



UNIVERSIDADE CATÓLICA DE PERNAMBUCO
PRÓ-REITORIA ACADÊMICA
COORDENAÇÃO GERAL DE PÓS-GRADUAÇÃO
MESTRADO EM DESENVOLVIMENTO DE PROCESSOS AMBIENTAIS

Francisco Sales de Albuquerque Filho

**Sistemas inteligentes aplicados à análise de riscos
ambientais**

Recife
2012

Francisco Sales de Albuquerque Filho

**Sistemas inteligentes aplicados à análise de riscos
ambientais**

Dissertação apresentada ao Programa de Pós-Graduação em Desenvolvimento em Processos Ambientais Universidade Católica de Pernambuco como pré-requisito para obtenção do título de Mestre em **Desenvolvimento de Processos Ambientais**.

Área de Concentração: Desenvolvimento em Processos Ambientais

Linha de Pesquisa: Modelagem, Simulação

Orientador: Sérgio Murilo Maciel Fernandes

Co-orientador: Francisco Madeiro Bernardino Junior

Recife

2012

de Albuquerque Filho, F. S.

Sistemas inteligentes aplicados à análise de riscos ambientais, 2012, 85.

Dissertação (Mestrado) - Universidade Católica de Pernambuco. Pró-reitoria Acadêmica. Curso de Mestrado em Desenvolvimento de Processos Ambientais, 2012.

1. Riscos Ambientais. 2. Poluição do ar. 3. Inteligência Artificial. 4. Redes Neurais Artificiais. 5. Otimização por Enxame de Partículas. 6. Lógica Nebulosa.

Programa de Pós-Graduação em Desenvolvimento de Processos Ambientais. Centro de Ciências e Tecnologia.

Sistemas inteligentes aplicados à análise de riscos ambientais

FRANCISCO SALES DE ALBUQUERQUE FILHO

Examinadores:

Prof. Dr. Sérgio Murilo Maciel Fernandes (Orientador)
Universidade Católica de Pernambuco – UNICAP

Prof. Dr. Emerson Alexandre de Oliveira Lima (Membro Interno)
Universidade Católica de Pernambuco – UNICAP

Prof. Dr. Tiago Alessandro Espínola Ferreira (Membro Externo)
Universidade Federal Rural de Pernambuco – UFRPE

AGRADECIMENTOS

À minha família e amigos.

SUMÁRIO

AGRADECIMENTOS	iii
LISTA DE FIGURAS	vi
LISTA DE TABELAS	viii
RESUMO	ix
ABSTRACT	x
CAPÍTULO I	
1.1 INTRODUÇÃO	14
1.2 OBJETIVOS.....	16
1.2.1 Objetivo Geral.....	16
1.2.2 Objetivos Específicos	16
1.3 Revisão da literatura	17
1.3.1 Poluição.....	17
1.3.1.1 Poluição Atmosférica.....	18
1.3.1.2 Indicadores da poluição atmosférica	18
1.3.1.2.1 Material particulado (PM)	20
1.3.1.2.2 Ozônio (O ₃)	20
1.3.1.2.3 Monóxido de carbono (CO)	20
1.3.1.2.4 Dióxido de enxofre (SO ₂)	21
1.3.1.2.5 Dióxido de nitrogênio (NO ₂)	21
1.3.1.3 Efeitos da poluição	21
1.3.1.4 Efeitos adversos na saúde humana	22
1.3.1.5 Ações mitigadoras.....	22
1.3.1.6 Modelos de previsão de poluentes gasosos.....	23
1.3.2 Séries temporais.....	25
1.3.2.1 Modelos estatísticos.....	25
1.3.2.2 Exemplos de métodos para identificar os parâmetros p e q	27
1.3.2.3 Atrasos temporais e o teorema de Takens	27
1.3.2.3.1 Atrasos temporais	27
1.3.2.3.2 Teorema de Takens.....	27
1.3.2.4 Modelos de previsão de séries temporais	28
1.3.2.5 A previsão como parte do processo de análise de riscos	28
1.3.3 Riscos.....	29
1.3.3.1 Métodos de análise de riscos.....	30
1.3.3.1.1 Diagramas de influência.....	30
1.3.3.1.2 Matriz de risco.....	31

1.3.4	Modelos baseados em inteligência artificial	35
1.3.4.1	Redes neurais artificiais	35
1.3.4.2	Otimização por enxame de partículas	39
1.3.4.3	Lógica nebulosa	41
1.3.4.3.1	Funções de pertinência	42
1.3.4.3.2	Variáveis linguísticas.....	44
1.3.4.3.3	Operadores lógicos.....	45
1.3.4.3.4	Sentenças condicionais – Regras (<i>Se-Então</i>)	45
1.3.4.3.5	Sistema de inferência nebulosa	46
1.3.4.4	Modelos híbridos.....	46
1.3.5	Simulações e resultados.....	47
1.3.5.1	Perceptron de multicamadas (MLP)	47
1.3.5.2	Otimização por enxame de partículas (PSO)	47
1.3.5.3	Métricas de desempenho	48
1.3.5.3.1	Métricas das previsões	50
1.3.5.3.2	Gráficos dos valores observados versus os valores previstos.....	53
1.3.5.4	Lógica nebulosa	55
1.4	REFERÊNCIAS BIBLIOGRÁFICAS	66
CAPÍTULO II		
TIME SERIES FORECASTING OF POLLUTANT CONCENTRATION LEVELS USING PARTICLE SWARM OPTIMIZATION AND ARTIFICIAL NEURAL NETWORKS		72
CAPÍTULO III		
CONCLUSÕES GERAIS.....		86
TRABALHOS FUTUROS		87

LISTA DE FIGURAS

CAPÍTULO I

Figura 1.1- Exemplos de metodologias de análise de riscos. Diagrama de influência (a) e Matriz de riscos (b).	31
Figura 1.2 – Fontes mais comuns de incertezas (Haimes, 2009).	33
Figura 1.3 – Exemplo de taxonomia de arquiteturas de redes neurais (Gardner, 1998).	36
Figura 1.4 – Exemplo de um neurônio simples (Haykin, 2001).	37
Figura 1.5 - Modelo de uma rede neural multicamadas.	38
Figura 1.6 - Funções de pertinência clássica (a) e nebulosa (b).	43
Figura 1.7 - Exemplos de funções de pertinência: triangular (a), trapezoidal (b), gaussiana (c) e, sigmóide (d).	44
Figura 1.8 - Exemplo distribuição da variável linguística (Risco) e seus modificadores (Baixo, Médio e, Alto).	44
Figura 1.9 – Comparação entre os valores observados e as previsões dos modelos PSO-MLP-LM e ARMA.	53
Figura 1.10 - Comparação entre os valores observados da série de NO ₂ e as previsões dos modelos PSO-MLP-LM e ARMA.	54
Figura 1.11 - Comparação entre os valores observados da série de SO ₂ e as previsões dos modelos PSO-MLP-LM e ARMA.	54
Figura 1.12 - Comparação entre os valores observados da série de CO e as previsões dos modelos PSO-MLP-LM e ARMA.	55
Figura 1.13 – Funções de pertinência das variáveis linguísticas que compõem o sistema nebuloso de qualidade do ar.	57
Figura 1.14 – Matriz de riscos.	61
Figura 1.15 – Superfícies de resposta em relação a Qualidade do Ar das seis possíveis combinações: CO x MP ₁₀ (a), NO ₂ x MP ₁₀ (b), SO ₂ x CO (c), SO ₂ x NO ₂ (d), SO ₂ x MP ₁₀ (e), NO ₂ x CO (f).	63
Figura 1.16 – Superfícies de resposta acerca das possíveis Consequências. Pessoas sensíveis x Densidade demográfica (a), Pessoas sensíveis x Sensibilidade do local (b), Densidade demográfica x Sensibilidade do local (c).	64
Figura 1.17 – Superfícies de resposta da variável Risco.	65

CAPÍTULO II

Figure 1 – A neuron example.....	74
Figure 2 – Artificial neural network model in this article.....	75
Figure 3 – Results of the last 100 values of the prediction made by the ARMA model and the PSO-MLP-LM model for the CO series (test set): normalized observed values, predicted MLP and ARMA values.....	80
Figure 4– Results of the last 100 values of the prediction made by the ARMA model and the PSO-MLP-LM model for the SO ₂ series (test set): normalized observed values, predicted MLP and ARMA values.....	81
Figure 5– Results of the last 100 values of the prediction made by the AR model and the PSO-MLP-LM model for the MP ₁₀ series (test set): normalized observed values, predicted MLP and AR values.	82
Figure 6 – Results of the last 90 values of the prediction made by the ARMA model and the PSO-MLP-LM model for the NO ₂ series (test set): normalized observed values, predicted MLP and ARMA values.....	83

LISTA DE TABELAS

CAPÍTULO I

Tabela 1.1 - Índice da qualidade do ar.....	19
Tabela 1.2 – Métricas das previsões da série de MP ₁₀	51
Tabela 1.3 – Métricas das previsões da série de NO ₂	51
Tabela 1.4 – Métricas das previsões da série de SO ₂	52
Tabela 1.5 – Métricas das previsões da série de CO.....	52
Tabela 1.6 – Divisão dos níveis de gases – µg/m ³	56
Tabela 1.7 – Parâmetros de consequências.....	59

CAPÍTULO II

Table 1 – Results for the CO series.....	79
Table 2 – Results for the SO ₂ series.....	80
Table 3– Results for the MP ₁₀ series.....	81
Table 4– Results for the NO ₂ series.....	82

RESUMO

Este trabalho aplica técnicas da área de inteligência artificial (IA) com o intuito de prever e classificar riscos ambientais, com o foco no problema da qualidade do ar. Para prever os riscos, dados foram adquiridos acerca das concentrações gasosas de determinados poluentes. Tais concentrações, denominadas de indicadores da qualidade do ar, são regulamentadas por várias legislações ao redor do mundo, inclusive a do Brasil. Estes dados foram empregados em um modelo que consiste de duas técnicas de IA: redes neurais artificiais e otimização por enxame de partículas. O resultado do modelo é a previsão de um dia adiante das concentrações gasosas dos indicadores da qualidade do ar. As previsões são usadas como entradas para a modelagem de riscos. A modelagem de riscos correlaciona as previsões dos poluentes observados para obter a qualidade do ar e o risco que tal qualidade oferece à saúde humana. O modelo de risco é baseado em lógica nebulosa, uma terceira técnica de IA. Ao término do trabalho, dois resultados foram alcançados. O primeiro foi o modelo de previsões que obteve resultados com um bom nível de acuidade. Em seguida, o modelo de riscos foi capaz de alcançar uma classificação coerente dos riscos ambientais.

Palavras-Chave: Riscos Ambientais, Poluição Atmosférica, Inteligência Artificial, Redes Neurais Artificiais, Otimização por Enxame de Partículas, Lógica Nebulosa.

ABSTRACT

In order to forecast and classify environmental risks, artificial intelligence (AI) techniques were applied to the air quality problem. Predetermined gaseous pollutant concentration data were acquired with the intent of predicting the risks. Such concentrations are denominated air quality indicators, and are regulated all around the world, including by Brazilian law. The data concerning these indicators were used in a model that consists of two AI techniques: artificial neural networks and particle swarm optimization. The air quality indicators concentration prediction resulted in one day ahead values. The risk modeling utilizes the predictions as input values, correlating them in order to obtain the resulting air quality and, the risk that such quality has upon the human health. The risk model is based on a third AI technique, called fuzzy logic. The present work obtained two main results. The first was the accurate forecasts made by the prediction model. The second was the achievement of a coherent classification of the risks.

Keywords: Environmental Risks, Atmospheric Pollution, Artificial Intelligence, Artificial Neural Networks, Particle Swarm Optimization, Fuzzy Logic.

CAPÍTULO I

1.1 INTRODUÇÃO

Poluição é um problema que vem recebendo ampla atenção da sociedade e de inúmeros campos da ciência. Esta crescente atenção vem desde os anos 60, quando a sociedade presenciava os efeitos nocivos da poluição (Cassell, 1969; Chittock, 2011). Nos dias atuais, a sociedade não mais aceita o uso não sustentável do meio ambiente sujeito a toda sorte de poluição, e exige mudanças. A mudança, sob o ponto de vista da sociedade, é obtida através de leis e pesquisas que ajudam na formulação de uma política mais consistente ao tratar do meio ambiente. Por isto, meios de prevenção e remediação da poluição são necessários para melhoria da qualidade de vida.

A poluição do ar, neste contexto, é de extrema importância por sua distribuição homogênea, pois atinge a todas as pessoas que estão sob seu efeito. Justamente por não fazer distinção entre credo, raça (cor) ou, nível de renda, atingindo a todos com a mesma intensidade. A poluição do ar é uma das fontes que mais degrada o meio ambiente e a saúde humana (Mukherjee, 2012). Sendo, por isto, do interesse da comunidade científica. Tal interesse pode ser constatado pelos numerosos trabalhos sobre o assunto (Şahin, 2011; Chan & Huang, 2003; Fanb & Lu, 2003; Jahn, 2011).

Indicadores da qualidade do ar são um meio de monitoramento do nível de poluentes na atmosfera. Os indicadores da qualidade do ar, representados pelos gases dióxido de enxofre (SO₂), monóxido de carbono (CO), material particulado (MP₁₀), dióxido de nitrogênio (NO₂) e ozônio (O₃), são empregados na formulação de políticas públicas e na aferição do grau de poluição do ar (Giannouli, 2010; Godish, 2003). Agências de meio ambiente os utilizam no monitoramento e no processo de tomada de decisões. As decisões variam de simples avisos à população, até a remoção das pessoas para outros locais. Estas ações são tomadas visando mitigar os efeitos da poluição atmosférica.

Os efeitos da poluição do ar na saúde das pessoas vêm sendo questionados em diversas pesquisas. Estudos (Curtis, 2006; Wilson A. M., 2004) comprovam que a poluição do ar possui efeitos adversos diretos à saúde humana. Os efeitos da poluição, desde os mais amenos, como asma, rinite, irritação nos olhos, até alguns tipos de câncer, além dos aumentos nas taxas de mortalidade e morbidade (Wilson & Kingham, 2005; Cairncross, 2007; Kassomenos & Papaloukas, 2008; Monks, 2009; Kampa, 2008).

Os possíveis efeitos da poluição, obtidos na análise de riscos, podem ser mitigados, por meio do gerenciamento de riscos. Para alcançar esta mitigação é importante a identificação dos riscos e sua classificação. Desta forma, ações podem ser direcionadas a riscos de maior importância. Os riscos são os resultados da exposição das pessoas à poluição e das

consequências que esta acarreta. Contudo, sendo a natureza a fonte dos dados, as formas de obtenção e processamento destes dados podem ser repletas de imprecisões (Darbra, 2008). Portanto, incertezas são inerentes aos meios de análise de riscos.

Um exemplo de técnica que lida com incertezas e imprecisões são os modelos baseados em lógica nebulosa. A lógica nebulosa é uma técnica da área de inteligência artificial (IA) fundamentada na forma imprecisa e complexa do pensamento humano.

Outro meio de prevenção é o prognóstico das concentrações gasosas, pois, ao saber com antecedência o nível de dado poluente, ações com maior grau de coordenação e precisão podem ser realizadas. Para alcançar tal predição, geralmente, são utilizados modelos estatísticos. Como fontes (entradas) destes modelos são empregadas séries temporais. As séries temporais são sequências de observações segundo algum índice cronológico sobre uma variável de interesse (ou fenômeno), que, neste caso, são as concentrações de poluentes.

Os modelos matemáticos e estatísticos usualmente aplicados à previsão de séries temporais são baseados em processos lineares (Box & Jenkins, 1994). Contudo, os processos no mundo real, geralmente, são complexos podendo, ainda, conter partes lineares e (ou) não-lineares (Zhang, 2003). Modelos que são mais adequados a processos complexos e não-lineares são amplamente utilizados, tais como as redes neurais artificiais, que são uma abstração do processo de raciocínio humano, que é paralelo, complexo e não-linear.

Uma dificuldade inerente ao uso de uma rede neural artificial (RNA) (Haykin, 2001) é a escolha do modelo e de sua estrutura, as quais serão empregadas para modelar o processo. Uma forma de suprir tal limitação, um algoritmo evolucionário, denominado de otimização por exame de partículas (PSO) (Eberhart, 1995) é incorporado ao processo. O PSO é capaz de determinar uma estrutura da RNA adequada ao problema. A combinação destes dois métodos dá origem ao que é intitulado de sistemas híbridos (Chen, 2008; Khashei & Bijari, 2012), que são a junção de duas ou mais técnicas em um mesmo sistema.

A dissertação está organizada da seguinte forma: na primeira seção é abordado um breve histórico sobre poluição, suas causas e efeitos. Na seção seguinte é apresentado o conceito de risco e meios de sua mitigação. Posteriormente, são apreciados os fundamentos de séries temporais e métodos de previsão. Nas últimas seções são apresentados conceitos de inteligência artificial e são discutidas três técnicas: redes neurais artificiais, otimização por exame de partículas e lógica nebulosa. E finalmente é apresentada a conclusão deste trabalho.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

A proposta deste trabalho consiste no desenvolvimento de um sistema de previsão de riscos ambientais baseado em sistemas inteligentes.

1.2.2 Objetivos Específicos

Para consolidação do objetivo geral, faz-se necessária a execução de etapas pertinentes ao trabalho proposto. Portanto, os seguintes objetivos específicos deverão ser alcançados:

- ✓ Estudar, aplicar e avaliar as técnicas de sistemas inteligentes, em particular, redes neurais artificiais, lógica nebulosa, e otimização por enxame de partículas, na determinação de riscos ambientais.

- ✓ Elaborar modelo(s) de previsão de riscos ambientais, tendo como foco a qualidade do ar, que possa(m) vir a ser utilizada(s) como ferramenta(s) de gestão ambiental.

1.3 Revisão da literatura

1.3.1 Poluição

Nos dias atuais a degradação do meio ambiente tem atingido níveis cada vez mais preocupantes. Os desastres naturais que causam inúmeras mortes e destruição são de difícil controle, pois são fenômenos incertos. Contudo, eles se tornam ainda mais devastadores pela ação indevida do homem sobre a natureza. Nesta particular ação devem ser ressaltados os efeitos nocivos da poluição sobre o meio ambiente e sobre as pessoas.

O avanço da tecnologia tem possibilitado inúmeros confortos à sociedade moderna. Entretanto, ao avaliar-se a qualidade de vida, pode-se chegar à conclusão de que a mesma possui outras características relevantes como, por exemplo, a saúde e os bem-estares humanos e ambientais, que se sobrepõem a suas aplicações como bens de consumo. (Hinrichs, 2010). Faz-se necessária uma harmonia entre a inovação tecnológica e a qualidade de vida, ou seja, um desenvolvimento sustentável, para que eventos indesejados não venham a ocorrer.

Ainda nos anos 60 foram constatadas várias consequências da poluição sobre a saúde humana. Tais adventos negativos foram provenientes de um começo de século com acelerada industrialização, caracterizado pelo uso indiscriminado de fontes de energia altamente tóxicas. Cidades inteiras, como Londres em 1952, ficaram sob uma nuvem formada por gases, cinzas e poeira (Cassell, 1969). Tal episódio londrino que durou cerca de duas semanas, ficou conhecido pelas quase 4.000 (quatro mil) mortes e inúmeros casos de asma, rinite e redução da capacidade cardiovascular da população (Monks, 2009).

Poluição antropogênica (causada pelo homem) pode ser definida como a introdução pelo homem, diretamente ou indiretamente, de substâncias ou energia no meio ambiente que venham a resultar em efeitos insalubres à saúde humana (UNEP, 1995). Tal definição possibilita a associação entre poluição e emissões ou presença de matéria ou energia, permitindo a medição de tal poluição. A medição das concentrações de poluentes torna possível a elaboração de pesquisas e estudos que venham a definir a capacidade de assimilação do meio. Posteriormente, os resultados destes estudos podem ser utilizados na elaboração de limites de concentração para estes poluentes, conhecidos como padrões ambientais (Sánchez, 2008).

Padrões ambientais delimitam quais são os níveis permitidos para os poluentes monitorados. Estes padrões são regulados por agências governamentais, tais como a Agência Americana de Proteção Ambiental (US-EPA), a organização mundial de saúde (WHO) e, a

Companhia Ambiental do Estado de São Paulo (CETESB-SP). O monitoramento destes padrões é feito através de indicadores de poluição, que são utilizados para verificar a presença de um determinado poluente ou condição ambiental. Como exemplos têm-se os níveis de pH da água, concentrações gasosas de monóxido de carbono e, níveis de metais no solo.

O monitoramento destes indicadores pode ser feito através de sensores biológicos (bioindicadores), como o líquen (vegetal que detecta a qualidade do ar ao seu redor) (Llop, 2012), ou sensores artificiais, como sensores de níveis de concentrações gasosas que indicam a concentração, por exemplo, do poluente dióxido de enxofre (SO₂) presente no ambiente. Os sensores provêm medidas que são utilizadas no monitoramento, possibilitando medições imediatas que posteriormente servirão de base para o cálculo de médias horárias e diárias de determinados poluentes (Osowski, 2007). A partir destas medições são realizadas comparações com os limites estipulados, e caso a(s) concentração(ões) ultrapassem os limites permitidos, avisos são emitidos e medidas de mitigação são colocadas em efeito. Tais medições são armazenadas, geralmente, de forma cronológica, nas agências realizadoras das ações de monitoramento.

1.3.1.1 Poluição Atmosférica

Dentre os diversos tipos de poluição, a poluição atmosférica é um assunto de crescente importância para a sociedade, devido a seus efeitos adversos à saúde humana. Inúmeros estudos realizados comprovam a associação entre poluentes e a redução de funções vitais do corpo humano. Problemas de redução aguda de funções pulmonares, agravamento de asma, rinite, recém-nascidos com pesos baixos, mortalidade e morbidade (Vlanchogianni, 2011; Wilson, 2004). Estes problemas estão relacionados a uma baixa qualidade do ar. A qualidade do ar pode ser determinada após a avaliação das concentrações gasosas dos poluentes atmosféricos, apresentadas na Tabela 1.1, tais como: NO₂, MP₁₀, SO₂, CO e O₃.

1.3.1.2 Indicadores da poluição atmosférica

Uma variedade de poluentes pode ser encontrada na atmosfera. Contudo, somente uma pequena parcela desta é utilizada como parâmetro do índice de qualidade do ar. Tal índice descreve a condição do ar em determinados locais e tempo, refletindo o grau de pureza do ar em relação aos níveis de concentração de determinados poluentes (Park, 2008). Estes poluentes, denominados de indicadores da qualidade do ar, são representados pelos seguintes gases: dióxido de enxofre (SO₂), material particulado (PM sigla inglesa para *Particulate Matter*),

óxidos de nitrogênio (NO_x), monóxido de carbono (CO) e ozônio (O_3). Os poluentes são divididos em duas categorias: os poluentes primários que são aqueles que são liberados na atmosfera a partir de fontes de processos industriais ou combustões, e os poluentes secundários, produzidos na atmosfera, e responsáveis por grande parte da neblina fotoquímica e irritação nos olhos, entre outras várias formas de danos a plantas e materiais associados à poluição do ar (Vallero, 2008).

Os poluentes são mensurados e suas concentrações são utilizadas como substitutas à exposição de indivíduos reais, uma vez que a exposição individual é extremamente difícil de ser medida e possui um custo oneroso. Por esta razão, estudos que focam na população e não nos indivíduos são mais frequentemente utilizados. Tais estudos são amplos e identificam mudanças na população, sendo chamados de estudos ecológicos (Wilson, 2004).

Tabela 1.1 - Índice da qualidade do ar.

Índice da qualidade do ar						
Categoria	Bom	Moderado	Insalubre para pessoas sensíveis	Insalubre	Muito insalubre	Perigoso
Valores indexados	[0-50]	(50-100]	(100-150]	(150-200]	(200-300]	(300-500]
Poluentes						
CO (ppm)	[0-4,5]	(4,5-9,5]	(9,5-12,5]	(12,5-15,5]	(15,5-30,5]	(30,5-50,4]
NO_2 ($\mu\text{g}/\text{m}^3$)	[0-10]	(10,5-20]	(20-30]	(30-40]	(40-60]	(61-100]
O_3 ($\mu\text{g}/\text{m}^3$)	[0-25]	(25-48]	(48-72]	(72-96]	(96-144]	(145-240]
MP_{10} ($\mu\text{g}/\text{m}^3$)	[0-15]	(15-30]	(30-45]	(45-60]	(60-90]	(90-150]
SO_2 ($\mu\text{g}/\text{m}^3$)	[0-15]	(15-30]	(30-45]	(45-60]	(60-90]	(90-150]

A classificação apresentada dos níveis das concentrações na Tabela 1.1 faz referência aos indicadores da qualidade do ar. A qualidade do ar é mensurada de acordo com o nível mais alto dentre todos os poluentes, somente considerando aquele cujo valor de sua concentração supera os demais. Exemplo: assumindo-se que a concentração do poluente SO_2 é a mais alta e seu valor seja $40 \mu\text{g}/\text{m}^3$. Como resultado, tem-se que a qualidade do ar seria classificada como Insalubre para pessoas sensíveis e seu índice estaria entre 100 e 150. Os valores indexados associam as concentrações dos poluentes às categorias. Com a elevação

da concentração dos poluentes, os valores indexados aumentam e as categorias são modificadas.

