

Desafios e oportunidades da aplicação de Sistemas Ciberfísicos no monitoramento da poluição urbana

Alessandro S. Santos¹, Leandro G. de Freitas¹, Igor C. Teixeira¹, Vagner L. Gava¹,
Gustavo R. Taira^{1,2}, Rosa V. Encinas Quille^{1,2}, Kelly R. Braghetto²

¹Instituto de Pesquisas Tecnológicas do Estado de SP – (IPT)
Av. Prof. Almeida Prado, 532 – Butantã – São Paulo – SP – Brasil

²Universidade de São Paulo (USP)
São Paulo, SP – Brazil

{alesan, lfreytas, igort, vlgava}@ipt.br,
{gustavo.taira, encinas, kellyrb}@usp.br

Abstract. *Human action has been causing changes in the environment, with urban pollution being one of the negative consequences for this ecosystem. With the technological development, new computational perspectives emerge for environmental monitoring. This paper presents key points of environmental phenomenology that can benefit from the evolution promoted by computing applied to environmental pollution studies, as well as make a review of research and technologies that have been applied in this context.*

Resumo. *A ação do homem tem provocado mudanças no meio ambiente, sendo a poluição urbana uma das consequências negativas aplicadas nesse ecossistema. Com o desenvolvimento tecnológico, novas perspectivas computacionais afloram para o monitoramento ambiental. Este artigo apresenta pontos-chaves da fenomenologia ambiental que podem se beneficiar da evolução promovida pela computação aplicada nos estudos da poluição, assim como realiza um levantamento de pesquisas e tecnologias utilizadas nesse contexto.*

1. Introdução

Com o crescimento da população urbana mundial, a poluição ambiental vem se mostrando um grande desafio para a conciliação das atividades humanas com a manutenção de um ambiente saudável nas cidades [Wong et al. 2006]. Sob uma perspectiva histórica, a poluição urbana intensificou-se com o crescimento industrial nos séculos XIX e XX, decorrente do aumento das emissões fabris. No Brasil, o rápido crescimento desordenado das cidades e a adoção tardia de instrumentos legais de controle da poluição implicaram na piora da qualidade ambiental nos centros urbanos, principalmente na segunda metade do século passado. Somente a partir da década de 80, com a promulgação das bases da política ambiental brasileira, é que o controle e o monitoramento da poluição passaram a ganhar relevância no cenário nacional [Günther 2006].

Com o advento das novas tecnologias da transformação digital, o monitoramento ambiental vem passando por mudanças significativas, utilizando novas formas de analisar grandes volumes de dados com a aplicação de inteligência artificial (IA) em um contexto de Internet das Coisas (IOT, de *Internet of Things*). Abrem-se novas perspectivas para o

monitoramento dos fenômenos ambientais por meio da abordagem de Sistemas Ciberfísicos (*Cyber-Physical Systems – CPSs*).

Os CPSs integram recursos de computação e controle das entidades no mundo físico que precisam ser executadas de maneira segura, eficiente e em tempo real. Esses sistemas consistem em sensores interconectados, coletando informações do mundo físico, e atuadores agindo sobre o meio ambiente, integrados sob uma plataforma digital de apoio de decisão inteligente [Sanislav et al. 2016].

Este artigo apresenta uma revisão da aplicação de novas tecnologias digitais para o monitoramento e gestão da poluição ambiental, com enfoque nos desafios e oportunidades computacionais, que devem estar integradas às fenomenologias. O trabalho está inserido no âmbito do projeto “Plano de Desenvolvimento Institucional na Área de Transformação Digital: Manufatura Avançada e Cidades Inteligentes e Sustentáveis (PDIp)”, financiado pela FAPESP (proc. 2017/50343-2) e desenvolvido pelo IPT. O trabalho possui a colaboração do Instituto Nacional de Ciência e Tecnologia (INCT) da Internet do Futuro para Cidades Inteligentes (InterSCity), financiado pelo CNPq proc. 465446/2014-0, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, FAPESP proc. 14/50937-1 e FAPESP proc. 15/24485-9.

2. Fenomenologia e o monitoramento ambiental urbano

Um dos temas mais frequentes nas discussões ambientais é a poluição urbana, que pode se referir a qualidade do ar, poluição sonora ou contaminação do solo. A ação do homem tem uma relação direta com o problema e as oportunidades computacionais dependem do entendimento do fenômeno e seus desafios de monitoramento. Além disso, as questões espaço-temporais são pré-requisitos em qualquer das avaliações ambientais apresentadas.

2.1. Poluição do ar: qualidade do ar

A questão da qualidade do ar está relacionada à concentração de poluentes, classificados em diferentes níveis de periculosidade. Para fins didáticos, a Companhia Ambiental do Estado de São Paulo (CETESB) estabeleceu faixas de concentração delineando níveis dos principais gases poluentes da qualidade do ar que são equacionados em índices e descrições dos possíveis efeitos nocivos à saúde, mostrados na Tabela 1.

