

# Previsão da qualidade do ar nas Áreas Metropolitanas de Lisboa e Porto por métodos estatísticos

Filipa Marques\*, Francisco Ferreira, Sandra Mesquita

Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa  
Departamento de Ciências e Engenharia do Ambiente  
Campus da Caparica, Quinta da Torre  
2829-516 CAPARICA  
Telefone/fax: 212 948 374, \*fpsm@fct.unl.pt

## Sumário

Os modelos de previsão baseiam-se na relação empírica entre as concentrações diárias dos poluentes atmosféricos e as condições meteorológicas locais. A relação entre os poluentes e a meteorologia é explorada recorrendo a classificação e regressão em árvore. Desta forma, seleccionam-se variáveis meteorológicas que melhor explicam o comportamento e variações das concentrações. Os modelos de previsão utilizam as variáveis predictoras anteriormente seleccionadas para desenvolver modelos de regressão linear múltipla.

A previsão do índice de qualidade do ar, por métodos estocásticos, para as áreas metropolitanas de Lisboa e Porto é uma aplicação operacional, emitida diariamente. A previsão de cada índice assenta na previsão das concentrações de 2 poluentes. Os poluentes previstos diariamente são o Ozono troposférico ( $O_3$ ) (valor máximo horário) e as partículas em suspensão com diâmetro aerodinâmico inferior a  $10 \mu m$  ( $PM_{10}$ ) (média diária), dado ter-se verificado pela análise do histórico que na maioria dos dias foram estes os poluentes responsáveis pela determinação de índices diários de qualidade do ar.

A validação dos modelos estocásticos de previsão operacional para as aglomerações Área Metropolitana de Lisboa (AML) Norte e Porto Litoral mostra resultados estatisticamente satisfatórios.

**Palavras-chave:** Previsão qualidade do ar, Ozono,  $PM_{10}$ , prevqualar, modelos estatísticos, regressão linear

## Introdução

A qualidade do ar está directamente relacionada com a emissão de poluentes para a atmosfera e com as condições meteorológicas locais. Os episódios de má qualidade do ar estão associados à maior quantidade de poluentes emitidos para a atmosfera e às condições meteorológicas que interagem com os respectivos poluentes. Cada um dos poluentes reage de forma própria às condições meteorológicas. A análise dos dados dos diversos poluentes (registados nas estações de monitorização de qualidade do, qualar) revela que cada um dos poluentes atinge maiores concentrações em alturas do ano diferentes. O ozono troposférico atinge concentrações mais elevadas nos meses de verão e as partículas inaláveis no fim do Inverno, principio da primavera. A relação e os modos próprios de cada poluente interagir com as condições atmosféricas pode ser identificada investigando a relação entre meteorologia e poluentes (Ziomas et al, 1995, Cogliani, 2001, Hubbard et al, 1998). Os

resultados destes autores mostram que relacionando a meteorologia local com a poluição, a previsão estatística é uma ferramenta simples e objectiva.

Os modelos são construídos recorrendo à técnica de análise multivariada de dados conhecida por Classificação e Regressão em Árvore, já usada para desenvolver modelos de previsão estatística das concentrações (Cassmassi, 1998, Marques et al., 2006). Para cada poluente (PM10 e Ozono), registado em cada estação da rede de monitorização de qualidade do ar (Figura 1), com um registo histórico diário de 2 a mais anos, desenvolveram-se modelos de previsão, testando inicialmente várias condições iniciais, tais como, modelos sazonais, diários, diferentes intervalos do factor persistência