1.3.1.2.1 Material particulado (PM)

Material particulado, ou aerossóis, são partículas líquidas ou sólidas presentes no ar e são uma mistura de partículas de diferentes tamanhos (diâmetro da partícula) e fontes, composições e propriedades. Os efeitos no sistema respiratório humano são oriundos dos diferentes diâmetros destas partículas (medida em micron ou micra). O tamanho das partículas pode variar de $PM_{2.5}$ e PM_{10} (diâmetros menores que 2.5 μm e 10 μm , respectivamente).

As fontes de material particulado são muitas e incluem poeira do solo e estradas, queimadas de florestas e biomassa, emissões de combustão e processos industriais, bioaerossóis e, cinzas vulcânicas (Curtis, 2006).

1.3.1.2.2 Ozônio (O_3)

O ozônio forma a camada que protege o planeta da radiação solar, em uma das camadas da atmosfera terrestre, a estratosfera. Neste âmbito, o gás é benéfico. Contudo, na camada mais próxima ao solo, há o ozônio troposférico, e nesta parte da atmosfera o gás produz efeitos prejudiciais ao seres vivos. A formação do ozônio ao nível do solo está associada a emissões de compostos voláteis orgânicos (COV) e óxidos de nitrogênio e, posteriormente a processos fotoquímicos (Curtis, 2006; Godish, 2003).

1.3.1.2.3 Monóxido de carbono (CO)

O monóxido de carbono é altamente tóxico para os seres humanos, pois se liga a hemoglobina, o que a impede de se associar a molécula de oxigênio (O_2), resultando em depleção da capacidade pulmonar. O monóxido de carbono é também um produto da combustão incompleta presente em motores de veículos. Não havendo oxigênio suficiente para formação de dióxido de carbono (CO_2), tem-se o CO como resultado. Emissões antropogênicas, resultantes de processos tecnológicos (combustão e indústrias) e queima de biomassa, são as principais fontes de monóxido de carbono na atmosfera (Godish, 2003; Curtis, 2006; Hinrichs, 2010).

1.3.1.2.4 Dióxido de enxofre (SO₂)

Um gás incolor, tóxico, que é formado pela queima de enxofre para a criação de óxidos. É mais comumente presente na queima de combustíveis fósseis que contém enxofre para a geração de eletricidade para estações de energia. A parte antropogênica é produzida pela queima de carvão, emissões veiculares e emissões de gás/óleo de refinarias e campos. A quantidade de dióxido de enxofre que entra na atmosfera por meio de atividades humanas é similar àquela advinda de processos naturais, não sendo, entretanto, tão amplamente distribuída. O dióxido de enxofre tem como principais fontes naturais os vulcões e os oceanos (Hinrichs, 2010; Park, 2008). Neste trabalho somente é considerada poluição àquela advinda do homem.

1.3.1.2.5 Dióxido de nitrogênio (NO₂)

Um gás colorido que varia sua coloração de amarela a vermelho dependendo de sua concentração na atmosfera. Produzido também pela combustão, quando o óxido nítrico emitido, por uma estação de energia ou veículos, é combinado ao oxigênio na atmosfera. Devido a sua alta taxa de oxidação é corrosivo, sendo tóxico em altas concentrações. Possui efeitos secundários, como, por exemplo, na formação da neblina fotoquímica (Godish, 2003; Hinrichs, 2010; Park, 2008).

1.3.1.3 Efeitos da poluição

Os efeitos da poluição são relativos a poluentes específicos e dependem diretamente do nível de concentração do poluente. Em relação aos seres humanos, inúmeros estudos associam poluentes atmosféricos a vários tipos de problemas relacionados à saúde humana (Curtis, 2006; Kampa, 2008; Wilson, 2004).

Os efeitos da poluição do ar são advindos da combinação de condições atmosféricas não favoráveis à dispersão de poluentes e níveis de concentração não saudáveis. A dispersão de poluentes, ou a habilidade da atmosfera em transportar, dispersar (diluir), e por fim remover resíduos de produtos liberados, é limitada por uma série de fenômenos atmosféricos, como a direção e a velocidade dos ventos, turbulência e estabilidade atmosférica (temperatura e umidade) (Vallero, 2008; Godish, 2003).

1.3.1.4 Efeitos adversos na saúde humana

Foram realizados inúmeros estudos que evidenciam que níveis elevados de muitos poluentes aéreos podem afetar o corpo humano de modo adverso, incluindo os sistemas respiratório, cardiovascular (incluindo coração e cérebro), reprodutivo/capacidade de desenvolvimento, e neurológico/neuropsiquiátrico (Wang X. , 2006). Uma constante “descoberta” é a de que os poluentes atmosféricos contribuem para o aumento da taxa de admissões hospitalares, mortalidade e morbidade (Kampa, 2008; Jahn, 2011).

Muitos estudos estabelecem uma forte associação entre efeitos adversos na saúde e poluentes atmosféricos prioritários (PM_{10} , $PM_{2.5}$, O_3 , NO_2 , SO_2 e, CO) (Curtis, 2006; Wilson, 2004) os quais se encontravam bem abaixo dos níveis padrões estipulados pela US-EPA, WHO e outras inúmeras agências ao redor do mundo. A combinação desses poluentes atmosféricos pode obter efeitos sinérgicos adversos à saúde humana.

1.3.1.5 Ações mitigadoras

A necessidade natural da sociedade, para que episódios como os de Donora em 1948 (Santos, 2007; Stacey, 2010), no estado da Pensilvânia, e Londres em 1952, não reincidissem fez com que várias medidas acerca da poluição fossem tomadas. Tais episódios, onde a poluição foi descontrolada, foram responsáveis pela morte de várias pessoas e por dezenas de milhares de pessoas debilitadas (Cassell, 1969; Monks, 2009).

Há duas principais formas de controle da poluição: o estratégico e o tático. O controle estratégico, considerado também como plano sistemático, é feito para que seja duradouro e com o intuito de alcançar diminuições progressivas das concentrações gasosas. Exemplos de métodos de mitigação são: mudanças no composto de produtos, pela limitação ou extinção do uso de clorofluorcarbonetos (CFC), considerados co-responsáveis pela destruição da camada de ozônio. Outro exemplo é o uso de filtros em chaminés de indústrias, pois reduzem a concentração do poluente que irá chegar à atmosfera (Vallero, 2008).

O controle da poluição a nível tático é realizado com o foco no que é imediato e extremamente necessário. A permanência por vários dias de condições climáticas não favoráveis à dispersão, geralmente, acarreta um aumento da concentração de poluentes na área (Vallero, 2008). As possíveis ações para a contenção de um episódio combinam o desligamento das indústrias (fontes poluidoras) pelo período em que as condições continuem desfavoráveis a dispersão, com o envolvimento da sociedade, que deve aderir às solicitações e diretivas do governo para sua própria segurança. O resultado possui obviamente um custo

econômico inerente a estas ações, mas considerando outros possíveis resultados, como mortes e sequelas, o de menor impacto é onde não há fatalidades.

Ações para mitigar os infortúnios da poluição atmosférica são de extrema importância. Portanto, a demora na percepção de um possível episódio pode ser a diferença entre poucas ou muitas consequências para a saúde humana. Por esta razão se fazem necessários o uso da previsão de concentrações de poluentes gasosos e o emprego da análise de riscos ambientais, como medidas de controle.

1.3.1.6 Modelos de previsão de poluentes gasosos

Modelos (numéricos, gaussianos, estatísticos ou físicos) usados na dispersão de poluentes são capazes de simular processos químicos e físicos nos momentos de dispersão e de reações que ocorrem na atmosfera, prevendo, desta forma, as concentrações de poluentes. Modelos são ferramentas importantes ao lidar com tarefas relacionadas ao gerenciamento da qualidade do ar e suas posteriores ações (Moreira, 2008).

Os modelos levam em consideração diferentes parâmetros para que as modelagens possam ser realizadas. Estes parâmetros dependem do foco que se deseja dar a determinado processo. Parâmetros como a velocidade e a direção dos ventos, temperatura do ar, concentrações gasosas de poluentes, umidade relativa do ar, dados topográficos etc., possuem seu papel na determinação da concentração de poluentes gasosos (Holmes, 2006; Aguilera, 2011).

Existe uma gama de modelos que podem ser utilizados para representar tais fenômenos. Como exemplos têm-se aqueles que são empregados na modelagem de dispersão de poluentes. Um dos modelos utilizados é o BOX (Holmes, 2006), que por ser simples é amplamente usado como uma primeira impressão do fenômeno em questão. Este modelo simplifica as conjecturas espaciais, assumindo que os poluentes estão uniformemente distribuídos (Moreira, 2008). Outro modelo é o gaussiano (Pelliccioni, 2006) que pressupõe que a nuvem de contaminantes (pluma) ao espalhar-se, leva consigo os poluentes (Godish, 2003). Tal modelo é baseado nas distribuições gaussianas da pluma nas direções vertical e horizontal, ao assumir condições (e.g.: temperatura, direção e umidade) estáveis (Holmes, 2006).

As concentrações dos poluentes gasosos disponibilizadas pelas agências reguladoras devem ser coletadas, de modo a fornecer uma sequência ordenada cronologicamente, dando origem a um conjunto de dados. Tal conjunto pode ser entendido como uma série temporal (Box & Jenkins, 1994). Esta percepção torna viável o uso de outras técnicas para a previsão de poluentes, antes não utilizadas, como, por exemplo, modelos estatísticos.

Modelos estatísticos como os modelos ARIMA (processo auto-regressivo integrado de média móvel) e ARMA (processo auto-regressivo e de média móvel) (Box & Jenkins, 1994) são utilizados na previsão de poluentes gasosos (Kumar U. , 2010; Anttila, 2010; Kumar A. , 2011). Tais modelos são baseados nos conceitos de séries temporais, que assumem que valores passados são regressões de valores futuros (Box & Jenkins, 1994; Gujarati, 2005). Contudo, modelos estatísticos precisam de conhecimento de certas características da série em questão, como sazonalidade, tendência e autocorrelação.

Modelos computacionais baseados em inteligência artificial, como redes neurais artificiais (Haykin, 2001), também são utilizados na previsão de poluentes (Kurt, 2008).

1.3.2 Séries temporais

Uma série temporal é considerada, em termos gerais, como uma sequência de observações sobre uma variável de interesse, contínua ou discreta, e indexada de acordo com a ordem cronológica na qual foi obtida. Tais observações são empregadas para inferir sobre o processo que gerou a série (Shumway, 2011).

Uma série temporal pode ser definida como

$$X_t = \{x_t \in \mathbb{R} \mid t \in \mathbb{N}\}, \quad (\text{Eq. 1})$$

onde X_t é a sequência de observações (x_1, x_2, \dots, x_t) e t é o índice cronológico que define a granularidade da série. A variação de t (Δt) determina a frequência com que as observações da série são obtidas.

Uma série temporal pode ser determinística ou probabilística. A última significa que as observações são parcialmente determinadas pelo estado interno do sistema. Desta forma, algumas informações estão fadadas a serem perdidas. Observações podem ser obtidas de forma discreta ou contínua. Há, em alguns casos, a necessidade de transformar tais observações em valores referentes a um determinado período de tempo, resultando em um valor médio que pode ser horário, diário, anual etc.

Em princípio, ao analisar séries temporais busca-se o desenvolvimento de modelos matemáticos que descrevam de modo plausível a sequência de dados observada. A análise da série pode ser no domínio da frequência ou no domínio do tempo, sendo esta última empregada com maior assiduidade.

1.3.2.1 Modelos estatísticos

Diversos modelos foram propostos ao longo dos anos. Nos anos 60 e 70 modelos baseados em equações simultâneas eram largamente aplicados a séries econômicas (Hollauer, 2008). Com mudanças advindas do meio econômico, um novo paradigma foi desenvolvido por Box e Jenkins (Box & Jenkins, 1994) e ganhou espaço neste meio. Apesar de não ser baseada em teorias econômicas, como os modelos de equações simultâneas, a metodologia de Box e Jenkins é amplamente aplicada a problemas relacionados à previsão de séries temporais econômicas, e também nas mais diversas áreas do conhecimento.

A metodologia de Box e Jenkins, frequentemente chamada de metodologia ARIMA (auto-regressivo integrado de médias móveis) é baseada em dois modelos, a saber, modelos AR (p) e MA (q). Tais modelos assumem que valores passados da própria série e de choques aleatórios, respectivamente, podem explicar os futuros valores da própria série.

O primeiro modelo é o modelo AR (autorregressivo) de grau p ou, AR (p), definido por

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + \dots + \alpha_p(Y_{t-p} - \delta) + u_t, \quad (\text{Eq. 2})$$

onde Y_t é a série original, δ é a média, α_p é um coeficiente constante associado ao termo de observação no tempo $t - p$ acrescido de uma perturbação (erro ou ruído branco) u_t . Caso a média seja zero ($\delta = 0$),

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + u_t, \quad (\text{Eq. 3})$$

O segundo é o modelo MA (média móvel) de grau q ou, MA (q), definido por

$$Y_t = u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_q u_{t-q}, \quad (\text{Eq. 4})$$

onde u_t é o erro ou ruído branco, β_q são os coeficientes do polinômio. Sendo considerada uma combinação linear dos termos de erro.

O modelo ARMA (autorregressivo e de média móvel) ou ARMA (p, q) que possui tanto uma parte autorregressiva de grau p quanto de médias móveis de grau q , é definido por

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_q u_{t-q}. \quad (\text{Eq. 5})$$

E por último, o modelo ARIMA (auto-regressivo integrado de médias móveis) ou ARIMA (p, d, q). Aplica-se d diferenciações à série homogênea no intuito de deixá-la estacionária, e à série resultante aplica-se o modelo ARMA (p, q). É dito que foi aplicado à série o modelo ARIMA (p, d, q) com d diferenciações.

O modelo ARIMA pode ser entendido como uma generalização dos demais modelos. A alteração dos parâmetros (p, d, q) do modelo ARIMA, evidencia qual o modelo está sendo aplicado à modelagem da série. Exemplos:

- ARIMA (1, 0, 0) – modelo puramente auto-regressivo, similar ao modelo AR (1), modelo AR de primeira ordem;
- ARIMA (0, 0, 2) – modelo puramente de médias móveis, similar ao modelo MA (2), modelo MA de segunda ordem;
- ARIMA (1, 0, 2) – modelo ARMA (1, 2), um elemento auto-regressivo e dois de médias móveis.
- ARIMA (2, 1, 2) – o modelo possui uma diferenciação, pois $d = 1$. Após a diferenciação, o modelo ARMA (2, 2) é aplicado à série diferenciada. Possui dois

elementos auto-regressivos ($p = 2$) e, possui dois elementos de médias móveis ($q = 2$).

1.3.2.2 Exemplos de métodos para identificar os parâmetros p e q

São empregadas, usualmente, a função de auto-correlação (FAC) e a função de auto-correlação parcial (FACP) às séries temporais para determinar os parâmetros p e q (Gujarati, 2005). Tais funções são importantes ferramentas para a análise de séries temporais no domínio do tempo. A utilização do gráfico de auto-correlação, ou correlograma, é feita para proporcionar melhor compreensão do processo evolutivo da série ao longo do tempo, através dos relacionamentos entre valores separados por um específico número de pontos temporais (atrasos temporais ou *lags*) presentes na série (Box & Jenkins, 1994).

1.3.2.3 Atrasos temporais e o teorema de Takens

1.3.2.3.1 Atrasos temporais

Atrasos temporais (*lags*) são os pontos passados de uma série temporal. Tais pontos são empregados como entradas dos modelos de previsão. O modelo gera a previsão baseando-se nas entradas. Desta forma, o modelo associa valores passados a valores futuros. Por esta razão, os resultados do modelo dependem de quais atrasos temporais são utilizados como parâmetros do mesmo.

Uma das técnicas utilizadas para identificar quais os atrasos temporais significativos é técnica de visualização dos correlogramas gerados pelas funções FAC e FACP (Gujarati, 2005). Contudo, tais técnicas são aplicadas a séries temporais geradas por processos lineares, não sendo muito aplicadas a processos não-lineares, e caso sejam, são somente concebidas como estimativas.

1.3.2.3.2 Teorema de Takens

Takens (Takens, 1980) afirma que é possível a reconstrução do processo gerador de uma série temporal a partir de um subconjunto de seus retardos temporais. O problema é a cardinalidade (c) deste subconjunto. Takens provou que se c for grande o suficiente, é possível tal reconstrução, ou seja, se c contiver todas as informações necessárias referentes à série a reconstrução é viável.

1.3.2.4 Modelos de previsão de séries temporais

Modelos de previsão, como os modelos de Box e Jenkins, em geral, fazem duas suposições. A primeira é que a série a ser modelada é uma série homogênea e, a segunda é que o processo que gerou a série é linear.

Contudo, os processos, que originam as séries são advindos de problemas complexos reais, e, como tais, estão sujeitos a uma infinidade de relacionamentos. Então, uma premissa válida é que modelos mais complexos que incorporem processos não-lineares possam obter melhores resultados ao modelar tais relações (Guoqiang Zhang, 1998). Por esta razão, modelos de processamento não-lineares vêm sendo aplicados, na esperança de que a precisão resulte em previsões melhores. No campo da computação, mais especificamente, na área de inteligência artificial, um dos modelos mais utilizados em modelagens não-lineares é a Rede neural artificial (RNA). Uma RNA é um modelo matemático que tem como inspiração o cérebro humano e suas capacidades de aprendizado e processamento paralelo (Haykin, 2001). As redes neurais artificiais serão abordadas posteriormente.

1.3.2.5 A previsão como parte do processo de análise de riscos

As técnicas de previsão, neste trabalho, são aplicadas com o intuito de obter valores futuros das concentrações de determinados poluentes, os indicadores da qualidade do ar. Uma vez adquiridos, tais valores são empregados como entradas em um modelo de riscos ambientais. O modelo aplica técnicas de análise de riscos para quantificar o grau da degradação do ar. A análise de riscos utiliza os indicadores para classificar a qualidade do ar, como, por exemplo, saudável, não-saudável, perigosa, etc.

Na seção seguinte, uma breve introdução sobre risco é apresentada, como também são abordadas algumas técnicas de análise de riscos.

1.3.3 Riscos

Risco pode ser definido de diversas formas. A definição de risco que este trabalho emprega pode ser explanada como o potencial de ocorrência de resultados adversos à saúde humana e ao meio ambiente (consequências). Sendo definido por

$$\text{Risco} = \text{Potencial} * \text{Consequências}, \quad (\text{Eq. 6})$$

onde a inexistência de qualquer um dos parâmetros indica a inexistência do risco (Covello, 1993; Lahr, 2010).

No contexto de modelagem, o termo risco é utilizado para descrever os resultados do modelo e o termo incerteza é o grau de confiança nos resultados do modelo (Covello, 1993; Sánchez, 2008). O grau de incerteza depende do problema, e do modelo, ou da técnica aplicada ao gerenciamento e à análise de riscos.

O gerenciamento de riscos é definido como o processo pelo qual decisões são tomadas acerca da necessidade de um risco ser administrado e, os meios para alcançar tal gerenciamento. O gerenciamento de risco (Darbra, 2008) engloba as seguintes etapas:

1. Estabelecimento do contexto em que o risco será analisado – especifica qual foco será dado (e.g.: análise de risco em aquíferos, solo ou, ar) e como será conduzida a análise, pois métodos de coleta de informações são distintos a depender do foco desejado;
2. Identificação das possíveis fontes de risco – quais serão os parâmetros levados em consideração ao avaliar-se o risco, como por exemplo: níveis da concentração de poluentes atmosféricos, intensidade de chuvas;
3. Previsão da frequência e consequências – define o risco como sendo o produto entre a frequência e as consequências;
4. Caracterização do risco em níveis (e.g.: aceitável, tolerável e intolerável);
5. Criação do plano de riscos (contenção e mitigação);
6. Desenvolvimento do plano de risco;
7. Revisão e avaliação do plano de risco.

A análise de riscos está exposta nas primeiras quatro etapas do gerenciamento de risco. Nesta dissertação é abordada, em particular, a análise de riscos, não sendo as demais etapas do processo de gerenciamento contempladas.

A análise de riscos pode ser definida como um processo sistemático usado para descrever e quantificar os riscos associados a substâncias perigosas, processos, ações ou, eventos (Covello, 1993). A análise de riscos, em particular, a análise de riscos ambientais, é um elemento essencial na tomada de decisão de qualquer processo, para que possam ser minimizados os efeitos das atividades humanas no ambiente. Ela oferece meios concretos para avaliação dos riscos potenciais à saúde humana e ao meio ambiente, o que possibilita, posteriormente, a classificação quanto à importância dos mesmos.

No contexto da qualidade do ar, a análise de riscos ambientais tem como um de seus focos a previsão de episódios, para que haja mitigação e minimização dos riscos associados.

A previsão de episódios é possível através da previsão dos indicadores da qualidade do ar, considerando-os como fontes de risco. O risco, quando associado à qualidade do ar faz, frequentemente, referência aos níveis de exposição aos quais estão sujeitos os seres humanos, animais e estruturas físicas. Os níveis de exposição são tratados como “frequência” e, as “consequências” são os efeitos adversos aos quais estão todos sujeitos. Desse modo, o risco pode ser considerado como o produto entre a exposição e os efeitos adversos. A caracterização (classificação) do risco é inferida a partir da medida de risco obtida, finalizando as etapas da análise de risco (Onkal-Engin, 2004; Sowlat, 2011).

1.3.3.1 Métodos de análise de riscos

Vários métodos são aplicados à análise de riscos (Haimes, 2009), como por exemplo:

- *Harzard and operations analysis* (HAZOP) (Dunjó, 2010);
- Mapas de riscos (Lahr, 2010);
- Árvore de decisão (Frohwein, 1999);
- Matriz de riscos (Markowski, 2008);
- Diagramas de influência (Qin, 2011).

As seguintes subseções explanam brevemente duas destas técnicas: os diagramas de influência e matriz de riscos.

1.3.3.1.1 Diagramas de influência

A técnica de diagramas de influência é tanto um método para representar o problema de decisão, quanto um modo de resolver o problema de decisão (Qin, 2011). Essencialmente, um diagrama de influência é construído com uso de uma série de nós que são interconectados por arcos. Os nós representam as variáveis incertas (estocásticas) do problema, e os arcos

representam as conexões (influências) que existem entre as variáveis. Ao utilizar diagramas de influência, procura-se explicar e compreender a cadeia de lógica complexa causal que se propaga a partir de decisões estratégicas e estruturadas (Diekmann, 1992; Haimes, 2009). Exemplo desta metodologia é apresentado na Figura 1.1 (a).

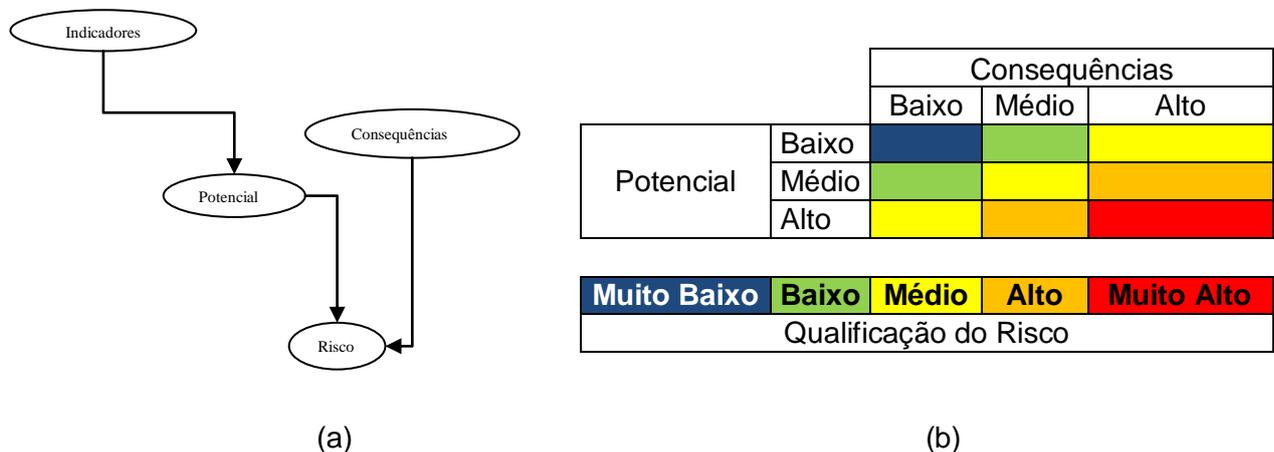


Figura 1.1- Exemplos de metodologias de análise de riscos. Diagrama de influência (a) e Matriz de riscos (b).