Tabela 1. Níveis de concentração e impactos na saúde [CETESB 2015]

Qualidade	Índice	MP ₁₀ (µg/m ³)	MP _{2,5} (µg/m ³)	O ₃ (µg/m ³)	CO (ppm)	NO ₂ (µg/m ³)	SO ₂ (µg/m ³)	Fumaça (µg/m ³)	Significado
N1 - Boa	0 - 40	0 - 50	0 - 25	0 - 100	0 - 9	0 - 200	0 - 20	0 - 50	
N2 - Moderada	41 - 80	> 50 - 100	> 25 - 50	> 100 - 130	> 9 - 11	> 200 - 240	> 20 - 40	> 50 - 100	Pessoas de grupos sensíveis (crianças, idosos e pessoas com doenças respiratórias e cardíacas) podem apresentar sintomas como tosse seca e cansaço. A população, em geral, não é afetada.
N3 - Ruim	81 - 120	> 100 - 150	> 50 - 75	> 130 - 160	> 11 - 13	> 240 - 320	> 40 - 365	> 100 - 150	Toda a população pode apresentar sintomas como tosse seca, cansaço, ardor nos olhos, nariz e garganta. Pessoas de grupos sensíveis (crianças, idosos e pessoas com doenças respiratórias e cardíacas) podem apresentar efeitos mais sérios na saúde.
N4 - Muito ruim	121-200	>150-250	>75-125	>160-200	>13-15	>320-1130	>365-800	>150-250	Toda a população pode apresentar agravamento dos sintomas como tosse seca, cansaço, ardor nos olhos, nariz e garganta e ainda falta de ar e respiração ofegante. Efeitos ainda mais graves à saúde de grupos sensíveis (crianças, idosos e pessoas com doenças respiratórias e cardíacas).
N5 - Péssima	>200	>250	>125	>200	>15	>1130	>800	>250	Toda a população pode apresentar sérios riscos de manifestações de doenças respiratórias e cardiovasculares. Aumento de mortes prematuras em pessoas de grupos sensíveis.

As projeções temporárias são agrupadas por faixas horárias de valores como: média, máximo, mínimo e desvio padrão. Esta questão está relacionada ao comportamento

típico da concentração de poluentes em ambiente urbano, que pode ser visto na Figura 1. Além disso, os órgãos ambientais estaduais apresentam os dados em forma de médias móveis das últimas 8 e 24 horas e médias diárias, mensais e anuais.

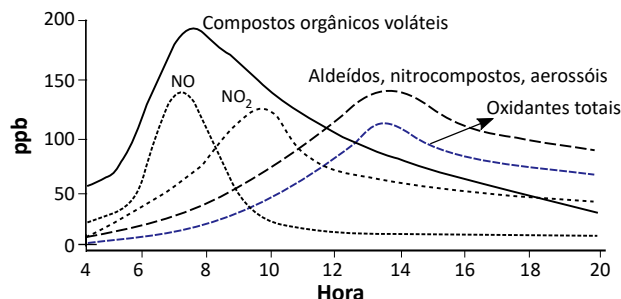


Figura 1. Perfil típico de concentração de poluentes em ambiente urbano.
Fonte: [Cardoso 2009].

No ambiente urbano, os veículos são um dos principais emissores de poluentes no ar, sendo que as emissões veiculares contêm diversas substâncias tóxicas que podem causar vários problemas ao sistema respiratório. Cada tipo de veículo contribui com as emissões de poluentes em diferentes proporções. Por exemplo, os veículos de passeio são os principais emissores de CO, enquanto NO_x são emitidos majoritariamente por ônibus e caminhões [CETESB 2015].

A problemática de controle de poluentes requer uma atividade sistemática de averiguação das causas e motivações da presença desses poluentes, assim como ações de mitigação dos efeitos nocivos. Neste caso, recomenda-se uma sistemática de gestão da qualidade do ar que incorpore instrumentos de gerenciamento ambiental integrado às questões sociais e econômicas [Council 2004].

2.1.1. Desafios e oportunidades computacionais sobre a qualidade do ar

De forma geral, existem duas abordagens adotadas na avaliação da poluição do ar. A primeira delas envolve cálculos de emissões de poluentes, os quais utilizam simulações com base em valores estimados de emissões veiculares e industriais que podem promover níveis de geração de poluentes. A segunda abordagem envolve a medida de concentração de poluentes, sendo esta mais aderente a sistemas ciberfísicos e utilizada por agências ambientais dos estados brasileiros.

As principais regiões metropolitanas brasileiras possuem mecanismos de monitoramento da qualidade do ar. No entanto, o primeiro diagnóstico das redes nacionais para este fim [IEMA 2014] sinaliza que partes importantes do Brasil não contam com monitoramento e que existem assimetrias estruturais significativas entre as redes de monitoramento. Além disso, as estações de monitoramento são instaladas com uma cobertura abrangente, mas que projeta uma região com uma baixa resolução espacial de dados, não permitindo que características de microrregiões sejam perceptíveis. A literatura já aponta que microrregiões com características da ocupação do solo, como túneis, cruzamentos viários e ruas flanqueadas por prédios em ambos os lados são pontos potenciais de concentração de poluentes [Vasconcellos et al. 2005, Goel and Kumar 2015].

Um dos grandes desafios com relação à qualidade do ar é aumentar as resoluções espaciais do monitoramento. No entanto, o custo de estações automáticas de monitoramento da qualidade do ar é elevado (350 mil dólares) [CETESB 2018]. Sendo assim, existe a demanda pela validação de equipamentos de monitoramento com baixo custo e sua difusão por várias áreas da cidade, como meio complementar aos mecanismos oficiais de alta confiabilidade.

2.2. Poluição sonora: ruído urbano

O ruído urbano é considerado o segundo maior tipo de poluição nas cidades, depois da poluição do ar. O ruído proveniente do tráfego urbano é hoje um dos maiores geradores de poluição sonora urbana e já vem sendo associado a importantes problemas que afetam a saúde dos cidadãos. A Figura 2 apresenta referenciais de níveis e algumas fontes geradoras de ruído.

