

IMPACTO DA POLUIÇÃO ATMOSFÉRICA E DAS ALTERAÇÕES CLIMÁTICAS NA SAÚDE POPULACIONAL UTILIZANDO REDES NEURAIIS ARTIFICIAIS

ATMOSPHERIC POLLUTION AND CLIMATE CHANGES IMPACT ON POPULATION'S HEALTH USING ARTIFICIAL NEURAL NETWORKS

Yara S. Tadano, Thiago Antonini Alves, Natalia S.S. Silva, Hugo S. Valadares

Universidade Tecnológica Federal do Paraná, Ponta Grossa, PR, Brasil

RESUMO

A poluição atmosférica tem sido investigada globalmente, devido aos riscos à saúde. A análise dos riscos à saúde devido às emissões de poluentes atmosféricos e às alterações climáticas é comumente realizada com o uso de regressões estatísticas. Entretanto, uma alternativa inovadora consiste no uso de Redes Neurais Artificiais (RNAs). Desta forma, o objetivo deste trabalho foi comparar o desempenho de diferentes RNAs na avaliação do impacto da poluição na saúde. Como estudo de caso foi analisado o impacto de material particulado com diâmetro aerodinâmico menor ou igual a 10 µm (PM₁₀) e alterações climáticas (temperatura e umidade) no número de internações por doenças respiratórias na cidade de Campinas/SP, Brasil. Os dados experimentais de poluição atmosférica e meteorológicos foram provenientes de estações de monitoramento da qualidade do ar localizadas na região de estudo. As RNAs, Máquinas de Aprendizado Extremo e Redes Neurais com Estados de Eco, foram aplicadas com sucesso.

ABSTRACT

Air pollution has been investigated worldwide, due to its health risks. The health risk analysis due to air pollutants emissions and climate changes are commonly performed using statistical regressions. However, an innovative alternative consists on using Artificial Neural Networks (ANNs). Thus, this work aims to compare the performance of different ANNs on the analysis of air pollution health impacts. As case study, it was analyzed the impact of particulate matter with aerodynamic diameter less or equal to 10 µm (PM₁₀) and climate changes (temperature and humidity) on the number of hospital admissions by respiratory diseases in Campinas city/Brazil. The experimental data of atmospheric pollution and meteorology were obtained from air quality monitoring stations located in the study area. The ANNs, Extreme Learning Machines and Echo State Networks was applied successfully.

1. INTRODUÇÃO

A poluição atmosférica tem sido um assunto chave, investigada globalmente, principalmente devido aos seus riscos à saúde humana, causando severos problemas respiratórios,

cardiovasculares, entre outros, podendo levar até à morte. A Organização Mundial da Saúde – OMS (*World Health Organization - WHO*) reportou mais de dois milhões de mortes por ano atribuídas à poluição do ar (WHO, 2006).

Geralmente, a análise dos riscos à saúde populacional devido às emissões de poluentes atmosféricos e às alterações climáticas é realizada com o uso de Métodos de Regressão Estatística, mais especificamente Modelos Lineares Generalizados e Modelos Aditivos Generalizados (Tadano *et al.*, 2012). Entretanto, uma metodologia alternativa inovadora consiste na utilização de Redes Neurais Artificiais (RNAs).

Desta forma, este trabalho tem o objetivo de comparar o desempenho de diferentes RNAs na avaliação do impacto da poluição atmosférica na saúde humana, como estudo de caso foi analisado o impacto de material particulado com diâmetro aerodinâmico menor ou igual à 10 μm (PM₁₀) e alterações climáticas (temperatura e umidade) no número de internações por doenças respiratórias na cidade de Campinas, estado de São Paulo, sudeste do Brasil.