Na Figura 1.1 (a) é um exemplo usual de diagramas de influência. Neste exemplo há quatro variáveis, as quais são interconectadas por arcos. A interpretação funcional de um diagrama de influência é a seguinte. A variável “Potencial” é influenciada (diretamente precedida) pela variável “Indicadores”. As variáveis “Indicadores” e “Consequências”, ambas são entradas do diagrama, pois estas não possuem nenhuma variável predecessora. A variável “Riscos” é diretamente influenciada pelas variáveis “Potencial” e “Consequências” e indiretamente influenciada pela variável “Indicadores”. Basicamente, diagramas de influência possuem três camadas: uma camada de entrada, uma camada intermediária e uma camada de saída. A camada de entrada contém todas as condições e suposições iniciais sobre as circunstâncias sob estudo (variáveis “Indicadores” e “Consequências”). A camada de saída contém as variáveis que representam as saídas de interesse (variável “Risco”). A camada intermediária contém todas as variáveis (variável “Potencial”) que interconectam as camadas de entrada e saída.

1.3.3.1.2 Matriz de risco

Matriz de risco é uma ferramenta de análise que é capaz de avaliar vários tipos de riscos baseando-se em uma série histórica (Markowski, 2008). Apesar de não ser muito precisa e da falta de uma base matemática sólida, as características de expressões gráficas intuitivas, fáceis de compreender e aplicar ainda a fazem ser aplicada a alguns setores da economia

como, por exemplo, a área de serviços (Ni, 2010). Exemplo de matriz de risco é apresentado na Figura 1.1 (b).

Para alcançar uma compreensão mais adequada de matriz de riscos alguns pontos-chaves são necessários (Ni, 2010):

1. A definição mais básica de risco diz: o risco é o resultado da combinação da severidade das consequências e de suas respectivas probabilidade ou, freqüências. Para construir uma matriz de riscos são somente necessárias duas variáveis: consequência e probabilidade;
2. As consequências, probabilidades e resultados (variável resposta) podem ser subdivididos em níveis, com qualificadores e escalas próprias;
3. O resultado do processo é representado por um conjunto de regras condicionais do tipo SE-ENTÃO (e.g.: SE as Consequências forem *altas* E a Probabilidade for *baixa* o Risco é *médio*).

A Figura 1.1 (b) classifica o *Potencial* e as *Consequências*. A matriz possui nove células e cinco zonas de forma irregular. Tais zonas são formadas pelas conjecturas das regras condicionais. A severidade das consequências e o potencial são classificados em três níveis: *baixo*, *médio* e *alto*. Contidas na matriz, as cores azul, verde, amarelo, laranja e vermelho estabelecem respectivamente os riscos *muito baixo*, *baixo*, *médio*, *alto* e *muito alto*. A descrição dos níveis (Khan, 2005):

- *Muito baixo* – risco mínimo ou passível de negligência;
- *Baixo* – risco de afetar uma parcela mais sensível da população, e, por isto, requer algum tipo de remediação
- *Médio* – risco que atinge uma parte considerável da população e com o potencial razoável de degradação da saúde humana;
- *Alto* – risco que aflige grande parte da população com consequências e alto potencial de devastação;
- *Muito alto* – risco que afeta a todos de modo extremo.

A análise de riscos é bastante útil no processo de tomada de decisão. Não obstante deste fato, tanto as medições (e.g.: adquiridas por meio de sensores) quanto o tratamento (e.g.: feito por analistas humanos) das informações estão sujeitos a incertezas e imprecisões (Darbra, 2008). Portanto, incertezas são inerentes à análise de riscos. Por estas razões, pesquisas vêm sendo desenvolvidas no sentido de suprir tais dificuldades. Um dos meios de superar as adversidades expostas é a aplicação de um método computacional adequado para trabalhar com incertezas. Um possível método é denominado de lógica nebulosa (Zadeh,

Fuzzy logic, 1988). A Figura 1.2 apresenta algumas fontes de incertezas, sendo o caminho contínuo da figura a caracterização que este trabalho emprega através das incertezas de parâmetros, mais especificamente, as incertezas ocasionadas pela(s) imprecisão(ões) linguística(s).

As incertezas advindas de parâmetros podem ser subdivididas em cinco categorias (Haimes, 2009):

- Erro aleatório em medições diretas – erro causado por limitações de conhecimento e técnicas;
- Erro sistemático – erro ocasionado por julgamentos subjetivos, ou interpretação equivocada das medições;
- Erro amostral – ocorre quando há uma inferência de uma representatividade limitada;
- Erro imprevisível ou choques aleatórios – faz referência a sensibilidade extrema exibida por sistemas, como, por exemplo, nos sistemas não-lineares (e.g.: sistemas ambientais);
- Erro de imprecisão linguística – afirmações imprecisas geram interpretações não desejadas. Alguns métodos tentam levar em consideração a imprecisão linguística, como, por exemplo, a lógica nebulosa.

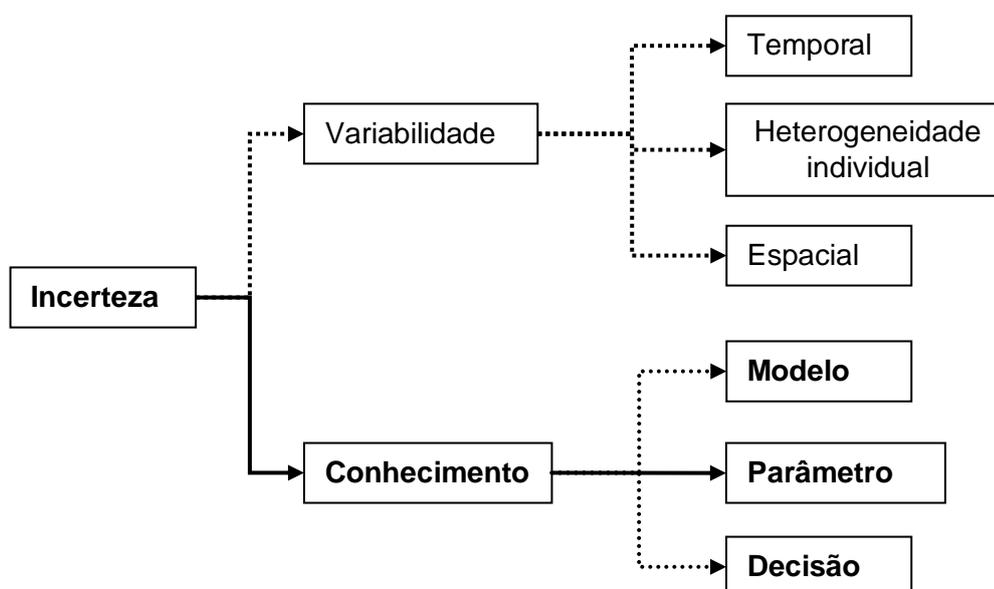


Figura 1.2 – Fontes mais comuns de incertezas (Haimes, 2009).

As incertezas causadas por parâmetros são posteriormente propagadas para o modelo que as utiliza e às decisões tomadas, quando baseadas nestes modelos.

Especialistas, engenheiros e gerentes ao avaliarem riscos em sistemas complexos, frequentemente, os percebem em termos linguísticos (variáveis linguísticas) como, por exemplo, baixo, médio, muito alto, etc. A justificativa do uso da lógica nebulosa é o fato desta lidar de forma adequada com estes tipos de incertezas, ao abranger termos vagos. As variáveis linguísticas podem ser aplicadas na racionalização do problema, e, portanto, podem ser empregadas na propagação das incertezas ao longo do processo de decisão (Sowlat, 2011; Khan, 2005). A lógica nebulosa será tratada nas próximas seções.

1.3.4 Modelos baseados em inteligência artificial

Inteligência artificial (IA) é o ramo da Ciência da Computação que busca nos sistemas naturais (suas estruturas e comportamentos), inspiração à resolução de problemas complexos do mundo real (Coppin, 2010). O resultado deste esforço é o desenvolvimento de modelos e métodos, como por exemplo:

- Algoritmos genéticos (Goren, 2010);
- Lógica nebulosa (Zadeh, Fuzzy logic, 1988);
- Colônia de formigas (Maniezzo, 2000);
- Redes neurais artificiais (Haykin, 2001);
- Otimização por enxame de partículas (Eberhart, 1995);
- Sistemas imunológicos artificiais (Forrest, 1999).

IA vem sendo há bastante tempo inserida no contexto ambiental, principalmente, devido ao fato desta lidar bem com problemas complexos não-lineares. Exemplos são as Redes neurais artificiais que têm encontrado grande aceitação nas ciências ambientais nas últimas décadas (Gardner, 1998; Maier, 2010). Também a Lógica nebulosa, nos últimos anos, vem sendo bastante aplicada a riscos, por lidar bem com imprecisões e incertezas (Chan & Huang, 2003). Contudo, ao conhecimento do autor, a Otimização por enxame de partículas não é largamente usada nas ciências ambientais. Nas seções subsequentes, as técnicas supracitadas são abordadas com mais detalhes.

1.3.4.1 Redes neurais artificiais

Redes neurais artificiais (RNA) são ferramentas de modelagem computacional que encontram uma variedade de aplicações em problemas complexos do mundo real. RNA pode ser definida como uma estrutura composta por elementos de processamento simples, adaptativos e, massivamente interconectados (denominados de neurônios ou nós) que são capazes de realizar processamento de dados e representação do conhecimento (Coppin, 2010; Haykin, 2001). Na Figura 1.3 é possível visualizar alguns exemplos de redes neurais artificiais.

Existe uma variedade de tipos de redes neurais, os quais podem ser classificados em duas categorias de treinamento: supervisionado e não-supervisionado. O treinamento supervisionado emprega a forma de um tutor (professor) para corrigir erros. O professor é capaz de fornecer um vetor (conjunto) de treinamento, que contém sinais correspondentes de entrada e da saída desejada, à rede, reproduzindo o ambiente que se deseja simular. A rede,

então, ajusta os seus parâmetros de acordo com o treinamento e o erro. O erro é a diferença entre a saída desejada e saída real da rede. Ao ser constatado que os vetores de entrada estão sendo corretamente previstos pela rede, é dito que o conhecimento para simular a rede foi alcançado e o treinamento é terminado. Exemplos de treinamento supervisionado são:

- Redes de Perceptron de Camada Única;
- Redes de Perceptron de Múltiplas Camadas;
- Redes de Função de Base Radial.

O treinamento não-supervisionado é amplamente empregado em situações onde se deseja classificar grupos baseando-se em suas características. As redes que utilizam tal treinamento designam determinados padrões com características similares ao mesmo grupo (Basheer, 2000). Mapas de Kohonen, redes de Hopfield, e redes de Base Radial são alguns tipos de redes neurais que aplicam o treinamento não-supervisionado (Basheer, 2000; Haykin, 2001; Coppin, 2010).

Neste trabalho apenas as redes neurais multicamadas são consideradas, pois um número extenso de artigos as mencionam como possíveis soluções para problemas na área ambiental, especificamente na área de previsão de poluentes gasosos.

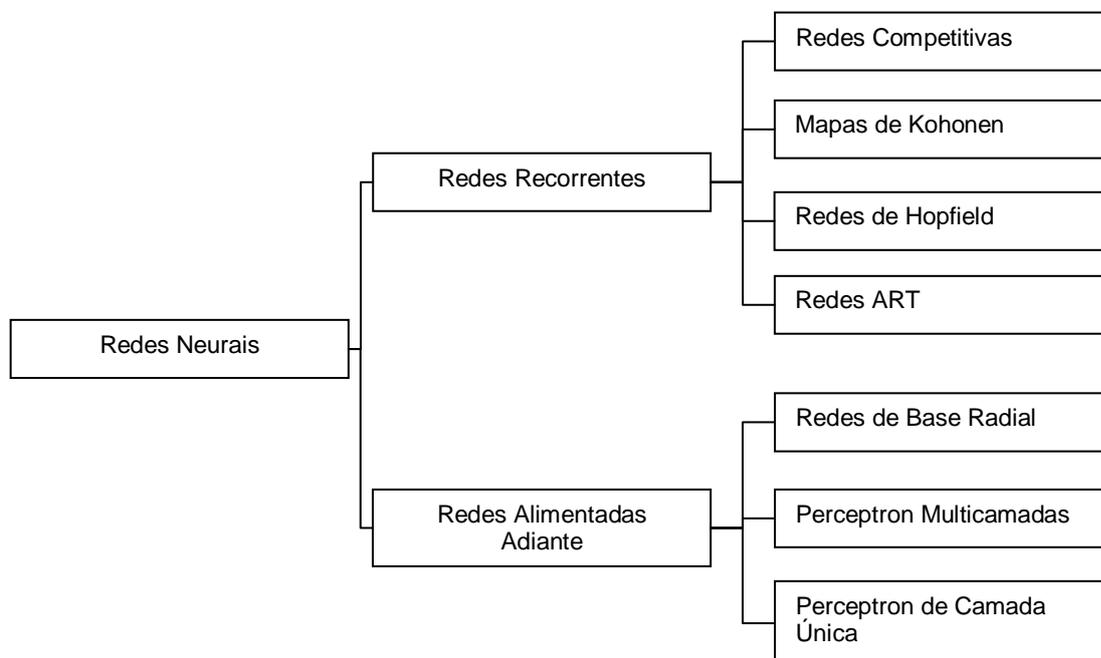


Figura 1.3 – Exemplo de taxonomia de arquiteturas de redes neurais (Gardner, 1998).

Um neurônio simples apresentado na Figura 1.4, ou perceptron, possui duas componentes principais:

- Uma soma ponderada

$$s = \sum_{i=1}^n w_i x_i + b \quad (\text{Eq. 7})$$

onde s a soma das entradas $(x_1, x_2, x_3, \dots, x_n)$, em que w_i é o respectivo peso (peso sináptico) da entrada x_i e b é o viés (*bias*) do neurônio;

- Uma função de ativação que pode ser linear, não-linear ou lógica, e é responsável por prover uma saída de acordo com a soma ponderada s . Exemplos de funções que podem assumir o papel de função de ativação são as funções sigmóide, linear, tangente hiperbólica e degrau. Usualmente, todos os neurônios presentes em uma dada camada possuem a mesma função de ativação (Basheer, 2000).

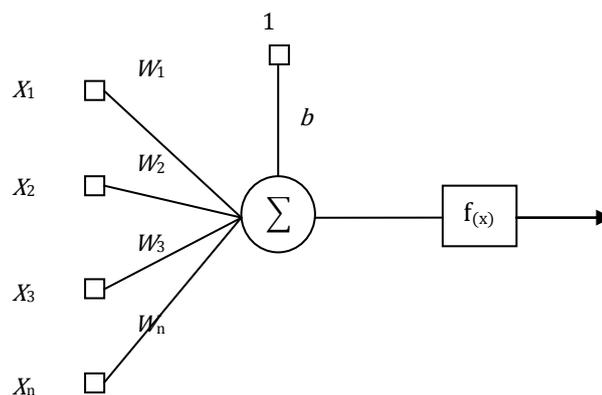


Figura 1.4 – Exemplo de um neurônio simples (Haykin, 2001).

Uma rede neural multicamadas, Figura 1.5, ou rede de perceptrons multicamada (MLP), consiste de duas ou mais camadas denominadas: camada de entrada, camada(s) escondida(s) e, camada de saída. Cada nó na camada de entrada recebe um sinal de entrada único. Na realidade, os nós nesta camada normalmente são passivos, o que significa que eles não processam dados, mas somente recebem os sinais em suas entradas e os propagam para a próxima camada (Haykin, 2001). Uma MLP formada por três ou mais camadas possui camada(s) entre as camadas de entrada e saída, denominada(s) camada(s) escondida(s). As camadas escondidas processam os sinais enviados pela camada anterior, que pode ser a camada de entrada ou outra camada escondida, por meio de uma soma ponderada e uma função de ativação específica. Posteriormente, a camada de saída recebe os sinais da camada anterior e, realiza uma soma ponderada e aplica uma função de ativação particular ao resultado.

Usualmente os dados utilizados para o treinamento, validação e teste de uma rede neural são divididos em três grupos. Primeiro, existe o conjunto de treino que geralmente corresponde à metade ou mais de todo o conjunto de dados. Este conjunto é usado pela MLP para ajustar seus pesos e vieses de acordo com os dados apresentados. Existem muitos algoritmos que são utilizados no treinamento de uma MLP. Os mais empregados são os algoritmos de gradiente descendente (Maier, 2010).

Os algoritmos de gradiente descendente (Mukherjee, 2012) podem ser subdivididos em métodos de primeira ordem, como o algoritmo de retro-propagação (Rumelhart, 1986), ou métodos de segunda ordem, como é o caso do algoritmo de Levenberg-Marquardt (Marquardt, 1963). Uma terceira opção para o treinamento da rede é o uso de um método de otimização global, como algoritmos genéticos e otimização por enxame de partículas, contudo tais algoritmos possuem um custo computacional alto (Ince, 2010).

O processo de treinamento é obtido através da atualização dos pesos sinápticos e vieses contidos na rede e, o mesmo depende de qual algoritmo está sendo empregado. O conjunto de validação é usado para validar o treinamento da rede. Este conjunto avalia a capacidade da rede de generalizar uma série de dados de entrada. O número de erros de validação deve ser menor que um limite estipulado, caso seja maior, o treinamento é interrompido e o melhor estado (pesos e vieses) anterior é recolocado. Este tipo de validação também é conhecido como validação cruzada. Finalmente, o conjunto de teste, uma série de dados que ainda não foram usados, representado o restante do conjunto inicial de dados, é empregado para avaliar a desempenho da rede (Haykin, 2001; Coppin, 2010).

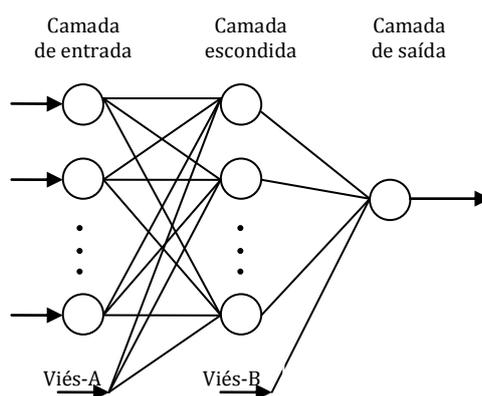


Figura 1.5 - Modelo de uma rede neural multicamadas.

O processo de treinamento é realizado até o momento em que algum critério de parada seja alcançado. Os erros associados com os critérios de parada são erros de validação (generalização) e erros de treino. O erro de validação faz referência ao número de vezes consecutivas que a previsão da rede dos pontos do conjunto de validação é mantida ou

agravada. O erro de teste é a medida de desempenho da rede ao trabalhar com o conjunto de teste. Os erros de treinamento, de validação e, de teste são avaliados por uma métrica de desempenho. O erro é a medida de comparação entre os dados observados e os previstos pela rede. Diversas métricas podem ser empregadas para esse fim, inclusive o MSE (erro médio quadrático), a estatística U de Theil, entre outros.

Redes neurais apresentam alguns impasses para garantir que a melhor rede possível será aplicada ao problema em questão. Tais dificuldades proporcionam a formulação de novas suposições e, posteriormente novas técnicas. Algumas dessas questões são consideradas neste trabalho.

Uma das mais importantes decisões no processo de desenvolvimento de uma rede neural é a escolha apropriada de seu conjunto de entradas. Grande parte da literatura escolhe o conjunto de entradas baseando-se em conhecimento especialista ou emprega o uso de sistemas de conhecimento geral. Tais escolhas podem resultar em muitas ou poucas entradas. O excesso de entradas pode acarretar em entradas sem a informação desejada ou informações redundantes. O último pode ainda trazer outros problemas como o *overfitting* (superajuste ou treinamento extremo). É quando a rede na fase de treinamento ao ajustar os pesos sinápticos para classificar melhor o conjunto de treino, acaba por perder a capacidade de generalização. Como resultado da perda da capacidade de generalização, ao serem apresentadas novas entradas, a rede as classifica de maneira equivocada (Gardner, 1998; Maier, 2010).

Existem opções para selecionar o conjunto de entradas, como nas técnicas *ad-hoc*, na qual o desenvolvedor decide quais entradas são pertinentes ao problema. Outra possível técnica é a de escalonamento, que consiste em se aumentar sistematicamente (construtivo) ou se remover (poda). Uma terceira opção, dentre possíveis outras, é um algoritmo de otimização global, como a otimização por enxame de partículas ou, algoritmos genéticos. Tais técnicas são formuladas para maximizar o desempenho do modelo (Maier, 2010).

1.3.4.2 Otimização por enxame de partículas

O algoritmo de otimização por enxame de partículas (PSO) foi desenvolvido por Kennedy e Eberhart (Eberhart, 1995). É um algoritmo evolucionário, o qual foi inspirado pelo movimento realizado por bandos de pássaros e cardumes ao procurarem por comida. O algoritmo procura otimizar uma população de soluções aleatórias. Cada partícula (indivíduo da população) possui uma posição e uma velocidade, representando a solução para o problema de otimização e a direção da procura no espaço de busca. A partícula ajusta a velocidade e a posição de acordo com as melhores experiências. O algoritmo determina o melhor local

encontrado pela própria partícula e o melhor global, o qual é de conhecimento de todo o enxame.

A teoria do algoritmo de otimização por enxame de partículas assume que em uma dada população de tamanho M , cada indivíduo ($1 \leq i \leq M$) possui uma posição atual X_i , uma velocidade V_i associada, uma melhor solução local l_i e cada partícula possui um valor referente à função de *fitness* (função de aptidão). Uma função de aptidão (*fitness*) é a medida de mérito da partícula e avalia a adequação da solução da partícula na solução do problema. O melhor local da partícula é, simplesmente, a melhor solução encontrada pela própria partícula até o presente momento. A melhor solução global X_g é a melhor solução dentre todas as partículas. Os parâmetros c_1 e c_2 são as constantes de aceleração das partículas, e geram uma tendência que contribui na quantificação da velocidade na qual as partículas irão se mover no espaço de busca. Para cada ciclo de iteração, as velocidades das partículas são aceleradas em direção aos mínimos local e global:

$$V_i(t) = wV_i(t-1) + c_1r_1(l_i - X_i(t)) + c_2r_2(X_g - X_i(t)), \quad (\text{Eq. 8})$$

e as posições das partículas são atualizadas empregando dinâmica discreta:

$$X_i(t) = X_i(t-1) + V_i(t), \quad (\text{Eq. 9})$$

em que o termo w é o peso de inércia, o qual é utilizado para balancear as habilidades de procura local e global do algoritmo, desta forma, controlando a influência de informações prévias na atualização de novos valores de velocidades. Na maioria dos casos, o termo w decresce linearmente, de 1 (um) a um valor próximo de 0 (zero), durante o processo. Os parâmetros r_1 e r_2 são duas sequências aleatórias entre 0 e 1, sendo assim, os componentes aleatórios do processo (Eberhart, 1995; Lu & Wang, 2006).

No algoritmo de PSO cada indivíduo pertencente à população evolui baseando-se em dois pontos. O primeiro ponto consiste das suas próprias experiências, aprimorando suas inerentes qualidades. No segundo, assumindo-se que os indivíduos são seres sociais, eles compartilham informações entre si. Estes dois tipos de informação correspondem ao aprendizado individual (cognitivo – conhecimento local) e transmissão cultural (social – conhecimento da melhor posição dentro do enxame) respectivamente.

