Fuzzy logic and hospital admission due to respiratory diseases using estimated values by mathematical model

Lógica fuzzy e internações por doenças respiratórias usando dados estimados por modelo matemático

Luciana Cristina Pompeo Ferreira da Silva Vieira [1]

Paloma Maria da Silva Rocha Rizol [2]

Luiz Fernando Costa Nascimento [1]

[1] Departamento de Energia, Faculdade de Engenharia de Guaratinguetá (FEG), UNESP. Av. Ariberto Pereira da Cunha 333, Pedregulho. 12516-410 Guaratinguetá SP Brasil. luiz.nascimento@pq.cnpq.br

[2] Departamento de Engenharia Elétrica, FEG, UNESP. Guaratinguetá SP Brasil.

Abstract *Hospitalizations due to respiratory diseases generate financial costs for the Health System in addition to social costs. Objective of this study was to develop and validate a fuzzy linguistic model for prediction of hospitalization due to respiratory diseases. We constructed a fuzzy model for prediction of hospitalizations due to pneumonia, bronchitis, bronchiolitis and asthma second exposure to fine particulate matter ($PM_{2.5}$) in residents of Volta Redonda, RJ, in 2012. The model contains two inputs, $PM_{2.5}$ and temperature, with three membership functions for each input, and an output with three membership functions for admissions, which were obtained from DATASUS. There were 752 hospitalizations in the period, the average concentration of $PM_{2.5}$ was 17.1 µg/m3 (SD = 4.4). The model showed a good accuracy with $PM_{2.5}$, the result was between 90% and 76.5% for lags 1, 2 and 3, a sensitivity of up to 95%. This study provides support for creating executable software with a low investment, along with the use of a portable instrument could allow number of hospital admission due to respiratory diseases and provide support to local health managers. Furthermore, the fuzzy model is very simple and involves low computational costs, an implementation making possible.*

Key words *Air pollution. Particulate matter. Respiratory diseases. fuzzy logic. Mathematical modeling*

Resumo *Internações por doenças respiratórias geram custos financeiros para o Sistema de Saúde além de custos sociais. O objetivo deste estudo foi elaborar e validar um modelo linguístico "fuzzy" para previsão do número de internações por doenças respiratórias. Foi construído um modelo "fuzzy" para predição de internações por pneumonias, bronquite, bronquiolite e asma segundo exposição ao material particulado fino ($PM_{2.5}$) em residentes de Volta Redonda, RJ, em 2012. O modelo contém duas entradas $PM_{2.5}$ e temperatura, com três funções de pertinência para cada entrada, e uma saída com três funções de pertinência para internações, que foram obtidas do DATASUS. Foram 752 internações no período, a concentração média do $PM_{2.5}$ foi 17,1 µg/m3 (dp = 4,4). O modelo mostrou uma boa acurácia com $PM_{2.5}$, o resultado foi entre 90% e 76,5% para os lags 1, 2 e 3, com sensibilidade de até 95%. Este estudo fornece subsídios para a criação de programa executável, que não exige um grande investimento, juntamente com o uso de um instrumento portátil pode permitir uma estimativa do número de internações e prestar apoio aos gestores municipais de saúde. Além disso, o modelo "fuzzy" é muito simples e implica em baixas despesas computacionais, tornando possível uma implementação.*

Palavras-chave *Poluição do ar, Material particulado, Doenças do aparelho respiratório, Lógica fuzzy, Modelagem matemática*

Introduction

In the year 2012, there were about 1.3 million hospitalizations due to respiratory diseases in Brazil, with nearly 60,000 cases in the state of Rio de Janeiro; they cost the Brazilian Unified Health System (SUS) a sum of approximately US\$ 6 billion, and around US\$ 28 million (1 US\$ ≈ R\$ 2.00) in Rio de Janeiro[1].

Active and passive smoking, allergic diseases, overcrowding and exposure to air pollutants stand out among the factors associated with respiratory diseases[2-9].

Air pollution is characterized by the presence of atmospheric components that are harmful to living beings, whose origin can be either natural or anthropogenic, the latter being the pollution generated by the vehicular fleet, industries and power plants.

Ozone (O3), carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO_2) and particulate matter (PM) are the most commonly studied air pollutants which are associated with relative humidity and temperature and can contribute to the onset and worsening of respiratory diseases[10].

