

PREVISÃO HORÁRIA DA CONCENTRAÇÃO DE OZÔNIO TROPOSFÉRICO USANDO APRENDIZAGEM PROFUNDA

Lucas Vilas Boas Alves¹; Erick Giovani Sperandio Nascimento²; Davidson Martins Moreira³

¹ Bolsista; Projeto de Implantação de Infraestrutura de Pesquisa em Simulação e Modelagem Computacional no Estado da Bahia Utilizando Processamento de Alto Desempenho – Fundação de Amparo à Pesquisa do Estado da Bahia - FAPESB; lucas.alves@fbter.org.br

² Doutorado; Centro Universitário SENAI CIMATEC; Salvador-BA; erick.sperandio@fieb.org.br

³ Doutorado; Centro Universitário SENAI CIMATEC; Salvador-BA; davidson.moreira@fieb.org.br

RESUMO

O objetivo deste trabalho é construir e avaliar um modelo baseado em aprendizagem profunda para estimar os níveis de ozônio troposférico por hora, até vinte e quatro horas à frente, utilizando dados do sistema automático de monitoramento da qualidade do ar na região metropolitana da cidade de Vitória, Espírito Santo, Brasil. Foram testados alguns tipos e arquiteturas de redes neurais artificiais aplicando como variáveis de entrada as concentrações de poluentes atmosféricos e parâmetros meteorológicos. O desempenho do modelo foi avaliado a partir das previsões sobre os dados de teste usando estatísticas como o coeficiente de correlação de Pearson (R), o coeficiente de regressão (R²), o fator de dois (Fac2) e o erro percentual absoluto médio (MAPE).

PALAVRAS-CHAVE: redes neurais artificiais, previsão, qualidade do ar, ozônio

1. INTRODUÇÃO

O ozônio troposférico (O₃) é um poluente secundário na troposfera e um dos oxidantes fotoquímicos que causam problemas de qualidade do ar. Ele é formado a partir de reações químicas entre gases emitidos por fontes naturais e antropogênicas, como óxidos de nitrogênio e compostos orgânicos voláteis, na presença de radiação solar. O ozônio pode causar impactos prejudiciais à saúde humana, irritando o sistema respiratório, reduzindo a capacidade pulmonar e agravando problemas de asma e doenças pulmonares crônicas.¹ Além disso, pode danificar plantas causando vários tipos de sintomas, incluindo clorose e necrose, e afetando seu crescimento.²

No final de 2018, o Conselho Nacional do Meio Ambiente (CONAMA) introduziu uma nova resolução estabelecendo padrões atualizados de qualidade do ar e, mais especificamente, limiares de concentração de ozônio para níveis de atenção, alerta e emergência.³ Portanto, é importante desenvolver um modelo de previsão poderoso que possa ajudar as autoridades e a população a tomarem medidas preventivas e evitar riscos iminentes à saúde, mesmo antes que os limites de atenção sejam atingidos.

Deste modo, o presente trabalho tem como propósito a elaboração e a validação de uma rede neural artificial para prever as concentrações horárias de ozônio até vinte e quatro horas adiante usando medições da Rede Automática de Monitoramento da Qualidade do Ar (RAMQAr), de propriedade do Instituto Estadual de Meio Ambiente e Recursos Hídricos do Espírito Santo (IEMA-ES).

2. METODOLOGIA

Existe atualmente na literatura uma série de modelos de redes neurais artificiais, sendo fundamental a realização de experimentos para examinar qual arquitetura melhor se encaixa ao problema que precisa ser resolvido. Deste modo, foram investigadas algumas arquiteturas de redes neurais MLP (*Multilayer Perceptron*) e LSTM (*Long short-term memory*) utilizando 60% dos dados obtidos para treinamento e validação dos modelos e os 40% restantes para testes. A implementação dos ensaios foi realizada com a linguagem *Python* utilizando a plataforma *TensorFlow* em conjunto com sua API (*Application Programming Interface*) de alto nível *Keras*.

Os dados utilizados nos experimentos foram medidos com frequência horária durante o ano de 2002 na estação de monitoramento RAMQAr 8, localizada no bairro Vila Capixaba, município de Cariacica, região metropolitana da cidade de Vitória, Espírito Santo, Brasil. As amostras recolhidas foram tratadas para eliminar medições invalidadas e selecionar apenas parâmetros meteorológicos e da qualidade do ar de interesse. Este procedimento resultou em um conjunto de dados com 1108 amostras divididas em três períodos contínuos referentes aos meses de fevereiro, março, outubro e novembro de 2002.