O algoritmo é formulado como se segue:

1. Inicializar os parâmetros, como as constantes de aceleração, peso de inércia, número de partículas, número máximo de iterações, limites de velocidade,

delimitação das velocidades e posições e, por fim, o limite do erro da função de aptidão;

2. Avaliar os valores das funções de aptidão de cada partícula, comparando-as e, portanto, fixando a melhor solução local e a melhor solução global;
3. De acordo com Eq. (8) e Eq. (9) calcular as novas velocidade e posições de cada partícula e, depois atualizá-las;
4. Para cada partícula, comparar o valor atual da função de aptidão com a melhor solução local, se o valor atual for melhor, atualizar o valor da função de aptidão da melhor solução local e sua posição com o valor atual;
5. Para cada partícula, comparar o valor atual da função de aptidão com a melhor solução global, se o valor atual for melhor, atualizar o valor da função de aptidão da melhor solução global e sua posição com o valor atual;
6. Se qualquer critério de parada for alcançado (número máximo de iterações ou o valor da função de aptidão alcançar o limite estipulado) então, o procedimento é parado e, os resultados exibidos. Caso contrário, retornar para o passo 2 (dois).

1.3.4.3 Lógica nebulosa

Para uma melhor compreensão da lógica nebulosa (*fuzzy logic*) deve-se começar pelo conceito de conjuntos nebulosos (Zadeh, Fuzzy sets, 1965). Um conjunto nebuloso é uma analogia de como o ser humano percebe e interage com o ambiente ao seu redor. Tal analogia permite que sejam trabalhados problemas incertos e imprecisos, que são inerentes ao pensamento humano. Para um melhor entendimento do problema, assume-se que conjuntos nebulosos, são como uma extensão da teoria de conjuntos aristotélicos. A teoria clássica considera os elementos de um conjunto de maneira dicotômica, ou seja, o elemento pertence ($x \in F$) ou não pertence ($x \notin F$) ao conjunto, também conhecida como regra do meio excluído. Para acomodar aspectos imprecisos do pensamento humano, os conjuntos nebulosos permitem o uso de graus de pertinência, o que significa que x pode pertencer parcialmente a F . Nos conjuntos nebulosos a pertinência é uma questão de grau.

De modo semelhante, aos conjuntos nebulosos, podemos conceituar a lógica nebulosa (Zadeh, Fuzzy logic, 1988) como uma ampliação dos conceitos clássicos. A lógica nebulosa pode ser entendida como uma generalização da lógica booleana. A lógica booleana admite que o valor 1 (um) seja designado para hipóteses verdadeiras e, 0 (zero) para hipóteses falsas. Entretanto, a lógica nebulosa classifica seus elementos através de funções de pertinência,

onde cada elemento possui um valor no intervalo fechado entre 0 e 1. O valor 0 (zero) é admitido para elementos que com certeza não pertencem ao conjunto e 1 (um) para elementos que completamente pertencem ao conjunto.

Um exemplo muito empregado para facilitar a compreensão de conjunto nebuloso é o de “pessoas altas” (Amini, 2011). Neste exemplo apresentado na Figura 1.6, é exposta a diferença entre os conceitos, clássico (discreto) e nebuloso (contínuo). Os elementos de A e B fazem referência à medida da altura de duas pessoas. De acordo com a Figura 1.6 (a), há uma disjunção entre os conjuntos de pessoas altas (PA) e pessoas não-altas ($\neg PA$), esta disjunção representa o conceito clássico de conjuntos, onde $A \notin PA$, logo, tem-se que $A \in \neg PA$. O conceito de conjuntos nebulosos é expresso na Figura 1.6 (b), onde não existe uma partição dos conjuntos bem definida, o que implica na seguinte constatação: $A \in PA$ e $A \in \neg PA$. O elemento A pertence com graus diferentes a ambos os conjuntos e ao mesmo tempo. A mesma constatação é válida de forma análoga para B . Em um conjunto nebuloso a pertinência pode ser considerada como uma questão de grau.

1.3.4.3.1 Funções de pertinência

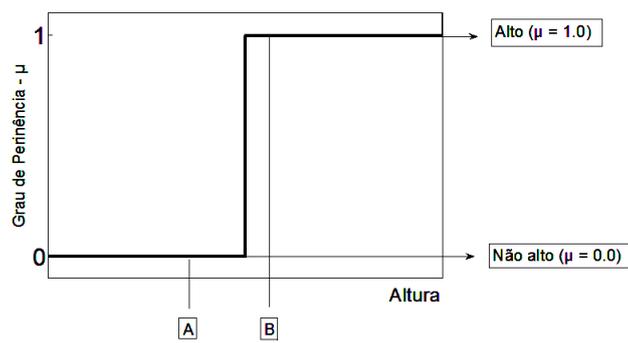
O conjunto de entradas, A e B na Figura 1.6, é frequentemente designado de universo de discurso. Assumindo que x denota os elementos presentes no universo de discurso, então um conjunto nebuloso F em X é definido como os pares ordenados,

$$F = \{x, \mu_A(x) \mid x \in X\} \quad (\text{Eq. 10})$$

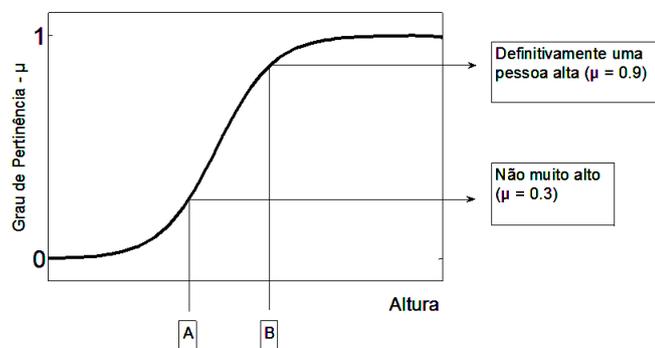
$\mu_A(x)$ é a função de pertinência de x em F . A função de pertinência mapeia cada elemento de X a um valor de pertinência entre 0 e 1.

Várias funções podem assumir o papel de função de pertinência. Tais funções são usadas para representar a distribuição de uma variável linguística. Entre as mais aplicadas, conforme a Figura 1.7, tem-se as funções:

- Triangular;
- Trapezoidal;
- Gaussiana;
- Sigmoide.



(a)



(b)

Figura 1.6 - Funções de pertinência clássica (a) e nebulosa (b).

Existem vários modos de atribuir funções de pertinência a variáveis linguísticas (nebulosas). Este processo de atribuição pode ser intuitivo ou pode ser baseado em algum algoritmo ou operação lógica (Ross, 2004). Exemplos de formas de atribuições são:

- Intuição – a capacidade de compreensão inata de cada pessoa de identificar características pertinentes ao problema;
- Inferência – capacidade de induzir ou inferir uma conclusão, quando informações e fatos são previamente conhecidos;
- Especialistas – conhecimento advindo de fontes externas, pessoas que possuem determinada informação a respeito do problema que se deseja modelar (e.g.: pesquisadores, engenheiros, gerentes, etc.).

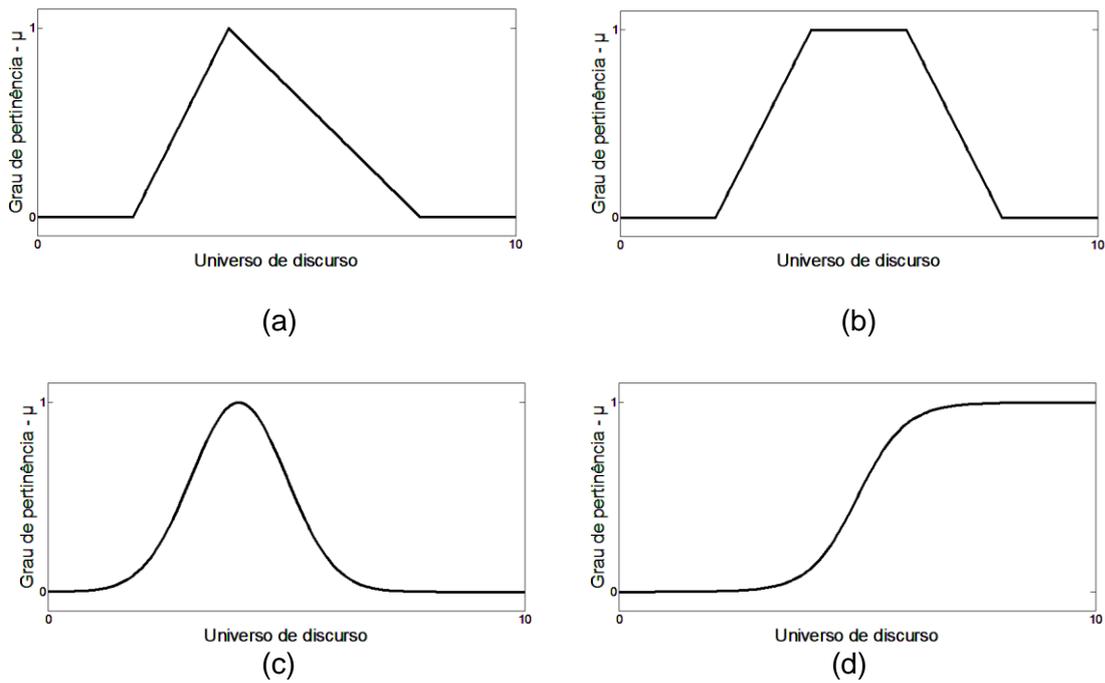


Figura 1.7 - Exemplos de funções de pertinência: triangular (a), trapezoidal (b), gaussiana (c) e, sigmóide (d).

1.3.4.3.2 Variáveis linguísticas

Variáveis linguísticas, como o nome sugere, são variáveis que possuem palavras ou sentenças como seus valores internos. Exemplos são:

- Altura;
- Idade - muito novo;
- Concentração alta;
- Risco médio.

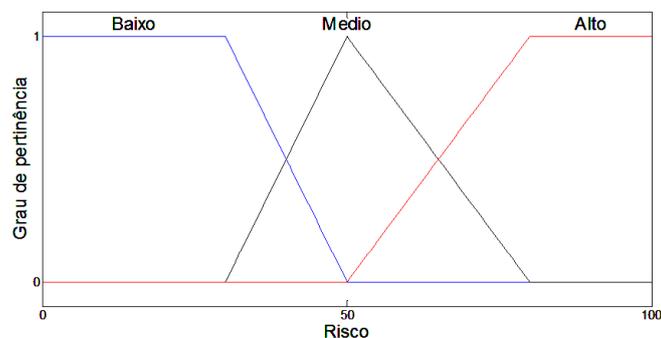


Figura 1.8 - Exemplo distribuição da variável linguística (Risco) e seus modificadores (Baixo, Médio e, Alto).

Variáveis linguísticas são compostas por um termo primário (e.g.: risco) e um conjunto de modificadores (e.g.: alto, médio, não muito baixo). Uma distribuição de uma variável linguística pode ser visualizada na Figura 1.8.

1.3.4.3.3 Operadores lógicos

As definições dos operadores lógicos empregados pela lógica nebulosa são uma extensão das definições clássicas correspondentes. As definições envolvendo operadores lógicos (E, OU e, Complemento) são estabelecidas a seguir:

O complemento de um conjunto nebuloso F é denominado F' e é definido por

$$f_{F'} = 1 - f_F. \quad (\text{Eq. 11})$$

A união de dois conjuntos nebulosos A e B com suas respectivas funções de pertinência $f_A(x)$ e $f_B(x)$ é um conjunto C , designado como $C = A \cup B$, o qual possui sua função de pertinência associada àquelas de A e B por

$$f_C = \text{Max} [f_A(x), f_B(x)], \quad x \in X. \quad (\text{Eq. 12})$$

Analogamente a união, a interseção de dois conjuntos nebulosos A e B com suas respectivas funções de pertinência $f_A(x)$ e $f_B(x)$ é um conjunto C , designado como $C = A \cap B$, o qual possui sua função de pertinência associada àquelas de A e B por

$$f_C = \text{Min} [f_A(x), f_B(x)], \quad x \in X. \quad (\text{Eq. 13})$$

1.3.4.3.4 Sentenças condicionais – Regras (Se-Então)

A lógica nebulosa consiste de sentenças condicionais que são formuladas a partir de regras *se-então*.

Uma regra *se-então* singular pode ser expressa da seguinte forma

$$\text{se } x \text{ é } A \text{ então } y \text{ é } B,$$

onde A e B são variáveis linguísticas definidas por conjuntos nebulosos com amplitude (universo de discurso) X e Y , respectivamente. A parte 'se' da regra " x é A " é denominada de antecedente ou premissa, enquanto a parte 'então' da regra " y é B " é denominada de consequente ou conclusão. Exemplos de regras condicionais são;

- Se concentração do poluente é baixa, então, qualidade do ar é boa;
- Se exposição é alta e o perigo é médio, então, o risco é alto.

1.3.4.3.5 Sistema de inferência nebulosa

Inferência nebulosa é o processo que formula o mapeamento de conjunto de determinadas entradas puras (discretas) (por meio de sensores ou, operadores humanos) em uma saída quantitativa ou ação. O mapeamento é realizado ao aplicar um conjunto de regras nebulosas a estes valores de entrada. Tal processo provê a base para que decisões possam ser tomadas ou, padrões possam ser discernidos. O processo de inferência nebulosa envolve todas as etapas anteriores: funções de pertinência, operadores lógicos e, regras se-então (Coppin, 2010).

Existem dois tipos de processamento de inferência nebulosa: o Mamdani (Mamdani, 1977) e o Sugeno (Sugeno, 1992), sendo o Mamdani o mais utilizado. A diferença entre os dois está na etapa final, relativa às funções de pertinência da saída. No método de Sugeno, tais funções são lineares ou constantes, enquanto, no método de Mamdani, estas podem assumir uma variedade de distribuições (e.g.: sigmóide, gaussiana, trapezoidal etc.).

1.3.4.4 Modelos híbridos

São denominados modelos híbridos, a combinação de duas ou mais técnicas (e.g.: estatística e computacionais) quando empregada com o intuito de trabalhar diferentes partes do mesmo problema, suprindo as deficiências umas das outras. São exemplos destes modelos:

- ARIMA e GARCH (ambos são modelos estatísticos) (Kumar U. , 2010);
- ARIMA, Redes neurais e, Lógica nebulosa (modelo estatístico e modelos computacionais) (Valenzuela, 2008);
- Redes neurais e PSO (ambos são modelos computacionais) (Fanb & Lu, 2003).

Ao combinar modelos distintos, diferentes aspectos do problema podem ser capturados, gerando uma melhor compreensão do mesmo. Um entendimento mais aprofundado da situação torna possível uma modelagem mais precisa e eficiente. Por estas razões, sistemas híbridos vêm ganhando cada vez mais espaço em aplicações no campo das ciências ambientais.

1.3.5 Simulações e resultados

A seguir serão abordados os parâmetros iniciais tanto da MLP quanto do PSO e os resultados alcançados ao fim das simulações, como também é abordado o modelo de lógica nebulosa.

1.3.5.1 Perceptron de multicamadas (MLP)

Neste trabalho uma rede de perceptrons multicamada (MLP) possui três características passíveis de mudanças:

- Sua estrutura, subdividida em duas partes, onde cada uma pode ter no máximo 10 (dez) elementos: a camada de entrada e a camada escondida;
- Seus pesos sinápticos;
- Os pesos de seus vieses (*biases*).

Como também possui três peculiaridades que são imutáveis:

- As funções de ativação da camada escondida e da camada de saída são as funções: sigmóide e a linear, respectivamente;
- A quantidade de elementos na camada de saída é fixa em 1 (hum) elemento;
- Os conjuntos de dados empregados no treinamento da MLP são divididos de acordo com o relatório Proben (Prechelt, 1994), que prevê a seguinte separação dos dados:
 - Treinamento – 50%;
 - Validação – 25%;
 - Teste – 25%.

Os critérios de parada do treinamento para todas as redes, aqui aplicadas, foram os mesmos. Estes critérios são:

- Número máximo de 1000 (hum mil) iterações;
- Algoritmo empregado: Levenberg-Marquardt;
- Erro de generalização acima de 5%;
- Erro do processo de treinamento menor que 10^{-4} .

1.3.5.2 Otimização por enxame de partículas (PSO)

O PSO possui uma quantidade predefinida de indivíduos (partículas), onde cada indivíduo é uma estrutura de dados e consiste de uma MLP e características determinadas, tais como:

- Características de uma MLP:

- Estrutura da MLP (camada de entrada e a camada escondida);
- Pesos sinápticos;
- Pesos dos vieses;
- Velocidades de cada uma das características;
- Estado (espaço) atual de cada uma das características;
- Melhor estado (melhor local) encontrado pela própria partícula;
- Conhecimento do melhor estado encontrado por uma partícula pertencente ao enxame (melhor global);

Adaptando a metodologia empregada por Lu & Wang (2006), foram realizados 10 (dez) experimentos, almejando-se alcançar uma estrutura de uma MLP adequada ao problema de previsão de séries temporais de poluentes gasosos. Em tais experimentos foram utilizados certos parâmetros fixos, tais como:

- O número máximo de 100 iterações;
- A partir da 70^a iteração há o treinamento da melhor MLP encontrada através do algoritmo Levenberg-Marquardt, caso não ocorra uma melhora de 0,1% a cada iteração (estagnação) do PSO;
- O tamanho da população (enxame) igual a 60 partículas.

Após o término dos experimentos, um dentre os dez é escolhido como o modelo mais adequado à previsão de determinada série. Tal escolha baseia-se na função de aptidão (*fitness*), a saber.

Com o interesse de comparar o modelo aplicado, outros modelos foram também utilizados na previsão destas séries. São eles os modelos ARMA e AR. Os resultados das previsões de todos os modelos são apresentados a seguir subdivididos por série temporal.

1.3.5.3 Métricas de desempenho

Para o problema de previsão de séries temporais não há na literatura uma métrica universal capaz de avaliar a adequação da previsão do modelo. Os seguintes trabalhos de (Brunelli, 2007; Zemouri, 2010) são apenas exemplos de autores que empregaram várias métricas com o propósito de avaliar o desempenho do modelo. Portanto, no presente trabalho, para possibilitar uma melhor apreciação do desempenho do modelo, são consideradas sete métricas. A primeira delas é o MSE (erro médio quadrático), a qual é uma das métricas de desempenho mais aplicadas em redes neurais artificiais.

$$\text{MSE} = \frac{1}{N} \sum_{j=1}^N (\text{observado}_j - \text{previsto}_j)^2, \quad (\text{Eq. 14})$$

onde N é o número de padrões, $observado_j$ é a previsão desejada (valor real da série) para o j -ésimo padrão, e $previsão_j$ (resposta do modelo) é o valor previsto para o j -ésimo padrão.

A segunda medida relevante é a estatística U de Theil, que é baseada no MSE normalizado pelo erro da previsão, sendo posteriormente comparando ao modelo *Random Walk*. Um modelo *Random Walk* assume que o valor mais adequado à previsão no tempo $t + 1$ é o valor obtido no tempo t acrescido de um termo de ruído (Mattos, 2009).

$$\text{Theil} = \frac{\sum_{j=1}^N (\text{observado}_j - \text{previsão}_j)^2}{\sum_{j=1}^N (\text{observado}_j - \text{observado}_{j-1})^2} \quad (\text{Eq. 15})$$

A terceira métrica relevante, a ARV (variância média relativa), é dada por

$$\text{ARV} = \frac{1}{N} \frac{\sum_{j=1}^N (\text{previsão}_j - \text{observado}_j)^2}{\sum_{j=1}^N (\text{previsão}_j - \text{observado})^2}, \quad (\text{Eq. 16})$$

a qual associa o desempenho do modelo com a média da série temporal. O termo $\overline{\text{observado}}$ é a média da série temporal. Se ARV for igual a 1, então, o modelo possui um desempenho similar ao da média da série temporal. Se ARV for maior que 1, então, o modelo apresenta um comportamento pior que o da média da série temporal, e se ARV for menor que 1, então, o modelo tem um desempenho melhor que a média da série temporal.

Outra métrica de desempenho é o MAPE (erro médio percentual absoluto), dado por

$$\text{MAPE} = \frac{1}{N} \sum_{j=1}^N \left| \frac{\text{observado}_j - \text{previsão}_j}{\text{observado}_j} \right|. \quad (\text{Eq. 17})$$

Uma métrica importante para comparação entre modelos é o IA (índice de adequação), o qual é capaz de expressar a diferença entre os dados observados e os previstos. Este índice pode variar dentro do intervalo fechado de 0 a 1. Valores mais altos indicam uma melhor adequação ao problema de previsão. O indicador é formulado por

$$\text{IA} = 1 - \frac{\sum_{j=1}^N (\text{observado}_j - \text{previsão}_j)^2}{\sum_{j=1}^N (|\text{observado}_j - \text{previsão}_j| + |\text{observado}_j - \overline{\text{observado}}|)^2}. \quad (\text{Eq. 18})$$

A última métrica de desempenho é o POCID (previsão de mudança na direção), o qual mapeia a tendência da série temporal, em outras palavras, mensura a capacidade do modelo de prever se valores futuros irão crescer ou decrescer baseando-se em valores passados. A métrica é formulada como segue:

$$\text{POCID} = 100 \frac{\sum_{j=1}^N D_j}{N}, \quad (\text{Eq. 19})$$

onde

$$D_j = \begin{cases} 1, & \text{se } (\text{observado}_j - \text{observado}_{j-1})(\text{previsão}_j - \text{previsão}_{j-1}) > 0, \\ 0, & \text{caso contrário.} \end{cases}$$

A função que avalia a adequação do modelo é denominada de função de aptidão (*fitness*), e pode ser uma dentre várias. Neste trabalho, tal função é dada por

$$\text{fitness} = \frac{1}{(1+\text{MSE})}. \quad (\text{Eq. 20})$$

A função considera o MSE. Contudo, o algoritmo é também avaliado baseando-se nas outras métricas. Em um sistema perfeito, todas as métricas de desempenho devem tender a zero, com as exceções do POCID que deve tender a 100 (cem), o IA e a função de aptidão que devem tender a 1 (um).

1.3.5.3.1 Métricas das previsões

A seguir são apresentados os resultados obtidos, através de simulações, das séries dos seguintes poluentes gasosos: material particulado, monóxido de carbono, dióxido de enxofre e dióxido de nitrogênio. As Tabela 1.2, Tabela 1.2, Tabela 1.2, Tabela 1.2 apresentam os resultados alcançados pelos modelos PSO-MLP-LM, AR e ARMA. As métricas exibidas do modelo que emprega o PSO são valores próximos àqueles alcançados pelos modelos AR e ARMA, o que significa, estatisticamente, que os modelos obtiveram desempenhos similares. Uma das possíveis explicações para tal resultado é que, os processos subjacentes (geradores) das séries observadas não puderam ser determinados (Khashei & Bijari, 2010). Apesar de não haver conseguido identificar o processo gerador das séries envolvidas neste estudo, o modelo PSO-MLP-LM foi capaz de inferir determinados padrões sem o tratamento prévio das séries de dados. Diferentemente, os modelos AR e ARMA. A priori, tais modelos necessitam de verificação prévia de determinadas informações (Box & Jenkins, 1994), acerca da série em estudo, antes de sua aplicação.

Tabela 1.2 – Métricas das previsões da série de MP₁₀

	Modelo AR	Modelo ARMA	PSO-MLP-LM
MÉTRICAS			
<i>MSE</i>	0,01928	0,01770	0,018253
<i>U de Theil</i>	1,50803	1,59015	1,280943
<i>POCID</i>	47,48603	45,25140	43,50282
<i>ARV</i>	0,00383	0,00446	0,004278
<i>MAPE</i>	1,34506	1,65346	1,622086
<i>IA</i>	0,77220	0,79755	0,798398
<i>FITNESS</i>	0,98108	0,98261	0,982074
<i>Camada escondida</i>			7
<i>Camada de entrada (Time lags)</i>			5

Tabela 1.3 – Métricas das previsões da série de NO₂

	Modelo AR	Modelo ARMA	PSO-MLP-LM
MÉTRICAS			
<i>MSE</i>	0,02370	0,02306	0,02723
<i>U de Theil</i>	0,93585	1,18699	1,11626
<i>POCID</i>	43,33333	44,44444	40,90909
<i>ARV</i>	0,01049	0,01448	0,01138
<i>MAPE</i>	0,55831	0,70137	0,62015
<i>IA</i>	0,73838	0,76161	0,73525
<i>FITNESS</i>	0,97684	0,97746	0,97349
<i>Camada escondida</i>			6
<i>Camada de entrada (Time lags)</i>			6

Tabela 1.4 – Métricas das previsões da série de SO₂.

