

PREVISÃO DA QUALIDADE DO AR PARA LISBOA - A ABORDAGEM ESTATÍSTICA

Jorge Neto⁽¹⁾⁽²⁾, Pedro M. Torres⁽¹⁾ e Francisco Ferreira⁽¹⁾

⁽¹⁾ Departamento de Ciências e Engenharia do Ambiente, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa; jorge.neto@meteo.pt, pmt@fct.unl.pt, ff@fct.unl.pt

⁽²⁾ Departamento de Observação e Redes, Instituto de Meteorologia.

Resumo

O cumprimento por parte de Portugal da legislação nacional e europeia no domínio da qualidade do ar é um dos objectivos fundamentais da política de ambiente, nomeadamente como salvaguarda da saúde pública e dos ecossistemas. Nas últimas épocas de Verão, Lisboa tem revelado concentrações elevadas de diversos poluentes atmosféricos, ultrapassando o estabelecido na legislação vigente. Este artigo apresenta o trabalho que tem sido realizado para se atingir um conhecimento mais profundo da relação entre a qualidade do ar e a meteorologia na região de Lisboa, mais precisamente no que se refere ao ozono (O_3) e às partículas de diâmetro inferior a $10\ \mu m$ (PM_{10}). Até à data, este estudo tem incidido fundamentalmente sobre o O_3 .

O objectivo final é a previsão da qualidade do ar usando métodos estatísticos (análise de regressão linear múltipla (RM) e regressão em árvore (CART)). A validação e a análise de sensibilidade ao modelo foram realizadas utilizando um período de dados distinto daquele que foi utilizado para o seu desenvolvimento. Daí resultou numa melhor compreensão de cada uma das situações meteorológicas e de qualidade do ar, permitindo já uma previsão com considerável fiabilidade em muitos dos cenários identificados.

Metodologia

A primeira parte do trabalho contemplou a análise do comportamento horário, diário e mensal do ozono (O_3) na cidade de Lisboa e área envolvente, bem como sobre as excedências aos limiares horários fixados de acordo com a legislação em vigor. As estações da qualidade do ar usadas para tal são nove (Figura 1): oito da rede de qualidade do ar da Comissão de Coordenação e Desenvolvimento Regional de Lisboa e Vale do Tejo (CCDR-LVT) (Beato, Entrecampos, Loures, Olivais e Reboleira situadas na cidade de Lisboa e Loures, Odivelas e Oeiras-Marquês localizadas na área suburbana de Lisboa) e uma estação do Instituto do Ambiente (IA) em Alfragide.

O período de dados usado para este fim foi de 3 anos (2001 a 2003). A eficiência de medição em todas as estações foi superior a 75%, excepto para o caso dos Olivais a qual teve uma eficiência de 55% (a medição de ozono em Olivais começou em Fevereiro de 2002).

A legislação nacional que regula os níveis de O₃, Decreto-Lei n.º 320/2003 que resulta da transposição da Directiva Europeia 2002/3/CE, estabelece limiares que não deverão ser excedidos:

- 180 µ.g.m⁻³ (médias horárias): designado como limiar de informação ao público;
- 120 µ.g.m⁻³ (médias octohorárias): designado limiar de protecção da saúde humana.



Figura 1 – Estações de monitorização do O₃ na região de Lisboa usadas para a análise do comportamento do O₃ (adaptado de Ferreira et al., 2004)

Numa primeira fase foram efectuadas correlações entre as concentrações de ozono e algumas variáveis meteorológicas de superfície da estação de Lisboa Gago Coutinho (assinalada pela estrela azul na Figura 1) do Instituto de Meteorologia (IM). Posteriormente foram desenvolvidos modelos estatísticos para a previsão da qualidade do ar no que se refere ao O₃ e PM₁₀.

As simulações iniciais foram realizadas com base em cinco anos de dados (1998-2002) sobre algumas variáveis meteorológicas à superfície da estação de Lisboa Gago Coutinho e concentrações de poluentes medidos nas estações da rede de qualidade do ar. Obtiveram-se modelos de regressão linear múltipla para algumas estações em diferentes períodos do ano.