Um total de dezessete parâmetros observacionais das concentrações de poluentes atmosféricos e propriedades meteorológicas foram utilizados como variáveis de entrada do modelo, uma vez que representam o

estado do fluido atmosférico em termos de suas características e composição química ao longo do tempo. Estas variáveis são listadas abaixo.

- **Parâmetros temporais:** Ano, mês, dia, hora e minuto da amostra;
- **Parâmetros de qualidade do ar:** Concentrações de partículas inaláveis, partículas totais em suspensão, dióxido de enxofre, monóxido de carbono, monóxido de nitrogênio, dióxido de nitrogênio, óxidos de nitrogênio e ozônio;
- **Parâmetros meteorológicos:** Velocidade e direção escalar do vento, temperatura e umidade relativa.

Como objetivo do modelo foi estabelecido a predição da concentração de ozônio nas 24 horas subsequentes de modo simultâneo. O desempenho das previsões foi mensurado com relação aos dados de teste usando estatísticas como o coeficiente de correlação de Pearson (R), o coeficiente de regressão (R2), o fator de dois (Fac2), o erro mínimo, o erro máximo e o erro percentual absoluto médio (MAPE).

3. RESULTADOS E DISCUSSÃO

A rede neural artificial que obteve os melhores resultados nas métricas avaliadas foi uma *Multilayer Perceptron* com 17 neurônios na camada de entrada, 8 camadas ocultas e camada de saída com 24 neurônios. A Tabela 1 resume a arquitetura e as características do modelo que alcançou o melhor desempenho nos ensaios realizados. O treinamento desta rede utilizou o algoritmo *Feed-forward Backpropagation* com taxa de aprendizado de 0.394, otimizador *Adadelta* e função de erro MSE (*Mean Squared Error*).

Tabela 1 - Arquitetura da Rede Neural Artificial

Camada	Quantidade de neurônios	Função de ativação	Número de parâmetros treináveis
Camada de entrada	17	N/A	0
1ª camada oculta	17	Linear	306
2ª à 8ª camada oculta	81	ReLU (Rectified Linear Unit)	41310
Camada de saída	24	Linear	1968

Considerando o modelo como um todo, levando em consideração todas as previsões de 1 a 24 horas à frente, os resultados apresentaram coeficiente R de 0,659, R2 de 0,435, Fac2 de 0,820 e MAPE de 47,83%, indicando um desempenho razoável para a rede neural artificial. A Tabela 2 resume, para alguns horizontes de previsão, as estatísticas do modelo utilizando o conjunto de dados de teste.

Tabela 2 - Estatísticas do modelo no conjunto de dados de teste

Previsão	Coeficiente de correlação de Pearson (R)	Coeficiente de regressão (R2)	Fator de dois (Fac2)	Erro Mínimo	Erro Máximo	Erro Percentual Absoluto Médio (MAPE)
Ozônio (t+1)	0,848	0,719	0,918	0,037470	26,313	33,98%
Ozônio (t+12)	0,650	0,423	0,822	0,010683	43,990	47,89%
Ozônio (t+24)	0,645	0,415	0,807	0,012179	53,512	48,16%
Geral	0,659	0,435	0,820	0,000025	54,822	47,83%

A precisão do modelo é apresentada graficamente nas Figuras 1 e 2, onde são exibidas as previsões de ozônio para 1 hora adiante e para 24 horas adiante realizadas com base em dados de fevereiro de 2002. Estas imagens expõem uma série de 141 amostras, ou aproximadamente 6 dias de medições, com as linhas azuis representando as concentrações reais de ozônio e as linhas laranjas representando as previsões da concentração de ozônio.

É importante apontar que as previsões a respeito da formação do ozônio tornam-se mais difíceis por este ser um poluente secundário e, conseqüentemente, depender de outros parâmetros como a intensidade da radiação solar, e a concentração de óxidos de nitrogênio e compostos orgânicos voláteis na atmosfera.