	Modelo AR	Modelo ARMA	PSO-MLP-LM
MÉTRICAS			
<i>MSE</i>	0,01928	0,01770	0,01027
<i>U de Theil</i>	1,50803	1,59015	1,37968
<i>POCID</i>	47,48603	45,25140	46,70659
<i>ARV</i>	0,00383	0,00446	0,00797
<i>MAPE</i>	1,34506	1,65346	0,71555
<i>IA</i>	0,77220	0,79755	0,74474
<i>FITNESS</i>	0,98108	0,98261	0,98983
<i>Camada escondida</i>			1
<i>Camada de entrada</i>			1
<i>(Time lags)</i>			

Tabela 1.5 – Métricas das previsões da série de CO.

	Modelo AR	Modelo ARMA	PSO-MLP-LM
MÉTRICAS			
<i>MSE</i>	0,00877	0,00725	0,00846
<i>U de Theil</i>	1,40684	1,49364	0,88470
<i>POCID</i>	51,68539	55,05618	55,36723
<i>ARV</i>	0,00755	0,00742	0,00756
<i>MAPE</i>	0,39724	0,39176	0,41614
<i>IA</i>	0,72593	0,75828	0,74389
<i>FITNESS</i>	0,99130	0,99279	0,99160
<i>Camada escondida</i>			5
<i>Camada de entrada</i>			5
<i>(Time lags)</i>			

1.3.5.3.2 Gráficos dos valores observados versus os valores previstos

Os gráficos apresentados nas Figura 1.9, Figura 1.10, Figura 1.11 e Figura 1.12 são comparações dos últimos 100 pontos da série temporal observada versus as previsões dos modelos PSO-MLP-LM e ARMA, com a exceção da série de dióxido de nitrogênio (Figura 1.10) que apresenta somente os últimos 80 pontos. Apesar das métricas, exibidas anteriormente, dos modelos PSO e ARMA estarem muito próximas, as curvas das previsões apresentadas pelo modelo PSO-MLP-LM, em princípio, são mais fidedignas aos valores observados. Pois, possuem maior amplitude, ao acompanhar os valores de máximo e de mínimo apresentados nas séries observadas, enquanto que as curvas das previsões apresentadas do modelo ARMA permanecem mais ao centro do gráfico.

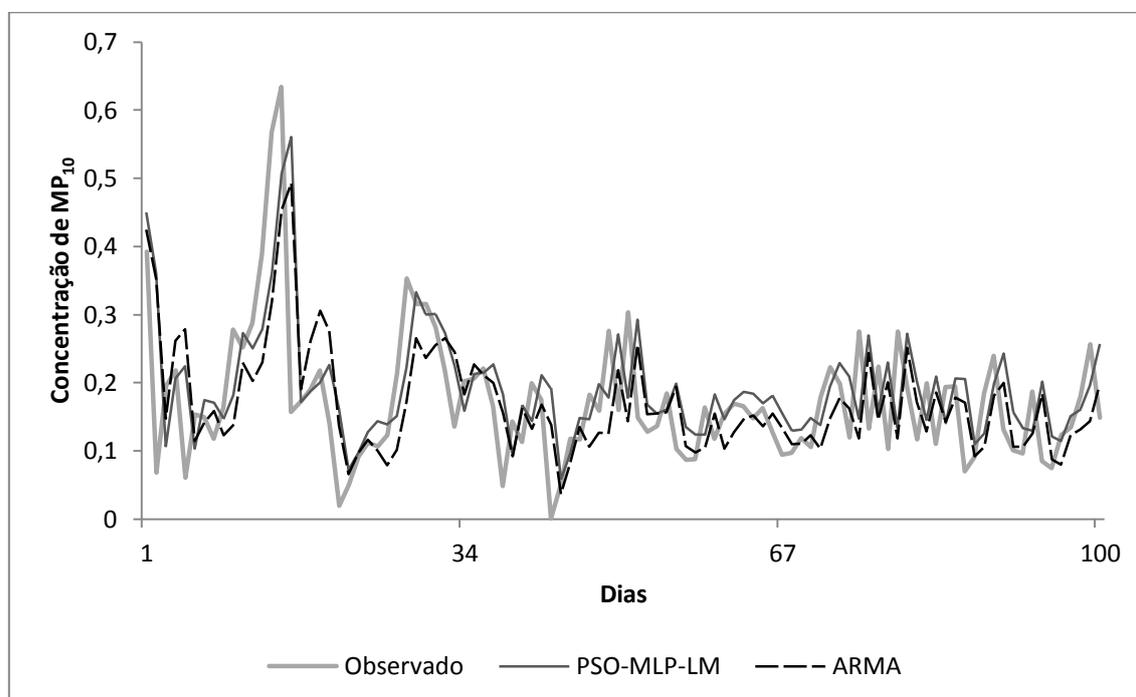


Figura 1.9 – Comparação entre os valores observados e as previsões dos modelos PSO-MLP-LM e ARMA.

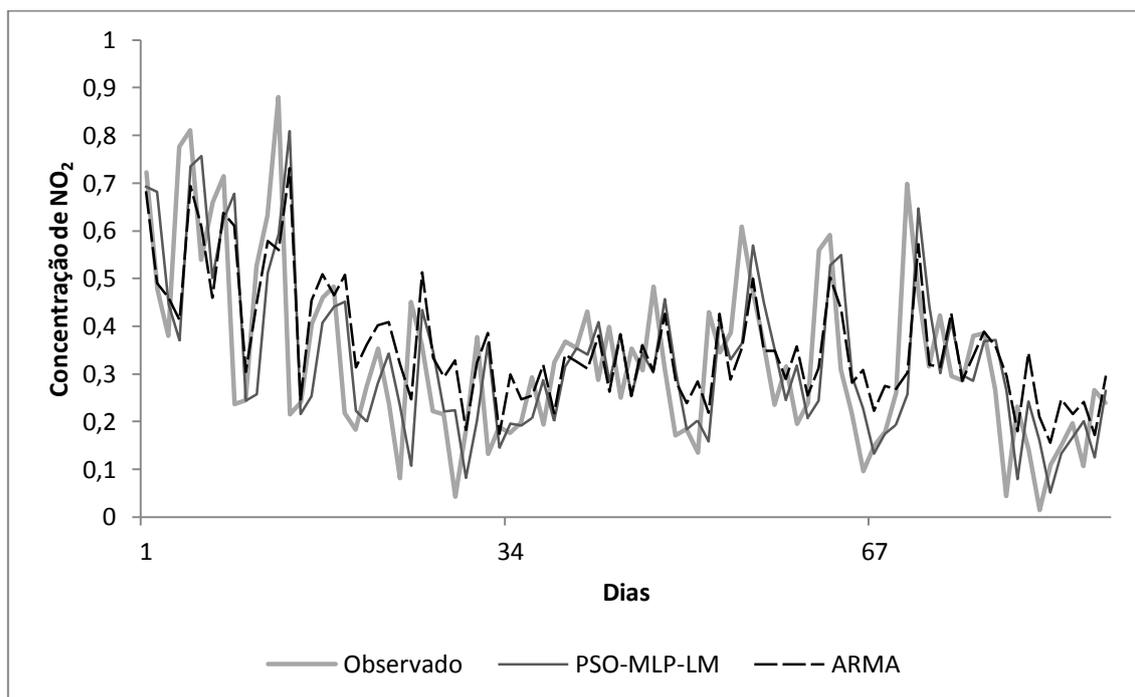


Figura 1.10 - Comparação entre os valores observados da série de NO_2 e as previsões dos modelos PSO-MLP-LM e ARMA.

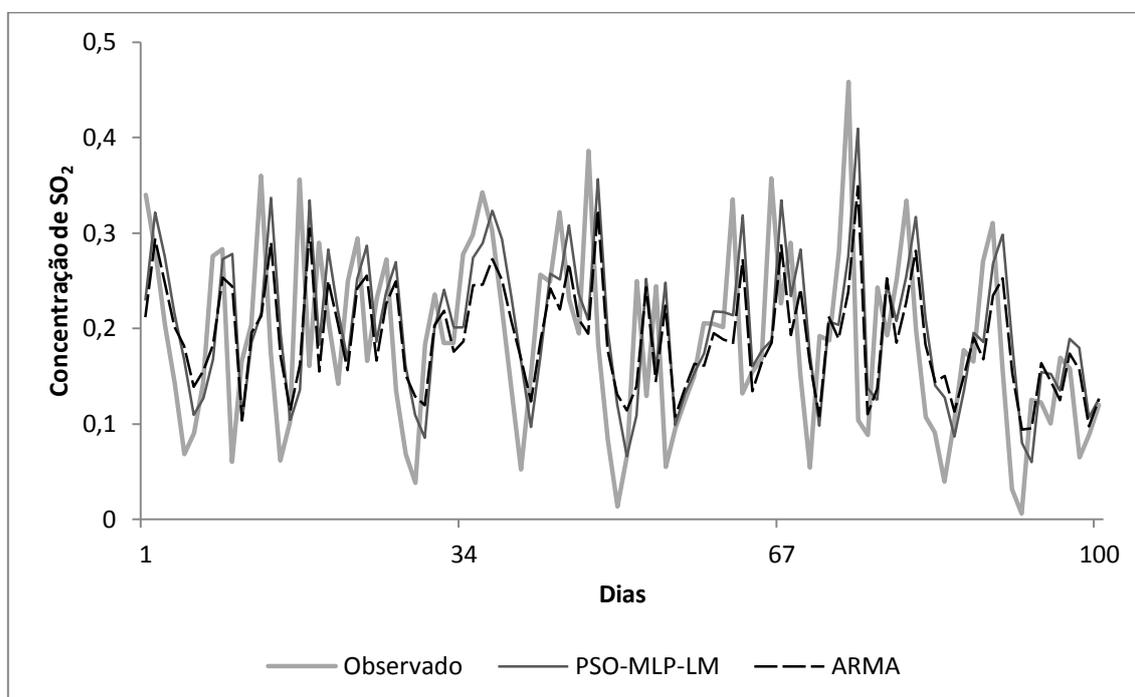


Figura 1.11 - Comparação entre os valores observados da série de SO_2 e as previsões dos modelos PSO-MLP-LM e ARMA.

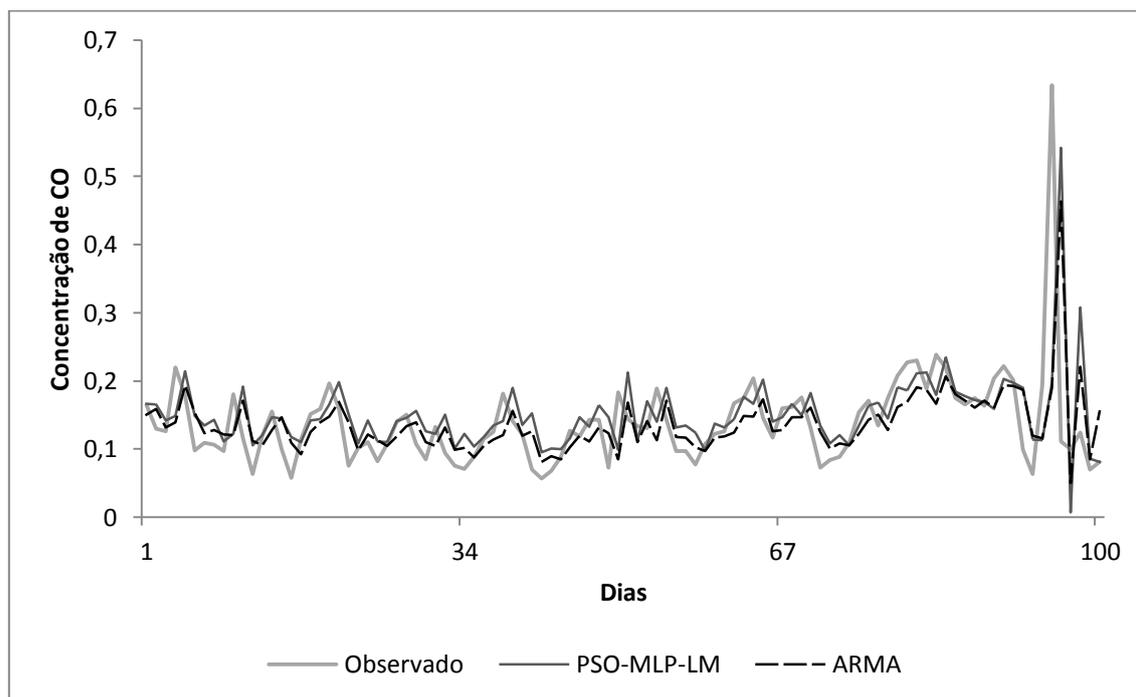


Figura 1.12 - Comparação entre os valores observados da série de CO e as previsões dos modelos PSO-MLP-LM e ARMA

1.3.5.4 Lógica nebulosa

O sistema nebuloso, que tem como variável de saída a qualidade do ar, é formulado a partir de quatro entradas referentes às concentrações dos poluentes gasosos coletadas. Após a computação destas concentrações (fuzzificação) através de um conjunto de funções de pertinência, um sistema de regras nebulosas é empregado para associar os níveis dos poluentes com a respectiva qualidade do ar. A defuzzificação, ou a transformação de valores linguísticos em valores numéricos, acontece na saída do sistema nebuloso resultando em uma saída quantitativa do processo. As etapas necessárias para a formulação do problema são:

- Formular a estrutura do sistema nebuloso:
 - Quantidade de entradas e saídas;
- Determinar as funções de pertinência:
 - Quantidade de subdivisões;
 - O modo como as subdivisões serão seccionadas;
- Estabelecer as regras nebulosas:
 - Quantidade de regras;
 - Qual a relação entre as variáveis de entradas e as variáveis de saída.

As variáveis de entrada do sistema nebuloso escolhidas são, neste caso, os níveis de concentrações de quatro poluentes. Têm-se, então, quatro variáveis de entrada representadas pelas concentrações dos poluentes:

- Monóxido de carbono;
- Dióxido de enxofre;
- Dióxido de nitrogênio;
- Material particulado.

Como variável de saída foi escolhida a qualidade do ar. Sendo esta responsável por indicar qual a degradação/qualidade que se encontra o ar, em resposta aos níveis de concentração dos poluentes atmosféricos.

Segundo a metodologia aplicada por (Onkal-Engin, 2004), as funções de pertinência de cada poluente foram balizadas nos níveis de concentrações referentes aos mesmos. Nesta dissertação são utilizadas somente as funções trapezoidais e triangulares (Figura 1.13). Na Tabela 1.6 são exemplificados valores que auxiliam na formulação das seguintes equações de funções de pertinência (Eq. 21):

Tabela 1.6 – Divisão dos níveis de gases – $\mu\text{g}/\text{m}^3$.

Níveis de critério		Qualidade do ar	SO ₂	MP ₁₀	CO _(56PM)	NO ₂
D_A	Bom	50	15	15	4,5	10
D_B	Moderado	100	30	30	9,5	20
D_C	Insalubre para pessoas sensíveis	150	45	45	12,5	30
D_D	Insalubre	200	60	60	15,5	40
D_E	Muito insalubre	300	90	90	30,5	60
D_F	Perigoso	500	150	150	50,4	100

$$\mu_A = \begin{cases} 1, & 0 \leq x \leq D_A \\ \frac{D_B - x}{D_B - D_A}, & D_A < x < D_B \\ 0, & x \geq D_B \end{cases}$$

$$\mu_B = \begin{cases} 0, & x \leq D_A \text{ ou } x \geq D_C \\ \frac{x - D_A}{D_B - D_A}, & D_A < x < D_B \\ 1, & x = D_B \\ \frac{D_C - x}{D_C - D_B}, & D_B < x < D_C \end{cases} \quad (\text{Eq. 21})$$

$$\mu_C = \begin{cases} 0, & x \leq D_B \text{ ou } x \geq D_D \\ \frac{x - D_B}{D_C - D_B}, & D_B < x < D_C \\ 1, & x = D_C \\ \frac{D_D - x}{D_D - D_C}, & D_C < x < D_D \end{cases}$$

$$\mu_D = \begin{cases} 0, & x \leq D_C \text{ ou } x \geq D_E \\ \frac{x - D_C}{D_D - D_C}, & D_C < x < D_D \\ 1, & x = D_D \\ \frac{D_E - x}{D_E - D_D}, & D_D < x < D_E \end{cases}$$

$$\mu_E = \begin{cases} 0, & x \leq D_D \text{ ou } x \geq D_F \\ \frac{x - D_D}{D_E - D_D}, & D_D < x < D_E \\ 1, & x = D_E \\ \frac{D_F - x}{D_F - D_E}, & D_E < x < D_F \end{cases}$$

$$\mu_F = \begin{cases} 0, & x \leq D_E \\ \frac{x - D_E}{D_F - D_E}, & D_E < x < D_F \\ 1, & x \geq D_F \end{cases}$$

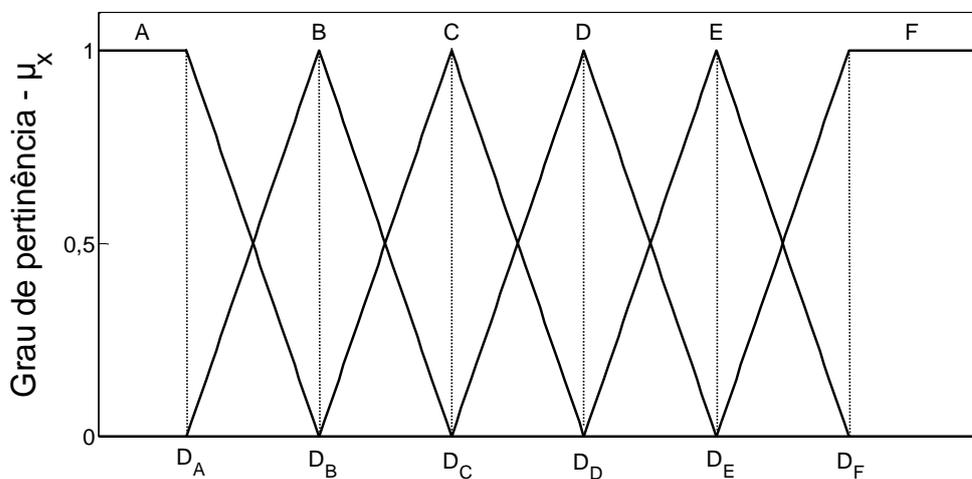


Figura 1.13 – Funções de pertinência das variáveis linguísticas que compõem o sistema nebuloso de qualidade do ar.

Regras nebulosas

Para cada poluente (CO, NO₂, SO₂ e MP₁₀) existe uma variável de entrada respectiva. Cada variável possui uma influência específica (peso) quando relacionada à variável de saída

(Qualidade do ar). Alguns exemplos de regras nebulosas são apresentados a seguir para o monóxido de carbono (CO) e tais regras são similares para os demais poluentes.

- Regras nebulosas para o CO:
 - SE Concentração de CO está no nível A ENTÃO Qualidade do ar É boa;
 - SE Concentração de CO está no nível B ENTÃO Qualidade do ar É moderada;
 - SE Concentração de CO está no nível C ENTÃO Qualidade do ar É insalubre para pessoas sensíveis;
 - SE Concentração de CO está no nível D ENTÃO Qualidade do ar É insalubre;
 - SE Concentração de CO está no nível E ENTÃO Qualidade do ar É muito insalubre;
 - SE Concentração de CO está no nível F ENTÃO Qualidade do ar É perigosa.
- Cada conjunto de regras possui um peso associado referente à influência de determinado poluente possui sobre a qualidade do ar. Tais pesos são:
 - CO – 0,35;
 - MP₁₀ – 0,35;
 - NO₂ – 0,1;
 - SO₂ – 0,2;

O segundo sistema nebuloso baseia-se na metodologia aplicada por Khan (2005). Tal sistema mensura as possíveis consequências de um determinado evento, e consiste das seguintes entradas:

- Densidade demográfica do local;
- Percentual de pessoas com sensibilidade maior aos poluentes no local;
- Sensibilidade do local – proximidade com estabelecimentos que englobam um número extenso de pessoas, tais como:
 - Escolas;
 - Hospitais;
 - Centros comerciais.

A variável de saída é a quantificação das consequências que as pessoas podem vir a sofrer, caso exista um nível baixo da qualidade do ar. A Tabela 1.7 detalha os parâmetros das consequências. As funções de pertinência das variáveis de entrada e saída são exemplificadas a seguir.

Tabela 1.7 – Parâmetros de consequências.

Parâmetros de consequências	Escala
Sensibilidade do local (SL)	
Sem estabelecimentos básicos	0
Estabelecimentos de baixo valor	2
Indústrias e fábricas	4
Residenciais, parques	6
Escolas, igrejas, patrimônios públicos	8
Hospitais e locais sensíveis	10
Densidade demográfica do local (DL)	
<20	1
20-49	2
50-99	3
100-300	4
>300	5
Percentual de pessoas com sensibilidade (PS)	
<10	1
10-25	2
26<40	3
41<50	4
>50	5

As funções de pertinência para os parâmetros de consequências (variáveis de entrada do sistema) são apresentadas abaixo:

- o Sensibilidade do local:

$$\begin{aligned}
 \mu_{Baixo} &= \begin{cases} 1, & x \leq 4 \\ \frac{x-4}{2}, & 4 < x < 6 \\ 0, & x \geq 6 \end{cases} \\
 \mu_{M\u00e9dio} &= \begin{cases} 0, & x \leq 4 \text{ ou } x \geq 8 \\ \frac{x-4}{2}, & 4 < x < 6 \\ \frac{8-x}{2}, & 6 < x < 8 \\ 1, & x = 6 \end{cases} \\
 \mu_{Alto} &= \begin{cases} 0, & x \leq 6 \\ \frac{8-x}{2}, & 6 < x < 8 \\ 1, & x \geq 8 \end{cases}
 \end{aligned} \tag{Eq. 22}$$

- Densidade demográfica:

$$\begin{aligned} \mu_{Baixo} &= \begin{cases} 1, & x \leq 2 \\ \frac{x-2}{1}, & 2 < x < 3 \\ 0, & x \geq 3 \end{cases} \\ \mu_{M\u00e9dio} &= \begin{cases} 0, & x \leq 2 \text{ ou } x \geq 4 \\ \frac{x-2}{1}, & 2 < x < 3 \\ \frac{4-x}{1}, & 3 < x < 4 \\ 1, & x = 3 \end{cases} \\ \mu_{Alto} &= \begin{cases} 0, & x \leq 3 \\ \frac{4-x}{1}, & 3 < x < 4 \\ 1, & x \geq 4 \end{cases} \end{aligned} \quad (\text{Eq. 23})$$

- Percentual de pessoas sens\u00edveis:

$$\begin{aligned} \mu_{Baixo} &= \begin{cases} 1, & x \leq 2 \\ \frac{x-2}{1}, & 2 < x < 3 \\ 0, & x \geq 3 \end{cases} \\ \mu_{M\u00e9dio} &= \begin{cases} 0, & x \leq 2 \text{ ou } x \geq 4 \\ \frac{x-2}{1}, & 2 < x < 3 \\ \frac{4-x}{1}, & 3 < x < 4 \\ 1, & x = 3 \end{cases} \\ \mu_{Alto} &= \begin{cases} 0, & x \leq 3 \\ \frac{4-x}{1}, & 3 < x < 4 \\ 1, & x \geq 4 \end{cases} \end{aligned} \quad (\text{Eq. 24})$$

Os pesos de cada vari\u00e1vel lingu\u00edstica contido no segundo sistema nebuloso em rela\u00e7\u00e3o \u00e0 vari\u00e1vel de sa\u00edda (Consequ\u00eancias):

- SL = 0,13;
- DL = 0,21;
- PS = 0,66;

As regras que expressam o sistema nebuloso denominado consequ\u00eancias s\u00e3o similares umas as outras, e as regras para a vari\u00e1vel Sensibilidade Local s\u00e3o dispostas a seguir:

- Regras nebulosas para o SL:
 - SE SL for baixa ENT\u00c3O consequ\u00eancia \u00c9 baixa;
 - SE SL for m\u00e9dia ENT\u00c3O consequ\u00eancia \u00c9 m\u00e9dia;
 - SE SL for alta ENT\u00c3O consequ\u00eancia \u00c9 alta.