Os modelos obtidos foram melhorados de acordo com as metodologias desenvolvidas para Los Angeles, Estados Unidos da América (Casmassi, 1987 e outros), recorrendo-se a um conjunto de variáveis mais alargado e diversificado. O período de dados utilizado foi diminuído para apenas dois anos (2001-2002) de modo criar uma base de dados mais consistente, sem as falhas existentes no intervalo mais longo, fazendo diminuir a capacidade de previsão dos modelos desenvolvidos. As variáveis utilizadas para a nova base de dados são apresentadas na Tabela 1. Das variáveis acrescentadas em relação ao ensaio inicial, destaca-se a diferença de pressão entre estações da rede de estações automáticas do IM que veio substituir a intensidade e rumo do vento usadas anteriormente. As variáveis em altitude são obtidas através das observações aerológicas (radiosondagens), das 12 horas UTC. Calcularam-se ainda variáveis retiradas indirectamente do perfil vertical de temperatura: a altura máxima da camada de mistura, segundo o método de Holzworth, 1964, a altura da camada de inversão e a espessura da mesma. Foi ainda efectuada uma análise qualitativa das situações sinópticas observadas à superfície e

aos 500 hPa para os dois anos acima mencionados (2001-2002), com base nas análises do modelo global de previsão do *European Centre for Medium-Range Weather Forecasting* (ECMWF). A análise permitiu identificar sete classes de três tipos de situações sinópticas distintas verificadas à superfície (Tabela 2). Em altitude, aos 500 hPa, obtiveram-se cinco classes (Tabela 3).

Tabela 1 – Variáveis utilizadas para a previsão do modelo

Variáveis diversas		Variáveis meteor. à superfície	
DW	dia semana	SSS	Situação sinóptica superfície
F	flag_tipo de dia	Tx_Lx	Temperatura Max. Lisboa
DD	Duração dia solar	vTx_Lx	Temperatura Max. Lisboa d ₁ -d ₂
Variáveis meteor. em altitude		RH_Lx	Humidade Relativa média Lisboa
S500	Situação sinóptica 500hPa	vRH_Lx	Humidade Relativa média Lisboa d ₁ -d ₂
mhML	Altura máxima camada mistura	Tx_St	Temperatura Max. Santarém
hILB	Altura da base camada inversão	vTx_St	Temperatura Max. Santarém d ₁ -d ₂
dIL	Espessura da camada inversão	HR_St	Humidade Relativa média Santarém
H1000	Alt.Geopotencial 1000hPa 12H	vHR_St	Humidade Relativa média Santarém d ₁ -d ₂
H850	Alt.Geopotencial 850hPa 12H	Tx_Ev	Temperatura máxima Évora
H700	Alt.Geopotencial 700hPa 12H	vTx_Ev	Temperatura máxima Évora d ₁ -d ₂
H500	Alt.Geopotencial 500hPa 12H	HR_Ev	Humidade Relativa média Évora
T925	Temperatura 925hPa 12H	vHR_Ev	Humidade Relativa média Évora d ₁ -d ₂
T850	Temperatura 850hPa 12H	Tx_Bj	Temperatura máxima Beja
T700	Temperatura 700hPa 12H	vTx_Bj	Temperatura máxima Beja d ₁ -d ₂
HR925	Humidade Relativa 925hPa 12H	HR_Bj	Humidade Relativa média Beja
HR850	Humidade Relativa 850hPa 12H	vHR_Bj	Humidade Relativa média Beja d ₁ -d ₂
HR700	Humidade Relativa 700hPa 12H	DP_LPo	Diferença Pressão Lisboa-Porto 12H
Variáveis da qualidade do ar		vDP_LPo	Diferença Pressão Lisboa-Porto 12H d ₁ -d ₂
PM-1	[PM ₁₀] d ₁	DP_LPa	Diferença Pressão Lisboa-Portalegre 12H
vPM	[PM ₁₀] d ₁ -d ₂	vDP_LPa	Diferença Pressão Lisboa-Portalegre 12H d ₁ -d ₂
O3-1	[O ₃] d ₁	DP_LEv	Diferença Pressão Lisboa-Évora 12H
vO3	[O ₃] d ₁ -d ₂	vDP_LEv	Diferença Pressão Lisboa-Évora 12H d ₁ -d ₂
O38-1	[O ₃] ₈ d ₁	DP_LFr	Diferença Pressão Lisboa-Faro 12H
vO38	[O ₃] ₈ d ₁ -d ₂	vDP_LFr	Diferença Pressão Lisboa-Faro 12H d ₁ -d ₂
NO-1	[NO] d ₁		
vNO	[NO] d ₁ -d ₂		
NO2-1	[NO ₂] d ₁		
vNO2	[NO ₂] d ₁ -d ₂		
CO	[CO] d ₁		
vCO	[CO] d ₁ -d ₂		