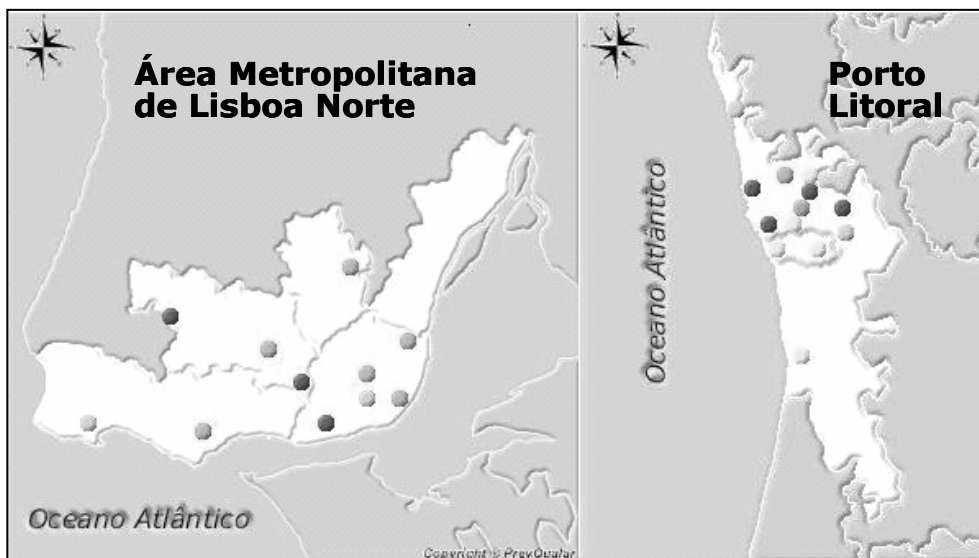


Figura 1 – Estações de monitorização de qualidade do ar na Área Metropolitana de Lisboa Norte (esquerda) e Porto Litoral (direita) com modelos de previsão de concentrações da média diária de PM10 e máximo horário diário de ozono (ver [www.prevqualar.org](http://www.prevqualar.org)).

O índice diário da qualidade do ar (IQAR) para a globalidade de cada zona é calculado de acordo com as regras definidas a nível nacional. Apresenta-se ainda a agregação do índice em combinações espaciais diferentes resultando numa informação localmente mais detalhada. Apresentam-se assim os seguintes índices: Índice por subzona (Figura 2), Índice por estação, Índice de tráfego e Índice de fundo. Os resultados desta abordagem revelam que a agregação de poluentes de estações com características ambientais iguais é fiável e apresenta índice de sucesso idêntico à agregação ao nível global da zona.

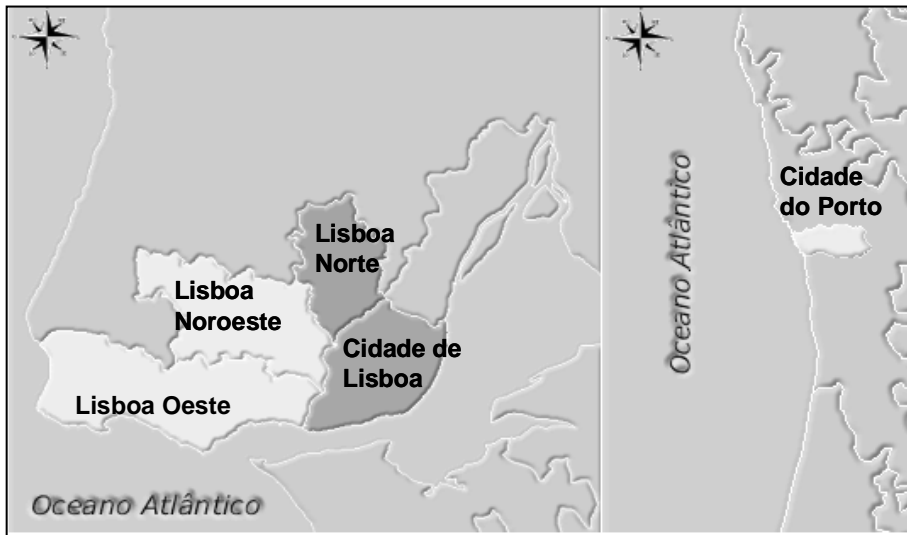


Figura 2 – Agregação do índice de qualidade do ar em sub zonas para a: 1) Área Metropolitana de Lisboa Norte (esquerda) resultando em 4 sub zonas: Cidade de Lisboa, Lisboa Norte, Lisboa Noroeste e Lisboa Oeste; 2) Porto Litoral (direita), resultando em apenas uma subzona, Cidade do Porto (ver [www.prevqualar.org](http://www.prevqualar.org)).

Para cada modelo de cada poluente desenvolvido em cada estação, é testada a *performance* em reproduzir os dados observados. Os resultados dos modelos de previsão são aqui apresentados. Estes revelam fiabilidade em simular os valores das concentrações observadas.