2. REDES NEURAS DESORGANIZADAS

De acordo com a literatura consultada, Nascimento *et al.* (2009) foram um dos pioneiros a utilizar RNA para avaliar o impacto da poluição atmosférica na saúde humana, sendo que os autores utilizaram a arquitetura clássica rede *Perceptron* de Múltiplas Camadas (*Multilayer Perceptron* – MLP). Entretanto, outras redes neurais podem ser aplicadas, como as Máquinas Desorganizadas (*Unorganized Machines* - UM), tais como: Máquinas de Aprendizagem Extremo (*Extreme Learning Machines* - ELM) e Redes Neurais com Estados de Eco (*Echo State Networks* - ESN) (Boccatto *et al.*, 2011).

2.1. Máquinas de Aprendizagem Extremo - ELM

Máquinas de Aprendizagem Extremo (ELM) são redes neurais *feedforward*, estruturalmente muito semelhantes às tradicionais MLP (Haykin, 1998; Huang *et al.*, 2006). A principal diferença entre tais estruturas é que, no processo de treinamento, as ELMs não ajustam os pesos da camada intermediária, a qual possui neurônios gerados de forma aleatória e

independente. Neste caso, o treinamento passa a ser resumido em encontrar os melhores pesos da camada de saída, por meio de uma solução de um problema de otimização via mínimos quadrados, baseada no paradigma supervisionado, isto é, com a utilização de um sinal de referência. Esta característica faz com que o ajuste dos pesos da rede seja rápido e eficiente computacionalmente (Huang *et al.*, 2015).

As ELMs foram propostas por Huang *et al.* (2006), os quais apresentaram por meio de rigorosa demonstração matemática que os pesos da camada intermediária podem ser escolhidos de forma arbitrária, desde que a função de ativação dos pesos seja infinitamente diferenciável. Por meio de uma abordagem construtiva, os autores mostraram a capacidade de aproximação universal da estrutura, ou seja, tais redes podem aproximar, com erro arbitrário, qualquer mapeamento não linear. A prerrogativa discutida é que o erro de aproximação sempre pode ser reduzido com a introdução de novos neurônios na camada escondida. Além disso, tais estruturas possuem capacidade de generalização, uma vez que podem dar respostas adequadas, mesmo frente a dados de entrada desconhecidos, desde que sejam treinadas devidamente (Huang *et al.*, 2006; Siqueira *et al.*, 2014). A Fig. 1 apresenta um esquema da ELM, na qual o vetor $\mathbf{u}_n = [u_n, u_{n-1}, \dots, u_{n-K+1}]^T$ contém o sinal de entrada.

Os dados são transmitidos para os neurônios de saída por meio de combinações não lineares, com os coeficientes especificados

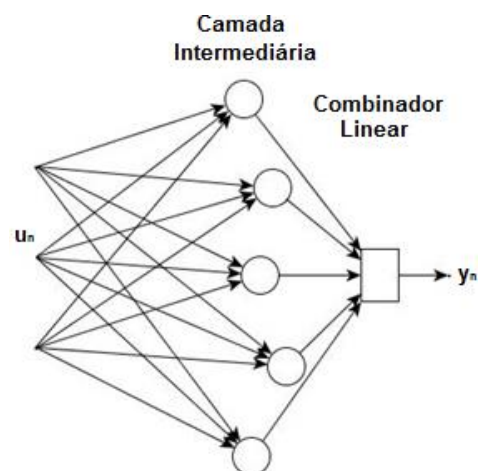


Fig. 1 – Máquina de Aprendizagem Extremo (ELM)

na matriz \mathbf{W}^h , cujas ativações são calculadas, de acordo com Huang *et al.* (2006), por:

$$\mathbf{x}_n^h = \mathbf{f}^h(\mathbf{W}^h \mathbf{u}_n + \mathbf{b}) \quad (1)$$

sendo que, \mathbf{b} é o *bias* de cada neurônio e $\mathbf{f}^h(\cdot)$ é a função de ativação dos neurônios intermediários. Tais ativações são linearmente combinadas para produzir a saída da rede:

$$\mathbf{y} = \mathbf{W}^{\text{out}} \mathbf{x}_n^h \quad (2)$$

onde \mathbf{W}^{out} é a matriz dos pesos da camada de saída.