Tabela 2 – Classificação das situações sinópticas à superfície

SSS	
Sistemas frontais	1 Sistemas frontais
Sistemas depressionários	2 Depressão/instabilidade
	3 Influência de depressão
Sistemas de Altas Pressões	4 Calma à superfície
	5 Circulação de N/NW
	6 Anticiclone e vale invertido desde o Norte de Africa
	7 circulação de NE/E

Tabela 3 – Classificação das situações sinópticas aos 500hPa

SS500	
1	Depressão
2	Vale
3	Aproximação de vale
4	Crista em desenvolvimento ou fluxo zonal
5	Crista

Relativamente às variáveis da qualidade do ar, utilizaram-se as concentrações dos poluentes do dia anterior (d_{-1}), do dia a prever (d) e também a diferença entre as referentes ao último (d_{-1}) e penúltimo dia (d_{-2}). Os diferentes poluentes considerados foram, para além do O_3 e PM_{10} , o monóxido de azoto (NO), dióxido de azoto (NO_2) e o monóxido de carbono (CO).

Os dados de qualidade do ar seleccionados foram os valores de concentração máximos obtidos no conjunto das estações em análise, de acordo com o seguinte período de integração:

- PM_{10} , NO e NO_2 : concentração média diária;
- O_3 : concentração média horária;
- CO e O_3 : concentração octo-horária.

As concentrações de PM_{10} foram medidas nas estações Avenida da Liberdade, Cascais, Entrecampos, Loures, Olivais e Reboleira. Os restantes poluentes foram medidos nas estações de Entrecampos, Loures, Alfragide, Beato, Olivais e Reboleira.

Os dados serviram para testar estatisticamente novos modelos usando RM e CART (*Classification and Regression Tree*). Construiu-se primeiro um modelo de RM usando todos os dados disponíveis na base de dados e, após obtenção da CART, desenvolveram-se para cada um dos grupos finais uma RM (RM+CART) (Casmassi, 1987). Deve referir-se que a RM foi efectuada passo a passo, em modo de incrementação de variáveis, eliminando aquelas que apresentam um nível de significância de superior a $\alpha=0,10$.

Após a obtenção dos modelos foram efectuadas uma validação e uma análise de sensibilidade utilizando os dados referentes ao período de Janeiro, Fevereiro, Julho e Agosto de 2000. Este período de dados não foi utilizado no desenvolvimento do modelo. Pode então comparar-se os diversos resultados e fazer uma análise minuciosa da correlação dos modelos obtidos.

Resultados

A análise da Figura 2 mostra que o maior número de excedências de O_3 ocorre durante a tarde, tal como era esperado, uma vez que é o período mais quente do dia (86% das excedências do limiar de alerta ao público ocorrem entre as 13h00 e 17h00). O tráfego rodoviário é de facto a maior fonte de poluição atmosférica na região de Lisboa (Ferreira *et al.*, 2000). Nesta figura visualizam-se os perfis do O_3 e NO_2 destacando-se o seu comportamento inverso.

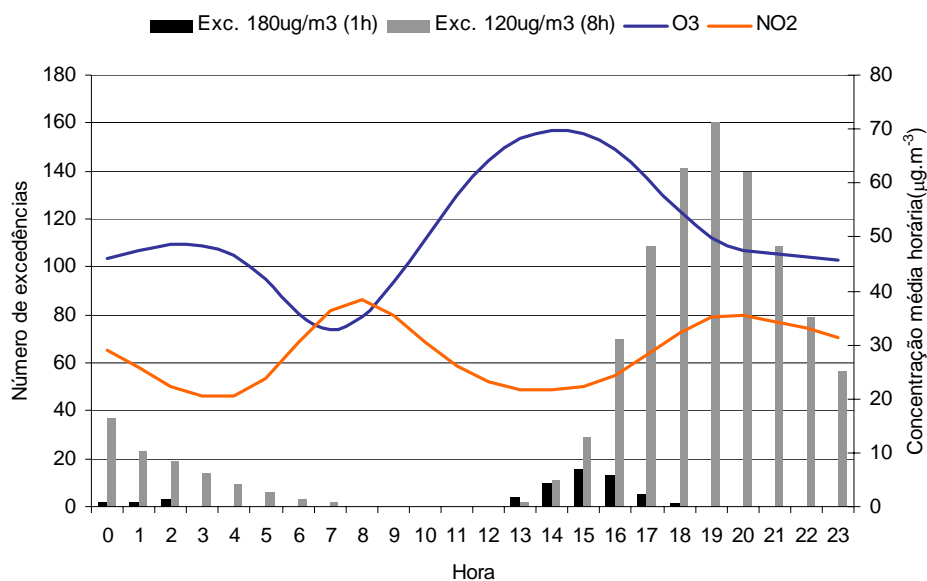


Figura 2 – Concentrações médias de O₃ e NO₂ e número de excedências horárias de O₃ para Lisboa no período de 2001-2003 (adaptado de Ferreira et al., 2004)

Na Figura 3 observa-se que o número de excedências de O₃ é superior nos meses mais quentes. No entanto, os valores médios horários mais elevados ocorrem na Primavera. Isto deve-se ao facto de amplitude diária do valor do O₃ ser menor nesta estação do ano.