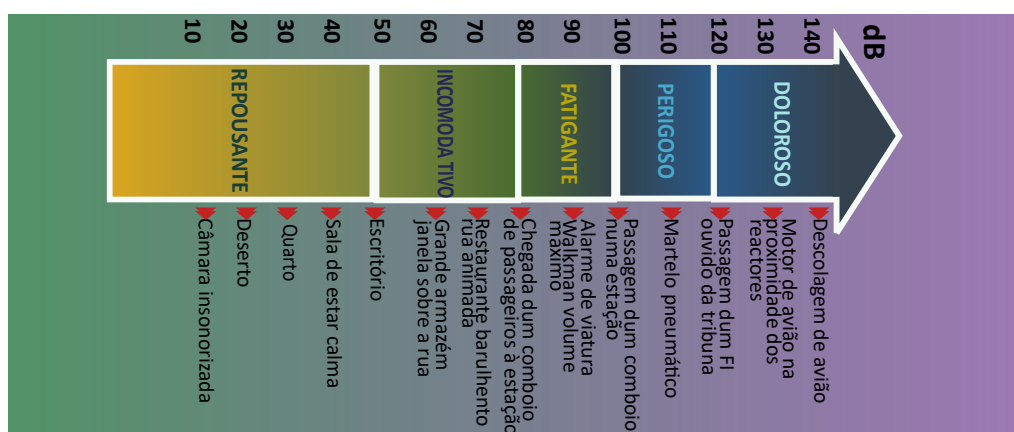


Figura 2. Níveis de Ruído. Fonte: Agência Portuguesa de Ambiente, 2019.

O ruído urbano é influenciado principalmente por fatores como volume de tráfego, tipo de veículos, características de ocupação do solo e velocidade de deslocamento dos veículos. Além disso, possui uma característica temporal bem particular, pois ao cessar a fonte geradora do ruído, o efeito significativo perdura por pouco tempo no meio urbano.

O mapeamento de ruído urbano envolve basicamente duas abordagens. A primeira são decibelímetros que captam os níveis de ruídos em decibels (dB), obtendo valores para realizar análises de impacto com base nas faixas da Figura 2. A segunda abordagem realiza a gravação de áudio, que é analisada com equipamentos especiais que podem classificar o tipo de ruído, seu nível e a direção de onde o ruído está sendo gerado (técnica chamada *Acoustic Beamforming*).

A criação de mapas de ruído é uma alternativa para diagnosticar a poluição sonora, provendo ao gestor mecanismos de planejamento do desenvolvimento urbano, principalmente em regiões com grande adensamento populacional. O mapeamento deve funcionar como diagnóstico e ferramenta de apoio ao planejamento e ordenamento urbano na gestão de ruído da cidade, tendo como uma de suas finalidades a identificação de áreas prioritárias para redução de ruídos e a preservação de zonas com níveis sonoros apropriados [Sales and Aquilino 2017].

2.2.1. Desafios e oportunidades computacionais sobre o ruído urbano

As cidades na Europa com mais de 200 mil habitantes seguem diretrizes e um guia de boas práticas para criar mapas de ruído urbano, entretanto são muito amplos e foram elaborados para cidades que têm um perfil diferente das cidades brasileiras. Santiago e Valdivia, no Chile estão entre as poucas cidades na América Latina que têm mapeamento. Na capital chilena, o primeiro estudo data de 1989 e duas atualizações foram feitas em 2011 e 2015. Já no Brasil, não existe qualquer mapeamento completo.

A maior cidade brasileira teve as discussões sobre a necessidade de se ter um mapa de ruído intensificadas a partir de 2014. Um dos resultados desta mobilização foi a Lei 16.499, de julho de 2016, que dispôs sobre a elaboração do Mapa do Ruído Urbano da Cidade de São Paulo.

A criação de mapas de ruído urbano pode usar metodologias que empregam campanhas com equipamentos tradicionais com medição de ruído ponto a ponto. No entanto, em uma cidade de grandes dimensões, o levantamento dos dados demandaria muito tempo, além de ser extremamente oneroso. Portanto, um dos desafios no mapeamento de ruído é aumentar a resolução dos mapeamentos, assim como obter mecanismos para simular ou prever os níveis de ruído em ambiente urbano, ou mesmo ter ferramentas de monitoramento que possam determinar o perfil sonoro de diferentes partes da cidade.

2.3. Poluição do solo: áreas contaminadas

Apesar de não ser visualmente evidente, a poluição do solo é um tema bastante desafiador para a gestão ambiental das cidades. Os contaminantes presentes no solo podem se propagar no ambiente, alterando sua qualidade e prejudicando a saúde pública e demais bens a proteger, tanto no terreno onde ocorre, como em seus arredores. O gerenciamento incorreto pode causar ainda a desvalorização dos imóveis e, em situações mais críticas, pode representar graves riscos à população, como os de incêndio e explosão [CETESB 2001].

Segundo a CETESB, as principais atividades geradoras de áreas contaminadas (ACs) estão relacionadas a postos de combustíveis, emissões industriais, áreas de descarte de resíduos, minerações e locais de acidentes, além de fontes difusas diversas, como a poluição por defensivos agrícolas e vazamentos de esgoto. Esses poluentes se infiltram no subsolo e podem atingir as águas subterrâneas, gerando as plumas de contaminação. O transporte pode ainda atingir aquíferos e águas superficiais, prejudicando a qualidade de mananciais estratégicos para o abastecimento da população. Um esquema ilustrativo desses processos é apresentado na Figura 3.

O monitoramento das ACs geralmente envolve a coleta e análises físico-químicas de amostras de solo, água (subterrânea e superficial) e gases do solo. Essas amostras são coletadas por meio de sondagens ou com a instalação de poços de monitoramento. Nas etapas iniciais de diagnóstico ambiental, uma ampla gama de parâmetros é analisada com dois objetivos principais: qualificar e quantificar quimicamente os contaminantes e caracterizar o meio físico que os aloja. Já nas etapas de remediação e pós intervenção, o monitoramento é focado apenas nos compostos alvo, com o intuito de comprovar a eficácia das medidas adotadas. Em certos casos, principalmente das plumas de gases e/ou da água subterrânea, o monitoramento ambiental pode ser estendido por longos períodos (anos ou décadas), a depender da estratégia de intervenção adotada.