Funções de pertinência da variável de saída – consequência expressa em percentual:

$$\begin{aligned} \mu_{Baixo} &= \begin{cases} 1, & x \leq 25 \\ \frac{x-25}{25}, & 25 < x < 50 \\ 0, & x \geq 50 \end{cases} \\ \mu_{Médio} &= \begin{cases} 0, & x \leq 25 \text{ ou } x \geq 75 \\ \frac{x-25}{25}, & 25 < x < 50 \\ \frac{75-x}{25}, & 50 < x < 75 \\ 1, & x = 50 \end{cases} \\ \mu_{Alto} &= \begin{cases} 0, & x \leq 50 \\ \frac{75-x}{25}, & 50 < x < 75 \\ 1, & x \geq 75 \end{cases} \end{aligned} \quad (\text{Eq. 25})$$

O terceiro e último sistema nebuloso é o de riscos e é formado pela combinação dos dois sistemas nebulosos previamente descritos: o sistema nebuloso de qualidade do ar e o sistema nebuloso de consequências. Portanto, o sistema resultante pode ser descrito como sendo:

$$\text{Risco} = \text{Potencial} * \text{Consequências}.$$

O risco pode ser apresentado na forma de matriz, como pode ser visualizado na Figura 1.14:

		Riscos		
		Consequências		
		Baixo	Médio	Alto
Potencial	Baixo	<i>Muito Baixo</i>	<i>Baixo</i>	<i>Médio</i>
	Médio	<i>Baixo</i>	<i>Médio</i>	<i>Alto</i>
	Alto	<i>Médio</i>	<i>Alto</i>	<i>Muito Alto</i>

Figura 1.14 – Matriz de riscos.

As funções de pertinência para as variáveis de entrada são:

- As variáveis Consequências e Riscos possuem funções de pertinência similares as apresentadas previamente na Eq. 25;
- A variável Potencial possui uma interpretação um pouco distinta da apresentada anteriormente. Tal diferença é exemplificada a seguir:

Funções de pertinência da variável Potencial:

$$\mu_{Baixo} = \begin{cases} 1, & x \leq 100 \\ \frac{x-100}{100}, & 100 < x < 200 \\ 0, & x \geq 200 \end{cases}$$

$$\mu_{M\u00e9dio} = \begin{cases} 0, & x \leq 100 \text{ ou } x \geq 300 \\ \frac{x-100}{100}, & 100 < x < 200 \\ \frac{300-x}{100}, & 200 < x < 300 \\ 1, & x = 200 \end{cases} \quad (\text{Eq. 26})$$

$$\mu_{Alto} = \begin{cases} 0, & x \leq 200 \\ \frac{300-x}{100}, & 200 < x < 300 \\ 1, & x \geq 300 \end{cases}$$

As regras para o sistema nebuloso de Risco s\u00e3o:

- o SE Potencial for baixo E Consequ\u00eancia for baixa ENT\u00c3O Risco \u00c9 muito baixo;
- o SE Potencial for baixo E Consequ\u00eancia for m\u00e9dia ENT\u00c3O Risco \u00c9 baixo;
- o SE Potencial for baixo E Consequ\u00eancia for alta ENT\u00c3O Risco \u00c9 m\u00e9dio;
- o SE Potencial for m\u00e9dio E Consequ\u00eancia for baixa ENT\u00c3O Risco \u00c9 baixo;
- o SE Potencial for m\u00e9dio E Consequ\u00eancia for m\u00e9dia ENT\u00c3O Risco \u00c9 m\u00e9dio;
- o SE Potencial for m\u00e9dio E Consequ\u00eancia for alta ENT\u00c3O Risco \u00c9 alto;
- o SE Potencial for alto E Consequ\u00eancia for baixa ENT\u00c3O Risco \u00c9 m\u00e9dio;
- o SE Potencial for alto E Consequ\u00eancia for m\u00e9dia ENT\u00c3O Risco \u00c9 alto;
- o SE Potencial for alto E Consequ\u00eancia for alta ENT\u00c3O Risco \u00c9 muito alto.

Os pesos das regras supracitadas pertencentes ao sistema de riscos possuem pesos iguais a 1 (um), uma vez que n\u00e3o foi poss\u00edvel estabelecer a influ\u00eancia entre as vari\u00e1veis de entrada (potencial e consequ\u00eancias) em rela\u00e7\u00e3o \u00e0 vari\u00e1vel de sa\u00edda (risco).

Resultados

Os resultados das aplica\u00e7\u00f5es dos modelos de l\u00f3gica nebulosa s\u00e3o superf\u00edcies de resposta que demonstram como as vari\u00e1veis se relacionam. As superf\u00edcies s\u00e3o dispostas de acordo com o sistema ao qual elas pertencem. No primeiro sistema, apresentado na Figura 1.15, tem-se quatro vari\u00e1veis de entrada (CO, MP₁₀, SO₂ e NO₂) e uma de sa\u00edda (Qualidade do Ar). A figura exibe as seis poss\u00edveis combina\u00e7\u00f5es entre as vari\u00e1veis e como as vari\u00e1veis de entrada influenciam a vari\u00e1vel resposta.

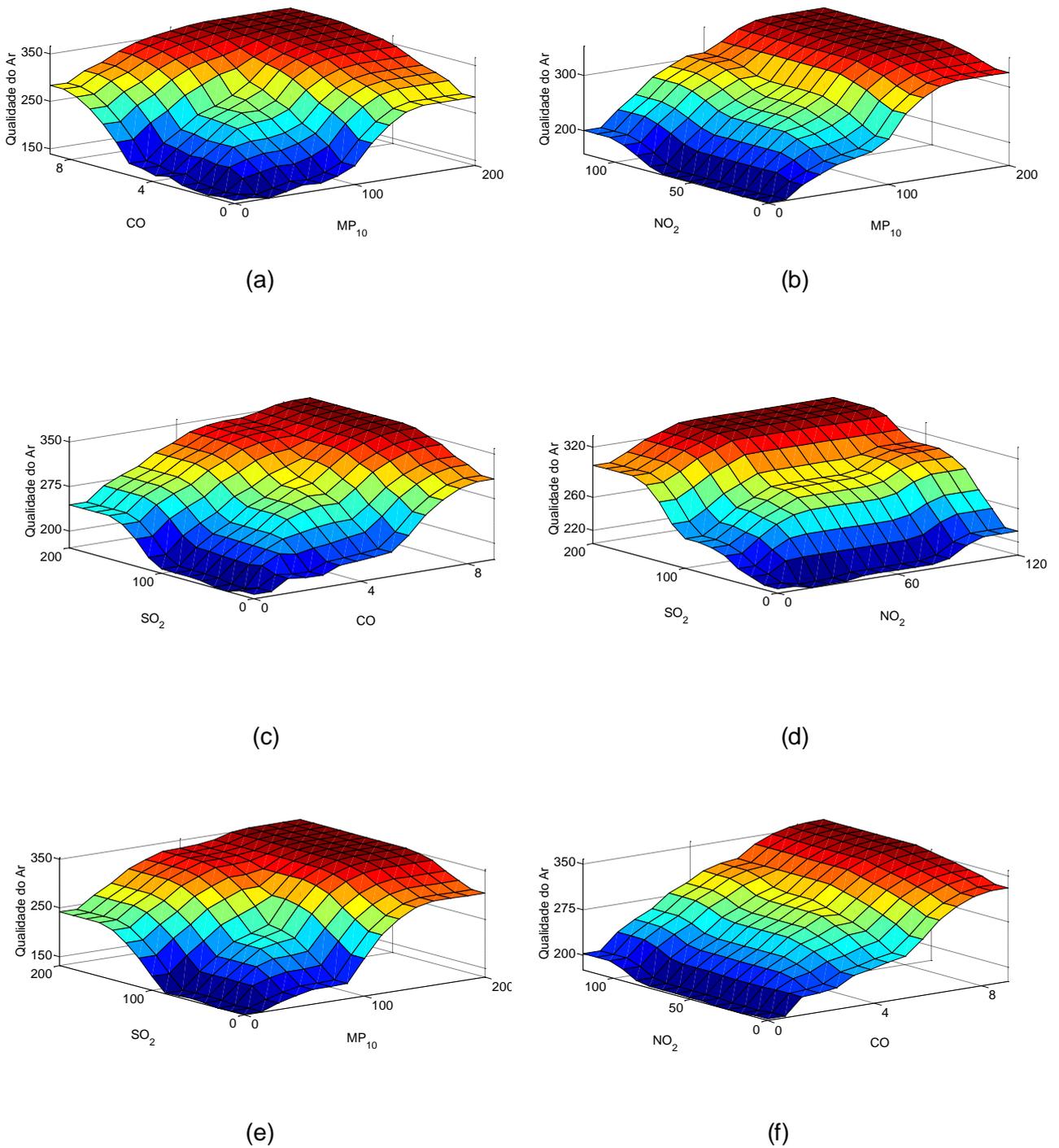


Figura 1.15 – Superfícies de resposta em relação a Qualidade do Ar das seis possíveis combinações: CO x MP₁₀ (a), NO₂ x MP₁₀ (b), SO₂ x CO (c), SO₂ x NO₂ (d), SO₂ x MP₁₀ (e), NO₂ x CO (f).

Na Figura 1.15 são apresentadas as superfícies de respostas para a qualidade do ar, e nota-se que cada poluente influencia a qualidade do ar de forma distinta. Os poluentes mais

influentes são aqueles que possuem maior peso, e, portanto, são identificados como os poluentes CO e MP₁₀. Outro aspecto importante é a concepção de sinergia que existe entre os poluentes, sendo observado que suas concentrações influenciam mais a variável resposta em conjunto do que de modo singular.

No segundo sistema nebuloso as consequências são mensuradas através de três pontos: a densidade demográfica do local (DL), a sensibilidade do local (SL) e o percentual de pessoas sensíveis a poluição presentes no local (PS). São apresentadas, na Figura 1.16, as superfícies de respostas destas variáveis visando mensurar as consequências para determinados local e grupo de pessoas.

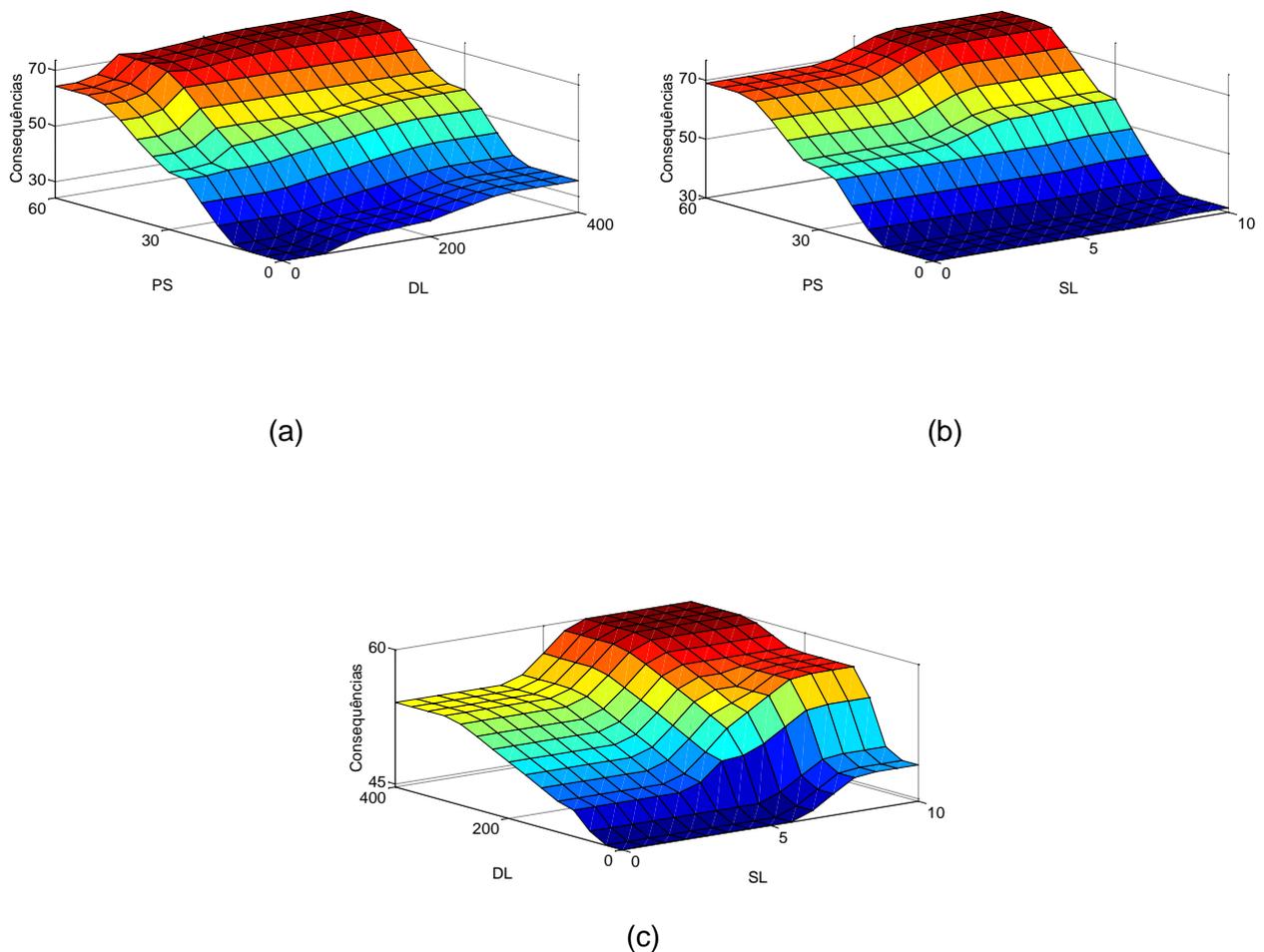


Figura 1.16 – Superfícies de resposta acerca das possíveis Consequências. Pessoas sensíveis x Densidade demográfica (a), Pessoas sensíveis x Sensibilidade do local (b), Densidade demográfica x Sensibilidade do local (c).

Na Figura 1.16 é perceptível que a variável “Pessoas Sensíveis” possui uma maior influência sobre as consequências, ou seja, quanto maior for o percentual de pessoas sensíveis em determinado local maior serão as consequências. Contudo, as três variáveis contribuem de forma conjunta na obtenção das consequências.

No terceiro sistema nebuloso, exibido na Figura 1.17, são empregados os dois sistemas anteriores (Qualidade do ar e as Consequências), como variáveis de entrada, para quantificar os riscos aos quais as pessoas estariam expostas. Esta forma de aplicação da análise de riscos permite que sejam explorados determinados pontos que enfatizam a contribuição de múltiplos poluentes à degradação da qualidade do ar, e as consequências advindas desta depleção. Permitindo, desta forma, caracterizar de modo mais conciso o risco ao qual a população estaria exposta, pois, neste trabalho um conjunto de quatro poluentes foi levado em consideração.

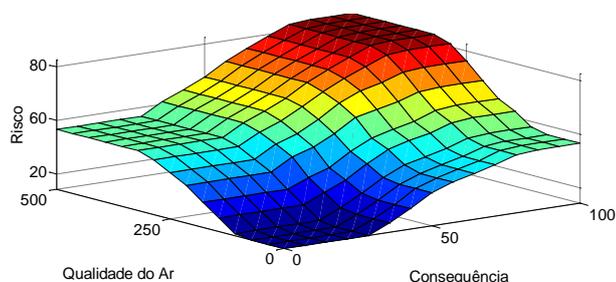


Figura 1.17 – Superfícies de resposta da variável Risco.

1.4 REFERÊNCIAS BIBLIOGRÁFICAS

Aguilera, P. Bayesian networks in environmental modelling (in press). **Environmental Modelling & Software**, p. 1-13, 2011.

Amini, S. A Conceptual Model Based on the Fuzzy Set Theory to Measure and Evaluate the Performance of Service Processes. **15th IEEE International Enterprise Distributed Object Computing Conference Workshops**, p. 122-131, 2011.

Anttila, P. Trends of primary and secondary pollutant concentrations in Finland in 1994–2007. **Atmospheric Environment**, p. 30-41, 2010.

Basheer, I. A. Artificial neural networks: fundamentals, computing, design, and application. **Journal of microbiological methods**, p. 3-31, 2000.

Box, G. E., & Jenkins, G. M. Time Series Analysis: Forecast and Control. New Jersey: **Prentice-Hall**, 1994.

Brunelli, U. Two-days ahead prediction of daily maximum concentrations of SO₂, O₃, PM₁₀, NO₂, CO in the urban area of Palermo, Italy. **Atmospheric Environment**, v. 41, p. 2967–2995, 2007.

Cairncross, E. K. A novel air pollution index based on the relative risk of daily mortality associated with short-term exposure to common air pollutants. **Atmospheric environment**, p. 8442-8454, 2007.

Cassell, E. J. Air pollution - the relationship between health effects and control philosophy. **IEEE Transactions on geoscience electronics**, p. 220-226, 1969.

CETESB-SP. (s.d.). Companhia Ambiental do Estado de São Paulo. Acesso em 03 de Dezembro de 2011, disponível em CETESB-SP: <http://www.cetesb.sp.gov.br/>

Chan, C. W., & Huang, G. H. Artificial intelligence for management and control of pollution minimization and mitigation process. **Engineering Applications of Artificial Intelligence**, v. 16, p. 75-90, 2003.

Chen, S. H. Artificial Intelligence techniques: An introduction to their use for modelling environmental systems. **Mathematics and Computers in Simulation**, v. 78, p. 379–400, 2008.

Chittock, D. G. (). A review of international practice in the design of voluntary Pollution Prevention Programs. **Journal of Cleaner Production**, v. 19, p. 542-551, 2011.

Coppin, B. Inteligência artificial. Rio de Janeiro: **LTC**, 2010.

Covello, V. T. Risk assessment methods - Approaches for assessing health and environmental risks. New York: **Plenum Press**, 1993.

Curtis, L. Adverse health effects of outdoor air pollution. **Environment international**, p. 815-830, 2006.

Darbra, R. M. How to measure uncertainties in environmental risk assessment. **Trends in Analytical Chemistry**, v. 27 (4), p. 377-385, 2008.

Diekmann, J. E. Risk analysis: lessons from artificial intelligence. **International Journal of Project Management**, v. 10 (2), p. 75-80, 1992.

Dunjó, J. Hazard and operability (HAZOP) analysis. A literature review. **Journal of Hazardous Materials**, v.173, p. 19-32, 2010.

Eberhart, R. Particle Swarm Optimization. **IEEE Proceedings – Neural Networks**, v.4, p. 1942-1948, 1995.

Fanb, H., & Lu, W. Z. Application of evolutionary neural network method in predicting pollutant levels in downtown area of Hong Kong. **Neurocomputing**, v. 51, p. 387-400, 2003.

Forrest, S. Architecture for an Artificial Immune System. **Evolutionary Computation**, v. 7 (1), p. 45-68, 1999.

Frohwein, H. Alternative measures of risk of extreme events in decision trees. **Reliability Engineering and System Safety**, v. 66, p. 69-84, 1999.

Gardner, M. W. Artificial neural networks (the multilayer perceptron) - A review of applications in the atmospheric sciences. **Atmospheric Environmet**, p. 2627-2636, 1998.

Giannouli, M. Impact of European emission control strategies on urban and local air quality. **Atmospheric environment**, p. 1-10, 2010.

Godish, T. Air quality. Florida: **Lewis publishers**, 2003.

Goren, H. G. A review of applications of genetic algorithms in lot sizing. **Journal of Intelligent Manufacturing**, v. 21, p. 575-590, 2010.

Gujarati, D. N. Econometria básica. São Paulo: **Pearson Education do Brasil**, 2005.

Gumrah, F. The application of artificial neural networks for the prediction of water quality of polluted aquifer. **Water Air and Soil Pollution**, p. 275-294, 2000.

Guoqiang Zhang, B. Forecasting with artificial neural networks: The state of the art. **International Journal of Forecasting**, 14, pp. 35-62, 1998.

Haimes, Y. Y. Risk modeling, assessment and management, 3rd ed. Charlottesville: **Wiley**, 2009.

Haykin, S. Redes Neurais: Princípios e Prática. Porto Alegre: **Bookman**, 2001.

Heidenberger, K. Dynamic project selection and funding under risk: A decision tree based MILP approach. **European Journal of Operational Research**, v. 95, p. 284-298, 1996.

Hinrichs, R. A. Energia e meio ambiente. São Paulo: **Cengage Learning**, 2010.

Hollauer, G. Prevendo o crescimento da produção industrial usando um número limitado de combinações de previsões. **Econ. aplic.**, v. 12 (2), p. 177-198, 2008.

Holmes, N. S. A review of dispersion modelling and its application to the dispersion of particles: An overview of different dispersion models available. **Atmospheric Environment**, p. 5902–5928, 2006.

Ince, T. Evaluation of global and local training techniques over feed-forward neural network architecture spaces for computer-aided medical diagnosis. **Expert Systems with Applications**, v. 37, p. 8450-8461, 2010.

Jahn, H. J. Particulate matter pollution in the megacities of the Pearl River Delta, China – A systematic literature review and health risk assessment. **International Journal of Hygiene and Environmental Health**, 2011.

Kampa, M. Human health effects of air pollution. **Environmental pollution**, p. 362-367, 2008.

Kassomenos, P., & Papaloukas, C. Assessment and prediction of short term hospital admissions: the case of Athens, Greece. **Atmospheric Environment**, v. 42 , p. 7078–7086, 2008.

Khan, F. I. Risk-based prioritization of air pollution monitoring using fuzzy synthetic evaluation technique. **Environmental Monitoring and Assessment**, v. 105, p. 261-283, 2005.

Khashei, M., & Bijari, M. A new class of hybrid models for time series forecasting. **Expert Systems with Applications**, 39, pp. 4344-4357, 2012.

Khashei, M., & Bijari, M.. An artificial neural network (p, d, q) model for timeseries forecasting. **Expert Systems with Applications**, v. 37, p. 479-489, 2010.

Kumar, A. Forecasting of daily air quality index in Delhi. **Science of the Total Environment**, p. 5517-5523, 2011.

Kumar, U. GARCH modelling in association with FFT-ARIMA to forecast ozone episodes. **Atmospheric Environment**, p. 4252-4265, 2010.

Kurt, A. An online air pollution forecasting system using neural networks. **Environment International**, p. 592-598, 2008.

Lahr, J. Environmental risk mapping of pollutants: State of the art and communication aspects. **Science of the Total Environment** , v. 408, p. 3899-3907, 2010.

Li, J. An integrated fuzzy-stochastic modeling approach for risk assessment of groundwater contamination. **Journal of environmental management**, p. 173-188, 2007.

Llop, E. The use of lichen functional groups as indicators of air quality in a Mediterranean urban environment. **Ecological Indicators**, v. 13, p. 215-221, 2012.

Lu, W. Z., & Wang, D. Forecasting of ozone level in time series using MLP model with a novel hybrid training algorithm. **Atmospheric Environment**, v. 40, p. 913–924, 2006.

Maier, H. R. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. **Environmental Modelling & Software**, v. 25, p. 891-909, 2010.

Mamdani, E. H. Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis. **IEEE Transactions on Computers**, C-26 (12), p. 1182-1191, 1977.

Maniezzo, V. An ANTS heuristic for the frequency assignment problem. **Future Generation Computer Systems**, v. 16, p. 927–935, 2000.

- Markowski, A. S. Fuzzy risk matrix. **Journal of Hazardous Materials**, v. 159, p. 152-157, 2008.
- Marquardt, D. W. An algorithm for least-squares estimation of nonlinear parameters. **Journal Society Industrial Applied Mathematics**, v. 11 (2), p. 431-441, 1963.
- Monks, P. Atmospheric composition change – global and regional air quality. **Atmospheric environment**, p. 5268-5350, 2009.
- Moreira, D. M. Meteorologia e poluição atmosférica. **Ambiente e sociedade - Campinas**, p. 1-13, 2008.
- Mukherjee, I. Comparing the performance of neural networks developed by using Levenberg–Marquardt and Quasi-Newton with the gradient descent algorithm for modelling a multiple response grinding process. **Expert Systems with Applications**, v. 39, p. 2397-2407, 2012.
- Mukhopadhyay, K. An empirical investigation of air pollution from fossil fuel combustion and its impact on health in India during 1973–1974 to 1996–1997. **Ecological Economics**, v. 55, p. 235-250, 2005.
- Ni, H. Some extensions on risk matrix approach. **Safety Science**, v. 48, p. 1269–1278, 2010.
- Onkal-Engin, G. Assessment of urban air quality in Istanbul using fuzzy synthetic evaluation. **Atmospheric Environment**, v. 48, p. 3809–3815, 2004.
- Osowski, S. Forecasting of the daily meteorological pollution using wavelets and support vector machine. **Engineering Applications of artificial intelligence**, p. 745-755, 2007.
- Park, C. Oxford - Dictionary of environment and conservation. New York: **Oxford university press**, 2008.
- Pelliccioni, A. Air dispersion model and neural network: A new perspective for integrated models in the simulation of complex situations. **Environmental Modelling & Software**, v. 21, p. 539-546, 2006.
- Prechelt, L. Proben1: A set of neural network benchmark problems and benchmarking rules. Technical Report 21/94, p. 1-38, 1994.
- Qin, W. Study of Software Requirement Risk Based on Fuzzy Influence Diagram. **Energy Procedia**, v. 11, p. 1540-1550, 2011.
- Ross, T. J. Fuzzy Logic with Engineering Applications 2nd Ed. Hoboken: **John Wiley & Sons**, 2004.
- Rumelhart, D. E. Learning representations by back-propagating errors. **Nature**, v. 323 (9), p. 533-536, 1986.
- Şahin, Ü. A. Application of cellular neural network (CNN) to the prediction of missing air pollutant data. **Atmospheric Research**, v. 101, p. 314–326, 2011.
- Sánchez, L. E. Avaliação de impacto ambiental - conceitos e métodos. São Paulo: **Oficina de textos**, 2008.