Como mencionado anteriormente, o treinamento da rede é bastante simplificado sendo, essencialmente, resumido a encontrar os coeficientes ótimos de um combinador linear, por meio de uma solução de mínimos quadrados. Para resolver esta tarefa, usualmente o Operador de *Moore-Penrose* é utilizado.

$$\mathbf{W}^{\text{out}} = (\mathbf{X}_h^T \mathbf{X}_h)^{-1} \mathbf{X}_h^T \mathbf{d} \quad (3)$$

sendo que, $\mathbf{X}_h \in \mathcal{R}^{T_s \times N_k}$ é a matriz contendo as saídas da camada intermediária, $(\mathbf{X}_h^T \mathbf{X}_h)^{-1} \mathbf{X}_h^T$ é a pseudoinversa de \mathbf{X}_h e $\mathbf{d} \in \mathcal{R}^{T_s \times N_k}$ é o sinal desejado.

2.2. Redes Neurais com Estado de Eco - ESN

Redes Neurais Recorrentes (RNNs) são munidas de laços de realimentação, originando uma memória intrínseca, o que pode ser particularmente favorável em problemas que envolvam dependência temporal entre as amostras, como tarefas dinâmicas ou de previsão de séries temporais. A presença da realimentação pode levar a respostas mais precisas da rede, uma vez que mais informações estão disponíveis para a formação da saída da rede (Jaeger, 2001).

Nesta categoria encontram-se as Redes Neurais com Estado de Eco (ESNs), propostas por *H. Jaeger* em 2001, que assim como as ELMs, também são aproximadores universais e possuem capacidade de generalização (Jaeger, 2001; Schafer & Zimmermann 2007).

A principal dificuldade na aplicação de Re-

Redes Recorrentes Clássicas é, sem dúvida, o seu treinamento. Neste caso, é necessário ajustar todos os pesos da rede, tarefa realizada por meio da aplicação de técnicas de otimização não linear irrestrita. Este processo pode levar a instabilidade, a dificuldades na manipulação da função custo a ser otimizada, a convergência local e, em geral, o custo computacional é bastante elevado (Yildiz *et al.*, 2012).

As ESNs, por outro lado, contornam estes problemas uma vez que, assim como no caso das ELMs, sua camada intermediária (ou reservatório de dinâmicas) composta por neurônios não lineares e que contém laços de realimentação, permanece sem ser ajustada. O processo de treinamento é semelhante ao caso das ELMs, ou seja, encontrar os coeficientes de um combinador linear baseado em um problema de mínimos quadrados com sinal de referência. Novamente, tal metodologia confere à estrutura um processo de treinamento simples, eficiente e com boas soluções para problemas dinâmicos ou estáticos (Boccatto *et al.*, 2012).

A Fig. 2 mostra a estrutura genérica de uma ESN, na qual o vetor de entradas é novamente $\mathbf{u}_n = [u_n, u_{n-1}, \dots, u_{n-K+1}]^T$, que são transmitidas da camada de entrada \mathbf{W}^{in} para o reservatório de dinâmicas por meio de combinações não lineares.

O vetor de estados de eco \mathbf{x}_n é formado pela saída dos neurônios intermediários presentes na matriz \mathbf{W} , a qual contém os laços

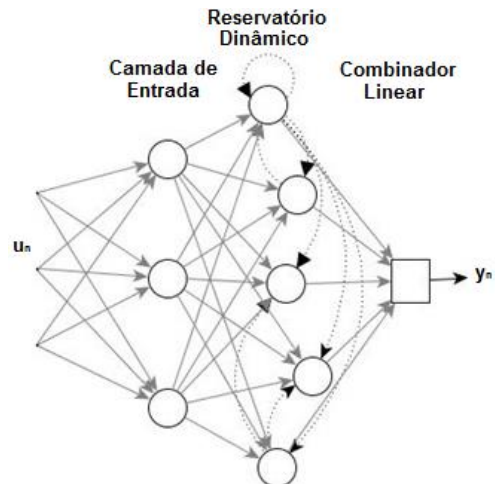


Fig. 2 – Redes Neurais com Estados de Eco (ENS)