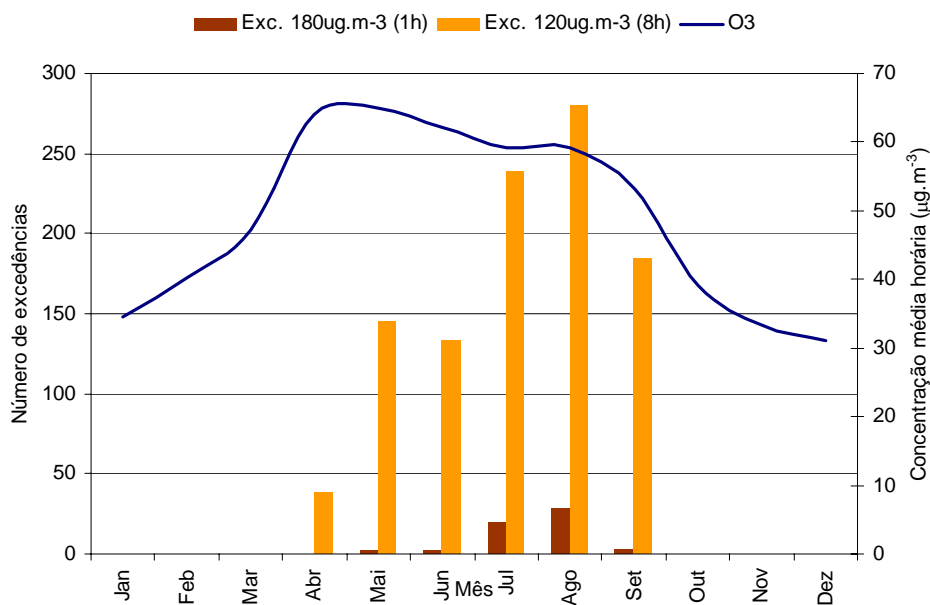


Figura 3 – Análise mensal de excedências e concentrações de O₃ para Lisboa no período de 2001-2003 (adaptado de Ferreira et al., 2004)

Na Figura 4 está representada a altura máxima da camada de mistura, constatando claramente a relação entre este parâmetro e as concentrações de O₃.

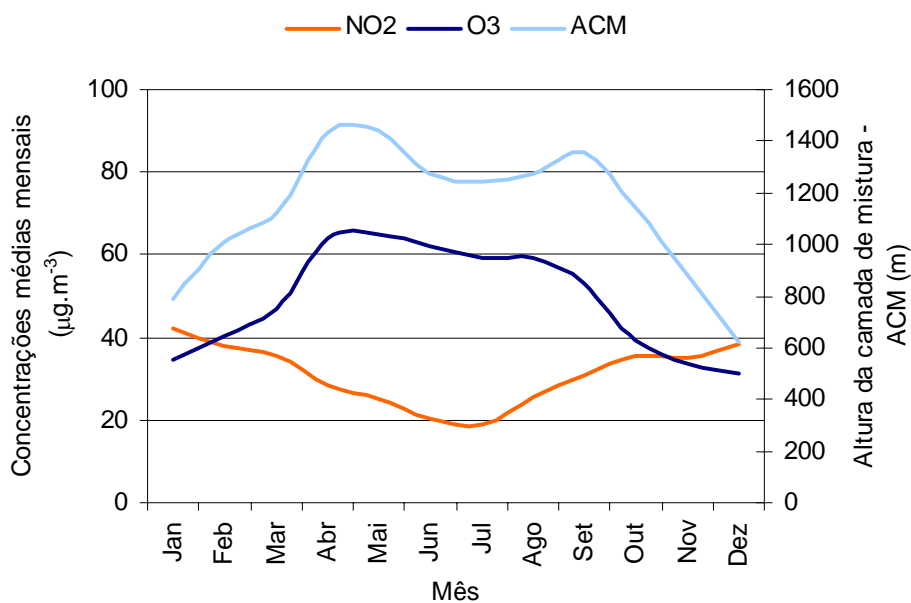


Figura 4 – Concentrações médias mensais de O₃ e NO₂ e altura máxima da camada de mistura em Lisboa (adaptado de Ferreira *et al.*, 2004)

O padrão semanal do O₃ mostra uma boa correlação com o volume de tráfego, medido indirectamente através das concentrações de CO (Figura 5) (Ferreira, 2002). As concentrações mais elevadas de O₃ ocorrem ao domingo e as mais reduzidas à sexta-feira.

