



***MODELAÇÃO DA DISTRIBUIÇÃO ESPACIAL
DA QUALIDADE DO AR EM LISBOA USANDO
SISTEMAS DE INFORMAÇÃO GEOGRÁFICA***

Sandra Maria Pereira Mesquita

Dissertação apresentada como requisito parcial para obtenção do
grau de Mestre em Ciência e Sistemas de Informação Geográfica

Instituto Superior de Estatística e Gestão de Informação
da Universidade Nova de Lisboa

MODELAÇÃO DA DISTRIBUIÇÃO ESPACIAL DA QUALIDADE DO AR EM LISBOA USANDO SISTEMAS DE INFORMAÇÃO GEOGRÁFICA

Dissertação orientada por:

Professor Doutor Francisco Manuel Freire Cardoso Ferreira (Orientador)

Professora Doutora Maria Júlia Fonseca Seixas (Co-Orientador)

Novembro 2009

AGRADECIMENTOS

Esta dissertação não teria sido possível sem as oportunidades de trabalho na área da poluição atmosférica que me foram concedidas pelo Professor Doutor Francisco Ferreira, orientador da presente tese, no âmbito de protocolos existentes entre o Departamento de Ciências e Engenharia do Ambiente da Faculdade de Ciência e Tecnologia da Universidade Nova de Lisboa (DCEA/FCT/UNL) e diversas instituições públicas como a Agência Portuguesa do Ambiente (APA) e a Comissão de Coordenação e Desenvolvimento Regional da Região de Lisboa e Vale do Tejo (CCDR-LVT).

O apoio prestado pela co-orientadora Professora Doutora Júlia Seixas, minha professora de licenciatura e pós-graduação, em termos técnicos e de metodologias de investigação constituíram um pilar sem o qual a presente tese não teria sido possível.

Agradeço aos técnicos das diversas instituições que disponibilizaram dados para este trabalho em particular à Eng. Luísa Nogueira cujo apoio técnico e contínua disponibilidade tem sido fundamental para o aprofundamento dos conhecimentos de todos quantos temos trabalhado com ela.

Agradeço ao Eng. Victor Góis, Eng. Pedro Torres e Eng. Hugo Maciel o desenvolvimento de um extenso trabalho relativo ao inventário de emissões para a LVT e sua desagregação espacial, sem o qual este trabalho ficaria definitivamente mais pobre.

Todo o grupo de trabalho da área da qualidade do ar do DCEA contribuiu de alguma forma com o seu trabalho de recolha e tratamento de dados, conhecimentos teóricos e técnicos e troca de experiências, para o desenvolvimento do presente trabalho. Assim agradeço aos meus colegas Ana Cristina Carvalho, Nuno Grosso, Joana Monjardino, Pedro Gomes, Filipa Marques e Hugo Tente. Agradeço também aos colegas da área da detecção remota e alterações climáticas, Nuno Pacheco, Nuno Carvalhais e João Pedro Nunes, sempre disponíveis para dar uma ajuda mais tecnológica e descobrir novas fontes de dados georreferenciados.

Finalmente agradeço o incentivo de toda a minha família, em particular do Daniel, cujo apoio na utilização de ferramentas SIG foi fundamental para o desenvolvimento do presente trabalho.

MODELAÇÃO DA DISTRIBUIÇÃO ESPACIAL DA QUALIDADE DO AR EM LISBOA USANDO SISTEMAS DE INFORMAÇÃO GEOGRÁFICA

RESUMO

Com a presente dissertação pretendeu-se demonstrar a importância e potencialidade dos Sistemas de Informação Geográfica (SIG) na avaliação da poluição atmosférica. Usando ferramentas SIG, foram aplicadas e comparadas várias metodologias para modelar a distribuição espacial anual do poluente atmosférico dióxido de azoto (NO_2) na cidade de Lisboa. Foram usados dados de NO_2 da rede fixa de estações de monitorização de qualidade do ar e de duas campanhas com cerca de 100 tubos de difusão realizadas em 2001 e 2002 aos quais foram aplicadas técnicas de interpolação mais comuns como o *Inverse Distance Weighting (IDW)*, *Ordinary Kriging* e *Ordinary CoKriging*, e ainda a Regressão Linear Múltipla (RLM). Como variáveis independentes para a RLM e *CoKriging* foram usados mapas de densidade populacional, ocupação do solo, topografia e inventário de emissões. A validação e comparação dos vários métodos testados, permitiu concluir que o método que produziu mapas mais realistas e com menor erro foi a RLM, recorrendo às estatísticas de emissões de tráfego e em área. Usando os mapas obtidos com RLM foi ainda possível estimar, em ambiente SIG, a população e áreas expostas a vários níveis de poluição.

MODELLING THE SPATIAL DISTRIBUTION OF AIR QUALITY IN LISBON USING GEOGRAPHIC INFORMATION SYSTEMS

ABSTRACT

The present dissertation aimed to demonstrate the importance and potential of Geographic Information Systems (GIS) in the evaluation of air pollution. GIS tools were applied to compare several methodologies to model the annual spatial distribution of nitrogen dioxide (NO_2) concentrations in the Lisbon city. NO_2 data from monitoring stations network and two one week diffusive sampling campaigns performed in 2001 and 2002 were used to apply interpolation widespread techniques like Inverse Distance Weighting (*IDW*), ordinary Kriging e ordinary CoKriging, and also the more recent multiregression mapping method. As independent data for multiregression mapping method and CoKriging population density, traffic emissions, land use and topography maps were tested. The validation and comparison of the tested methods identified multiregression mapping method, using traffic and area emissions statistics, as the method producing the most realistic maps and with less error. This maps allowed to estimate, in GIS environment, the population and areas exposed to several levels of pollution.

PALAVRAS- CHAVE

Dióxido de azoto
Geoestatística
Interpolação
Lisboa
Mapeamento
Modelação espacial
Monitorização
Qualidade do ar
Regressão linear

KEYWORDS

Nitrogen dioxide
Geostatistics
Interpolation
Lisbon
Mapping
Spatial modelling
Monitoring
Air quality
Linear regression