de recursão e são calculados por:

$$\mathbf{x}_{n+1} = \mathbf{f}(\mathbf{W}^{\text{in}} \mathbf{u}_{n+1} + \mathbf{W} \mathbf{x}_n), \quad (4)$$

sendo que, $\mathbf{f}(\cdot) = (f_1(\cdot), f_2(\cdot), \dots, f_N(\cdot))$ contém as ativações dos neurônios do reservatório. A saída da rede é o vetor \mathbf{y}_n calculado por:

$$\mathbf{y}_{n+1} = \mathbf{f}^{\text{out}}(\mathbf{W}^{\text{out}} \mathbf{x}_{n+1}), \quad (5)$$

onde $\mathbf{f}^{\text{out}}(\cdot) = (f_1^{\text{out}}(\cdot), f_N^{\text{out}}(\cdot), \dots, f_N^{\text{out}}(\cdot))$ são as funções de ativação dos neurônios da camada de saída.

O cálculo dos pesos da camada de saída pode ser realizado por meio da aplicação do Operador de *Moore-Penrose*, substituindo \mathbf{X}_h por \mathbf{X} na Eq. (3).

O projeto de reservatório dinâmico segue a proposta inicial de Jaeger (2001):

$$\mathbf{W} = \begin{cases} 0,4 \text{ com probabilidade de } 0,025 \\ -0,4 \text{ com probabilidade de } 0,025 \\ 0,0 \text{ com probabilidade de } 0,950 \end{cases} \quad (6)$$

O elemento teórico que garante a presença de memória nas ESNs é conhecido por propriedade de estados de eco (Jaeger, 2001). Desta propriedade, em condições específicas no reservatório definido na matriz \mathbf{W} , o histórico recente das entradas é que rege a dinâmica interna dos neurônios desta camada. Desta forma, tais estados se tornam transformações não lineares e diz-se que a rede possui eco. Como mencionado anteriormente, a consequência imediata é que os pesos da rede podem ser definidos previamente uma vez que a memória dinâmica está assegurada (Jaeger, 2001; Yildiz *et al.*, 2012).

3. ESTUDO DE CASO

No intuito de comparar o desempenho de Redes Neurais Artificiais em estudos epidemiológicos, o estudo de caso considerado consistiu na avaliação do impacto do material particulado com diâmetro aerodinâmico menor ou igual a $10 \mu\text{m}$ (PM_{10}) e de alterações climáticas, como variação de temperatura e umidade relativa do ar, no número de internações por doenças respiratórias na população da cidade de Campinas/ Brasil, de 2007 a 2009.

A cidade de Campinas está situada no estado de São Paulo, sudeste do Brasil. Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE, 2016), em 2009, a população de Campinas era de aproximadamente 1,1 milhões de habitantes distribuída em uma área de mais de 790 km^2 . Campinas apresenta clima tropical de altitude, com verões quentes e chuvosos, com temperaturas médias no verão de $22 \text{ }^\circ\text{C}$ e de $18 \text{ }^\circ\text{C}$ no inverno. O índice pluviométrico anual é da ordem de 1.400 mm , com chuvas concentradas principalmente de outubro a março (Estado de São Paulo, 2011).

3.1. Base de dados utilizada

Os dados experimentais de poluição atmosférica (concentração de PM_{10}) e meteorológicos (temperatura ambiente e umidade relativa do ar), de 01 de janeiro de 2007 a 31 de dezembro de 2009, foram provenientes de estações de monitoramento da qualidade do ar administradas pela Companhia Ambiental do Estado de São Paulo (CETESB, 2010). Os dados de internações por doenças respiratórias foram obtidos junto ao Sistema Único de Saúde Brasileiro (DataSUS, 2010).