Santos, U. P. Poluição, aquecimento global e repercussões na saúde. **Revista da Associação Médica Brasileira**, v. 53 (3), p. 193-194, 2007.

Shumway, R. H. Time series analysis and Its applications, 3rd Ed. New York: **Springer**, 2011.

Sowlat, M. H. A novel, fuzzy-based air quality index (FAQI) for air quality assessment. **Atmospheric Environment**, v. 45, p. 2050-2059, 2011.

Stacey, C. E. Donora (Images of America). San Francisco: **Arcadia Publishing**, 2010.

Sugeno, M. Stability analysis and design of fuzzy control systems. **Fuzzy Sets and Systems**, v. 45, p. 135-156, 1992.

Takens, F. Detecting strange attractor in turbulence. **Dynamical Systems and Turbulence**, v. 898, pp. 366-381, 1980.

UNEP. Barcelona Convention 1995. Acesso em 02 de Dezembro de 2011, disponível em United Nations Environment Programme: http://www.unep.ch/regionalseas/regions/med/t_barcel.htm

US-EPA. (s.d.). *Environmental Protection Agency*. Acesso em 03 de Dezembro de 2011, disponível em United States Environmental Protection Agency: www.usepa.gov

Valenzuela, O. Hybridization of intelligent techniques and ARIMA models for time series prediction. **Fuzzy Sets and Systems**, p. 821-845, 2008.

Vallero, D. Fundamentals of air pollution. Londres: **Academic press**, 2008.

Vlachogianni, A. Evaluation of a multiple regression model for forecasting of the concentrations of NO_x and PM₁₀ in Athens and Helsinki. **Science of the total environment**, p. 1559-1571, 2011.

Wang, W., & Lu, W. Z. Prediction of maximum daily ozone level using combined neural network and statistical characteristics. **Environment International**, v. 29, p. 555– 562, 2003.

Wang, X. Evaluating impacts of air pollution in China on public health: Implications for future air pollution and energy policies. **Atmospheric environment**, p. 1706-1721, 2006.

WHO. (s.d.). *World Health Organization*. Acesso em 03 de Dezembro de 2011, disponível em World Health Organization: www.who.int/en

Wilson, A. M. Air pollution and the demand for hospital services: a review. **Environment international**, p. 1109-1118, 2004.

Wilson, J. G., & Kingham, S. A review of intraurban variations in particulate air pollution: Implications for epidemiological research. **Atmospheric Environment**, v. 39, p. 6444–6462, 2005.

Zadeh, L. A. Fuzzy logic. **Computer**, v. 21 (4), p. 83-93, 1988.

Zadeh, L. A. Fuzzy sets. **Information and control**, v. 8, p. 338-353, 1965.

Zemouri, R. Defining and applying prediction performance metrics on a recurrent NARX time series model. **Neurocomputing**, v. 73, p. 2506–2521, 2010.

Zhang, G. P. Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing**, v. 50, p. 159-175, 2003.

CAPÍTULO II

TIME SERIES FORECASTING OF POLLUTANT CONCENTRATION LEVELS USING PARTICLE SWARM OPTIMIZATION AND ARTIFICIAL NEURAL NETWORKS

Manuscrito

Time series forecasting of pollutant concentration levels using particle swarm optimization and artificial neural networks

F. S. de Albuquerque Filho¹, P. S. G. de Mattos Neto², F. Madeiro¹ and S. M. M. Fernandes¹

¹– Universidade Católica de Pernambuco – UNICAP
{francisco.saf, smmfast}@gmail.com
madeiro@dei.unicap.br

²– Universidade Federal de Pernambuco – UFPE
psgmn@cin.ufpe.br

Abstract – In this paper, the application of an intelligent hybrid system for the time series forecasting problem of atmospheric pollutants concentration levels is evaluated. The proposed method consists of an Artificial Neural Network (ANN) combined with a Particle Swarm Optimization (PSO) algorithm. The used method searches not only the relevant time lags for the correct characterization of the time series, but also the best neural network architecture. An experimental analysis is performed using four real time series and the results are shown in terms of six performance measures. For comparison purposes statistical methods applied to time series prediction were also used. Experimental results demonstrate that the applied methodology achieves a fair prediction of the presented pollutant time series.

Keywords – Particle Swarm Optimization, Artificial Neural Network, pollutants concentration, time series.

1 Introduction

Pollution is one of the most relevant problems of the metropolitan regions. With the population growth and the economical increase, environmental health problems have awakened the society interest. Problems that have affected the ecosystem, such as, atmospheric pollution, noise pollution, garbage and the disposal of it and, in particular, the air pollution, have direct effect on people's health [1].

There are numerous indicators that show the pollution effects on people's health [2]. Some of the most important are: particle matter (MP₁₀), carbon monoxide (CO), sulfur dioxide (SO₂), and nitrogen dioxide (NO₂). When the indicators' concentration level exceeds the threshold of the established air quality, severe problems may affect the human health [3]. There are many national environmental agencies around the globe that make their own policies and have established standards and indicators regarding the air quality and the allowed levels for atmospheric pollutants. Environmental agencies observe the indicators as a monitoring measure, using a network of pollution and atmospheric sensors. The measurement results are observations equally spaced and ordered in time (e.g. hourly, daily, monthly), resulting in a time series of pollutant concentration, for at least the pollutants mentioned above.

Methods used for time series prediction are native of the Statistics field, such as the AR (autoregressive) model and the ARMA (autoregressive moving average) model [9-10] [15] [24]. However, intelligent models have been proposed to deal with such problem, such as: artificial neural networks (ANN) [1] [3] [5] [6] [9] [14], particle swarm optimization (PSO) [5] [6] [13] [25], genetic algorithms [20] [21] and fuzzy logic [12] [17].

In the forecasting field, in particular for the air pollution problem [1] [3] [11] [14] [20] [21], neural networks have been applied with success. However, to use such a technique, the choice of an appropriate architecture for the problem solution is fundamental. Therefore, intelligent hybrid models [1] [5] [6] [9] [16], that are the combination of different methods, are proposed with the purpose of achieving better prediction performance, with the intent of aggregating the strong points of several algorithms in a single system.

The focus of this article is to employ a PSO for tuning and training an ANN. The resulting model is applied to the time series forecasting problem. To the authors' knowledge, in the literature, PSO has not been an extensively used algorithm to predict time series of gaseous pollutants.

The paper is organized as follows: firstly, a brief introduction about time series, neural networks, and particle swarm optimization (sections 2, 3, and 4, respectively) is presented. Later, in section 5, the performance measures used in this work are described, followed by section 6, which describes the models used in this article. Section 7 presents the simulation results for four time series: carbon monoxide (CO) and sulfur dioxide (SO₂), particle matter (MP₁₀), and nitrogen dioxide (NO₂). Finally, section 8 presents the conclusions of this paper.

2 Time series

A time series is a set of observations of a variable of interest, representing a sequence of observations ordered in time. The variable is observed in discrete temporal points, usually equally spaced, and the analysis of such temporal behavior involves the description of the process or the phenomenon that generated the sequence [15] [24]. The time series can be defined by:

$$X_t = X_t \in \mathbb{R} \mid t = 1, 2, 3 \dots N,$$

where t is the temporal index and N is the number of observations. Therefore, X_t is the sequence of observations ordered in time and equally spaced. The application of forecasting techniques rely on the ability of identifying underlying regular patterns in the data set to make it possible to create a model capable of generating the next temporal patterns [6].

A crucial point regarding the time series prediction problem is to capture the temporal relationship between the given series data. Hence, through the observation of the correct time lags, it should be possible to reconstruct the dynamics that generated the real series. This hypothesis was proposed and theoretically justified by Takens [23], which stated that such reconstruction could be possible. However, one of the major problems is the correct choice of the relevant time lags. Therefore, the methodology used in this paper is based on the search process of the relevant time lags for the correct characterization of the process that governs the temporal series [5] [6].

3 Artificial Neural Networks (ANN)

The artificial neural networks (ANN) are computational modeling tools that find a wide variety of applications in modeling complex real world problems [7] [8] [13]. ANN can be defined as structures composed of simple, adaptive, and massively interconnected processing elements (referred to as artificial neurons or nodes) that are capable of accomplishing data processing and knowledge representation.

A single neuron [Figure 1], or perceptron, has two main components. First, a weighted sum $s = \sum_{i=1}^n w_i x_i + b$ that achieves a weighted summation of the inputs ($x_1, x_2, x_3, \dots, x_n$), where w_i is the respective weight (synaptic weight) of the input x_i and b is the bias of the neuron. Second, an activation function which can be a linear, nonlinear or logic. This function is responsible for given an output accordingly to the result of s . A range of functions can assume the role of an activation function, including sigmoid, linear and, hyperbolic tangent. Usually, every neuron in a specific layer has the same activation function.

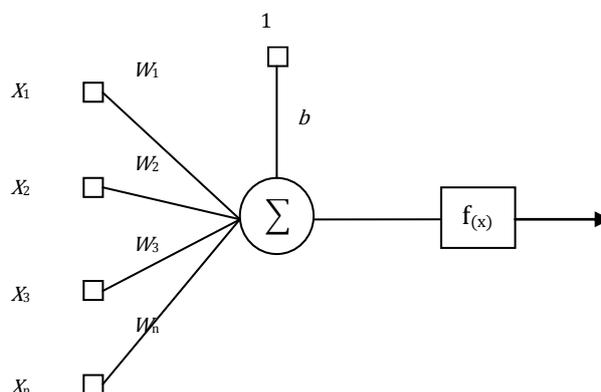


Figure 1 – A single model of a neuron.

A multilayer neural network or multilayer perceptron network (MLP), consisting of two or more layers, as follows: input layer, hidden layer(s) and output layer [Figure 2]. When formed by two layers, it only possesses the input layers and the output layer. Each node in the input layer receives a unique input signal. Actually, the nodes in this layer usually are passive, meaning they do not process data, but only receive signals on their input and passes them to the nodes in the next layer [7] [8]. When the network is formed by three or more layers, the layer(s) between the input and output layers is (are) denominated hidden layer(s). A hidden layer processes the signals sent to it by calculating a weighted summation and using a specific activation function. Later on, the output layer receives the signal from the previous layer, performing a weighted summation and applying a particular activation function to the resultant output.

In this article, feed-forward neural networks are used, in which the sigmoid function ($f(x) = 1/(1 + \exp^{-kx})$) and the linear function are used as the activation functions of the hidden and output layers, respectively.

The used data set for training, validating, and testing the neural network is divided in three groups. Firstly, there is the training set and it corresponds usually to half or more of the whole data. It is used by the ANN to adjust its' weights and biases accordingly to the presented data. There are many algorithms that are used for training an ANN. The most used ones are the back-propagation and the Levenberg-Marquardt algorithms. The training process is obtained through the update of the synaptic weights and biases within the network, and it depends on the algorithm used. The validation set, a part of the data set yet to be presented, is used for validating the network training. It checks the network's capability to generalize a series of input data. The number of validation errors should be less than a stipulated limit, if it is higher, the training is stopped and the previous best network state (weights and biases) is reset. This type of validation is also known as cross-validation. Finally, the test set, which is an unused part of the initial set, is used to evaluate the network performance [7] [25].

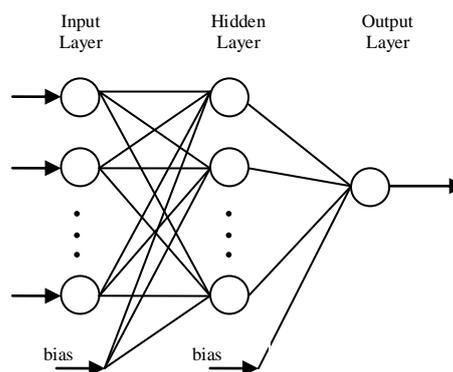


Figure 2 – Artificial neural network model in this article.

The training process is performed until any stop criteria is achieved. The errors associated with the stop criteria are the validation (generalization) and training errors. The validation error makes reference to the number of consecutive times that the network's prediction of the validation set is maintained or worsens. The test error is the network's performance measurement of the test set. The training, validation and test errors are evaluated by the MSE (mean square error) (see Eq. (3)). The error is the measurement of comparison between the observed data and the predicted ones.

The stop criteria for the neural network training algorithm in this paper are:

- Maximum number of iterations (1000);
- Validation or generalization error further than 5%;
- Training process error below 10^{-4} .

4 Particle Swarm Optimization (PSO)

The Particle Swarm Optimization (PSO) algorithm was developed by Kennedy and Eberhart in 1995 [4]. It is an evolutionary algorithm, which is inspired by flocks of birds or schools of fish searching for food. The algorithm seeks to optimize a population of random solutions. Each particle (individual) has a position and velocity, representing the solution to the optimization problem and the direction of the search in the search space. The particle adjusts the speeds and positions according to the best experiences. The algorithm determines the local best found by the particle itself and the global best, which is known by the entire swarm.

The particle swarm optimization theory assumes that in a population with size M , each individual ($1 \leq i \leq M$) has a present position X_i , an associated velocity V_i , a local best solution l_i , and each particle has a fitness function as well. A fitness function is the particle's measure of merit and it evaluates the particle's solution adequacy in solving the problem. The particle local best is simply the best solution obtained by it until the present moment. The global best solution X_g is the best solution among all the particles. The $c1$ and $c2$ parameters are the acceleration constants of the particles, and they dictate the velocity in which the particles move in the search space. For each iteration cycle, the particles velocities are updated (accelerated) in direction to the local and global minima:

$$V_i(t) = wV_i(t-1) + c1r1(l_i - X_i(t)) + c2r2(X_g - X_i(t)), \quad \text{Eq. (1)}$$

and the particles positions are updated using simple discrete dynamic:

$$X_i(t) = X_i(t-1) + V_i(t), \quad \text{Eq. (2)}$$

in which the term w is the inertia weight, the one that is used to balance the local and global search abilities of the algorithm, controlling the influence of previous information in the update of the new velocity. Usually, the term w decreases linearly from 1 to nearly 0 during the process [13]. The parameters $r1$ and $r2$ are two uniform random sequences between 0 and 1.

In the PSO algorithm each individual in the population evolves according to its own experience, estimating its own quality. Since the individuals are social beings, they also exchange information with their neighbors. These two kinds of information correspond to individual learning (cognitive – local knowledge) and cultural transmission (social – knowledge of the best position in the swarm), respectively. The algorithm [4] is formulated as follows:

1. Initialize the parameters, such as the acceleration constants, inertia weigh, number of particles, maximum number of iterations, velocity boundaries, initial and constrained velocities and positions and eventually the error limit for the fitness function;
2. Evaluate the particles' fitness function value, comparing with each other, therefore setting the local best and global best;
3. In accordance with Eq. (1) and Eq. (2) calculate the particles' new speed and position, and then update each particle.
4. For each particle, compare the current fitness value with the local best, if the current value is better, update the local best fitness value and the position with the current one;
5. For each particle, compare the current fitness value with the global best, if the current value is better, update the global best fitness value and the position with the current one;
6. If any stop criteria is achieved (maximum number of iterations or the fitness function has reached the stipulated limit) then stop the procedure, output the results. Otherwise, return to step 2.

5 Performance Measures

For the problem of time series forecasting, in the literature, there is no single metric universally adopted to evaluate the model's adequacy to predict. The following are some examples of authors that used several metrics with the purpose of assessing the model's performance: [5] [14] [22]. Therefore, in the present article, 6 (six) metrics are considered to allow a better appreciation of the forecasting system performance. The first of them is the MSE (mean square error), which is one of the most used performance measures for neural networks [5],

$$\text{MSE} = \frac{1}{N} \sum_{j=1}^N (\text{target}_j - \text{output}_j)^2, \quad \text{Eq. (3)}$$

where N is the number of patterns, target_j is the desired output (series real value) for the j -th pattern and output_j (model response) is the predicted value for the j -th pattern.

A second relevant measure is NMSE (normalized mean square error), or Theil's U statistic, that is based on the MSE, which is normalized by the prediction error, comparing it to a Random-Walk model. A Random-Walk model assumes that the most adequate value for the prediction at time $t + 1$ is the value obtained at time t , plus a noise term [5].

$$\text{NMSE} = \frac{\sum_{j=1}^N (\text{target}_j - \text{output}_j)^2}{\sum_{j=1}^N (\text{output}_j - \text{output}_{j-1})^2}. \quad \text{Eq. (4)}$$

A third relevant evaluation measure, the ARV (average relative variance), is given by

$$\text{ARV} = \frac{1}{N} \frac{\sum_{j=1}^N (\text{output}_j - \overline{\text{target}})^2}{\sum_{j=1}^N (\text{output}_j - \text{target})^2}, \quad \text{Eq. (5)}$$

which associates the model performance with the time series average. The term $\overline{\text{target}}$ is the time series average. If ARV equals 1, the model possesses a performance similar to the time series average. If ARV is greater than 1, then

the model possesses a performance worse than the time series average, and if ARV is smaller than 1, then the model possesses a better performance than the time series average.

Another performance measure is the MAPE (mean absolute percentage error), given by

$$\text{MAPE} = \frac{1}{N} \sum_{j=1}^N \left| \frac{\text{target}_j - \text{output}_j}{\text{target}_j} \right|. \quad \text{Eq. (6)}$$

A yet important measure for model comparison is the IA (index of agreement), which is able to express the difference among the observed and the forecasted data. This index can range from 0 to 1. Higher values indicate a better adequacy to the prediction problem. The statistical indicator is stated by

$$\text{IA} = 1 - \frac{\sum_{j=1}^N (\text{target}_j - \text{output}_j)^2}{\sum_{j=1}^N (|\text{target}_j - \text{output}_j| + |\text{target}_j - \text{target}|)^2}. \quad \text{Eq. (7)}$$

The last performance measure is the POCID (prediction of change in direction), which maps the time series tendency, in other words, the model's capacity of predicting if future values will increase or decrease based on past values. The measure is formulated as follows:

$$\text{POCID} = 100 \frac{\sum_{j=1}^N D_j}{N}, \quad \text{Eq. (8)}$$

where

$$D_j = \begin{cases} 1, & \text{if } (\text{target}_j - \text{target}_{j-1})(\text{output}_j - \text{output}_{j-1}) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

The function that evaluates the adequacy of the model, is called fitness function, and can be a variety of functions. In this paper, is given by

$$\text{fitness} = \frac{1}{(1+\text{MSE})}. \quad \text{Eq. (9)}$$

The function takes into consideration the MSE (mean square error). Nevertheless, the algorithm is also evaluated regarding the other metrics. In a perfect system, all performance measures must tend to zero, with the exceptions of the POCID that must approximate to 100, the IA that must tend towards 1, and the fitness function, that must approximate to 1.

6 MODELS

This section describes the specifics from each of the proposed models within this article.

The PSO is used to find the best possible MLP structure (input and hidden layers size) and to train the MLP characteristics (weights, biases). The following remarks are necessary to understand how the PSO is employed to train the MLP model, using Eq. (1) and Eq. (2):

- One multi-layer perceptron neural network corresponds to one particle in PSO (with position and velocity properties). Each particle has as components the network structure (input layer and hidden layer sizes), the weights matrix, and the biases (hidden and output layers);
- The number of particles, in other words, the population size defines how many MLPs will be used to search the optimal network and its characteristics;
- Any MLP individual characteristics (weights, bias and structure) are updated one by one during the search process for the optimal solution. The updates are made based upon Eq. (1) and Eq. (2).

The proposed hybrid system combines two intelligent techniques, PSO and ANN [5]. The idea is that each individual of the population, which is an ANN, be adjusted and have its structure dimensioned by the PSO. Therefore, the PSO will determine:

- The minimum number of time lags for the characterization of the series: initially a maximum number of time lags (MAX_LAGS), so that the PSO can choose any value in the interval [1, MAX_LAGS] for each individual in the population.
- The number of hidden neurons: initially a maximum number of neurons (MAX_NHIDDEN), so that the PSO can choose any value in the interval [1, MAX_NHIDDEN] for each individual of the population.

With this methodology it is possible to minimize the search time for the most compact network structure capable of representing the data.

The PSO, when calculating new velocities and positions, uses these values in the update of the number of units in the input layer, the number of neurons in the hidden layer, the weights (of the hidden layer and output layer) and the biases (of the hidden layer and output layer). In this work, four combinations of the PSO algorithm with Multi-layer Perceptron (MLP) artificial neural network (ANN) were explored.

The PSO updates the network's parameters within these models. Then, a network simulation is conducted with the intent to acknowledge the network's adequacy to the presented data. This process is ended if any criteria are met. The model variations are as follows:

In the first two models only the PSO updates the networks' weights and biases matrixes:

1. PSO-MLP – The PSO algorithm adjusts the network parameters.
2. PSO-MLP-with non-consecutive lags (NCL) – The PSO algorithm has the ability to choose a non-consecutive sequence of time lags. This characteristic is applied aiming at a more compact network structure, therefore, making a lower computational cost possible.

In the last two models, both PSO and MLP itself are able to adjust the networks' weights and biases matrixes:

3. PSO-MLP-LM – At every predetermined stage in the algorithm (e.g.: every 100 iterations), or stagnation point (e.g.: when the PSO is no longer able to enhance its prediction), the MLP trains the network. The training is performed by the Levenberg-Marquardt (LM) gradient descent algorithm. Such training is capable of improving the model's prediction.
4. PSO-MLP-LM-with non-consecutive lags (NCL) – This model has two abilities. The first, it can choose a non-consecutive sequence of time lags and. The second: it uses the gradient descent algorithm to enhance the prediction. This model combines both abilities. Therefore, it is capable of performing more accurate prediction with a smaller network structure.

For comparison reasons the AR (autoregressive) model and the ARMA (autoregressive moving average) model are also applied to the presented series. The models' parameters were established after performing statistical procedures on the series in this work, such as: correlation, autocorrelation and, variability.

7 Simulation and Results

The main purpose for the use of the PSO combined with an ANN is to combine the exploitation and the exploration characteristics. The PSO has the exploratory behavior. This means that it can look for a solution in many locations of the search space and probably find a global minimum. One of the main difficulties with the PSO is that once it finds a possible global minimum, its performance does not improve fast enough, taking a long time to reach the best result possible for that particular search space area. Thus, when the PSO is unable to substantially enhance the performance, the ANN can exploit the search space in a local manner, making the search and achieving a local minimum at that specific area in less time. However the ANN has a well known problem: it can be stuck in local minimum, without the capability to get out of it. Therefore, after those two methods are combined, two results are expected. First, the ANN would reduce the PSO's time to search for a best possible solution in a particular search area. Second, the PSO would search a large part of the search area, not allowing the ANN to be locked in a local minimum [1] [13] [16] [25].

In this paper, four time series are used to evaluate the methodology adequacy to the problem. All of the series correspond to natural phenomena. They are gaseous concentrations of: carbon monoxide (CO), sulfur dioxide (SO₂), nitrogen dioxide (NO₂) and particular matter (MP₁₀). The data are from the city of Sao Caetano, located in the state of Sao Paulo, in Brazil, and were obtained by the CETESB-SP (Environmental Company of the State of Sao Paulo) [18]. The series were normalize to the interval [0,1] and divided in three groups, following the Proben1 technical report [19]: training set with 50% of the time series data, validation set with 25% of the time series data and test set with 25% of the time series data. The PSO parameters were the same for all the experiments. The number of iterations was 1000, the parameters $c1$ and $c2$ were established as 1.8 and 1.5, respectively. The terms $r1$ e $r2$ are random numbers between [0, 1].

Ten particles were used, where each PSO individual was an artificial neural network with maximum architecture of 10 – 10 – 1, which makes reference to a multilayer perceptron network, that denotes 10 units in the

input layer, 10 units in the hidden layer and 1 unit in the output layer (prediction horizon of one step forward). For each time series, 10 experiments were performed with the combined algorithms, in which the experiments of each algorithm with the greatest fitness function were chosen as the representative of the respective model.

In the next subsections the experimental results achieved are shown for the four time series. With the purpose of comparing results of the used models, the results of the AR and ARMA model are also presented. The neural network with the architecture $Z - Y - 1$, consists of Z units in the input layer, Y units in the hidden layer and 1 unit in the output layer (prediction horizon of one step forward).