### Metodologia

Os parâmetros meteorológicos utilizados são observações registadas nas estações meteorológicas da rede Nacional do Instituto de Meteorologia, e as observações de radiossondagem da estação de Lisboa/Gago Coutinho. Recorrendo aos dados históricos das concentrações diárias dos poluentes atmosféricos PM10 e Ozono troposférico, e das variáveis meteorológicas à superfície e em altitude, investigou-se a relação diária entre estes parâmetros (em cada estação de monitorização) utilizando a técnica de Classificação e Regressão em Árvore (Cassmassi, 1998). Esta é uma técnica de classificação objectiva que selecciona variáveis predictoras a partir do conjunto de dados meteorológicos/poluentes.

Uma das variáveis incluída no conjunto de variáveis predictoras a testar é o valor das concentrações de cada poluente registado no dia anterior. Esta variável é utilizada como predictor de forma a introduzir nos modelos o factor persistência/tendência das concentrações na atmosfera. A classificação em árvore divide as variáveis em grupos binários que melhor explicam o valor observado do poluente. A selecção binária apresenta uma divisão em forma de árvore. O resultado é um conjunto de variáveis predictoras que 'decidem' quais as condições que estão a contribuir para o valor observado do poluente, Ryan (1995). O passo seguinte é construir uma equação de regressão linear múltipla utilizando como variáveis predictoras as variáveis resultantes da análise anterior. Utilizou-se para a AML Norte dados diários de 2000 a 2004, sendo a validação aqui apresentada para o ano de 2005. Para o Porto Litoral o conjunto de dados utilizado compreende os anos de 2000 a 2005.

Obtém-se assim para cada poluente, em cada estação de monitorização de qualidade do ar, um ou mais modelos de previsão das concentrações para o dia seguinte. Para cada poluente, é atribuído um índice correspondente à gama de concentrações definidas na classificação do índice. O índice final é divulgado de acordo com as regras definidas a nível nacional, onde o índice global atribuído é dado pelo poluente com pior nível previsto.

Apresentamos como exemplo para a AML Norte, o resultado da previsão da média diária de PM10 (Figura 3 a) e o valor da média diária observada e na figura 3 b, o resultado da previsão do valor máximo horário diário de ozono previsto e observado

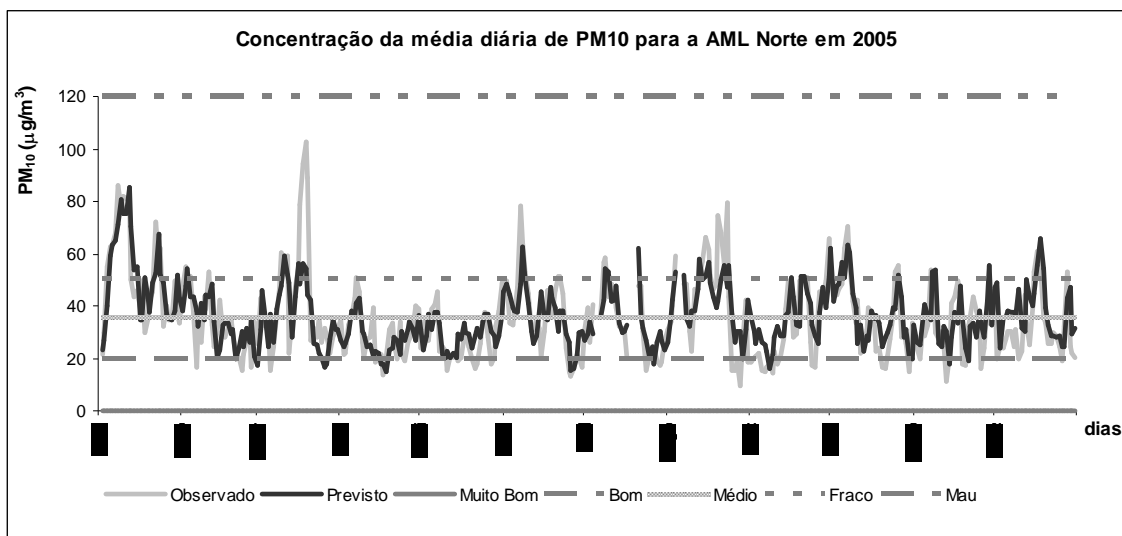


Figura 3 a) – Comparação entre o valor previsto e observado da média diária de PM10 para a AML Norte. As barras horizontais indicam o limite inferior do índice de qualidade do ar associado a este poluente.