Uma análise descritiva dos dados foi realizada e foi observado que a concentração máxima diária de PM_{10} ($129 \mu\text{g}/\text{m}^3$) não ultrapassou os padrões primário e secundário de qualidade do ar ($150 \mu\text{g}/\text{m}^3$). Quanto aos dados meteorológicos, não houve nenhum dia que apresentou características de eventos extremos, estando sempre dentro da normalidade da região, temperaturas entre 15 e $30 \text{ }^\circ\text{C}$ e umidade relativa do ar entre 38 e 99% .

Para verificar a relação existente entre a variável resposta (internações por doenças respiratórias) e as variáveis explicativas consideradas neste estudo (concentração de PM_{10} , temperatura e umidade relativa do ar) foi construída uma matriz de correlação de *Pearson* (Quadro 1), onde estão identificadas as correlações estatisticamente significativas. A correlação entre internações por doenças respiratórias (DR) e concentração de PM_{10} foi positiva e estatisticamente significativa, indicando que o número de internações por DR aumenta conforme a concentração de PM_{10} aumenta.

Quadro 1 - Matriz de correlações de *Pearson* entre internações por doenças respiratórias (DR), concentração de PM₁₀ e variáveis meteorológicas

	DR	PM ₁₀	Temp.	Umid.
DR	1,00	-	-	-
PM ₁₀	0,35*	1,00	-	-
Temp.	-0,08	0,10	1,00	-
Umid.	-0,19*	-0,57*	-0,33*	1,00

O Quadro 1 mostra ainda correlações negativas do número de internações por DR com a temperatura (temp.) e com a umidade (umid.), indicando que quanto menor a temperatura e a umidade, maior a quantidade de internações por DR, apesar do coeficiente não se apresentar estatisticamente significativo para a temperatura. A correlação entre a concentração de PM₁₀ e a umidade relativa do ar se mostrou negativa e estatisticamente significativa, indicando que as concentrações de PM₁₀ são maiores em dias mais secos.

Em estudos epidemiológicos, existe um padrão que é frequentemente observado; as internações hospitalares tendem a diminuir durante os finais de semana e feriados. Portanto, também foram considerados dois fatores de confusão potenciais: dia da semana e feriados.

3.1. Aplicação de Redes Neurais Desorganizadas

O problema proposto a ser resolvido com a utilização de Redes Neurais Desorganizadas é a previsão de internações hospitalares por problemas respiratórios, utilizando como dados de entrada uma base de dados. Este problema pode ser encarado como uma tarefa de mapeamento não linear estático, passível de ser resolvido por meio da aplicação de Redes Neurais Artificiais *Feedforward* ou Recorrentes, as quais são aproximadores universais (Siqueira *et al.*, 2014).

Para o treinamento da rede neural foram utilizados os dados do período entre 01/01/2007 a 30/09/2008; a validação dos dados foi realizada para o período entre 01/10/2008 a 31/12/2008, e o processo de teste utilizou dados diários de todo o ano de 2009.

A etapa inicial para a obtenção dos resultados computacionais é treinar as ELMs

e ESNs. Cinco entradas foram consideradas: concentração diária de PM₁₀, temperatura ambiente, umidade relativa do ar, dia da semana e feriados. A resposta desejada é o número de internações por doenças respiratórias correspondente.

Além disso, é muito importante observar que o efeito na saúde provocado pela poluição atmosférica pode ocorrer alguns dias após a exposição. Desta forma, em estudos epidemiológicos de poluição atmosférica, é muito comum investigar o impacto da poluição até sete dias após a exposição (Tadano *et al.*, 2012). Esta característica foi levada em consideração neste estudo.