Figura 1 - Concentração de ozônio observada e prevista 1 hora adiante utilizando dados de teste

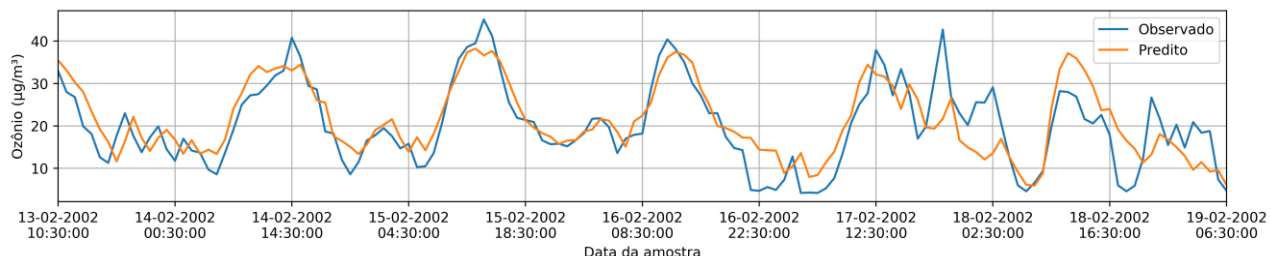
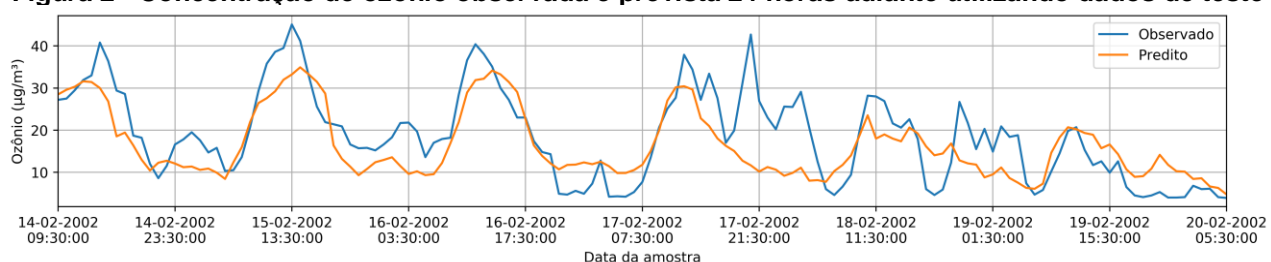


Figura 2 - Concentração de ozônio observada e prevista 24 horas adiante utilizando dados de teste



4. CONSIDERAÇÕES FINAIS

Os resultados indicaram um desempenho razoável para o modelo de previsão e mostraram que modelos de aprendizagem profunda podem ser aplicados com sucesso à previsão horária da concentração de ozônio para áreas urbanas. Uma vez que tais modelos são treinados e ajustados aos dados, o processo de inferência, ou seja, o procedimento de previsão, tem um custo computacional muito baixo, o que significa que ele pode ser usado como uma abordagem alternativa em comparação com sistemas de modelagem numérica, que exigem muito mais poder computacional.

Agradecimentos

Este trabalho contou com o apoio da FAPESB (Fundação de Amparo à Pesquisa do Estado da Bahia), na forma de uma bolsa de Desenvolvimento de Inovação Tecnológica no projeto "Implantação de Infraestrutura de Pesquisa em Simulação e Modelagem Computacional no Estado da Bahia Utilizando Processamento de Alto Desempenho", desenvolvido no Centro de Supercomputação para Inovação Industrial do SENAI CIMATEC.

5. REFERÊNCIAS

¹ NETTO, Fabiola et al. **Formação de ozônio troposférico: UMA REVISÃO DA LITERATURA**. In: XXV Simpósio Internacional de Engenharia Automotiva. São Paulo, 2017. Disponível em: https://www.researchgate.net/publication/319969924_FORMACAO_DE_OZONIO_TROPOSFERICO_UMA_REVISAO_DA_LITERATURA. Acesso em: 25 de mar. 2019.

² ESPOSITO, Marisia. **Estudo da relação entre necroses foliares indicadoras de ozônio atmosférico e defesas antioxidativas em folhas de Nicotiana tabacum "Bel W3"**. Dissertação (Mestrado) - Instituto de Botânica da Secretaria de Estado do Meio Ambiente. São Paulo, p. 102. 2008. Disponível em: http://arquivos.ambiente.sp.gov.br/pgibt/2013/10/Marisia_Pannia_Esposito_MS.pdf. Acesso em: 25 de mar. 2019.

³ BRASIL. Ministério do Meio Ambiente. Conselho Nacional do Meio Ambiente. **Resolução nº 491, de 19 de novembro de 2018**. Disponível em: <http://www2.mma.gov.br/port/conama/legiabre.cfm?codlegi=740>. Acesso em: 25 de mar. 2019.