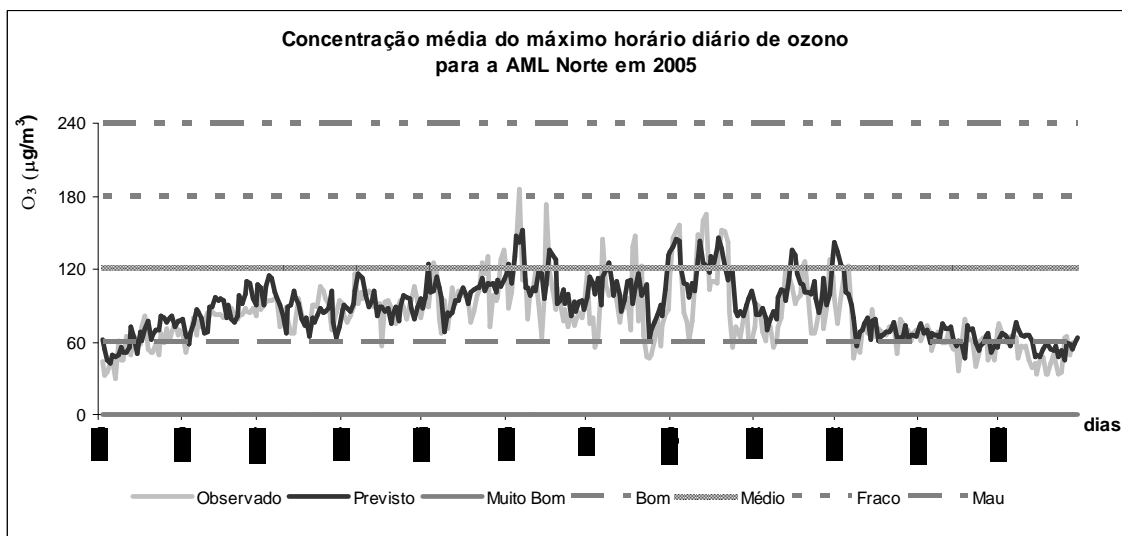


Figura 3 b) – Comparação entre o valor previsto e observado do máximo horário diário de ozono para a AML Norte. As barras horizontais indicam o limite inferior do índice de qualidade do ar associado a este poluente.

A análise destes gráficos mostra que para ambos os poluentes, a previsão simula o comportamento da variação diária.

A previsão das concentrações de cada poluente é diariamente calculada, utilizando as variáveis previstas dadas pelo modelo do 'European Centre for Medium-Range

Weather Forecasts - ECMWF'. A validação dos modelos de previsão foi feita utilizando os dados de observações meteorológicas e as previsões do modelo ECMWF, de forma a testar os resultados deste modelo. O critério de validação aplicado baseia-se num conjunto de características estatísticas, detalhadas em EPA (2003).

Apresentamos em baixo a tabela de contingência do índice global previsto e observado para a AML Norte, tabela 1, Porto Litoral, tabela 2, (resultante do pior nível previsto entre os 2 poluentes).

Índice Global AML Norte (PM10 & Ozono)		Índice Observado					Total
		Mau	Fraco	Médio	Bom	Muito Bom	
Índice Previsto	Mau	<b>0</b>	0	0	0	0	<b>0</b>
	Fraco	0	<b>38</b>	11	4	0	<b>53</b>
	Médio	0	23	<b>60</b>	42	1	<b>125</b>
	Bom	0	3	24	<b>147</b>	4	<b>178</b>
	Muito Bom	0	0	0	0	<b>0</b>	<b>0</b>
	Total	<b>0</b>	<b>64</b>	<b>95</b>	<b>193</b>	<b>5</b>	<b>357</b>

Tabela 1 – Tabela de contingência do índice global previsto para a AML Norte. A validação aqui apresentada foi feita para o ano de 2005 com os dados disponíveis de poluentes e com as previsões hindcast (previsões históricas que não sofreram alterações após o dia da previsão) das variáveis meteorológicas do ECMWF. Desta forma apresenta-se aqui os resultados de uma validação de previsões semi-hindcast pois os dados de poluentes utilizados diariamente tem por vezes falhas na hora a que a previsão é feita, que são mais tarde completadas.