Para comparar o desempenho dos dois tipos de redes neurais utilizados (ELM e ESN), como métrica de erro foram adotados o Erro Quadrático Médio (*Mean Square Error* - MSE) e o Erro Percentual Absoluto Médio (*Mean Absolute Percentual Error* - MAPE), dados respectivamente por:

$$MSE = \frac{1}{N_s} \sum_{t=1}^{N_s} (d_t - y_t)^2, \quad (7)$$

e

$$MAPE = \frac{1}{N_s} \sum_{t=1}^{N_s} \left| \frac{d_t - y_t}{d_t} \right|, \quad (8)$$

sendo N_s o número de amostras, y_t a resposta das redes e d_t a saída desejada.

O Quadro 2 mostra o desempenho das Máquinas Desorganizadas (ELM e ESN) de zero a sete dias após a exposição para o conjunto de testes. Os resultados representam a média de 30 simulações. O anacronismo “NN” é relativo ao número de neurônios na camada intermediária (ou no reservatório de dinâmicas) obtidos no melhor desempenho (menor MSE) dentre todas as simulações. Um exemplo de como este processo foi realizado é mostrado na Fig. 3, onde é apresentado o MSE em função do número de neurônios para um caso considerando cinco dias de atraso com o uso da ESN. Como observado, o mínimo MSE de testes é alcançado com 30 neurônios na camada escondida.

Analisando o Quadro 2, pode ser observado que o menor MSE em dois casos (quatro e seis dias após a exposição) não coincide com o me-

Quadro 2 – Erro Quadrático Médio (MSE) e Erro Percentual Absoluto Médio (MAPE)

Dias após a exposição	RNA	NN	MSE	MAPE [%]
0	ELM	30	39,7793	35,1728
	ESN	30	41,5387	35,9539
1	ELM	20	37,3725	34,6598
	ESN	40	39,1007	35,1124
2	ELM	40	39,4254	35,4855
	ESN	70	40,9656	35,7968
3	ELM	20	39,1909	35,1523
	ESN	70	39,1695	34,8004
4	ELM	20	38,6538	35,5980
	ESN	40	38,6574	35,5083
5	ELM	30	39,1538	36,1703
	ESN	50	38,5759	34,9216
6	ELM	20	36,9819	34,8838
	ESN	40	38,1070	34,6889
7	ELM	20	38,9004	35,4133
	ESN	20	39,6671	36,1381

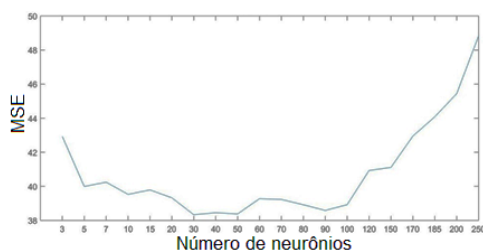


Fig. 3 - Definição do número de neurônios com base no MSE

nor MAPE. Neste trabalho, será considerado o MSE como o parâmetro para atestar o melhor desempenho, uma vez que é esta métrica que procura-se reduzir no processo de treinamento de uma rede neural.

Considerando a capacidade de aproximação e generalização das redes, em seis dos oito casos analisados, a ELM alcançou melhores resultados para o conjunto de testes, inclusive alcançando o melhor resultado geral - para o impacto seis dias após a exposição. A exceção foi para três e cinco dias após a exposição, onde a ESM apresentou melhores resultados. Este resultado era esperado, pois em estudos deste tipo o efeito do PM₁₀ que resulta em internações hospitalares costuma se manifestar em uma janela temporal entre cinco e seis dias, corroborando a importância de considerar os dias de atraso.

Ressalta-se ainda que um estudo preliminar foi feito com base na análise do MSE de uma Rede Neural MLP para seis dias de atrasos, visando analisar o desempenho da mesma frente às Redes Desorganizadas. Neste caso, o melhor resultado alcançado foi para 60 neurônios, sendo encontrado MSE = 38,7797 e MAPE = 33,5667, valores superiores àqueles encontrados com as demais redes.

A dispersão dos resultados das 30 simulações realizadas foi semelhante àquela mostrada pelo Gráfico *Boxplot* apresentado na Fig. 4. Como pode ser observado, as ELMs apresentaram uma dispersão menor do que a ESN na média das simulações.