ACRÓNIMOS

CoKrig – *Ordinary co-Kriging*

IDW – *Inverse Distance Weighting*

Krig – *Ordinary Kriging*

MDP – Método de Difusão Passiva – tubos de difusão

MR – Método de Referência

Mt – Margem de tolerância

NO₂ – Dióxido de azoto

NO_x – Óxidos de azoto

R – Coeficiente de correlação

R² – Coeficiente de determinação

RLM – Regressão Linear Múltipla

RMSE – *Root Mean Square Error (Raiz quadrada da média do quadrado dos erros)*

SIG – Sistemas de Informação Geográfica

VL – Valor limite

ÍNDICE DO TEXTO

AGRADECIMENTOS	i
RESUMO	ii
ABSTRACT	iii
PALAVRAS- CHAVE	iv
KEYWORDS	iv
ACRÓNIMOS	v
ÍNDICE DE FIGURAS	viii
ÍNDICE DE TABELAS	xi
1 INTRODUÇÃO	1
2 OBJECTIVOS	4
3 ENQUADRAMENTO	5
3.1 QUALIDADE DO AR: QUÍMICA NA ATMOSFERA, FONTES E EFEITOS NA SAÚDE HUMANA E NO AMBIENTE ATMOSFÉRICO DO DIÓXIDO DE AZOTO.....	5
3.2 QUADRO LEGISLATIVO	7
3.2.1 <i>Valores legais estabelecidos para o NO₂</i>	8
3.2.2 <i>Requisitos de avaliação, ferramentas e objectivos de qualidade</i>	9
3.2.3 <i>Verificação do cumprimento da legislação</i>	11
3.3 MAPEAMENTO DAS CONCENTRAÇÕES DE NO ₂	12
3.3.1 <i>Medição</i>	13
3.3.2 <i>Modelação</i>	15
3.3.3 <i>Interpolação</i>	16
3.4 ESTIMATIVA DAS ÁREAS E POPULAÇÃO EXPOSTAS À POLUIÇÃO ATMOSFÉRICA.....	26
3.5 AVALIAÇÃO DA INCERTEZA	27
3.6 UTILIZAÇÃO DE FERRAMENTAS DE SIG NA AVALIAÇÃO DA QUALIDADE DO AR E ESTIMATIVA DA POPULAÇÃO EXPOSTA.....	29
4 METODOLOGIA	31
4.1 PRÉ-TRATAMENTO DE DADOS DE NO ₂	31
4.1.1 <i>Estações de monitorização fixa</i>	32
4.1.2 <i>Campanhas de monitorização</i>	32
4.1.3 <i>Obtenção de indicadores anuais extrapolando os dados das campanhas usando regressão linear simples</i>	34
4.2 PRÉ-TRATAMENTO DE VARIÁVEIS AUXILIARES.....	35
4.3 MODELAÇÃO DA DISTRIBUIÇÃO ESPACIAL	37
4.3.1 <i>Inverso da distância ponderada (IDW)</i>	37
4.3.2 <i>Ordinary Kriging</i>	37
4.3.3 <i>Ordinary CoKriging</i>	37
4.3.4 <i>Regressão linear múltipla</i>	38
4.4 AVALIAÇÃO DA INCERTEZA	40
4.4.1 <i>Avaliação da incerteza das campanhas de monitorização</i>	41
4.4.2 <i>Avaliação da incerteza na obtenção de indicadores anuais</i>	41
4.4.3 <i>Avaliação da incerteza da modelação da distribuição espacial</i>	42
4.4.4 <i>Seleção do melhor mapa e produção de mapas de erro</i>	43
4.5 ESTIMATIVA DAS ÁREAS E DA POPULAÇÃO EXPOSTA A POLUIÇÃO ATMOSFÉRICA	44
4.6 DIAGRAMA GERAL: OPERAÇÕES EM AMBIENTE SIG.....	44
5 RESULTADOS E DISCUSSÃO	47
5.1 PRÉ-TRATAMENTO DE DADOS DE NO ₂	47

5.1.1	<i>Estações de monitorização fixa</i>	47
5.1.2	<i>Campanhas de monitorização</i>	49
5.1.3	<i>Obtenção de indicadores anuais extrapolando os dados das campanhas usando regressão linear simples</i>	54
5.2	PRÉ-TRATAMENTO DAS VARIÁVEIS AUXILIARES.....	60
5.2.1	<i>Ocupação do solo</i>	60
5.2.2	<i>Emissões de tráfego rodoviário</i>	62
5.2.3	<i>Emissões em área</i>	63
5.3	MODELAÇÃO DA DISTRIBUIÇÃO ESPACIAL.....	65
5.3.1	<i>Inverso da distância ponderada (IDW)</i>	65
5.3.2	<i>Ordinary Kriging</i>	67
5.3.3	<i>Ordinary CoKriging</i>	69
5.3.4	<i>Regressão linear múltipla</i>	70
5.3.5	<i>Comparação e avaliação da incerteza dos métodos de interpolação usados</i> 73	
5.4	ESTIMATIVA DAS ÁREAS E DA POPULAÇÃO EXPOSTA A POLUIÇÃO ATMOSFÉRICA.....	84
5.5	UTILIZAÇÃO DE FERRAMENTAS DE SIG.....	86
6	CONCLUSÕES E DESENVOLVIMENTOS FUTUROS	87
	REFERÊNCIAS BIBLIOGRÁFICAS	91
	ANEXOS	100
1	PARÂMETROS DE MODELAÇÃO DA DISTRIBUIÇÃO ESPACIAL	101

ÍNDICE DE FIGURAS

Figura 1. Diagrama simplificado do ciclo dos compostos de azoto (CE, 1997).....	7
Figura 2. Representação esquemática do enquadramento legislativo da avaliação e gestão qualidade do ar ambiente na UE e em Portugal.	8
Figura 3. Método de difusão passiva (FSM, 2007).	15
Figura 4. Determinação do valor da função expoente (<i>Power</i>) óptimo (Johnston et al., 2001)	19
Figura 5. Forma característica de um semivariograma (Johnston et al., 2001).....	21
Figura 6. Exemplos de modelos matemáticos aplicados a variogramas (Johnston et al., 2001)	22
Figura 7. Diagrama geral da metodologia	31
Figura 8. Tubos de difusão standard da Radiello	32
Figura 9. Localização dos pontos de monitorização, estações fixas e tubos de difusão em Lisboa.....	33
Figura 10. Colocação de tubos de difusão em triplicado junto a cada estação de monitorização da qualidade do ar fixa para controlo de qualidade	34
Figura 11. Anéis usados para calcular as estatísticas das variáveis auxiliares na envolvente dos locais amostrados na cidade de Lisboa durante a campanha com tubos de difusão.....	36
Figura 12. Grelha de 100m aplicada à cidade de Lisboa.....	39
Figura 13. Esquema geral para a obtenção e selecção dos mapas de distribuição das concentrações	45
Figura 14. Esquema geral para a obtenção das áreas e população expostas a determinados níveis de poluição partindo dos mapas de distribuição das concentrações seleccionados.....	46
Figura 15. Verificação do cumprimento da legislação relativa ao NO ₂ dentro da Região de Lisboa e cidade de Lisboa	48
Figura 16. Distribuição das emissões de NO ₂ (média 2001-2002) por sector de actividade na cidade de Lisboa (Gois et al., 2006).....	49
Figura 17. Concentrações médias da semana de 03/07 a 10/07 (campanha Verão 2001). 50	
Figura 18. Concentrações médias da semana de 29/01 a 05/02 (campanha Inverno 2002)	50
Figura 19. Regressão linear estabelecida para o local das estações fixas entre a média das campanhas usando tubos de difusão (MDP) e a média anual 2001-2002 medida usando o método de referência (MR)	57
Figura 20. Regressão linear estabelecida, para o local das estações fixas, entre a média das campanhas usando tubos de difusão (MDP) e o 19º máximo horário usando o método de referência (MR)	57
Figura 21. Estimativa da média das médias anuais de 2001-2002 para os pontos de amostragem da campanha MDP usando a relação linear estabelecida na Figura 19	57
Figura 22. Estimativa do 19º máximo horário (média dos anos 2001-2002) para os pontos de amostragem da campanha MDP usando a relação linear estabelecida Figura 20.....	58
Figura 23. Anéis usados para extrair as estatísticas das variáveis auxiliares da vizinhança dos pontos de amostragem a diferentes distâncias (exemplo: emissões em área e de tráfego).....	60
Figura 24. Ocupação do solo da cidade de Lisboa (CORINE <i>Land cover</i> 2000), média anual de NO ₂ e vias de tráfego	61
Figura 25. Variação do coeficiente de correlação da relação linear, entre as concentrações do NO ₂ e da contagem do número de células de cada tipo de ocupação, dentro de cada anel de influência.	62
Figura 26. Emissões anuais de NO _x resultantes do tráfego rodoviário em Lisboa desagregadas ao nível da via (Gois et al., 2006)	62