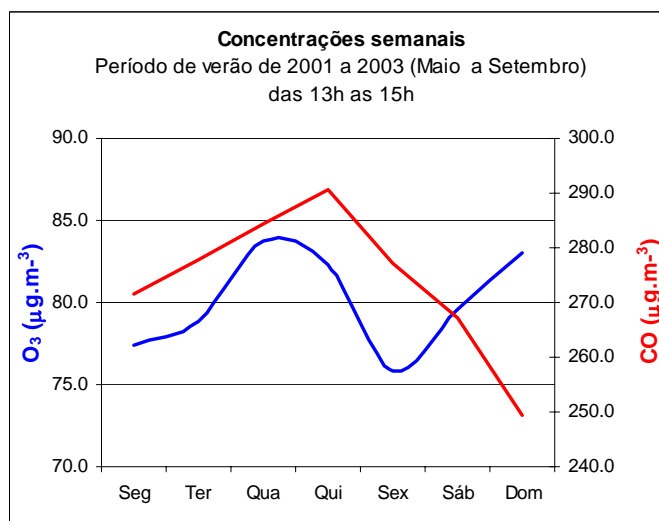


Figura 5 – Padrão semanal das concentrações do O₃ e CO em Lisboa no período de verão de 2001-2003 entre as 13 e 15 h

A Figura 6 realça uma boa relação entre a temperatura do ar à superfície (2 metros) observada e os valores de O₃, tal como era esperado.

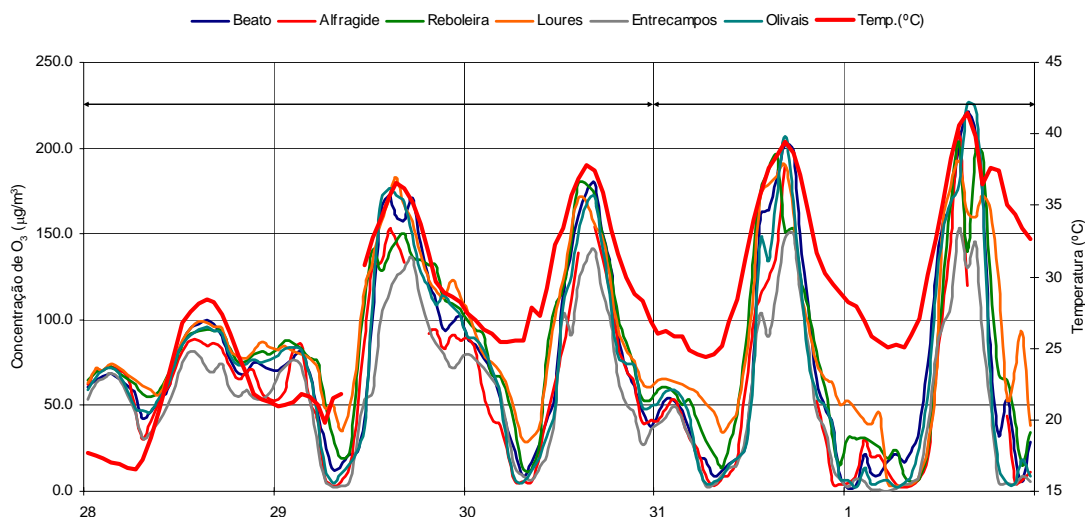


Figura 6 – Relação dos valores da temperatura do ar com concentrações de ozono em diferentes estações de Lisboa (adaptado de Ferreira et al., 2004)

Ao desenvolver modelos estatísticos e tendo construído a base de dados referente aos anos de 2001-2002, foram obtidos diversos modelos de RM para o O_3 e para PM_{10} : um obtido utilizando todos os dados e outros construídos com as hipóteses finais obtidas na CART para os dois poluentes como variáveis dependentes, tal como já foi referido na metodologia (Figura 7). Realça-se o facto de que, para a estimação da concentração de O_3 , a temperatura máxima em Santarém acaba por ter um papel mais importante que a temperatura máxima de Lisboa. Este resultado parece ser um indicador de que a massa de ar, quando proveniente do quadrante leste, e portanto quente no Verão, está associado a valores altos de O_3 .