Figura 3. Fontes de poluição do solo e sua migração. Fonte: [CETESB 2001].

2.3.1. Desafios e oportunidades computacionais sobre áreas contaminadas

Apenas no Estado de SP, 6.285 áreas estão cadastradas como contaminadas ou reabilitadas [CETESB 2020], podendo este número ultrapassar os 30.000 casos [Teixeira et al. 2016]. Segundo o órgão ambiental, hidrocarbonetos de petróleo são os principais contaminantes. A remediação dessas áreas vem se mostrando um grande desafio para agentes públicos e privados, tanto em função da complexidade técnica, como por restrições financeiras e administrativas. Além dos projetos serem bastante custosos, muitas vezes o monitoramento tem que ser mantido por longos períodos para se garantir um uso seguro da área.

A gestão das ACs é feita de forma independente pelo órgão ambiental, não existindo um sistema computacional unificado que agregue o crescente volume de dados existente. Além disso, há a dificuldade de acessibilidade aos relatórios e o formato de disponibilização dos laudos, que anteriormente a 2017 eram protocolados em formato impresso.

Outra oportunidade é a automatização do monitoramento ambiental. Considerando que as metodologias convencionais são usualmente executadas por técnicos em campo, com equipamentos para medições *in situ* e posterior análises laboratoriais, essas atividades se mostram custosas, demoradas e de alta periculosidade aos executores.

Nesse contexto, o desenvolvimento de plataformas digitais para integração e disponibilização dos dados do gerenciamento de ACs, com base em protocolos, IoT e ferramentas de Ciência de Dados, abrem novas perspectivas para a remediação ambiental.

2.4. Fenomenologia e desafios

Com base na discussão dos fenômenos e desafios computacionais do monitoramento da poluição ambiental, a Tabela 2 lista as principais fontes geradoras dos fenômenos, enquanto a Tabela 3 mostra os principais parâmetros ambientais para o seu monitoramento.

Tabela 2. Principais fontes geradoras dos fenômenos

Fontes geradoras	Veículos	Emissões e resíduos industriais	Lixo urbano	Comércio de combustíveis	Acidentes e vazamentos
Fenômeno					
Poluição do Ar	✓	✓	✓		
Poluição sonora	✓	✓			
Poluição do solo		✓	✓	✓	✓

Tabela 3. Principais parâmetros ambientais para monitoramento

Parâmetros	Gases		Água	Aúdio	Pluvio- metria	Temperatura, Umidade do ar, Pressão atm.
	<i>SO₂, NO₂, CO, O₃, MP</i>	<i>CH₄, O₂, CO₂ LEL, COV_s</i>	Subterrânea*			
Fenômeno						
Poluição do ar	✓					✓
Poluição sonora				✓	✓	
Poluição do solo		✓	✓		✓	✓

(*) Parâmetros físico-químicos, concentração de contaminantes e nível d' água subterrânea

3. Abordagem ciberfísica e tecnologias computacionais aplicadas ao monitoramento ambiental

A abordagem de uso do conceito ciberfísico para o monitoramento ambiental passa por diversas fases. O modelo conceitual apresentado por [Santos 2018], mostrado na Figura 4, aborda o monitoramento da qualidade do ar e suas fases, o qual pode ser ampliado para os demais tipos de poluição. O modelo propõe o uso de dispositivos IoT e sua conectividade e tecnologias para processamento computacional e análise de dados (como aprendizado de máquina e *Big Data*). Esse ecossistema cria uma nova perspectiva de abordagem da problemática para análises de poluição ambiental.

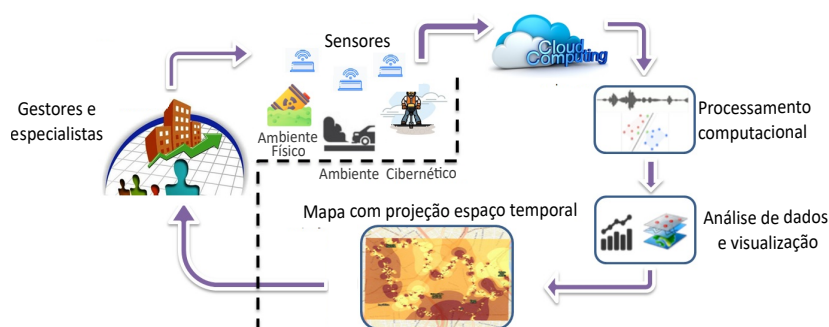


Figura 4. Modelo conceitual para utilização de tecnologias de sensoriamento. Adaptado de: [Santos 2018].

3.1. Internet das Coisas (IoT, de *Internet of Things*)

No contexto deste artigo, os dispositivos IoT são aqueles que podem ser utilizados para monitorar parâmetros relevantes para a fenomenologia da poluição urbana, para a avaliação da qualidade do ar, ruído urbano ou áreas contaminadas.

3.1.1. IoT na qualidade do ar

O uso de dispositivos IoT de baixo custo para monitoramento da qualidade do ar vem crescendo e com boas perspectivas para uso em grande escala [Kumar et al. 2015]. [Morawska 2018] fizeram um levantamento de 17 projetos de grande porte, conduzidos por consórcios internacionais que lidam com experimentos de IoT de baixo custo para monitorar a qualidade do ar. A Figura 5 mostra os projetos que duraram mais que o tempo típico de financiamento público, que é de 3 a 4 anos. Dois dos projetos mais antigos têm financiamento coletivo e estão ativos até hoje, diferentemente da maior parte dos projetos financiados por governos.