Figura 27. Variação do coeficiente de determinação da relação linear, entre as concentrações do NO ₂ e das estatísticas das emissões de tráfego dentro de cada anel de influência	63
Figura 28. Emissões de NO ₂ em área em Lisboa desagregadas para o nível da subsecção estatística (Gois et al., 2006; INE, 2001)	64
Figura 29. Variação do coeficiente de correlação da relação linear entre as concentrações do NO ₂ e as estatísticas das emissões em área para cada anel de influência	65
Figura 30. Interpolação da média de NO ₂ durante de 2001 e 2002 usando <i>IDW</i> . Parâmetros do método e resultados da validação cruzada incluindo todos os pontos	66
Figura 31. Interpolação do 19º máximo horário de NO ₂ médio de 2001 e 2002 usando o <i>IDW</i> . Parâmetros do método e resultados da validação cruzada incluindo todos os pontos	67
Figura 32. Interpolação da média de NO ₂ durante os anos de 2001 e 2002 usando o método <i>ordinary kriging</i> e estatísticas de erro da previsão fazendo validação cruzada	68
Figura 33. Interpolação do 19º máximo horário de NO ₂ médio de 2001 e 2002 usando <i>ordinary kriging</i> . e estatísticas de erro da previsão fazendo validação cruzada	68
Figura 34. Interpolação da média anual de NO ₂ para a média de 2001 e 2002 usando <i>cokriging</i> com as variáveis auxiliares: soma das emissões NO _x devidas ao tráfego rodoviário na envolvente de 50m, entre 50 e 100m e entre 100 e 250m	69
Figura 35. Interpolação do 19º máximo horário de NO ₂ médio de 2001 e 2002 usando <i>ordinary cokriging</i> com as variáveis auxiliares: soma das emissões derivadas do NO _x do tráfego na envolvente de 50m, entre 50 e 100m e entre 100 e 250m	70
Figura 36. Interpolação da média de NO ₂ em 2001 e 2002 usando RLM com as variáveis auxiliares e estatísticas de erro da validação cruzada	72
Figura 37. Interpolação do 19º máximo horário de NO ₂ médio em 2001 e 2002 usando RLM com as variáveis auxiliares e estatísticas de erro da validação cruzada	72
Figura 38. Ajustamento dos modelos de interpolação aos dados obtidos da estimativa dos resultados das campanhas para a média anual, usando os resultados das estações fixas	74
Figura 39. Ajustamento dos modelos de interpolação aos dados obtidos da estimativa dos resultados das campanhas para o 19º máximo horário, usando os resultados das estações fixas	75
Figura 40. Ajustamento dos resultados dos modelos obtidos através da validação cruzada com os valores reais de cada uma das estações (média anual)	75
Figura 41. Ajustamento dos resultados dos modelos obtidos através da validação cruzada com os valores reais de cada uma das estações (19º máximo horário)	75
Figura 42. Erro global obtido para as estações fixas usando cada um dos métodos de interpolação (média anual)	80
Figura 43. Erro global obtido para as estações fixas usando cada um dos métodos de interpolação (19º máximo horário)	80
Figura 44. Valor real da média anual (2001-2002) de NO ₂ nas 6 estações fixas comparado com os resultados da validação cruzada para cada um dos métodos de interpolação	80
Figura 45. Valor real do 19º máximo horário (2001-2002) de NO ₂ nas 6 estações fixas comparado com os resultados da validação cruzada para cada um dos métodos de interpolação	81
Figura 46. Concentrações de NO ₂ , expressas pela média anual, (a) e respectivos mapas de erro (b) fazendo validação cruzada para os quatro métodos de interpolação em análise	82
Figura 47. Concentrações de NO ₂ , expressas pelo 19º máximo horário, e respectivos mapas de erro fazendo validação cruzada para os quatro métodos de interpolação em análise	83
Figura 48. Extracção das grelhas de 100m que o modelo de RLM estimou estarem acima do VL e ao VL+MT para a média anual	85
Figura 49. Extracção das grelhas de 100m que o modelo de RLM estimou estarem acima do VL para o 19º máximo horário	85

Figura 50. Sobreposição das grelhas de 100m, que o modelo de RLM estimou estarem acima do VL ($40\mu\text{g}/\text{m}^3$) e ao VL+MT ($56\mu\text{g}/\text{m}^3$) para a média anual, com o mapa da densidade populacional 86

ÍNDICE DE TABELAS

Tabela 1. Níveis e valores definidos na legislação para NO ₂ para a protecção da saúde humana definidos (DL 111/2002 e Directiva 2008/50/CE).....	9
Tabela 2. Objectivos de qualidade para a avaliação do dióxido de azoto (NO ₂) (Directiva 2008/50/CE).	10
Tabela 3. Tipo de modelo ou processos requeridos normalmente utilizados para cada escala e para o NO ₂ (Denby et al., 2008).	16
Tabela 4. Variáveis auxiliares usadas.....	35
Tabela 5. Estatísticas das variáveis auxiliares usadas como variáveis independentes.	38
Tabela 6. Avaliação do cumprimento do DL n.º 111/2002 para o NO ₂ relativa à protecção da saúde humana em 2001 e 2002	48
Tabela 7. Avaliação da incerteza associada aos dados de NO ₂ obtidos na campanha de Verão 2001 usando os indicadores precisão, exactidão e incerteza global.	52
Tabela 8. Avaliação da incerteza associada aos dados de NO ₂ obtidos na campanha de Inverno 2002 usando os indicadores precisão, exactidão e incerteza global.....	53
Tabela 9. Resultados de NO ₂ nas 9 estações fixas durante o período das campanhas (Verão 2001 e Inverno 2002) obtidos usando tubos de difusão e os equipamentos das estações, e o período anual (2001 e 2002) usando os equipamentos das estações.	54
Tabela 10. Coeficientes de determinação do ajustamento das regressões lineares, para estações existentes, entre os resultados de NO ₂ usando tubos de difusão (MDP) durante o período das campanhas (Verão 2001, Inverno 2002 e média das duas) e os equipamentos das estações fixas (MR) no período anual (para 2001, 2002 e média dos dois).....	56
Tabela 11. Cálculo do erro da estimativa dos resultados da campanha para média anual média dos anos 2001 e 2002 usando a validação cruzada.....	59
Tabela 12. Cálculo do erro da estimativa dos resultados da campanha para o 19º máximo horário médio dos anos 2001 e 2002 usando a validação cruzada	59
Tabela 13. Coeficiente de correlação (R) da relação linear entre as concentrações do NO ₂ e da contagem do número de células dentro de cada anel relativamente à ocupação do solo	61
Tabela 14. Coeficiente de correlação (R) da relação linear entre as concentrações do NO ₂ e as estatísticas para cada anel de influência relativamente às emissões do tráfego rodoviário.....	63
Tabela 15. Coeficiente de correlação (R) da relação linear entre as concentrações do NO ₂ e as estatísticas para cada anel de influência relativamente às emissões em área	64
Tabela 16. Coeficientes do modelo de regressão linear múltipla obtido para a média anual e 19º máximo horário (método <i>forward</i>).....	71
Tabela 17. Sumário do ajustamento do modelo de regressão linear múltipla obtido para a média anual e para o 19º máximo horário (método <i>forward</i>).....	71
Tabela 18. Resultados da validação cruzada por estação para a média anual. Avaliação de cada um dos métodos de interpolação isoladamente (erro) e em conjunto com a estimativa dos resultados da campanha (erro global).	78
Tabela 19. Resultados da validação cruzada por estação para o 19º máximo horário. Avaliação de cada um dos métodos de interpolação isoladamente (erro) e em conjunto com a estimativa dos resultados das campanhas (erro global).....	79
Tabela 20. Áreas e população em excelência ao VL e VL+MT anual e para o VL horário ..	85
Tabela 21. Ajustamento dos resultados dos métodos de interpolação e estatísticas de erro usando a técnica de validação cruzada relativamente aos valores reais obtidos nas seis estações de monitorização – Média anual	88
Tabela 22. Ajustamento dos resultados dos métodos de interpolação e estatísticas de erro usando a técnica de validação cruzada relativamente aos valores reais obtidos nas seis estações de monitorização – 19º Máximo horário.....	88