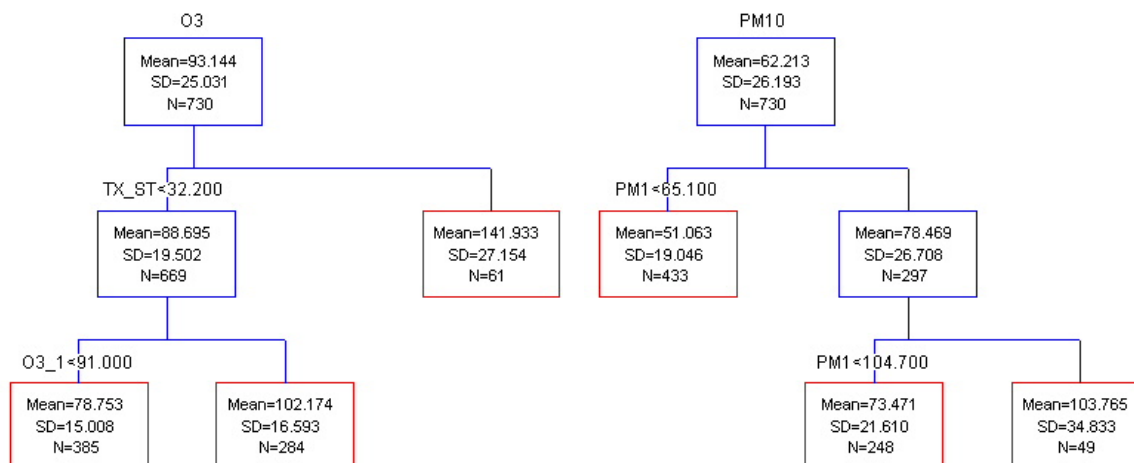


Figura 7 – Modelos de regressão em árvore obtidos para o O_3 e PM_{10} do período de 2001-2002

Para o modelo do O_3 , outra variável encontrada é o valor do dia anterior. Relativamente aos resultados obtidos com PM_{10} , a única variável encontrada com maior importância foi o valor do dia anterior. Não se encontrou discriminação de dados com base em variáveis meteorológicas.

A variância explicada pelos modelos obtidos (Tabela 4 e Tabela 5), variam entre 97.9 e 98.9% para o O₃ e 93.6 e 96.9% para as PM₁₀.

Tabela 4 – Resultados obtidos para os diversos modelos de regressão múltipla para o O₃

O ₃	Geral	Tx_St<32.2 e O3-1<91	Tx_St<32.2 e O3-1>=91	Tx_St>=32.2
R	0.990	0.989	0.994	0.996
R²	0.990	0.979	0.989	0.992
Estimação erro padrão	13.819	11.805	11.269	13.871
N	729	385	282	60

Os resultados dos modelos obtidos são considerados bons, não sendo apresentadas as equações correspondentes aos seis modelos seleccionados (três para o ozono e três para partículas, de acordo com a discriminação em árvore apresentada), porque tal foi considerado demasiado extenso em termos de apresentação.

Tabela 5 – Resultados obtidos para os diversos modelos de regressão múltipla para os PM₁₀

PM ₁₀	Geral	PM_1<65.1	65.1<=PM_1<104.7	PM_1>=104.7
R	0.971	0.968	0.978	0.984
R²	0.942	0.936	0.956	0.969
Estimação erro padrão	16.379	13.895	16.463	21.569
N	727	432	246	49

Deve salientar-se que, para ambos os poluentes, a variável que representa a situação sinóptica à superfície (G) é apenas usada no modelos de previsão para ambos os poluentes quando as concentrações são mais elevadas. Ao contrário do que era esperado, os modelos parecem não atribuir muito peso aos efeitos dessa variável. A introdução da variável G para todas as hipóteses poderá passar pela construção de modelos CART com essa variável. Outra característica obtida para o estudo de O₃ como variável dependente no último modelo é que, ao contrário de outras, as variáveis meteorológicas em altitude não são utilizadas. Já para as PM₁₀, refira-se que as variáveis meteorológicas em altitude são muito mais utilizadas do que para os modelos de O₃.

A Figura 8 e a Figura 9 mostram os resultados obtidos através da utilização dos modelos no período de 2 de Julho a 31 de Agosto de 2000, para ambos os poluentes estudados. Os coeficientes de correlação obtidos entre as observações e os resultados dos modelos são apresentados na O coeficiente de correlação mais elevado para o O₃ foi de 0.850. Para o PM₁₀ o coeficiente de correlação mais elevado é de apenas 0.740. O modelo para PM₁₀ apresenta uma maior dificuldade em prever grandes alterações de concentrações ao contrário do O₃. Isto pode dever-se ao facto da concentração de O₃ estar muito condicionada pela temperatura, enquanto as PM₁₀ não têm uma relação forte com nenhuma variável meteorológica. Daí que o estudo realizado das situações meteorológicas de forma mais integrada possa vir a ser fundamental para o desenvolvimento de modelos mais eficazes na previsão da concentração de PM₁₀.