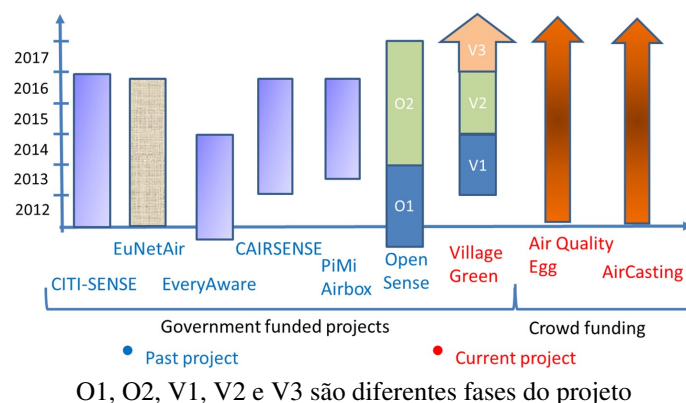


Figura 5. Duração de projetos de IoT de baixo custo para monitorar a qualidade do ar. Fonte: [Morawska 2018].

O sensor de gás acoplado ao dispositivo IoT pode ter diferentes princípios de operação: eletroquímicos, óxidos metálicos ou sensores ópticos. A Organização Meteorológica Mundial realizou um levantamento de limitações observadas de cada tipo de sensor, assim como a interferência por parâmetros meteorológicos, efeito de envelhecimento (*Drift*), sensibilidades cruzadas com outros poluentes e deriva [Lewis et al. 2018].

Outro fator a ser observado é em relação ao desempenho de cada tipo de sensor. Neste sentido, [Karagulian et al. 2019] apresentam resultados de testes em campo e em laboratórios que ajudam no momento da escolha da montagem do kit IoT mais apropriado para realização de experimentos.

3.1.2. IoT no ruído urbano

Os dispositivos IoT aplicados em projetos de monitoramento da poluição sonora envolvem basicamente duas abordagens. A primeira são sensores que captam os níveis de ruídos em decibels (dB), como um decibelímetro, identificando os valores e enquadrando-os dentro de faixas, como apresentado na Figura 2. A segunda abordagem realiza a gravação de áudio, que é analisada para classificar o tipo de ruído e seu nível.

Um grande projeto de uso de sensores IoT para mapeamento da poluição sonora é apresentado pelo projeto *SONYC - Sound of New York* [Bello et al. 2018], que propõe uma plataforma de detecção inteligente de baixo custo, capaz de monitorar de forma contínua e em tempo real uma fonte específica de ruído. Esta solução adiciona novas camadas de métodos avançados de Ciência de Dados para análise de ruídos em larga escala, incluindo modelagem preditiva de ruídos, usando estatísticas espaciais e modelagem física. Além disso, permite visualizações 3D interativas da atividade de ruído em função do tempo e do espaço, permitindo uma melhor compreensão dos padrões de ruído.

3.1.3. IoT na poluição do solo

A aplicação de IoT nos estudos de poluição do solo ainda apresenta poucas referências e muitas oportunidades de desenvolvimento. Especificamente em relação à adoção de sensores, os avanços tecnológicos em termos de precisão, durabilidade e redução de custos

vêm viabilizando a implantação desses dispositivos no monitoramento de áreas contaminadas. Esses equipamentos podem tanto ser implantados diretamente no solo, para sensoriamento de contato direto, como podem ser implantados em sistemas de amostragem automatizada, que levam as amostras do meio até os dispositivos analíticos na superfície.

Um exemplo apresentado por [Horst et al. 2017] é um sistema de monitoramento inteligente de poços, apelidado de *Internet of Well*, que integra a instalação de sensores e microcontroladores nos poços, alimentados por energia solar e transmissão remota de dados para *dashboards* na *web*, possibilitando a visualização do resultado em tempo real.

3.2. Conectividade em IoT para poluição ambiental

O monitoramento ambiental por dispositivos de baixo custo utiliza tecnologias sem fio devido aos requisitos de implementação em campo. Neste sentido, pode-se categorizá-las conforme a maneira como as informações fluem do nó sensor até a borda de integração com a infraestrutura digital, sendo esta borda chamada de sorvedouro, *gateway*, roteador, concentrador ou *datalogger*, dependendo da tecnologia e fornecedor.

Duas categorias são mais frequentes, cooperativa e não cooperativa. Na cooperativa, o nó sensor se comunica com outros nós para obter a informação de interesse, em um esquema de retransmissão multi-salto, enquanto no modelo não cooperativo o dado flui sem a necessidade de utilização e cooperação de outros nós. A comunicação não cooperativa é o modelo mais frequente para as implementações de monitoramento de poluição ambiental, aproveitando-se de infraestrutura já existente nas cidades. Nesse contexto, destacam-se as tecnologias celulares e LPWAN.

Celulares Com o crescimento massivo de dispositivos móveis em ambiente urbano, a comunicação celular expandiu-se, estando presente em praticamente todo o território nacional, permitindo a transferência de dados em tecnologias celulares, como 3G, 4G, 5G, etc. A combinação dessas tecnologias com aplicações IoT permitem o uso de sistemas de monitoramento ambientais mais robustos e eficientes, com segurança integrada, além de recursos de gerenciamento de tráfego escaláveis no ambiente urbano. Por outro lado, essas tecnologias são oferecidas por operadoras de telefonia e o seu custo e contexto mercadológico devem ser avaliados.