Studies carried out in medium-sized Brazilian cities also show the harmful effects of the exposure to such pollutants on human health, and that these do not occur exclusively in large urban centers(3, 6, 8). Association analyses between exposure to air pollutants and hospitalizations typically use statistical techniques, e.g. Poisson's regression or the Generalized Linear Model.

The majority of studies that estimated the effects of being exposed to air pollutants involve pollution concentrations that are quantified by state environmental agencies. Nevertheless, in many municipalities, there are no meteorological stations and only a few pollutants are monitored in some cases. The utilization of data estimated by mathematical models, such as the Coupled Chemistry Aerosol-Tracer Transport model to the Brazilian developments on the Regional Atmospheric Modeling System (CCATT - BRAMS) is a viable alternative[11,12].

CCATT - BRAMS is a mathematical model that performs numerical simulations of weather and climate, explicitly solves large spatial distribution phenomena and parameterizes processes that occur at smaller scales than the model's spatial resolution. It is run operationally by the Weather Prediction and Climate Studies Center of the National Institute Space Research - CPTEC/INPE[13] which produces daily patterns and forecasts throughout South America for up to three days . Its horizontal resolution is 25 km X 25 km, being the first level at 40 meters above the ground. These estimates are carried out every three hours or 8 times a day, which has already been validated[4,11].

By this approach, concentration is considered homogeneous throughout a city and there may be uncertainty which can be solved by using the fuzzy logic.

This logic, presented in 1964 by Lotfi A. Zadeh[14], is also known as the fuzzy logic which is an extension of classical logic that accepts intermediate values between 0 and 1 (true and false), while in classical logic, an element has only two degrees of membership (belongs or not belong) of a certain set; the fuzzy logic can check the same element's degree of membership of one or more sets, i.e. this element can have different degrees of membership. Thus, it is aimed at developing a predictive model with the Fuzzy Inference System (FIS) in order to estimate the number of hospitalizations from respiratory diseases caused by exposure to pollutant $PM_{2.5}$ which had been estimated by the mathematical model CCATT-BRAMS, comprising local residents of Volta Redonda, Brazil.

Methods

A fuzzy linguistic model was built with a fuzzyfier block, an inference block, a set of rules and a defuzzifier block, whose inference system is shown in Figure 1. The fuzzyfier block turns the actual input value (crisp) into its corresponding fuzzy value[15] and the inference block is the central block where the rules that form the database are processed and used by the processor to solve a particular problem, thus generating the fuzzy output values.

The set of rules is a set of propositional rules (IF → THEN), that were elaborated with assistance of an expert. The fuzzy inference system operates according to the set of rules, and the defuzzifier block performs the opposite of what the fuzzyfier block does, i.e. it turns the fuzzy value into a real value.

A uni-pollutant model was developed with $PM_{2.5}$ and temperature (TEMP) as input variables, and its output variable was the number of hospitalizations due to respiratory diseases ranging from zero to 11; the diagnosis were those corresponding to the codes of the ICD's 10th review (J12.0 to J18.9, J20.0 to J20.9, J21.0 to J21.9, J45.0 to J45.9) involving subjects of both sexes, regard-
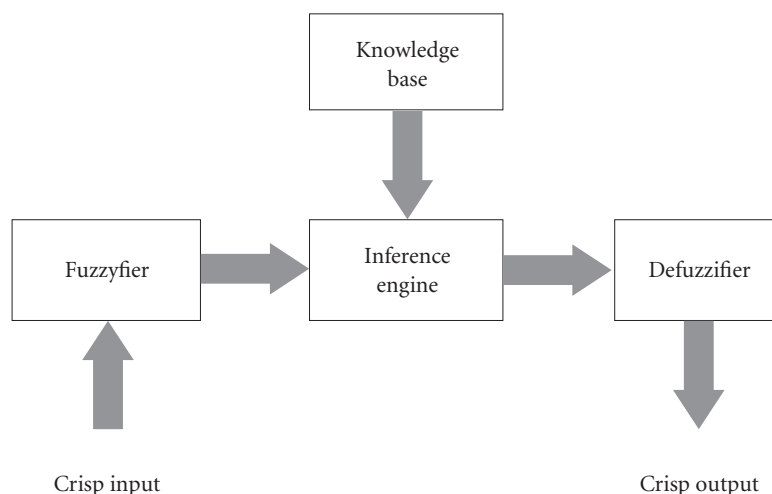
Figure 1. Fuzzy knowledge basic structure.

less of age, that were currently living in Volta Redonda, Brazil.