Tabela 6.

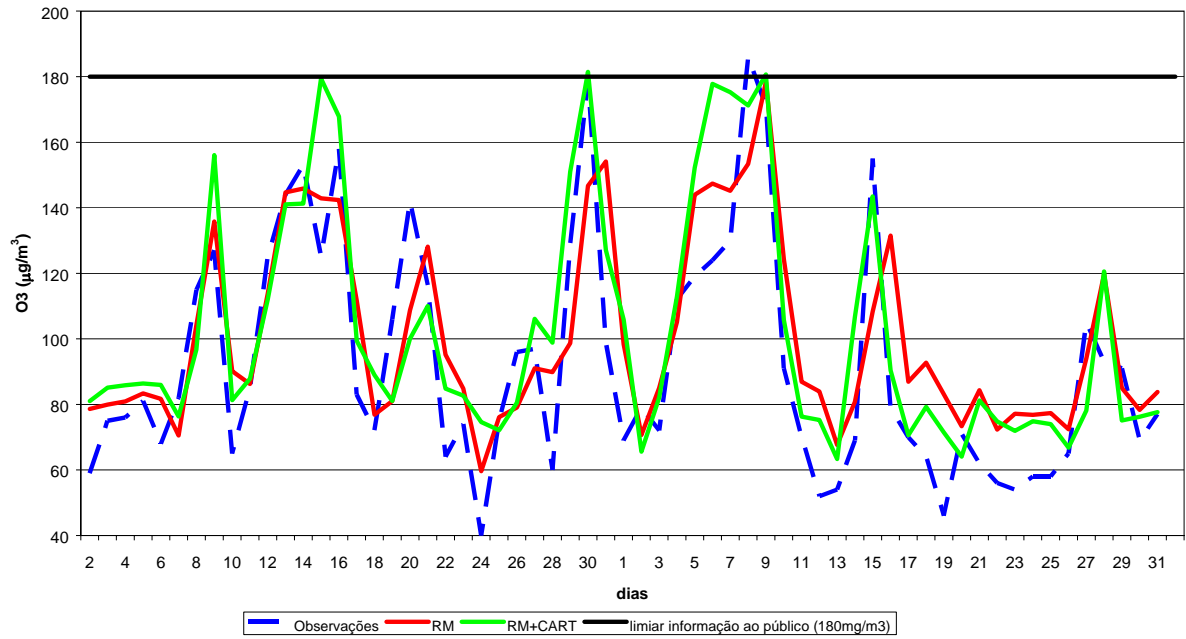


Figura 8 - Valores observados e simulados de concentrações máximas de O₃ usando os modelos de regressão múltipla e regressão em árvore+regressão múltipla (2 de Julho a 31 de Agosto de 2000)