1 Introdução

Actualmente, na generalidade dos países europeus, a maior fonte de poluição atmosférica são as emissões de tráfego rodoviário. Os veículos com motores a gasolina ou a diesel emitem uma variedade de poluentes, principalmente monóxido de carbono (CO), óxidos de azoto (NO_x), compostos orgânicos voláteis (COVs) e partículas em suspensão, que têm um crescente impacte na qualidade do ar em áreas urbanas. A poluição provocada pelo tráfego rodoviário tem vindo a piorar um pouco por todo o mundo. Ao contrário do que se verifica em relação ao impacte na qualidade do ar das fontes poluentes domésticas e industriais que tende a manter-se ou a melhorar.

De acordo com a Estratégia Temática sobre poluição do ar de Comissão Europeia (CE, 2005) na Europa ocorrem anualmente várias centenas de milhar de mortes prematuras, aumento das admissões hospitalares, medicação extra, e milhões de dias de trabalho perdidos causadas pela poluição atmosférica. Deste modo, os custos dos cuidados de saúde decorrentes da poluição atmosférica são muito elevados dentro da União Europeia. Os poluentes que acarretam maiores preocupações para a saúde humana são as partículas ("poeiras finas") transportadas pelo ar e o ozono troposférico – na verdade, não foram ainda identificados, para ambos, níveis abaixo dos quais não se verifiquem danos para a saúde. A exposição a estes poluentes pode conduzir a impactes que variam de perturbações ligeiras no sistema respiratório à morte prematura.

O dióxido de azoto, apesar de não ser o poluente que maior preocupação acarreta em termos da saúde humana, pontualmente, em áreas urbanas, ultrapassa os limites legais. É também, bastante útil para a avaliação da qualidade do ar uma vez que é considerado um bom marcador da poluição relacionada com o tráfego rodoviário (Briggs, 1997; WHO, 2004).

A avaliação da qualidade do ar em todo o território, principalmente em áreas urbanas, é desde 1996, uma das solicitações das directivas de qualidade do ar publicadas pela Comissão Europeia (Directiva 96/62/CE). A elaboração de mapas anuais de distribuição das concentrações dos poluentes atmosféricos é uma necessidade, nomeadamente para conhecer a representatividade espacial da rede monitorização de qualidade do ar existente, e avaliar e estimar a exposição da população e dos ecossistemas à poluição atmosférica. Os mapas de qualidade do ar são importantes para qualquer avaliação espacial dos efeitos da qualidade do ar designadamente no que diz respeito à exposição humana e dos ecossistemas (Horálek et al., 2005).

A tarefa mais importante de um sistema de informação de qualidade do ar é oferecer a informação mais completa acerca desta numa dada região. A informação relativa à distribuição regional de poluentes atmosféricos pode ser obtida por modelação, que dá uma

boa cobertura espacial, mas tem associada uma incerteza elevada, ou por medições, cuja cobertura espacial é mais escassa mas que se assume terem maior certeza. Tradicionalmente a avaliação da qualidade do ar é baseada nos dados de monitorização, mas quando é necessária informação entre as estações, é necessário utilizar métodos de interpolação (Denby et al., 2005).

Em Portugal, até ao final da década de 90, o conhecimento da qualidade do ar limitava-se a uma reduzida rede de estações de monitorização nas principais áreas urbanas e a algumas áreas industriais, a campanhas de medição pontuais e a algumas aplicações de modelos. A partir de 2000, com a implementação das novas directivas comunitárias (Directiva 96/62/CE; Directiva 1999/30/CE), o conhecimento sobre a distribuição das concentrações dos principais poluentes atmosféricos melhorou substancialmente com o aumento gradual do número das estações da rede nacional e com a realização, por várias entidades, de campanhas com recurso a tubos de difusão e estações ou equipamentos móveis (Ferreira et al., 2002).

A avaliação dos resultados da monitorização dos últimos anos, disponibilizados na base de dados *on-line* sobre qualidade do ar - QualAr (APA, 2008), revela que actualmente, tal como acontece na Europa, os principais problemas de qualidade do ar em Portugal estão relacionados com o tráfego rodoviário nas grandes áreas urbanas, principalmente nas áreas metropolitanas de Lisboa e do Porto. As elevadas emissões de tráfego nas áreas urbanas têm-se reflectido nas excedências aos limites impostos nas directivas comunitárias, verificadas pontualmente para o dióxido de azoto (NO₂) e frequentemente para as partículas (PM₁₀). Para além destes problemas de poluição, são ainda de salientar níveis elevados do ozono troposférico durante o Verão em áreas suburbanas e rurais.

Os mapas de concentração de poluentes disponíveis para Portugal, e para algumas áreas urbanas, têm sido obtidos pela aplicação de modelos determinísticos da qualidade do ar, como o CHIMERE (Monteiro et al., 2005; Monteiro et al., 2007), ou pela interpolação de resultados de campanhas de medição com tubos de difusão recorrendo a métodos como o *Inverse Distance Weighted (IDW)* ou o *Ordinary Kriging* (Ferreira et al., 2001; Ferreira et al., 2003; Mesquita et al., 2004). O recurso a novas metodologias como a detecção remota (Grosso, 2005; Grosso et al., 2006) ou a tecnologias envolvendo sistemas de informação geográfica (SIG) combinadas com métodos de regressão linear múltipla (RLM) (Briggs et al., 1997, Briggs et al., 2000), designados correntemente por modelos de regressão de ocupação do solo (*Land Use Regression models - LUR*) ou mapeamento usando regressão (*Regression Mapping*), constituem uma alternativa, ou um complemento, aos métodos de mapeamento mais tradicionais e começam agora a ser desenvolvidos em Portugal.

A utilização de modelos de regressão linear múltipla para o mapeamento da qualidade do ar tem sido frequente nos últimos anos. Estes modelos têm sido geralmente aplicados com sucesso à média anual de NO₂, NO_x, PM₁₀, COV e PM_{2,5} em cidades norte americanas e europeias. A performance destes modelos tem sido de uma forma geral superior ou equivalente a métodos como o *Kriging* e a modelos determinísticos (Hoek et al., 2008).