Low Power Wide Area Network – LPWAN A tecnologia LPWAN [Raza et al. 2017] atende as necessidades de monitoramento de poluição ambiental, uma vez que permite ampla área de cobertura, baixo consumo de bateria e uso de dispositivos de baixo custo. Inclui subtipos tais como NB-IoT (*Narrow-Band IoT*), LoRaWAN (*Low Power WAN*) e Sigfox. Por outro lado, opera em um espectro de frequência não licenciado, podendo gerar problemas de escalabilidade devido ao congestionamento do espectro, além da baixa largura de banda.

3.3. Aprendizado de máquina

Esta seção apresenta um levantamento de aplicações de aprendizado de máquina para fenômenos de poluição, baseado em buscas de publicações científicas realizadas em bases indexadas, tais como a Web of ScienceTM Core Collection, da Clarivate Analytics, e Scopus[®], da Elsevier. As buscas limitaram-se a artigos em inglês publicados entre os anos de 2015 e 2020. Para cada tipo de fenômeno, utilizou-se uma *string* de busca específica, de forma a obter as publicações mais relevantes em cada área.

3.3.1. Aprendizado de máquina aplicado à poluição do ar

Utilizou-se a *string* de busca: (“air pollut*”) AND (“machine learn*” OR “deep learn*” OR “artificial intelligen*”). Após filtragem e remoção de duplicatas, restaram as 284 publicações que foram analisadas.

No contexto de poluição de ar, nos últimos cinco anos, os métodos de aprendizado de máquina foram aplicados na criação de modelos de previsão, tanto de concentrações de poluentes [Li et al. 2017], quanto de índices de qualidade do ar [Wang et al. 2017], na criação de modelos classificadores da qualidade do ar [Rybarczyk and Zalakeviciute 2016] e na elaboração de modelos interpoladores/simuladores de valores de concentração de poluentes em regiões não monitoradas [de Hoogh et al. 2018]. Além disso, os métodos foram amplamente utilizados para calibração de sensores de baixo custo [Zimmerman 2018].

As técnicas mais aplicadas no contexto de poluição do ar foram: florestas aleatórias (*random forests*) [Zimmerman 2018] (68), redes neurais artificiais (ANN, de *artificial neural networks*) [Ni et al. 2017] (55), redes neurais recorrentes *long-short term memory* (LSTM) [Li et al. 2017] (43), modelos de conjunto (*ensemble*) [Siwek and Osowski 2016] (32), máquinas de vetores de suporte (SVM, de *support vector machines*) [de Hoogh et al. 2018] (28), regressão [Weichenthal et al. 2016] (25), regressão de vetores de suporte (SVR, de *support vector regression*) [Hu et al. 2017] (25), redes neurais convolucionais (CNN, de *convolutional neural networks*) [Huang and Kuo 2018] (19), árvores de decisão (*decision trees*) [Rybarczyk and Zalakeviciute 2016] (19) e máquinas de aprendizado extremo (ELM, de *extreme learning machines*) [Wang et al. 2017] (14).

Nos artigos analisados, observou-se vários parâmetros, incluindo dados de meteorologia (temperatura, umidade relativa, pressão atmosférica, direção e velocidade do vento, precipitação), de satélite AOD (*Aerosol Optical Depth*), de uso da terra, temporais, topográficos, de tráfego de veículos, de vegetação, de densidade populacional e espaciais. Majoritariamente, o principal parâmetro utilizado nos trabalhos envolveu a concentração de poluentes. Os poluentes atmosféricos mais analisados com as técnicas de aprendizado foram os materiais particulados $MP_{2.5}$ (149), os materiais particulados MP_{10} (67), dióxido de nitrogênio NO_2 (59), ozônio O_3 (56), dióxido de enxofre SO_2 (43), monóxido de carbono CO (35), óxidos de nitrogênio NO_x (23), dióxido de carbono CO_2 (8), monóxido de nitrogênio NO_2 (7) e compostos orgânicos voláteis VOC (5).

3.3.2. Aprendizado de máquina aplicado à poluição sonora

Utilizou-se a *string* de pesquisa: (“noise pollut*”) AND (“machine learn*” OR “deep learn*” OR “artificial intelligen*”). Após filtragem e remoção de duplicatas, 37 publicações foram selecionadas para análise.

As principais funcionalidades dos métodos nos últimos 5 anos foram a classificação de tipos de ruídos [Kumari et al. 2019], a predição de indicadores de níveis de ruído [Torija and Ruiz 2015], a classificação de níveis de ruído [Torija and Ruiz 2016], avaliação do incômodo causado por ruídos [Bravo-Moncayo et al. 2019] e a detecção de eventos anômalos [Alías et al. 2018]. Para essas aplicações, as técnicas mais

usadas foram: máquinas de vetor de suporte (SVM) [Torija and Ruiz 2016] (12), redes neurais artificiais (ANN) [Torija and Ruiz 2016] (11), redes neurais convolucionais (CNN) [Kumari et al. 2019] (11), florestas aleatórias [Liu et al. 2020] (6), árvores de decisão [Ali Khalil et al. 2019] (5), regressão [Bravo-Moncayo et al. 2019] (3), regressão de vetores de suporte (SVR) [Giannakopoulos et al. 2019] (3), processos gaussianos [Torija and Ruiz 2015] (2), modelos de conjunto [Nourani et al. 2020] (2) e redes neurais recorrentes LSTM [Navarro et al. 2020] (2).

3.4. *Big Data*

A integração de diferentes tipos de dados (i.e., ambientais, demográficos, econômicos, meteorológicos, energéticos e de transporte), em plataformas de *Big Data* aumenta a precisão das estimativas da poluição ambiental e melhor subsidia a tomada de decisão no seu gerenciamento. Segundo [Ma et al. 2020], a maior parte dos estudos na literatura sobre poluição de ar analisam uma só cidade (ou pequena região) e a influência de fatores de um mesmo tipo (e.g., meteorologia). O uso de tecnologias de *Big Data* para o armazenamento e processamento eficientes de grandes volumes de dados possibilita a realização de análises multivariadas abrangentes, que podem considerar até centenas de fatores de diferentes tipos em uma escala nacional.