Índice Global Porto Litoral (PM10 & Ozono)		Índice Observado					Total
		Mau	Fraco	Médio	Bom	Muito Bom	
Índice Previsto	Mau	0	0	0	0	0	<b>0</b>
	Fraco	1	<b>21</b>	9	6	1	<b>38</b>
	Médio	0	9	<b>21</b>	35	1	<b>66</b>
	Bom	0	2	6	<b>54</b>	4	<b>66</b>
	Muito Bom	0	0	0	2	<b>0</b>	<b>2</b>
	Total	<b>1</b>	<b>32</b>	<b>36</b>	<b>97</b>	<b>6</b>	<b>172</b>

Tabela 2 – Tabela de contingência do índice global previsto para o Porto Litoral. A validação aqui apresentada foi feita para com os resultados da previsão operacional deste Julho 2006 até Dezembro 2006.

Apesar de apresentarmos intervalos de validação diferentes\*, os modelos para a AML Norte apresentam uma taxa de sucesso (68%) superior ao Porto Litoral (55%). A probabilidade de detecção de episódios de IQAR fraco é superior para o Porto Litoral e a taxa de falso alarme deste nível de índice ligeiramente superior para a AML Norte.

\* Os modelos de cada uma das zonas foram construídos com diferentes períodos de tempo, sendo a validação apresentada para o ano seguinte de dados independentes aos que constituem a base de dados que serviu de desenvolvimento dos modelos.

## **Conclusões**

A metodologia desenvolvida para a previsão do índice de qualidade do ar, através da previsão das concentrações de cada poluente mostra resultados satisfatórios. Esta metodologia permite compreender a relação entre as diferentes variáveis meteorológicas e as concentrações de poluentes. A previsão do índice de qualidade do ar para as aglomerações apresenta resultados razoáveis. A validação do Porto Litoral será estendida a todo o ano de 2006.

A previsão das concentrações de ozono troposférico apresenta resultados melhores do que a previsão de PM10 tanto para a AML Norte como para o Porto litoral (ver [www.prevqualar.org](http://www.prevqualar.org)). Esta diferença remete para o próximo passo de desenvolvimento que consiste em calibrar os modelos com dados mais recentes de forma a actualizar os modelos.

## **Agradecimentos**

Este trabalho foi financiado pelo Instituto do Ambiente e contou com o apoio e colaboração do Instituto de Meteorologia, IP., e das Comissões de Coordenação e Desenvolvimento Regional que disponibilizaram os dados históricos e os dados horários diários, necessários ao desenvolvimento deste trabalho.

## **Referências**

Cassmassi, J. (1998): Objective Ozone Forecasting in the South Coast Air Basin: Updating the Objective Prediction Models for the Late 1990's and the Southern California Ozone Study (SCOS97-NARSTO) Applications. Presented at the *10th Joint Conference on Air Pollution Meteorology*, AMS Annual Meeting, Phoenix, AZ, January.

Cogliani, E., (2001): Air pollution forecast in cities by an air pollution index highly correlated with meteorological variables. *Atmospheric Environment*, 35, 2871-2877.

EPA (2003): Guidelines for developing an Air Quality (ozone and PM2.5) Forecasting Program, U.S Environmental Protection Agency, Office of Air Quality Planning and Standards, North Carolina, EPA-456/R-03-002.2

Hubbard, M. and Cobourn, W., (1998): Development of a regression model to forecast ground level ozone concentration in Louisville, KY. *Atmospheric Environment*, 32, 2637-2647.

Marques F., Cassmassi, J., Ferreira, F. E Mesquita, S. (2006) : Air quality forecast for Lisbon, Portugal: from data analysis to outreach. , *2006 National Air Quality Conferences: Air Quality Forecasting, Mapping, and Monitoring and Communicating Air Quality and Communities in Motion* "Quality of Air Means Quality of Life", February 5-8, 2006, San Antonio, Texas.

Ryan, W., (1995): Forecasting severe ozone episodes in the Baltimore Metropolitan Area. *Atmospheric Environment*, 29, 2387-2398

Ziomas, I., Melas, D., Zerefos, C., Bais, A. and Paliatsos, A., (1995) Forecasting peak pollutant levels from meteorological variables. *Atmospheric Environment*, 29, 3703- 3711