A estimativa da população exposta à poluição do ar é usualmente obtida pela sobreposição de mapas de distribuição das concentrações com o mapa da população residente (Stedman et al., 1997; Stedman et al., 2002).

A utilização de ferramentas de SIG, nomeadamente as técnicas de espacialização, facilita a avaliação da qualidade do ar e a estimativa da população exposta no espaço e no tempo. O potencial dos SIG para esta área assume diferentes formas:

- Podem ajudar a melhorar a integração e consistência dos dados, por disponibilizarem meios de obtenção e ligação de informação espacial dentro de uma mesma estrutura geográfica;
- Fornecem ferramentas explícitas para avaliar exposições, por fornecerem meios para combinar poluição e população;
- Ajudam também a preencher falhas nos dados ambientais e ultrapassar problemas de escala, por modelarem padrões espaciais da poluição do ar (Briggs, 2005).

2 Objectivos

A presente tese teve como objectivo global demonstrar como os SIG e as suas capacidades analíticas, disponibilizam ferramentas fundamentais para responder a uma das maiores necessidades em termos de avaliação da qualidade do ar: a obtenção de mapas de elevada resolução ao nível das áreas urbanas para os poluentes relacionados com o tráfego automóvel, de modo a estimar as áreas e população expostos a excedências de concentrações desses poluentes. Para atingir este objectivo foi necessário efectuar várias tarefas:

- Desenvolver e testar metodologias, em ambiente SIG, recorrendo, como caso de estudo, à distribuição de concentrações de NO₂ na cidade de Lisboa que permitam o mapeamento de dois indicadores do NO₂ (anual e horário), para os anos de 2001 e 2002. Em primeiro lugar, foi testada a possibilidade de recorrer a resultados de campanhas de monitorização com muitos pontos mas de curta duração para complementar a informação dada pelas estações de monitorização fixa, que disponibilizam informação em contínuo mas para poucos locais. Em segundo lugar, foram aplicadas técnicas de mapeamento usando os métodos de interpolação determinísticos, o *Inverse Distance Weighted (IDW)*, e geoestatísticos, o *Ordinary Kriging (Krig)* e *Ordinary Co-Kriging (CoKrig)*. Por último, foi testada a utilização da Regressão Linear Múltipla (RLM) como um método de interpolação válido para áreas urbanas.
- Desenvolver uma metodologia objectiva e coerente para avaliação da incerteza e comparação das vantagens e desvantagens dos vários métodos de mapeamento testados.
- Testar a adequação dos mapas produzidos para estimar a área e a população exposta a determinados níveis de poluição numa área urbana, apresentando-se a cidade de Lisboa como um exemplo de aplicação.

3 Enquadramento

No capítulo de enquadramento da presente dissertação abordam-se vários aspectos da avaliação da qualidade do ar. O enquadramento incide sobre: o que se entende por qualidade do ar, quais os principais problemas a ela associados, o seu enquadramento legislativo e como pode ser obtido o mapeamento das concentrações de poluentes atmosféricos. Aborda-se, ainda, a importância e a forma como pode ser feita a estimativa da área e população exposta à poluição atmosférica. Contextualiza-se o tema da incerteza, nomeadamente no âmbito de um SIG. Finalmente é explicada a importância dos SIG para a avaliação da qualidade do ar.

3.1 Qualidade do ar: Química na atmosfera, fontes e efeitos na saúde humana e no ambiente atmosférico do dióxido de azoto

A qualidade do ar é a presença de substâncias indesejáveis no ar, em quantidades suficientemente elevadas para causar efeitos prejudiciais (Nevers, 1995).

As fontes emissoras de poluentes atmosféricos são numerosas e variáveis, podendo ser antropogénicas ou naturais. As fontes antropogénicas são as que resultam das actividades humanas, como a actividade industrial ou o tráfego automóvel, enquanto as fontes naturais englobam fenómenos da natureza tais como emissões provenientes de erupções vulcânicas, fogos florestais de origem natural e emissão e transporte de partículas de regiões desérticas, nomeadamente, do Norte de África. As principais fontes antropogénicas de poluição atmosférica estão associadas aos sectores dos transportes, produção de electricidade, indústria, agricultura e aquecimento doméstico. Estes sectores são responsáveis pela emissão de vários poluentes atmosféricos – dióxido de enxofre, óxidos de azoto, amoníaco, compostos orgânicos voláteis e partículas – muitos dos quais interagem com outros formando novos poluentes. Em particular, as emissões de óxidos de azoto são produzidas pelos sectores dos transportes, energia, doméstico e industrial e podem ter influência nas concentrações de PM, na deposição ácida, na eutrofização e no ozono troposférico. Ou seja, este poluente, directa ou indirectamente tem influência negativa sobre a saúde humana, sobre os materiais, sobre as culturas agrícolas e biodiversidade.

O clima e, em particular, a ventilação, fluxos de ar e altura da camada de mistura, são tão importantes, na determinação das concentrações atmosféricas em superfície, como as taxas de emissão. Assim, as características da estrutura dos campos de ventos e a evolução dinâmica da altura da camada de mistura, são elementos chave para compreender a poluição atmosférica em áreas urbanas (CE, 1997).

Nas áreas urbanas, os processos de ventilação são de difícil descrição porque são modificados pela complexidade do terreno e pelos gradientes térmicos. Por exemplo, estes processos podem originar ilhas de calor, que ocorrem em volta das áreas urbanas, e que podem constituir uma barreira a uma eficiente troca de ar com a envolvente. Nestas condições, os poluentes atmosféricos são recirculados na camada de mistura e as concentrações podem ficar mais altas. Nas massas de ar recirculadas podem ocorrer reacções fotoquímicas e piorar a situação. Isto acontece particularmente na região mediterrânea, onde brisas marítimas que causam recirculação são fortes e comuns (CE, 1997).

O monóxido de azoto (NO) e o dióxido de azoto (NO₂) são designados por óxidos de azoto (NO_x). A maior quantidade de NO₂ presente na atmosfera é resultado da oxidação do NO. O NO é um gás sem cor e sem cheiro que é maioritariamente produzido pela combustão do fuel, a altas temperaturas, como acontece nos motores dos veículos rodoviários e na indústria de produção eléctrica. Estima-se que o tráfego rodoviário seja responsável por mais de metade das suas emissões na Europa (CE, 1997).

Assim, as concentrações de NO e NO₂ são mais elevadas nas áreas urbanas, onde o tráfego é mais intenso do que nas áreas rurais. Os episódios ocorrem quando surgem situações de grande estabilidade atmosférica e a massa de ar não consegue dispersar. As concentrações horárias em áreas urbanas (normalmente entre os 20 e os 80 µg/m³ mas que podem atingir os 400 µg/m³) variam significativamente ao longo do dia com picos duas vezes por dia em consequência das horas de ponta do tráfego rodoviário.