Além disso, o Teste de *Friedman* foi aplicado para avaliar se os resultados se apresentaram significativamente diferentes. Os valores de *p* encontrados foram próximos a zero, o que indica que alterar o preditor pode levar a resultados distintos (Luna & Ballini 2011).

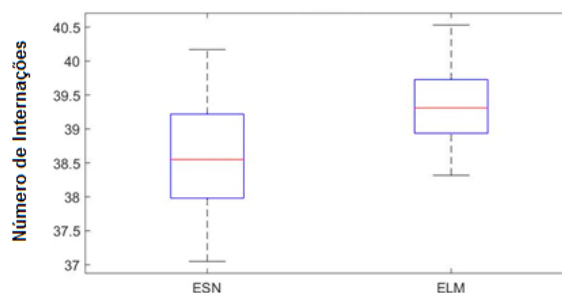


Fig. 4 – Gráfico *Boxplot* mostrando a dispersão dos resultados para o impacto 5 dias após a exposição

A análise dos resíduos para seis dias de atraso é apresentada no Quadro 3. Comparando o desvio padrão (Quadro 3) com os MSE de testes (Quadro 2), pode ser observado que o desvio padrão está uma ordem de grandeza abaixo do MSE de testes, assim como o Erro Absoluto Médio (*Mean Absolute Error* - MAE), enquanto a média do conjunto se aproxima de zero. Destes resultados, pode ser concluído que os dois modelos apresentaram uma capacidade de previsão apropriada.

Finalmente, a Fig. 5 apresenta o melhor resultado, que ocorreu para o impacto de seis dias após a exposição, o que permite uma melhor visualização dos resultados alcança-

Quadro 3 – Métricas dos resíduos

	Desvio Padrão	Média	MAE
ELM	4,9015	-0,0125	3,7890
ESN	4,7005	-0,1076	4,0805

dos neste estudo. A diferença ocorrida entre o número de internações observado e ajustado era esperada, pois existem outros fatores que podem levar à internação por doenças respiratórias.

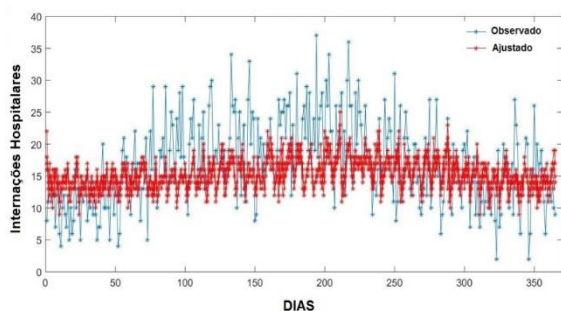


Fig. 5 – Riscos à saúde observados *versus* resultados do melhor ajuste, ocorridos 6 dias após a exposição

4. CONCLUSÃO

O presente estudo consistiu em um desafio na área de riscos à saúde populacional devido às emissões de poluentes atmosféricos e às alterações climáticas, pois a metodologia empregada, Redes Neurais Artificiais (RNAs), consiste em uma técnica inovadora no âmbito deste tópico. A metodologia frequentemente utilizada neste tipo de estudo consiste em modelos de regressão estatística. A vantagem do uso de Máquinas Desorganizadas (*Unorganized Machines* – UM) consiste na sua fácil implementação e operação, além do processo de treinamento ser rápido e eficiente. De acordo com os resultados obtidos, as Máquinas de Aprendizado Extremo (*Extreme Learning Machines* – ELM) apresentaram melhor desempenho geral do que as Redes Neurais com Estados de Eco (*Echo State Networks* – ESN).

Desta forma, conclui-se que o uso desta nova ferramenta poderá facilitar e melhorar as análises dos riscos da poluição atmosférica na saúde humana. A utilização de RNAs irá permitir às autoridades governamentais

prever impactos à saúde da população e estar preparado para eventos futuros.