Place of study: Volta Redonda is a medium-sized city of Rio de Janeiro State which is an important steel industry center. It is located at coordinates 22º 29 'S and 44º 05' W, in the lower middle section of the Paraiba do Sul River valley, 390 m above sea level. Its climate is humid and tropical, it is located between the Mar and Mantiqueira Mountains and crossed by two major Brazilian highways: Via Dutra (BR-116), connecting the city of Rio de Janeiro to São Paulo, and Lúcio Meira Highway (BR-393), both with heavy traffic conditions. Its population was estimated at 260,000 inhabitants with a vehicular fleet of 100 thousand vehicles in 2010[16].

Statistical analysis

Daily data on hospitalization, pollutant concentrations and temperatures were collected from January 1, 2012 to December 31, 2012. These data were obtained from the Department of Information of the Unified Health System - DATASUS. $PM_{2.5}$ data about its daily averages were estimated by CCATT - BRAMS.

The model was developed by using the Mamdani inference method and implemented through the fuzzy toolbox of Matlab 7.1.

Two input variables were used: $PM_{2.5}$ (μg/m³) and temperature (TEMP) (° C), with the universe of discourse ranging from 10 to 36 μg/m³

for $PM_{2.5}$ and 10 to 32 ° C for temperatures. The output variable is the number of hospitalizations from respiratory diseases, whose variation was from zero to 11.

Nine rules were obtained (3² - two inputs with three membership functions each), elaborated by an expert, which define the outputs of the fuzzy inference system.

For example:

• **IF** $PM_{2.5}$ is LOW **AND** temperature is HIGH **THEN** hospitalization is LOW.

• **IF** $PM_{2.5}$ is HIGH **AND** temperature is LOW **THEN** hospitalization is HIGH.

In order to rank the input variables, an analysis of the database was performed to determine the range in which the levels of concentration of these variables were low, medium or high, according to an expert. Figure 2 shows the input and output of the model.

The real values of hospitalization were compared with the results estimated by the model through the ROC curve for lags 1, 2, and 3 and a cut-off point of up to two hospitalizations.

Results

During the time period considered herein, there were 752 hospitalizations of individuals due to respiratory diseases in Volta Redonda, RJ. The descriptive analysis of this study is in Table 1 which shows the mean values and standard de-
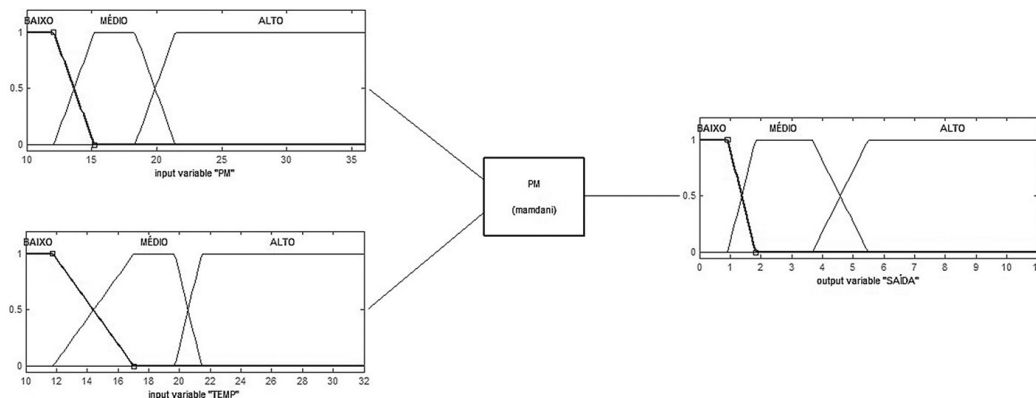
Figure 2. Fuzzy model structure for input variables PM$_{2,5}$ and temperature (TEMP) and output variable number of hospital admission (SAIDA), as low (BAIXO), medium (MÉDIO) and high (ALTO), Volta Redonda, Brazil, 2012.

Table 1. Mean values and standard deviation (sd), maximun and minimum values of fine particulate matter (PM$_{2,5}$), temperature (TEMP) and hospital admission due to respiratory diseases (INTER), Volta Redonda, Brazil, 2012.