Quanto aos efeitos na saúde humana do NO₂, estes fazem-se sentir ao nível do tracto respiratório, ao ter pouca afinidade com a água das mucosas, alcança as regiões profundas do tracto respiratório e inibe algumas funções dos pulmões, tais como a resposta imunológica, podendo irritar os pulmões e diminuir a resistência às infecções respiratórias. É tóxico quando ocorre uma exposição continuada, ou frequente, a concentrações mais elevadas do que as normalmente encontradas no ar ambiente, podendo causar um aumento na incidência de doenças respiratórias agudas em crianças.

Assim, os seus efeitos traduzem-se no aumento da susceptibilidade a doenças respiratórias, principalmente em crianças, e também no aumento da possibilidade de ataques de asma. O NO não é considerado perigoso para a saúde. No entanto, uma vez libertado na atmosfera é rapidamente oxidado pelo oxigénio na forma de ozono (O₃), ou radicais como o HO₂ ou RO₂ formando NO₂.

O NO₂ é um forte agente oxidante que reage no ar para formar ácido nítrico corrosivo, assim como nitratos orgânicos tóxicos. Tem um papel importante nas reacções atmosféricas que

produzem o ozono e o *smog*. Na presença da luz solar, reage com hidratos de carbono para produzir poluentes fotoquímicos como o ozono. Os óxidos de azoto têm um tempo de vida de aproximadamente um dia, no que diz respeito à conversão em ácido nítrico. Este ácido nítrico é removido da atmosfera por deposição directa no solo, ou transferido para o meio aquoso (nuvens ou água da chuva), contribuindo assim para a deposição ácida (CE, 1997). Na **Figura 1**, encontra-se um diagrama simplificado do ciclo dos compostos de azoto.

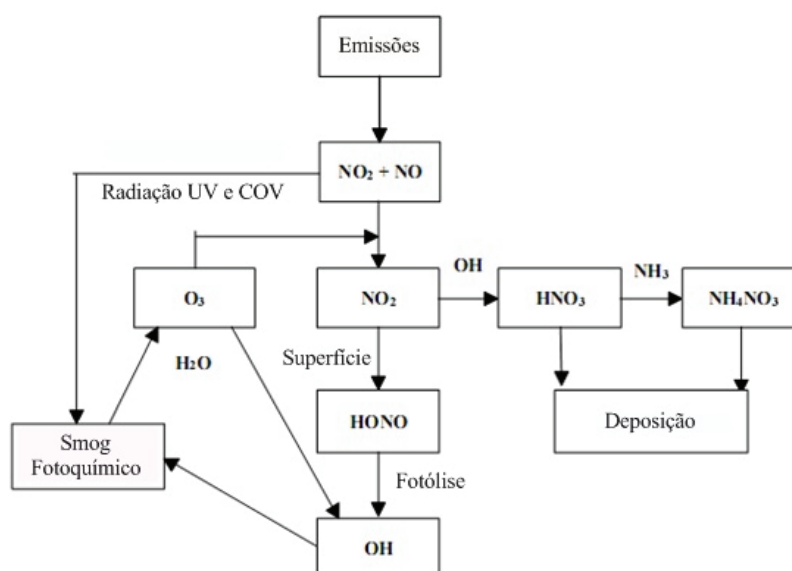


Figura 1. Diagrama simplificado do ciclo dos compostos de azoto (CE, 1997)

Apesar de não ser tão preocupante como as partículas (PM_{10} e $PM_{2,5}$), o NO_2 é um poluente com níveis e consequências graves em áreas urbanas. A sua monitorização é de grande utilidade porque pode ser usado como indicador da poluição provocada pelo tráfego rodoviário e como indicador dos níveis de partículas (WHO, 2004). Em Portugal, é o poluente que tem um maior número de locais de monitorização, tanto de fundo, como de tráfego ou industriais (APA, 2008). Para além deste aspecto, pode ser monitorizado usando métodos indicativos como os tubos de difusão passiva ao contrário das partículas que exigem métodos mais dispendiosos e menos versáteis (Ferreira et al., 2001; Ferreira et al., 2003; Mesquita et al., 2004).

3.2 Quadro legislativo

A Directiva 1996/62/CE, de 27 de Setembro, relativa à avaliação e gestão do ar ambiente, também denominada Directiva-Quadro da qualidade do ar, veio definir um novo quadro legislativo e estabelecer as linhas de orientação da política de gestão da qualidade do ar ambiente no seio da União Europeia (UE).

Um dos princípios base introduzidos por esta Directiva assenta no estabelecimento de objectivos de qualidade do ar ambiente na UE, os quais visam evitar, prevenir ou limitar efeitos nocivos sobre a saúde humana e sobre o ambiente. O diploma estabeleceu também que a avaliação da qualidade do ar deverá ser feita com base em métodos e critérios comuns em todos os Estados-Membros. Esta avaliação deverá dotar todo e cada Estado-Membro de informações adequadas sobre a qualidade do ar ambiente, para que delas o público seja informado.

Os princípios estabelecidos na Directiva 96/62/CE foram transpostos para a ordem jurídica interna pelo Decreto-Lei n.º 276/99, de 23 de Julho. As “Directivas-filhas” têm vindo a estabelecer os valores limite para diversos poluentes atmosféricos (**Figura 2**). A primeira Directiva desta natureza foi a Directiva 99/30/CE, do Conselho de 22 de Abril que estabeleceu os valores limite para o dióxido de enxofre, dióxido de azoto e óxidos de azoto, partículas em suspensão e chumbo no ar ambiente. A Directiva 2008/50/CE publicada em 21 de Maio de 2008, ainda não transposta para o direito interno, veio integrar e substituir a Directiva Quadro e as três primeiras Directivas Filhas.

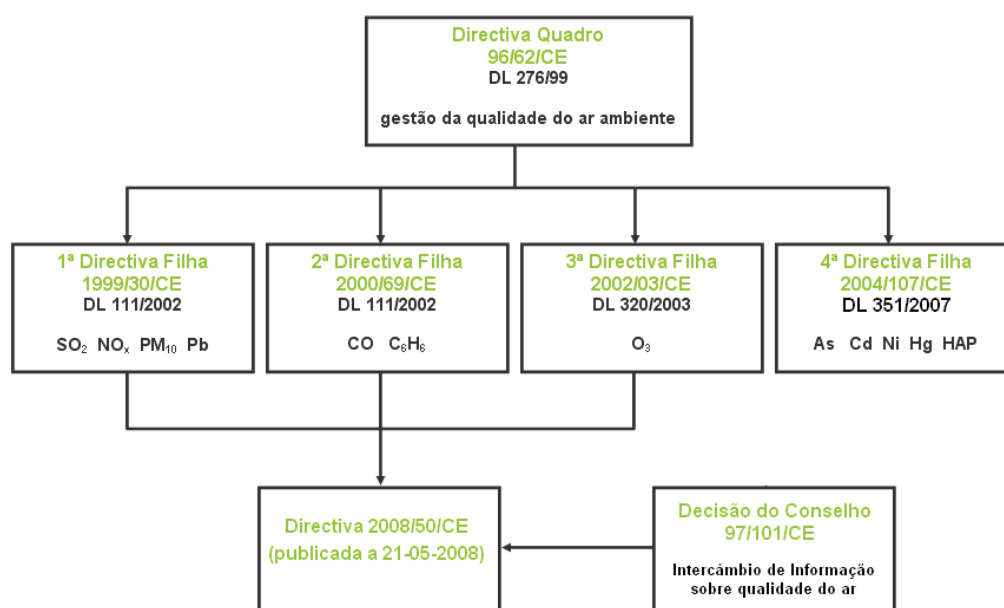


Figura 2. Representação esquemática do enquadramento legislativo da avaliação e gestão qualidade do ar ambiente na UE e em Portugal.