Um exemplo dessa abordagem é a plataforma de *mobile Big Data*¹ desenvolvida pela Telefônica Brasil (Vivo) em parceria com a prefeitura de São Paulo. A plataforma prediz os níveis de poluição na cidade com antecedência de 24h a 48h. Para a análise dos padrões de mobilidade e os níveis de qualidade de ar associados, a plataforma usa algoritmos de aprendizado de máquina, dados anonimizados e agregados da rede de telefonia celular, dados de sensores de trânsito e de qualidade de ar e dados de estações meteorológicas. Os algoritmos extrapolam as previsões para as regiões que não são cobertas pela infraestrutura de monitoramento. O projeto é parte da *Big Data for Social Good*, uma iniciativa que congrega organizações públicas e privadas para a adequação da indústria de tecnologia móvel à Agenda 2030 de Objetivos de Desenvolvimento Sustentável da ONU.

4. Conclusões

A avaliação da qualidade do ar já é monitorada no Brasil pelos órgãos ambientais estaduais há muito tempo, e a incorporação de dispositivos IoT de baixo custo é uma inovação recente, ampliando as análises espaço-temporais que hoje não são possíveis de se obter no ambiente urbano. No entanto, esse tipo de aplicação é complementar à rede metrológica pré-existente e não irá substituir estes mecanismos, mas sim apoiá-los com a ampliação do conhecimento sobre o comportamento e a fenomenologia da poluição ambiental em escalas diferentes das tradicionalmente apresentadas.

O ruído urbano é o segundo maior problema de poluição urbana e nem sempre seus riscos à saúde são bem percebidos pela população, com a inexistência de uma rede de monitoramento no Brasil e poucos mapas de ruídos. Assim, a perspectiva da utilização de IoT e aprendizado de máquina pode revolucionar o planejamento urbano nas cidades brasileiras, de forma a reorganizar planos de mobilidades dessas cidades e motivar ações de políticas públicas que promovam menor emissões de ruídos, como o emprego de veículos elétricos.

¹<https://www.gsma.com/betterfuture/resources/telefonica-case-study-predicting-air-pollution-levels-24-to-48-hours-in-advance-in-sao-paulo-brazil>

Na questão da poluição do solo, a aquisição de dados poderá ser otimizada por meio de estações de monitoramento automatizadas, dotadas de sensores, *data loggers* e outras ferramentas de IoT instaladas *in-situ*, com transmissão de dados para sistemas de armazenamento em nuvem, por redes celulares ou LPWAN. Quanto ao processamento dos dados, técnicas de aprendizado de máquina e *Big Data* vem demonstrando grande potencial para modelagem de sistemas multivariados, na previsão do transporte dos contaminantes no subsolo e para o estabelecimento de redes de monitoramento otimizadas.

Referências

- Ali Khalil, M., Hamad, K., and Shanableh, A. (2019). Developing machine learning models to predict roadway traffic noise: An opportunity to escape conventional techniques. *Transportation Research Record*, 2673(4):158–172.
- Alías, F., Alsina-Pagès, R. M., Orga, F., and Socoró, J. C. (2018). Detection of anomalous noise events for real-time road-traffic noise mapping: The Dynamap’s project case study. *Noise Mapping*, 5(1):71–85.
- Bello, J. P., Silva, C., Nov, O., DuBois, R. L., Arora, A., Salamon, J., Mydlarz, C., and Doraiswamy, H. (2018). SONYC: A system for the monitoring, analysis and mitigation of urban noise pollution. *arXiv preprint arXiv:1805.00889*.
- Bravo-Moncayo, L., Lucio-Naranjo, J., Chávez, M., Pavón-García, I., and Garzón, C. (2019). A machine learning approach for traffic-noise annoyance assessment. *Applied Acoustics*, 156:262–270.
- Cardoso, J. (2009). *Introdução à Química Ambiental*. Artmed Editora.
- CETESB (2001). Manual de gerenciamento de Áreas contaminadas. Technical report, CETESB.
- CETESB (2015). Qualidade do ar no estado de SP 2014. Technical report, CETESB.
- CETESB (2018). Emissões veiculares no estado de SP 2017. Technical report, CETESB.
- CETESB (2020). Relatório de áreas contaminadas e reabilitadas no estado de São Paulo. Technical report, CETESB.
- Council, N. R. (2004). *Air Quality Management in the United States*. The National Academies Press, Washington, DC.
- de Hoogh, K., Héritier, H., Stafoggia, M., Künzli, N., and Kloog, I. (2018). Modelling daily PM_{2.5} concentrations at high spatio-temporal resolution across Switzerland. *Environmental Pollution*, 233:1147–1154.
- Giannakopoulos, T., Orfanidi, M., and Perantonis, S. (2019). Athens urban soundscape (ATHUS): A dataset for urban soundscape quality recognition. In *International Conference on Multimedia Modeling*, pages 338–348. Springer.
- Goel, A. and Kumar, P. (2015). Characterisation of nanoparticle emissions and exposure at traffic intersections through fast-response mobile and sequential measurements. *Atmospheric Environment*, 107:374 – 390.
- Günther, W. M. R. (2006). Áreas contaminadas no contexto da gestão urbana. *São Paulo em Perspectiva*, 20:105–117.