	Mean (sd)	**Mínimum - maximum**
PM$_{2,5}$ (μg/m³)	17.1 (4.4)	11.7 – 35.2
TEMP (ºC)	21.3 (3.0)	12.6 – 28.3
INTER	3.6 (2.3)	0 - 11

viations, and minimum and maximum values for PM$_{2.5}$, temperature and hospitalizations. The time series of concentrations of PM$_{2.5}$, temperature and hospitalizations during 2012 in Volta Redonda - RJ are shown in Figures 3 (a), 3 (b) and 3 (c), respectively.

Figure 3 (a) shows that the air quality exceeded 25μg/m³ in 22 days (days 6%), i.e. a value considered acceptable by the WHO, but it was good along the following days. Temperatures ranged between 13°C and 31°C, with the maximum being reached in September and the minimum in July (Figure 3b).

There were a larger number of hospitalizations in July than in January and December which were the months with the smallest number of hospitalizations, thus suggesting seasonal distribution (Figure 3c).

The structure generated by the fuzzy model have used three membership functions for each

input: fine particulate matter (PM$_{2,5}$), temperature (TEMP) and three output membership functions (number of hospitalizations) (Figure 2).

After developing the fuzzy inference system, it was possible to obtain the response surface of the fuzzy model for the inputs (fine particulate matter (PM$_{2,5}$) and temperature (TEMP) and the output (number of hospitalizations). It can be observed that the higher the concentration of PM$_{2,5}$ is, the lower the temperature becomes, thus a larger number of hospitalizations, and the lower the concentration of PM$_{2,5}$ is, the higher the temperature becomes, thence a smaller number of hospitalizations. (Figure 4)

The values obtained for the ROC curve as the area under the standard curve in percentage terms were 90.0% (95% CI 85.8% - 94.2%), 80.8% (95% CI 73.8% - 87.9%) and 76.5% (95% CI 68.7% - 84.3%) for lags 1, 2 and 3, respectively.

The presented data show that the developed model is highly accurate at predicting the number of hospitalizations due to pollutant PM$_{2.5}$ using a cut-off point of up to two hospitalizations. It can be shown that the best accuracy for PM$_{2.5}$ was in lag 1, which corresponds to sensitivity of around 95% and specificity of 60% for a concentration of 14 ug/m³.

The model allows predicting, for example, 2.67 hospitalizations at a PM$_{2,5}$ concentration of 14 μg/m³ and temperature of 21 ° C and the actual value was three hospitalizations. The model estimated 4.85 hospitalizations, and the actual data on was five hospitalizations in the case of
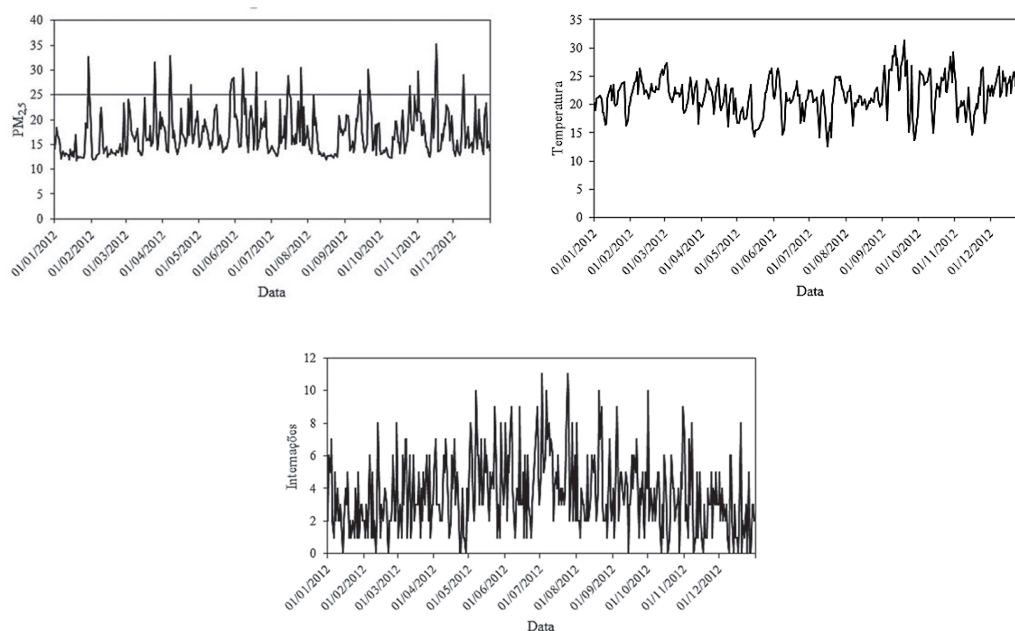
Figure 3. Daily distribution for concentrations of (a) $PM_{2,5}$ (µg/m³) , (b) temperature (⁰C) and (c) hospital admissions, Volta Redonda, Brazil, 2012.