3.2.1 Valores legais estabelecidos para o NO₂

Na **Tabela 1**, dado no presente trabalho apenas serem analisados resultados de NO₂, para a cidade de Lisboa, em 2001 e 2002, encontra-se apenas um resumo dos valores estabelecidos pelo Decreto-Lei n.º 111/2002 relativamente a este poluente para a protecção

da saúde humana, para os anos de 2001 e 2002. Estes limites não sofreram alterações com a nova Directiva para a qualidade do ar (Directiva 2008/50/CE).

	Média horária ($\mu\text{g}/\text{m}^3$)	Média anual ($\mu\text{g}/\text{m}^3$)
Valor Limite (VL)	200 a não ultrapassar mais do que 18 vezes no ano (a cumprir em 2010) ⁽¹⁾	40 (a cumprir em 2010)
Valor Limite + Margem de tolerância (VL+Mt)	290 (2001) 280 (2002) ⁽¹⁾ a não ultrapassar mais do que 18 vezes no ano	58 (2001) 56 (2002)
Limiar de Alerta ⁽²⁾	400	
Limiar Inferior de Avaliação (LIA) ⁽³⁾	100 ⁽¹⁾ a não ultrapassar mais do que 18 vezes no ano	26
Limiar Superior de Avaliação (LSA) ⁽³⁾	140 ⁽¹⁾ a não ultrapassar mais do que 18 vezes no ano	32

⁽¹⁾ É avaliada usando o indicador **19º máximo horário**.

⁽²⁾ Se for ultrapassado numa área superior a 100 km² ou numa aglomeração consoante a que apresentar menor área, durante 3 ou mais horas consecutivas.

⁽³⁾ De acordo com o estabelecido no Anexo IX do DL n.º 111/2002, a avaliação dos resultados em função dos limiares de avaliação para os últimos 5 anos em conjunto com a população da zona permite definir ou rever a estratégia de avaliação a implementar numa zona.

Tabela 1. Níveis e valores definidos na legislação para NO₂ para a protecção da saúde humana definidos (DL 111/2002 e Directiva 2008/50/CE).

3.2.2 Requisitos de avaliação, ferramentas e objectivos de qualidade

Uma das maiores preocupações das directivas comunitárias para a qualidade do ar é a necessidade de avaliar a qualidade do ar ambiente nos vários Estados Membros de uma forma comum, uniforme e coerente. O Artigo 3º da Directiva Quadro e o Anexo III da primeira Directiva-Filha definem uma série de requisitos para assegurar a harmonização e controlo de qualidade dos dados de qualidade do ar. Estes requisitos, relativamente à monitorização, incluem por exemplo: critérios para o desenho da rede e selecção dos locais; objectivos de qualidade dos dados em termos do mínimo de exactidão, recolha de dados, cobertura dos dados das medições e a definição de métodos standardizados de medição de referência ou equivalentes.

Dados os encargos financeiros inerentes à instalação e manutenção das estações fixas de monitorização da qualidade do ar, não é possível que estas existam em número suficiente para se conhecer a qualidade do ar em todo o território e para permitir o mapeamento da distribuição das concentrações de cada poluente como é objectivo da legislação actual. Deste modo, a legislação sobre qualidade do ar a nível europeu tem vindo a sugerir outras técnicas, designadas por técnicas de avaliação suplementar, que podem ser usadas

isoladamente ou em conjunto. Estas podem ser agrupadas em: medições indicativas, modelação e estimativas objectivas.

A legislação actual, relativamente à avaliação das concentrações, indica que para as zonas e aglomerações nas quais forem usadas outras fontes para além da medição, para completar os dados obtidos por esta ou como o único processo de avaliação da qualidade do ar, devem ser recolhidas, nomeadamente, as seguintes informações:

- Uma descrição dos resultados, incluindo a margem de erro e, em especial, a extensão de todas as áreas ou, se necessário, o comprimento da estrada dentro da zona ou aglomeração em que as concentrações ultrapassam o(s) valor(es) limite ou, se for caso disso, o(s) valor(es) limite acrescido(s) da(s) respectiva(s) margem(ns) de tolerância, bem como de todas as áreas em que as concentrações ultrapassassem o limiar superior ou o limiar inferior de avaliação;
- Para os VL que têm como objectivo a protecção da saúde humana, a população potencialmente exposta a concentrações que excedam o valor limite (adaptado da Secção II Anexo X do Decreto-Lei n.º 111/2002 de 16 de Abril e Directiva 2008/50/CE).

Em termos dos objectivos de qualidade para a avaliação e avaliação suplementar da qualidade do ar, Directiva 2008/50/CE (que resume a Directiva Quadro e as 1ª, 2ª e 3ª Directivas-filhas) define para o NO₂ os critérios estabelecidos na **Tabela 2**.

Dióxido de azoto	
Medições fixas	
Incerteza	15%
Mínimo de captura de dados	90%
Medições indicativas	
Incerteza	25%
Mínima recolha de dados	90%
Cobertura mínima de tempo	14% ⁽¹⁾
Incerteza modelação	
Hora	50%
Média anual	30%
Incerteza estimativas objectivas	75%

⁽¹⁾ Um dia de medição por semana aleatoriamente, distribuídos igualmente pelo ano, ou 8 semana distribuídas uniformemente pelo ano.

Tabela 2. Objectivos de qualidade para a avaliação do dióxido de azoto (NO₂) (Directiva 2008/50/CE).

A incerteza (expressa em relação a um grau de confiança de 95%) dos métodos de avaliação deve ser avaliada de acordo com os princípios do CEN - *Guide to the Expression of Uncertainty in Measurement* (ENV 13005-1999), da metodologia da ISO 5725:1994 e do guia disponibilizado pelo relatório CEN 'Air Quality – Approach to Uncertainty Estimation for Ambient Air Reference Measurement Methods' (CR 14377:2002E). As percentagens para

incerteza na tabela acima são dadas para medições individuais sobre o período de tempo considerado pelo valor limite, para o intervalo de confiança de 95%. A incerteza para as medições fixas pode ser interpretada como sendo aplicável na região do valor limite apropriado.

A incerteza associada ao modelo é definida como o desvio máximo entre as concentrações medidas e calculadas para 90% das estações de medição individuais, durante o período considerado, em relação ao VL, independentemente da cronologia das ocorrências. A incerteza associada ao modelo será interpretada como sendo aplicável na região do valor limite pertinente. As medições fixas que forem seleccionadas para comparação com os resultados da modelação devem ser representativas da escala abrangida pelo modelo.

O cálculo da incerteza associada ao modelo pode ser formulado como, o que se designou por, erro relativo da directiva (RDE) (Denby et al, 2008):

$$RDE = \frac{|O - M|}{VL}$$

Em que: VL – valor limite, O – valor observado e M – valor modelado.

Os requisitos para o mínimo de captura de dados e de cobertura de tempo de amostragem não incluem perdas de dados devido às calibrações ou manutenções normais dos instrumentos.