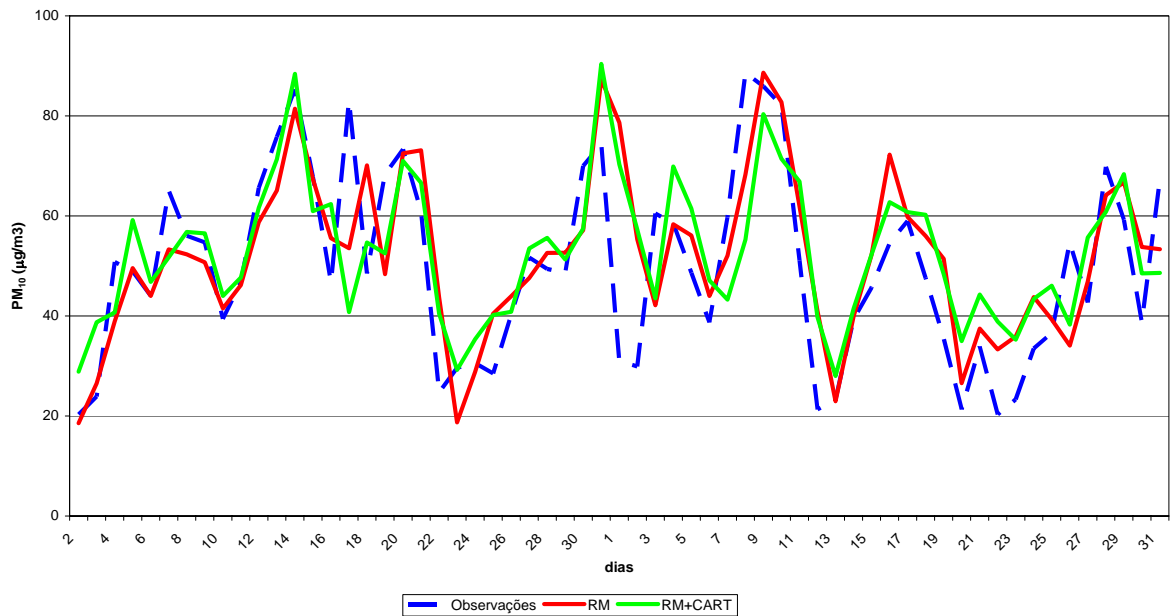


Figura 9 - Valores observados e simulados de concentrações médias de PM₁₀ usando os modelos de regressão múltipla e regressão em árvore+regressão múltipla (2 de Julho a 31 de Agosto de 2000)

O coeficiente de correlação mais elevado para o O₃ foi de 0.850. Para o PM₁₀ o coeficiente de correlação mais elevado é de apenas 0.740. O modelo para PM₁₀ apresenta uma maior dificuldade em prever grandes alterações de concentrações ao contrário do O₃. Isto pode dever-se ao facto da concentração de O₃ estar muito condicionada pela temperatura, enquanto as PM₁₀ não têm uma relação forte com nenhuma variável meteorológica. Daí que o estudo realizado das situações meteorológicas de forma mais integrada possa vir a ser fundamental para o desenvolvimento de modelos mais eficazes na previsão da concentração de PM₁₀.

Tabela 6 – Coeficientes de correlação entre os dados observados e os diferentes resultados dos modelos

R	RM	RM+CART
O₃	0.811	0.850
PM₁₀	0.740	0.685

Foram igualmente desenvolvidos outros métodos estatísticos como a análise de clusters e análise de componentes principais que não são aqui apresentados, apesar de terem sido importantes para um melhor conhecimento das variáveis. Desenvolvimentos futuros, como a utilização de redes neuronais, poderão revelar-se métodos eficazes na determinação de melhores modelos de previsão (Wang *et al.*, 2003).

Conclusões

O presente estudo mostra que a utilização de modelos de regressão linear múltipla desenvolvidos no quadro duma estrutura em árvore, como forma de melhor perceber as relações entre variáveis de qualidade do ar e meteorológicas e permitir a previsão de concentrações de ozono e partículas (PM₁₀) para a área de Lisboa, tem resultados muito satisfatórios. A variância explicada pelos modelos é muito elevada e sempre superior a 85%, permitindo a sua aplicação em termos de previsão. A aplicação dos modelos num período de tempo diferente do usado para o estabelecimento das regressões apresenta correlações inferiores, cujos valores são no entanto suficientemente significativos para demonstrar a robustez da metodologia aplicada.

Referências

Casmassi, J. (1987): Development of an objective ozone forecast model for the south coast air basin. 80th Annual Meeting of APCA (The Association Dedicated to Air Pollution Control and Hazardous Waste Management). New York – June 21-26, 1987.

ECMWF (2003): Analyses Charts, European Centre for Medium-Range Weather Forecasts, 2000 a 2002.

EPA (2003): Guidelines for Developing an Air Quality (Ozone and PM_{2.5}) Forecasting Program, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, North Carolina, EPA-456/R-03-002.

Ferreira, F., Tente, H., Torres, P., Cardoso, S. & Palma-Oliveira, J.M. (2000): Air Quality Monitoring and Management in Lisbon, Environmental Monitoring and Assessment 2000, 65, pp. 443-450.

Ferreira, F., Torres, P., Neto, J. & Tente, H. (2004): Ozone Levels in Portugal: the Lisbon Region Assessment. Em Proceedings of Air & Waste Management's 97th Annual Conference & Exhibition. June 22-25, 2004, Indianapolis, Indiana. CD-ROM, pp. 18.

Holzworth, G.C. (1967): Mixing depths, wind speeds and air pollution potential for selected locations in the United States, *J. Appl. Meteor.*, 1967, 5, 1039-1044.

IM (2003): Boletim diário, Instituto de Meteorologia. 2000 a 2002

Wang, W., Weizhen, L., Wang, X., Leung, A. (2003): Prediction of maximum daily ozone level using combined neural network and statistical characteristics. *Environment International* 29, pp. 555–562.