31 µg/m³ $PM_{2,5}$ concentration and temperature at 21 °C.

Discussion

To the best of our knowledge, this is the first study in Brazil that uses data estimated by a mathematical model and fuzzy logic tools. The model showed a good degree of accuracy at predicting hospitalizations due to pneumonia, bronchitis, bronchiolitis and asthma in Volta Redonda, Brazil, by exposure to fine particulate matter, as shown by the values of ROC curves.

The importance of carrying out the present study in Volta Redonda is that it is a major steel industry center in Brazil. The average values found for $PM_{2,5}$ correspond to 28 mg/m³ of PM_{10}, and exceeded the limits established by the WHO in 22 opportunities. These effects are usually experienced later and this piece of information is important for city managers to predict possible hospitalizations in upcoming days by considering the concentrations of particulate matter in previous days.

During the study period, the fuzzy model had a good accuracy at predicting hospitalizations

due to respiratory diseases by exposure to $PM_{2,5}$ in one, two and three days afterwards.

Studies on the effects of exposure to air pollutants use regression models, such as Poisson regression (3, 8) or linear regression[17]. However, these regression models estimate the *risk* of hospitalizations occurring due to an increase in the concentrations of certain pollutants, and the fuzzy model estimates a *possible* number of hospitalizations.

In São José dos Campos, a southeastern Brazilian city, in a cumulative effect of eight days, there was a 9.8% increase in children hospitalizations with a 24.7 µg/m³ increase in PM_{10} concentration. (3) In São Paulo, with an increase of 10 µg/m³ in particulate matter concentrations, there was an increase in 4.6% of children hospitalizations from asthma, 4.3% of elderly hospitalizations from chronic obstructive pulmonary disease and a 1 5% of elderly hospitalizations from ischemic heart disease, thus showing that the pollutants concentration is associated with increased hospitalizations[18].

A mean value of 28.6 µg/m³for $PM_{2,5}$ was found in Piracicaba, higher than that found in Volta Redonda, which had possibly been gen-
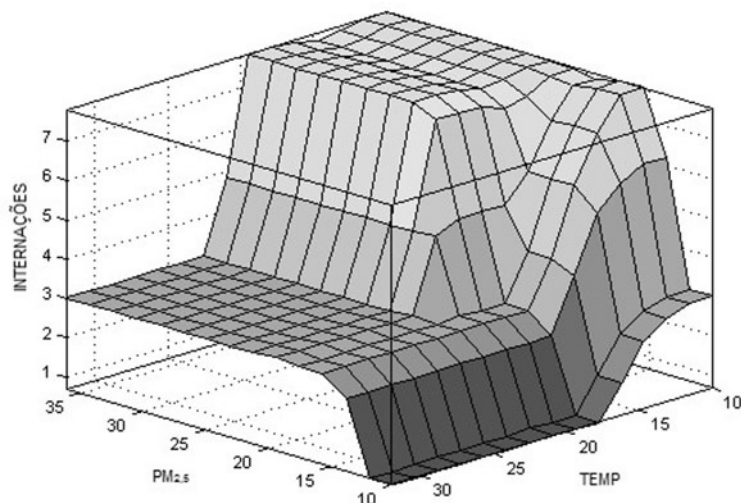
Figure 4. Fuzzy model surface for input PM2,5 and temperature (TEMP) and output hospital admission (INTERNAÇÕES), Volta Redonda, Brazil, 2012.

erated by the vehicle fleet and sugarcane straw burning because this municipality is one of the leading producers of sugar and ethanol in Brazil. An increase of 10 μg/m³ in $PM_{2,5}$ concentrations resulted in higher relative risk that ranged between 7.9 and 8.6 percent, thus concluding that the exposure to this pollutant is associated with children hospitalizations from respiratory diseases in one to three days afterwards, a fact that coincides with the findings of this study for lag 1, point at which the best accuracy was obtained[8].