7.1 Carbon Monoxide Series (CO)

The measurements available for the CO pollutant were collected between 2000 until the year of 2001. With the purpose of reducing data fluctuation, the data set of hourly observations was transformed in a set of 713 daily averages.

The best MLP structure was chosen after conducting ten experiments for all four models. The most adequate model was the PSO-MLP-NCL, based upon the fitness function. The architecture selected consists of a window of two time lags, [1 and 6], thus it has 2 (two) elements in the input layer, 10 units in the hidden layer. Therefore, the MLP architecture the best fit the problem at hand was defined as 2 – 10 – 1. After analyzing Table 1 it was observed that the PSO-MLP-NCL had predicted with a small margin of error the time series, outperforming all other models. Thus, it was selected as the most suited for the prediction of the concentration of CO. The results for all the performance measurements are shown in Table 1.

Table 1 – Results for the CO series.

	AR model	ARMA model	PSO-MLP	PSO-MLP- NCL	PSO-MLP-LM	PSO-MLP-LM- NCL
MEASURES						
<i>MSE</i>	0.00877	0.00725	0.00791	0.00686	0.00728	0.00884
<i>NMSE</i>	1.40684	1.49364	1.51043	2.08514	0.92525	1.31084
<i>POCID</i>	51.68539	55.05618	55.61798	58.19209	54.23729	52.24719
<i>ARV</i>	0.00755	0.00742	0.00983	0.00929	0.00755	0.01038
<i>MAPE</i>	0.39724	0.39176	0.41014	0.39432	0.41446	0.43161
<i>IA</i>	0.72593	0.75828	0.76003	0.78052	0.77216	0.74616
<i>FITNESS</i>	0.99130	0.99279	0.99216	0.99319	0.99278	0.99124
<i>Hidden Layer size</i>			7	10	5	8
<i>Input Layer (Time lags)</i>			[1 2]	[1 6]	[1 2 3 4]	[3 7]

The results presented in Table 1 also indicate that the PSO-MLP-NCL model has a better performance than the statistical models. The selected model is capable of relating past values to predicted values (POCID greater than 50%), and it is also the one with the most adequacy to solve the problem according to the IA that is greater than any other model. The Figure 3 represents the models' prediction and the observed data for the CO series. The figure shows how the models behaved on the predictions' last 100 consecutive points.

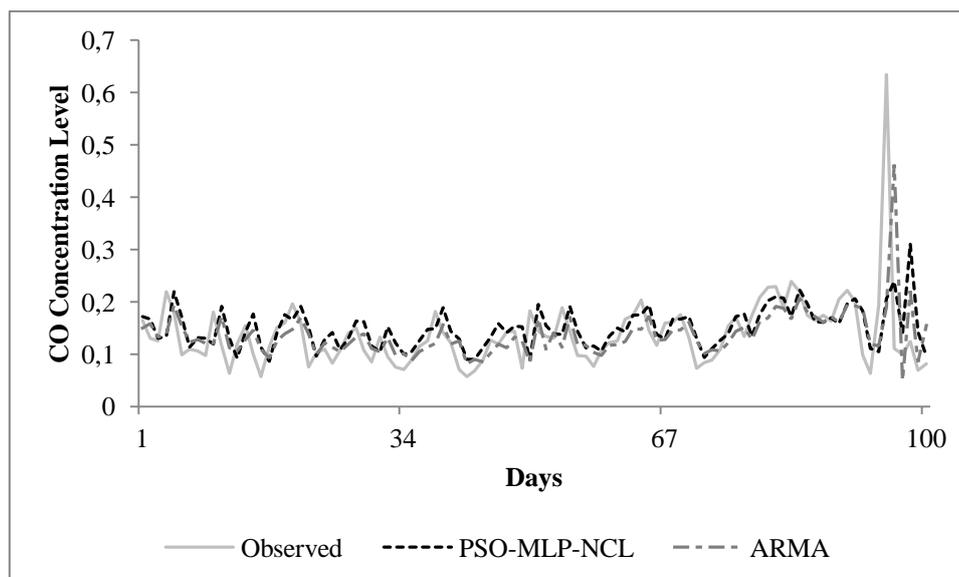


Figure 3 – Results of the last 100 values of the prediction made by the ARMA model and the PSO-MLP-LM model for the CO series (test set): normalized observed values, predicted MLP and ARMA values.

7.2 Sulfur Dioxide Series (SO₂)

The hourly values obtained by the network of sensors for the SO₂ pollutant were converted in daily averages, with the intention of reducing the variability of the concentration values. The final data set has a number of 699 points.

Each model was submitted to ten experiments. As a result the PSO-MLP-LM was selected, since it achieved better performance than all others, hence it was able to forecast such series with more accuracy. Such model has the temporal window of three inputs, being [1 2 3] as the time lags and the hidden layer has nine processing units. Therefore, the architecture that was selected was identified as 3 – 9 – 1. The results for all the performance measurements are shown in Table 2.

Table 2 – Results for the SO₂ series.

	AR model	ARMA model	PSO-MLP	PSO-MLP-Non consecutive lag	PSO-MLP-LM	PSO-MLP-LM-Non consecutive lag
MEASURES						
<i>MSE</i>	0.01068	0.00903	0.00927	0.00913	0.00883	0.01046
<i>NMSE</i>	1.45835	1.73154	2.62792	2.76644	2.57693	2.28375
<i>POCID</i>	46.70659	49.70060	48.50299	47.90419	47.90419	43.97590
<i>ARV</i>	0.00743	0.00918	0.01572	0.01498	0.01533	0.01241
<i>MAPE</i>	0.68479	0.70887	0.77784	0.75895	0.77051	0.82050
<i>IA</i>	0.71636	0.75644	0.77110	0.77286	0.77797	0.75025
<i>FITNESS</i>	0.98944	0.99105	0.990815	0.99095	0.991244	0.98965
<i>Hidden Layer size</i>			5	1	9	8
<i>Input Layer (Time lags)</i>			[1 2]	[1 5]	[1 2 3]	[1 4 6 8]

In Table 2 is possible to visualize that the PSO-MLP-LM model reached a better performance, according to the fitness, than those of the statistical models. In comparison with the other models, included the statistical ones,

the observed model's results indicate a better adequacy to the problem (greater IA). PSO-MLP-LM and ARMA models predictions are shown in Figure 4.

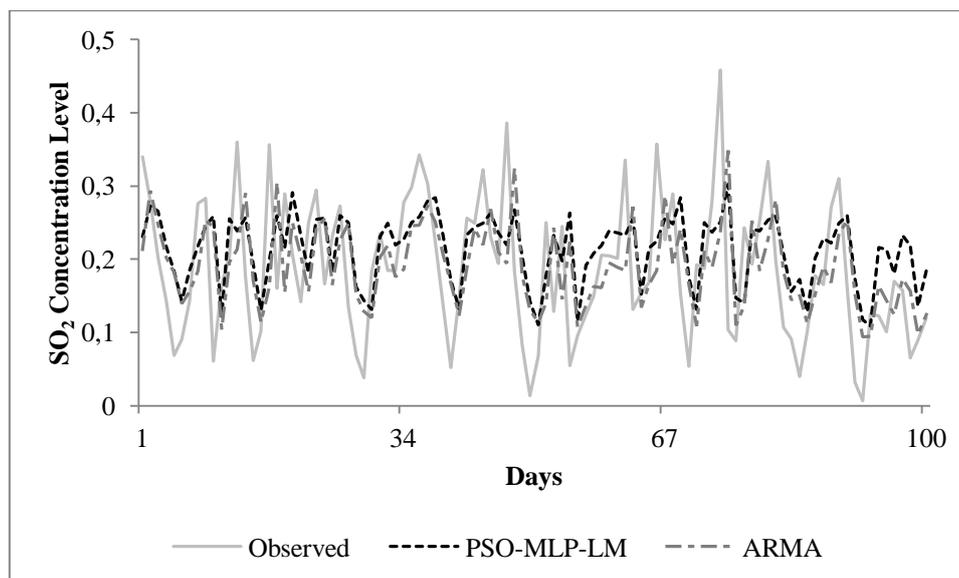


Figure 4– Results of the last 100 values of the prediction made by the ARMA model and the PSO-MLP-LM model for the SO₂ series (test set): normalized observed values, predicted MLP and ARMA values.

7.3 Particle matter series (MP₁₀)

The MP₁₀ series was obtained by the network monitoring sensors. These sensors provided the data set used as the time series for the models' prediction. The data set has 716 valid daily observations from the years 2002 and 2003.

The most adequate model for the forecasting of the MP₁₀ time series was chosen after ten experiments. It was observed that the PSO-MLP-LM-NCL model outperforms the other models, with the exception of the ARMA model. The selected model has a high adequacy to the series modeling. The results for all the performance measurements are shown in Table 3. Figure 5 shows the forecasts made by the PSO-MLP-NCL model and the AR model.

Table 3– Results for the MP₁₀ series.

	AR model	ARMA model	PSO-MLP	PSO-MLP-NCL	PSO-MLP-LM model	PSO-MLP-LM-NCL
MEASURES						
<i>MSE</i>	0.01928	0.01770	0.01794	0.01795	0.01913	0.01791
<i>NMSE</i>	1.50803	1.59015	2.52676	2.79771	2.12474	2.85436
<i>POCID</i>	47.48603	45.25140	45.25140	45.25140	44.94382	45.81006
<i>ARV</i>	0.00383	0.00446	0.00751	0.00708	0.00847	0.00689
<i>MAPE</i>	1.34506	1.65346	1.64014	1.59574	1.61379	1.56692
<i>IA</i>	0.77220	0.79755	0.81868	0.81767	0.81364	0.81800
<i>FITNESS</i>	0.98108	0.98261	0.982378	0.982365	0.981228	0.982408
<i>Hidden Layer size</i>			6	8	2	7
<i>Input Layer (Time lags)</i>			[1 2]	[1 4]	[1 2 3 4 5 6]	[1 4]

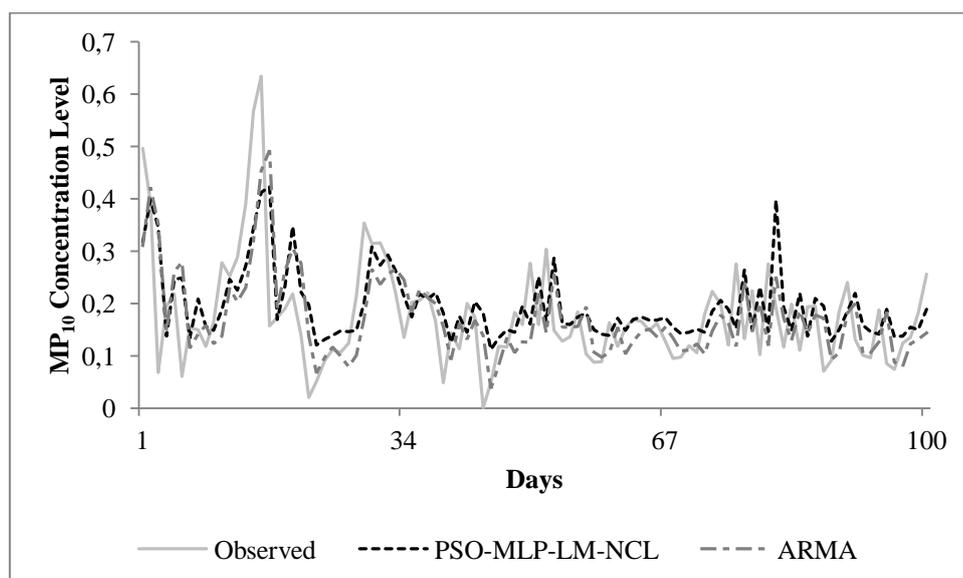


Figure 5– Results of the last 100 values of the prediction made by the AR model and the PSO-MLP-LM-NCL model for the MP_{10} series (test set): normalized observed values, predicted MLP and AR values.

7.4 Nitrogen dioxide series (NO_2)

The measurements made in the year 2000 for the nitrogen dioxide (NO_2) series were used to assess the model validity. The NO_2 series has 355 daily averages and it corresponds to the used data set.

The assessment made from the Table 4 is that the best suited model for the representation of the time series, chosen after ten experiments, was the PSO-MLP-LM. It was also observed that the PSO-MLP-LM model with non-consecutive lags was able to forecast the time series with a reduced MLP architecture, even though it had a shortcoming in its performance when compared to the PSO-MLP-LM model. The selected model had a better performance than that of the random walk model (NMSE less than 1), finally, it was capable of reaching a better performance than the time series average (AVR less than 1). The results for all the performance measurements are shown in Table 4. Figure 6 presents the observed data versus the predicted data.

Table 4– Results for the NO_2 series.

	AR model	ARMA model	PSO-MLP	PSO-MLP-NCL	PSO-MLP-LM	PSO-MLP-LM-NCL
MEASURES						
<i>MSE</i>	0.02370	0.02306	0.02676	0.02551	0.02718	0.02605
<i>NMSE</i>	0.93585	1.18699	1.74967	1.78871	0.77516	1.04540
<i>POCID</i>	43.33333	44.44444	45.55556	44.44444	41.57303	44.44444
<i>ARV</i>	0.01049	0.01448	0.02498	0.02587	0.01196	0.01605
<i>MAPE</i>	0.55831	0.70137	0.72517	0.70461	0.75287	0.62954
<i>IA</i>	0.73838	0.76161	0.76262	0.76856	0.75205	0.76189
<i>FITNESS</i>	0.97684	0.97746	0.97394	0.97512	0.97354	0.97462
<i>Hidden Layer size</i>			9	10	9	4
<i>Input Layer (Time lags)</i>			[1 2]	[1 2 5]	[1 2]	[3]

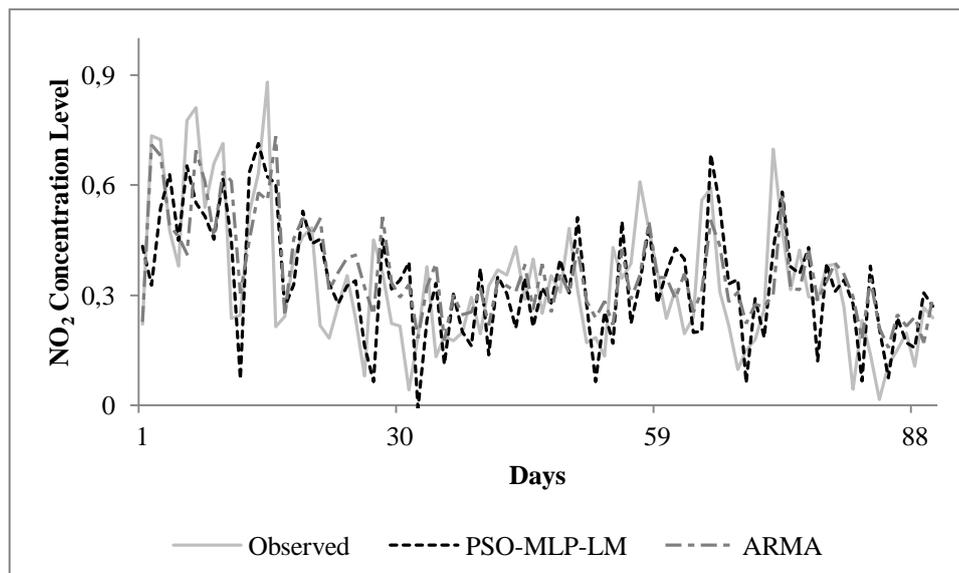


Figure 6 – Results of the last 90 values of the prediction made by the ARMA model and the PSO-MLP-LM model for the NO₂ series (test set): normalized observed values, predicted MLP and ARMA values.

8 Conclusions

In this paper, a hybrid system for time series forecasting of concentration levels of air pollutants was analyzed. It consisted of an intelligent hybrid model composed of an algorithm, Particle Swarm Optimization (PSO) and a Multi-layer perceptron (MLP) Artificial Neural Network (ANN). The PSO searched for a minimum number of time delays for proper characterization time series and a better structure of a neural network in terms of input units, hidden processing units, initial weights, and biases of the artificial neural network.

The results were presented in terms of six measures: the mean square error (MSE), the normalized mean square error (NMSE), the average relative variance (ARV), the mean absolute percentage error (MAPE), the prediction of change in direction (POCID) and, the index of agreement (IA). The four combined methods: PSO-MLP, PSO-MLP with non-consecutive lags, PSO-MLP-LM (Levenberg-Marquardt gradient descent algorithm) and PSO-MLP-LM with non-consecutive lags have been applied to four series of natural phenomena (carbon monoxide, sulfur dioxide, particulate matter, and nitrogen dioxide), with all its external dependencies to the models. For comparison purposes the statistical models AR (autoregressive) and ARMA (autoregressive moving average) were also used.

Despite the fact that the statistical models were able to achieve a good prediction, such a kind of model needs preliminary information about the time series, such as: tendency, stationarity, and correlation. The proposed methodology does not need previous knowledge about the time series, to be able to predict it. Therefore, it can be applied to modeling a not fully understood phenomenon.

The results show that the combination between the PSO and the MLP is a valid option for time series forecasting of concentration levels of air pollutants. Because it obtains a fair and accurate forecast with an acceptable computational cost.

9 Acknowledgements

We would like to thank the Foundation of Support to Science and Technology of the State of Pernambuco (FACEPE) for the financial support to the research.

10 References

- [1] W. Z. Lu et al, Application of evolutionary neural network method in predicting pollutant levels in downtown area of Hong Kong, **Neurocomputing**, 51 (2003), 387-400.
- [2] M. L. Bell et al, Environmental health indicators and a case study of air pollution in Latin American cities, **Environmental Research**, 111 (2011), 57-66.

- [3] A. Kurt et al, An online air pollution forecasting system using neural networks, **Environmental International**, 34 (2008), 592-598.
- [4] R. Eberhart, J. Kennedy, Particle Swarm Optimization, **IEEE Proceedings – Neural Networks**, 4 (1995), 1942-1948.
- [5] P. S. G. de Mattos Neto, A. R. Lima Junior, Tiago A. E. Ferreira. Time series forecasting using a perturbative intelligent system. **In Proceedings of GECCO**, (2009), 1477-1478.
- [6] P. S. G. de Mattos et al, Combining Artificial Neural Network and Particle Swarm System, **Proceedings of International Joint Conference on Neural Networks**, (2009), 2230-2237.
- [7] S. Haykin, Neural Networks: Principles and Practices, 2^a Edition, **Bookman**, (2001).
- [8] B. Coppin, Artificial Intelligence, 1^a Edition, **LTC**, (2010).
- [9] P. P. Balestrassi et al, Design of experiments on neural network's training for nonlinear time series forecasting, **Neurocomputing**, 72 (2009), 1160-1178.
- [10] M. Khashei, M. Bijari, A novel hybridization of artificial neural networks and ARIMA models for time series forecasting, **Applied Soft Computing**, 11 (2011), 2664-2675.
- [11] U. Kumar, K. De Ridder, GARCH modelling in association with FFT – ARIMA to forecast ozone episodes, **Atmospheric Environment**, 44 (2010), 4252-4265.
- [12] O. Valenzuelaa et al, Hybridization of intelligent techniques and ARIMA models for time series prediction, **Fuzzy Sets and Systems**, 159 (2008), 821-845.
- [13] L. Zhao, Y. Yang, PSO-based single multiplicative neuron model for time series prediction, **Expert Systems with Applications**, 36 (2009), 2805-2812.
- [14] U. Brunelli, Two-days ahead prediction of daily maximum concentrations of SO₂, O₃, PM₁₀, NO₂, CO in the urban area of Palermo, Italy, **Atmospheric Environment**, 41 (2007), 2967-2995.
- [15] G. E. P. Box, G. M. Jenkins, Time Series Analysis: Forecast and Control, 3th Edition, Prentice Hall, (1994).
- [16] D. Wang, W. Z. Lu, Ground-level ozone prediction using multilayer perceptron trained with an innovative hybrid approach, **Ecological Modelling**, 198 (2006), 332-340.
- [17] D. Graves, Witold Pedrycz, Fuzzy prediction architecture using recurrent neural networks, **Neurocomputing**, 72 (2009), 1668-1678.
- [18] CESTEB – SP, Companhia Ambiental do Estado de São Paulo, [Online]. Disponível: <http://www.cetesb.sp.gov.br>.
- [19] L. Prechelt, Proben1: A set of neural network benchmark problems and benchmarking rules, Technical Report 21/94, 1994. [Online]. Disponível: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.115.5355>.
- [20] G. Grivas, A Chaloulakou, Artificial neural network models for prediction of PM₁₀ hourly concentrations, in the Greater Area of Athens, Greece, **Atmospheric Environment**, 40 (2006), 1216-1229.
- [21] Y. Feng et al, Ozone concentration forecast method based on genetic algorithm optimized back propagation neural networks and support vector machine data classification, **Atmospheric Environment**, 45 (2011), 1979-1985.
- [22] R. Zemouri et al, Defining and applying prediction performance metrics on a recurrent NARX time series model, **Neurocomputing**, 73 (2010), 2506-2521.
- [23] F. Takens, Detecting strange attractor in turbulence, **Dynamical Systems and Turbulence**, 898 (1980), 366-381.
- [24] R. H. Shumway, D. S. Stoffer, Time series analysis and its applications, 3th Edition, **Springer**, (2011).
- [25] D. Dang, W. Z. Lu, Forecasting of ozone level in time series using MLP model with novel hybrid training algorithm, **Atmospheric Environment**, 40 (2006), 913-924.

CAPÍTULO III

CONCLUSÕES GERAIS

Neste trabalho foram aplicadas técnicas da área de inteligência artificial (IA) aos problemas de previsão e classificação de riscos ambientais, com o foco na qualidade do ar. As técnicas empregadas na previsão de concentrações de poluentes gasosos foram as Redes neurais artificiais (RNA) e o algoritmo de Otimização por enxame de partículas (PSO). Tais técnicas ao serem combinadas suprimiram uma conhecida dificuldade inerente ao problema de modelagem utilizando uma RNA, a escolha da arquitetura da rede. O modelo advindo da combinação das duas técnicas (PSO-MLP-LM) obteve resultados (previsões) similares às apresentadas pelos modelos estatísticos (AR e ARMA). Tais resultados demonstram que o modelo aplicado ao problema é uma opção viável à previsão de séries temporais de concentrações de poluentes gasosos. Apesar do modelo PSO-MLP-LM ter alcançando previsões próximas às dos modelos AR e ARMA, o mesmo o fez sem necessitar de tratamento prévio dos dados.

O problema de classificação de riscos ambientais emprega variáveis linguísticas para descrever a avaliação do risco como, por exemplo, risco alto, risco médio, risco muito baixo etc. Estas descrições são inerentemente vagas o que dificulta a avaliação das mesmas. Com o intuito de tratar estas incertezas a lógica nebulosa foi aplicada ao problema. As considerações pontuais são que a lógica nebulosa por lidar bem com as incertezas deste tipo de processo foi capaz de obter uma classificação quantitativa e não só qualitativa, sendo também apta a tratar relações de riscos causais, remetendo a uma classificação mais adequada e coerente dos riscos ambientais.

O modelo apresentado neste trabalho é passível de desenvolvimento, caso as entidades gerenciais do estado de Pernambuco optem por tal. Contudo, uma das dificuldades encontradas neste trabalho foi a aquisição dos dados, pois, no Estado de Pernambuco tais dados não estão disponíveis por não haver, até a presente data, uma rede de monitoramento funcional de poluentes atmosféricos. Por esta razão, os dados empregados neste estudo são advindos da cidade de São Caetano, no estado de São Paulo. Os dados foram adquiridos da CETESB-SP, por meio eletrônico.

Uma vantagem desta metodologia é a possibilidade de saber com antecedência, ou seja, hoje, quais serão as concentrações dos poluentes amanhã. Partindo deste pressuposto, tem-se mais tempo para o planejamento de ações, possibilitando, assim, maior qualidade nas ações realizadas pelas entidades responsáveis.

TRABALHOS FUTUROS

- No campo da previsão, uma das possíveis vertentes a ser perseguida seria a da procura por uma função de adequação que minimizasse os erros obtidos, proporcionando, desta forma, uma melhora no comportamento da previsão;
- Aplicação de uma metodologia que visasse múltiplos objetivos, desta forma minimizando não um, mas, um conjunto de erros, com a finalidade de melhorar a previsão de determinada série;
- No campo de análise de riscos seria a aplicação do modelo a realidade local (Recife-PE), caso os dados necessários sejam passíveis de serem adquiridos.

