreira, T. de J. N., Morais, J. H. de A., Caetano, R., & Osorio-de-Castro, C. G. S. (2023). Tratamento de dados do Sistema Nacional de Gerenciamento de Produtos Controlados para estudos de utilização de medicamentos com antimicrobianos. *Cadernos de Saúde Pública, 39*(5), e00173922. https://doi.org/10.1590/0102-311XPT173922

Giebułtowicz, J., & Nałęcz-Jawecki, G. (2014). Occurrence of antidepressant residues in the sewage-impacted Vistula and Utrata rivers and in tap water in Warsaw (Poland). *Ecotoxicology and Environmental Safety, 104*, 103–109. https://doi.org/10.1016/j.ecoenv.2014.02.020

He, K., Borthwick, A. G., Lin, Y., Li, Y., Fu, J., Wong, Y., & Liu, W. (2020). Sale-based estimation of pharmaceutical concentrations and associated environmental risk in the Japanese wastewater system. *Environment International, 139*, 105690. https://doi.org/10.1016/j.envint.2020.105690

Jagadeesan, K. K., Barden, R., & Kasprzyk-Hordern, B. (2023). PERK: An R/Shiny application to predict and visualise concentrations of pharmaceuticals in the aqueous environment. *Science of the Total Environment, 875*, 162352. https://doi.org/10.1016/j.scitotenv.2023.162352

Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data—Supplement 1 to the "Guide to the expression of uncertainty in measurement"—Propagation of distributions using a Monte Carlo method (JCGM 101:2008)*.

Montagner, C. C., Vidal, C., & Acayaba, R. D. (2017). Contaminantes emergentes em matrizes aquáticas do Brasil: Cenário atual e aspectos analíticos, ecotoxicológicos e regulatórios. *Química Nova, 40*(9), 1094–1110. https://doi.org/10.21577/0100-4042.20170091

Moreira, J. C., & Gonçalves, E. S. (2013). Contaminantes emergentes. *Revista Química*

Industrial, 733, 4–13. https://www.abq.org.br/rqi/2013/738/RQI-738-pagina4-Contaminantes-Emergentes.pdf

Organização das Nações Unidas. (2015). *Transformando nosso mundo: A Agenda 2030 para o desenvolvimento sustentável*. https://sdgs.un.org/2030agenda

Patel, M., Kumar, R., Kishor, K., Mlsna, T., Pittman, C. U., Jr., & Mohan, D. (2019). Pharmaceuticals of emerging concern in aquatic systems: Chemistry, occurrence, effects, and removal methods. *Chemical Reviews, 119*(6), 3510–3673. https://doi.org/10.1021/acs.chemrev.8b00299

Reis, E. O., Foureaux, A. F. S., Rodrigues, J. S., Moreira, V. R., Lebron, Y. A. R., Santos, L. V. S., Amaral, M. C. S., & Lange, L. C. (2019). Occurrence, removal and seasonal variation of pharmaceuticals in Brasilian drinking water treatment plants. *Environmental Pollution, 250*, 773–781. https://doi.org/10.1016/j.envpol.2019.04.102

Rodrigues, F., Durães, L., Simões, N. E. C., Pereira, A. M. P. T., Silva, L. J. G., & Feio, M. J. (2025). Pharmaceuticals in urban streams: A review of their detection and effects in the ecosystem. *Water Research, 268*, 122657. https://doi.org/10.1016/j.watres.2024.122657

Sales, P. T. F., Campos, L. C., Schimidt, F., Valadares, M. C., & Santiago, M. F. (2012). Estudo da tratabilidade de efluente da indústria farmacêutica por meio dos fungos Pycnoporus sanguineus, Schizophyllum comnune e fotocatálise. *Revista Eletrônica de Engenharia Civil, 5*(1), 56–74. https://doi.org/10.5216/reec.v5i1.20257

UNESCO. (2023). *Emerging pollutants: Protecting water quality for the health of people and the environment*. https://www.unesco.org/en/articles/event-emerging-pollutants-protecting-water-quality-health-people-and-environment

Valdez-Carrillo, M., Abrell, L., Ramírez-Hernández, J., Reyes-López, J. A., & Carreón-Diazconti, C. (2020). Pharmaceuticals as emerging contaminants in the aquatic environment of Latin America: A review. *Environmental Science and Pollution Research, 27*, 44863–44891. https://doi.org/10.1007/s11356-020-10842-9

Verlicchi, P., & Zambello, E. (2016). Predicted and measured concentrations of pharmaceuticals in hospital effluents: Examination of the strengths and weaknesses of the two approaches through the analysis of a case study. *Science of the Total Environment, 565*, 82–94. https://doi.org/10.1016/j.scitotenv.2016.04.165

Welch, S. A., Olsen, K., Nouri Sharikabad, M., Tollefsen, K. E., Grung, M., & Moe, S. J. (2022). Pharmaceutical pollution: Prediction of environmental concentrations from national wholesales data. *Open Research Europe, 2*, 71. https://doi.org/10.12688/openreseurope.14129.2