Como futura perspectiva, é interessante um estudo semelhante considerando dados de outros grandes centros populacionais do Brasil e do mundo; além da comparação com metodologias de regressão estatística e outras arquiteturas de Redes Neurais.

REFERÊNCIAS

- Boccatto, L., Lopes, A., Attux, R., Von Zuben, F.J. 2012. An extended echo state network using volterra filtering and principal component analysis, *Neural Networks* 32, p. 292-302.
- Boccatto, L., Soares, E.S., Fernandes, M.M.L.P., Soriano D.C., Attux, R. 2011. Unorganized machines: from Turing's ideas to modern connectionist approaches, *International Journal of Natural Computing Research*, 2 (4), p. 1-16.
- CETESB (Companhia Ambiental do Estado de São Paulo). 2010. Qualidade do ar no estado de São Paulo, 2007, 2008 e 2009. Disponível em: <http://www.cetesb.sp.gov.br/ar/qualidade-do-ar/31-publicacoes-e-relatorios>. Acesso em 04 de fevereiro de 2010.
- DataSUS (Departamento de Informática do Sistema Único de Saúde). 2010. SIHSUS Reduzida. Disponível em: <http://www2.datasus.gov.br/DATASUS/index.php?area=0701&item=1&acao=11>. Acesso em 04 de fevereiro de 2010.
- Estado de São Paulo. 2011. Secretaria dos Transportes Metropolitanos. 2011. Disponível em: http://www.stm.sp.gov.br/index.php?option=com_content&view=article&id=2022&Itemid=200. Acesso em 31 de agosto de 2011.
- Haykin, S. 1998. *Neural networks: a comprehensive foundation*, 2nd Edition, Prentice-Hall, New York. ISBN 01-327-3350-1.
- Huang, G., Huang, G.B., Song, S., You, K. 2015. Trends in extreme learning machines: a review, *Neural Networks*, 61, p. 32-48.
- Huang, G.B., Zhu Q.Y., Siew C.-K. 2006. Extreme learning machine: theory and applications, *Neurocomputing*, 70 (1-3), p. 489-501.
- IBGE (Instituto Brasileiro de Geografia e Estatística). 2016. Disponível em: <http://cidades.ibge.gov.br>. Acesso em 16 de abril de 2016.
- Jaeger, H. 2001. The echo state approach to analyzing and training recurrent neural

- networks, Bremem: German National Research Center for Information Technology, Tech. Rep. GMD Report 148.
- Luna, I., Ballini, R. 2011. Top-down strategies based on adaptive fuzzy rule-based systems for daily time series forecasting, *International Journal of Forecasting*, 27 (3), p. 708-724.
- Nascimento, E.M., Pereira, B.B., Seixas, J.M. 2009. Artificial neural networks: an application in the study of air pollution and its adverse health effects, *Revista Brasileira de Biometria*, 27 (1), p.37-50.
- Schafer, A.M., Zimmermann, H.G. 2007. Recurrent neural networks are universal approximators, *International Journal of Neural Systems*, 17(4), p. 253-263.
- Siqueira, H.V., Boccato, L., Attux, R., Lyra, C. 2014. Unorganized machines for seasonal streamflow series forecasting, *International Journal of Neural Systems*, 24, p. 1430009.1-1430009.16.
- Tadano Y.S., Ugaya C.M.L., Franco, A.T. 2012. Methodology to assess air pollution impact on human health using the generalized linear model with Poisson regression, *Air Pollution – Monitoring, Modelling and Health*, Dr. Mukesh Khare (Ed.), ISBN: 978-953-51-0424-7, inTech, p. 281-304.
- Yildiz, B., Jaeger H., Kiebel, S.J. 2012. Revisiting the echo state property, *Neural Networks*, 35, p. 1-9.
- WHO (World Health Organization). WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: Global update 2005, summary of risk assessment. Disponível em: http://apps.who.int/iris/bitstream/10665/69477/1/WHO_SDE_PHE_OEH_06.02_eng.pdf, 2005, Acesso em 14 de abril de 201