- Horst, J., Welty, N., Schnobrich, M., Sinha, P., and Kulkarni, P. (2017). Digital innovation: The next disruptive but transformative remediation frontier. *Groundwater Monitoring and Remediation*, 37(3):19–27.
- Hu, K., Rahman, A., Bhrugubanda, H., and Sivaraman, V. (2017). HazeEst: Machine learning based metropolitan air pollution estimation from fixed and mobile sensors. *IEEE Sensors Journal*, 17(11):3517–3525.
- Huang, C.-J. and Kuo, P.-H. (2018). A deep CNN-LSTM model for particulate matter (PM_{2.5}) forecasting in smart cities. *Sensors*, 18(7):2220.
- IEMA (2014). 1º diagnóstico da rede de monitoramento da qualidade do ar no Brasil. Technical report, Instituto de Energia e Meio Ambiente, São Paulo, SP, Brasil.
- Karagulian, F., Barbieri, M., Kotsev, A., Spinelle, L., Gerboles, M., Lagler, F., Redon, N., Crunaire, S., and Borowiak, A. (2019). Review of the performance of low-cost sensors for air quality monitoring. *Atmosphere*, 10(9):506.
- Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Sabatino, S. D., Bell, M., Norford, L., and Britter, R. (2015). The rise of low-cost sensing for managing air pollution in cities. *Environment International*, 75:199 – 205.
- Kumari, S., Roy, D., Cartwright, M., Bello, J. P., and Arora, A. (2019). EdgeL³: Compressing L³-Net for mote scale urban noise monitoring. In *2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, pages 877–884. IEEE.
- Lewis, A., von Schneidmesser, E., and Peltier, R. (2018). Low-cost sensors for the measurement of atmospheric composition: overview of topic and future applications. Technical report, World Meteorological Organization.
- Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., and Chi, T. (2017). Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. *Environmental pollution*, 231:997–1004.
- Liu, Y., Goudreau, S., Oiamo, T., Rainham, D., Hatzopoulou, M., Chen, H., Davies, H., Tremblay, M., Johnson, J., Bockstael, A., et al. (2020). Comparison of land use regression and random forests models on estimating noise levels in five Canadian cities. *Environmental Pollution*, 256:113367.
- Ma, J., Ding, Y., Cheng, J. C., Jiang, F., Tan, Y., Gan, V. J., and Wan, Z. (2020). Identification of high impact factors of air quality on a national scale using big data and machine learning techniques. *Journal of Cleaner Production*, 244:118955.
- Morawska, L. (2018). Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone? *Environment international*, 116:286–299.
- Navarro, J. M., Martínez-España, R., Bueno-Crespo, A., Martínez, R., and Cecilia, J. M. (2020). Sound levels forecasting in an acoustic sensor network using a deep neural network. *Sensors*, 20(3):903.
- Ni, X., Huang, H., and Du, W. (2017). Relevance analysis and short-term prediction of PM_{2.5} concentrations in Beijing based on multi-source data. *Atmospheric environment*, 150:146–161.

- Nourani, V., Gökçekeş, H., and Umar, I. K. (2020). Artificial intelligence based ensemble model for prediction of vehicular traffic noise. *Environmental research*, 180:108852.
- Raza, U., Kulkarni, P., and Sooriyabandara, M. (2017). Low power wide area networks: An overview. *IEEE Communications Surveys & Tutorials*, 19(2):855–873.
- Rybarczyk, Y. and Zalakeviciute, R. (2016). Machine learning approach to forecasting urban pollution. In *2016 IEEE Ecuador Technical Chapters Meeting (ETCM)*, pages 1–6. IEEE.
- Sales, E. and Aquilino, M. (2017). Desenvolvimento de método para mapeamento sonoro da cidade de São Paulo. Technical report, IPT, São Paulo, SP, Brasil.
- Sanislav, T., Mois, G., and Miclea, L. (2016). An approach to model dependability of cyber-physical systems. *Microprocessors and Microsystems*, 41:67–76.
- Santos, A. S. D. (2018). *Análise espaçotemporal da qualidade do ar em vias urbanas por meio de Redes de Sensores com nós embarcados em ônibus coletivos*. PhD thesis, Universidade de São Paulo.
- Siwek, K. and Osowski, S. (2016). Data mining methods for prediction of air pollution. *International Journal of Applied Mathematics and Computer Science*, 26(2):467–478.
- Teixeira, C. E., Motta, F. G., and de Moraes, S. L. (2016). *Panorama GAC: mapeamento da cadeia de gerenciamento de áreas contaminadas*. IPT.
- Torija, A. J. and Ruiz, D. P. (2015). A general procedure to generate models for urban environmental-noise pollution using feature selection and machine learning methods. *Science of the Total Environment*, 505:680–693.
- Torija, A. J. and Ruiz, D. P. (2016). Automated classification of urban locations for environmental noise impact assessment on the basis of road-traffic content. *Expert Systems with Applications*, 53:1–13.
- Vasconcellos, P. A. C., Carvalho, L. R. F., and Pool, C. S. (2005). Volatile organic compounds inside urban tunnels of São Paulo City, Brazil. *Journal of the Brazilian Chemical Society*, 16:1210 – 1216.
- Wang, D., Wei, S., Luo, H., Yue, C., and Grunder, O. (2017). A novel hybrid model for air quality index forecasting based on two-phase decomposition technique and modified extreme learning machine. *Science of The Total Environment*, 580:719–733.
- Weichenthal, S., Van Ryswyk, K., Goldstein, A., Bagg, S., Shekharizfard, M., and Hatzopoulou, M. (2016). A land use regression model for ambient ultrafine particles in Montreal, Canada: A comparison of linear regression and a machine learning approach. *Environmental research*, 146:65–72.
- Wong, C. S. C., Li, X., and Thornton, I. (2006). Urban environmental geochemistry of trace metals. *Environmental Pollution*, 142:1–16.
- Zimmerman, N. e. a. (2018). A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. *Atmospheric Measurement Techniques*, 11(1).