The mean value of $PM_{2,5}$ would be of around 28.6 μg/m³ for the city of Volta Redonda considering that the $PM_{2,5}$ concentration correspond to about 60% of PM_{10}[19].

This estimated mean value is lower than that found in Sorocaba[20], with a concentration of 34.4 mg/m³ for PM_{10}, which is possibly due to a larger number of vehicles and industries as Sorocaba is crossed by two major highways, Castelo Branco and Raposo Tavares, in which there is heavy traffic.

Fisher[21] makes a review by proposing a fuzzy approach to study the effects of air pollution. The author sets examples of its application under uncertainty of air pollution data and illustrates the flexibility of this approach by taking examples of air quality management, prevention and integrated pollution control, and human health evaluation.

A similar approach was adopted in a fuzzy study on the average time of hospitalization from pneumonia, in which the output of exposure effects was the time of hospitalization in days, showing that, as pollutants concentrations increase, the average time of hospitalization is longer[22]. The fuzzy logic was used to make the diagnosis of pneumonia according to clinical signs through fuzzy relations and by estimating the time of hospitalization from cardiovascular disease[23,24].

In the fuzzy linguistic model presented in this study, the output is the possible number of hospitalizations according to the concentration of $PM_{2,5}$ adjusted according to temperature from one to three days after exposure.

The present model performance is not able to be easily compared to other studies, given that there are no Brazilian studies on the effects of air pollutants exposure and hospital respiratory diseases by using the fuzzy logic.

The advantages of this model are that it can be applied in any location where data is available on the same air pollutants and the ease of obtaining predicted results. The use of data estimated by this system, validated by Ignotti et al.[4], which has also been used in a study conducted in São José do Rio Preto, Brazil, on cardiovascular diseases[25] allows reducing research costs and streamlining the process of analyzing the effects of air pollution on human health.

This study has a few limitations. One of them is that the pollutants concentrations are estimated by data obtained from mathematical modeling, and not from actual data quantified by monitors. Another possible limitation is the fact that hospitalizations are restricted to users of the Unified Health System, not incorporating health plan and private care users. Moreover, those treated in an outpatient regime are not included, and many patients are also treated in health centers, that are usually treated and then released, thus not being included in the hospitalization statistics. Datasus, hospitalization data does not include co-morbidities and other conditions associated with respiratory diseases, such as smoking. Another limitation can be associated with diagnostic errors of respiratory diseases.

This study provides support for designing an executable computer program that does not require large investment, along with the use of a portable instrument that is able to estimate hospitalization time and provide support to local health managers. Furthermore, the fuzzy model is very simple and involves low computational costs, which makes its implementation possible.

Collaborations

LCPFS Vieira, PMSR Rizol and LFC Nascimento participated in all phases of the study.

Acknowledgements

References

1. Brasil. Ministério da Saúde (MS). Departamento de Informática do SUS - DATASUS. *Portal da Saúde*. [cited 2014 May 18]. Available from: www2.datasus.gov.br/

2. Cançado JED, Saldiva PHN, Pereira LAA, Lara LBLS, Artaxo P, Martinelli LA, Arbex MA, Zanobetti A, Braga AL. The Impact of Sugar Cane-Burning Emissions on the Respiratory System of Children and the Elderly. *Environ Health Perspect* 2006; 114(5):735-739.

3. Nascimento LFC, Pereira LAA, Braga ALF, Módolo MCC, Carvalho Júnior JA. Efeitos da poluição atmosférica na saúde infantil em São José dos Campos, SP. *Rev Saude Publica* 2006; 40(1):77-82.

4. Ignotti E, Valente JG, Longo KM, Freitas SR, Hacon SS, Netto PA. Impact on human health of particulate matter emitted from burning in the Brazilian Amazon region. *Rev Saude Publica* 2010; 44(1):121-130.

5. Oliveira MS, Leon AP, Mattos IE, Koifman S. Differential susceptibility according to gender in the association between air pollution and mortality from respiratory diseases. *Cad Saude Publica* 2011; 27(9):1827-1836.

6. Arbex MA, Santos UP, Martins LC, Saldiva PHN, Pereira LAA, Braga ALF. Air pollution and the respiratory system. J Bras. Pneumol. 2012; 38(5):643-655.

7. Gouveia N, Habermann M. Tráfego veicular e mortalidade por doenças do aparelho circulatório em homens adultos. *Rev Saude Publica* 2012; 46(1):26-33.

8. Cesar ACG, Nascimento LFC, Carvalho JJA. Associação entre exposição ao material particulado e internações por doenças respiratórias em crianças. *Rev Saude Publica* 2013; 6(47):1209-1212.

9. Nascimento LFC, Francisco JB. Particulate matter and hospital admission due to arterial hypertension in a medium-sized Brazilian city. *Cad Saude Publica* 2013; 29(8):1565-1571.

10. São Paulo. Companhia Ambiental do Estado de São Paulo (Cetesb). Qualidade do ar. [cited 2014 May 05]. Available from: www.cetesb.sp.gov.br/ar/qualidade-do-ar.

11. Freitas SR, Longo KM, Silva Dias MAF, Chatfield R, Silva Dias P, Artaxo P, Andreae MO, Grell G, Rodrigues LF, Fazenda A, Panetta J. The Coupled Aerosol and Tracer Transport model to the Brazilian developments on the Regional Atmospheric Modeling System (CATT-BRAMS). Part 1: Model description and evaluation. *Atmospheric Chemistry and Physics Discussions* 2009; 9(7):2843-61.

12. Longo KM, Freitas SR, Pirre M, Marécal V, Rodrigues LF, Panetta J, Alonso MF, Rosário NE, Moreira DS, Gácita MS, Arteta J, Fonseca R, Stockler R, Katsurayama DM, Fazenda A, Bela M. The Chemistry CATT-BRAMS model (CCATT-BRAMS 4.5):a regional atmospheric model system for integrated air quality and weather forecasting and research. *Geoscientific Model Development* 2013; 13(6):1389-405.

13. Instituto Nacional de Pesquisas Espaciais (CPTEC/INPE). [cited 2015 July 03]. Available from: http://meioambiente.cptec.inpe.br/index.php?lang=pt.

14. Zadeh LA. "Fuzzy Sets". *Information and Control* 1965; 8(3):338-53.

15. Rocha PMS. *Projeto de um defuzificador analógico integrado em tecnologia CMOS* [dissertação]. São José dos Campos: Instituto Tecnológico de Aeronáutica; 2005.

16. Instituto Brasileiro de Geografia e Estatística (IBGE). [cited 2014 June 11]. Available from: http://www.ibge.gov.br/home/.

17. Andrade Filho VS, Artaxo P, Hacon S, Carmo CN, Cirino G. Aerossóis de queimadas e doenças respiratórias em crianças, Manaus, Brasil. *Rev Saude Publica* 2013; 47(2):239-247.

18. Gouveia N, Freitas CU, Martins LC, Marcilio IO. Hospitalizações por causas respiratórias e cardiovasculares associados à contaminação atmosférica no Município de São Paulo, Brasil. *Cad Saude Publica* 2006; 22(12):2669-2677.

19. São Paulo. Companhia Ambiental do Estado de São Paulo (Cetesb). *Relatório de qualidade do ar 2014*. [cited 2014 Nov 14]. Available from: http://ar.cetesb.sp.gov.br/publicacoes-relatorios/.

20. Negrisoli J, Nascimento LFC. Poluentes atmosféricos e internações por pneumonia em crianças. *Rev Paulista Pediatria* 2013; 31(4):501-506.

21. Fisher B. Fuzzy environmental decision-making: applications to air pollution. *Atmospheric Environment* 2003; 37(14):1865-1877.

22. Nascimento LFC, Rizol PMSR, Peneluppi AP. Estimating the average length of hospitalization due to pneumonia: a fuzzy approach. *Braz J Med Biol Res* 2014; 47(11):977-981.

23. Pereira JCR, Ortega NRS, Tonelli PA, Barros LC. Clinical signs of pneumonia in children: association with and prediction of diagnosis by fuzzy sets theory. *Braz J Med Biol Res* 2004; 37(5):701-709.

24. Coutinho KMV, Rizol PMSR, Nascimento LFC, Medeiros APP. Fuzzy model approach for estimating time of hospitalization due to cardiovascular diseases. *Cien Saude Colet* 2015; 20(8):2585-2590.

25. Mantovani KCC, Nascimento LFC, Moreira DS, Vieira LCPFS, Vargas NP. Air pollutants and hospital admissions due to cardiovascular diseases in São José do Rio Preto, Brazil. *Cien Saude Colet* 2016; 21(2):509-516.