3.2.3 Verificação do cumprimento da legislação

De acordo com as directivas europeias, anualmente os Estados Membros têm a obrigação legal de enviar à Comissão Europeia os resultados da avaliação realizada no ano anterior. Esta resposta é feita sob a forma de um questionário (2004/461/EC: Decisão da Comissão de 29 Abril de 2004 que substituiu a 2001/839/EC//EC: Decisão da Comissão de 15 de Fevereiro de 2002) em que uma das solicitações é o envio de mapas de distribuição da qualidade do ar e estimativas das áreas e população expostas a excedências aos VL.

Recorrendo a modelos, em princípio, é possível obter valores de concentrações para toda a área de zona. Quando os resultados da modelação são comparados com os limites legais para a qualidade do ar, a resolução espacial adquire importância já que elevadas resoluções tendem a resultar em máximos mais elevados, encontrados em *hotspots* muito localizados. Consequentemente o estado de excedência de uma zona depende da resolução (CE, 2002b).

Relativamente à resolução dos mapas para verificação do cumprimento da legislação, apesar dos requisitos relativos à avaliação da qualidade do ar, mapeamento e estimativas da população e áreas expostas, presentes nas directivas, serem escassos, existem algumas indicações na Directiva 2008/50/CE relativamente à localização dos pontos de amostragem (secção B do Anexo III) que permitem inferir a resolução dos mapas adequados a cada tipo de área:

- Zonas de tráfego - a avaliação deve ser representativa de um segmento de 100 m de estrada;
- Áreas urbanas de fundo - as avaliações devem ser representativas de grelhas de vários quilómetros i.e. 1–5 km;
- Áreas de ecossistemas - as grelhas devem ter 30 km, ou abaixo de 30 km, quando o terreno é complexo ou existem pequenas áreas de ecossistemas;
- Zonas industriais - a resolução deve ser representativa de 250x250 m (Denby et al., 2008).

3.3 Mapeamento das concentrações de NO₂

Obter mapas da poluição atmosférica relacionada com o tráfego automóvel não é uma tarefa simples. Os níveis de poluição do tráfego rodoviário variam substancialmente a uma distância de metros e o padrão da poluição em áreas urbanas é complexo (Laxen e Noordally, 1987; Hewitt, 1991). Para além destes factos, os dados de monitorização da poluição do ar em áreas urbanas são reduzidos e esparsos. A maioria das redes urbanas, apesar de respeitarem o número mínimo fixado pela legislação, ou até esse número ser superior, incluem poucos locais de medição, e os que existem raramente podem ser considerados representativos da exposição da totalidade da população. A utilização de uma técnica de modelação robusta é essencial para se obterem mapas de poluição do ar em áreas urbanas, com maior grau de confiança (Briggs et al., 2000).

A Directiva 1999/30/EC encoraja os Estados Membros a representar em mapas a distribuição das concentrações nas zonas, usando métodos como modelos, técnicas de interpolação ou outras técnicas matemáticas, sendo um dos objectivos ajudar a dar resposta ao questionário anual da comissão (CE, 2004) no que diz respeito à avaliação das excedências existentes em termos de áreas e população exposta.

Os mapas são geralmente o resultado de cálculos de modelos, interpolação de medições ou outras técnicas matemáticas. Cada mapa deve ser representado por um conjunto adequado de localizações (designadas por “pontos receptores”) e pelas concentrações nestes pontos.

No guia para resposta ao questionário anual (CE, 2002a) é referido que, para uso local, a resolução especial destes mapas deve ser suficiente para distinguir *hotspots*, como sejam as estradas. No entanto, para uso ao nível da Comunidade, esta alta resolução não é uma alternativa devido ao elevado tamanho dos ficheiros que teriam que ser processados e enviados.

Para efeitos de resposta às solicitações da Comissão Europeia, a escolha de distâncias entre receptores é deixada ao critério dos Estados Membros. Este pode usar uma grelha de receptores com densidade fixa ou uma distribuição de pontos receptores com densidade variável, onde as baixas densidades são usadas para áreas de baixo gradiente espacial e a alta densidade onde existem altos gradientes espaciais. A resolução espacial dos mapas será tipicamente de 10-100 km, com resolução incrementada localmente até 1 km. Os mapas definidos neste contexto são adequados para dar uma imagem da qualidade do ar de larga escala na UE, mas têm uma resolução consideravelmente mais baixa do que a necessária para identificar pequenos *hotspots* (CE, 2002a).

Neste capítulo apresentam-se vários tipos de ferramentas de avaliação e mapeamento da qualidade do ar, com maior incidência nas que serão utilizadas no decorrer do presente trabalho. São apresentados os requisitos e objectivos de qualidade definidos na legislação comunitária para cada tipo de ferramenta de avaliação, assim como as indicações disponíveis em guias comunitários para a elaboração de mapas e verificação do cumprimento dos limites legais para a protecção da saúde humana.

3.3.1 Medição

Tradicionalmente, a avaliação da qualidade do ar resulta das medições feitas nas estações fixas recorrendo a equipamentos de monitorização em contínuo, usando os métodos de amostragem e de análise de referência definidos na legislação ou equivalentes. No entanto, outro tipo de medições podem ser consideradas, nomeadamente as que usam outros métodos que não o de referência ou medições de curto prazo (campanhas) em vez de medições em contínuo.

O método de referência para análise de óxidos de azoto (monóxido e dióxido de azoto) definido na Directiva 2008/50/CE é o método automático por quimiluminescência (EN 14211:2005). Na directiva, é ainda referido que poderá ser usado outro método desde que se possa demonstrar que os resultados são equivalentes aos do método de referência.

No âmbito da avaliação da qualidade do ar as medições indicativas são entendidas como medições usando métodos mais simples ou efectuadas durante períodos curtos. Elas são menos rigorosas que as medições fixas e contínuas de alta qualidade mas podem ser

utilizadas para explorar e verificar a qualidade do ar onde os níveis de poluição são relativamente baixos, e como avaliação suplementar às medições de alta qualidade em outras áreas (CE, 2002).

As medições indicativas podem ser baseadas no uso de laboratórios móveis (ou qualquer outra plataforma móvel ou transportável) com medições efectuadas durante períodos de tempo relativamente curtos ou por métodos de medição manuais como a técnica de difusão passiva. Esta última técnica é de particular interesse para campanhas de medição da qualidade do ar de larga escala com elevada resolução, devido ao seu custo relativamente baixo e forma de operação simples em comparação com as estações de monitorização fixas (De Saeger et al., 1995).

Os métodos de amostragem podem ser divididos em métodos passivos e activos. Nos métodos activos a recolha da amostra é feita por bombagem de ar, enquanto nos métodos passivos a amostragem é feita, por exemplo, por difusão dos compostos numa substância absorvente (caso dos tubos de difusão).

A amostragem por difusão é efectuada através da utilização de tubos, os quais se colocam por um tempo determinado nas áreas a amostrar. Estes tubos contêm no seu interior um cilindro de material adsorvente específico que fixa o poluente. Durante o período de amostragem o ar flui para o interior do tubo a uma taxa controlada por difusão molecular, e o poluente é adsorvido no cilindro interior fornecendo assim um valor médio de concentração desse poluente no período de exposição (**Figura 3**) (Van Aalst et al., 1998).

O princípio do método de difusão passiva para a amostragem do dióxido de azoto assenta no facto deste poluente atmosférico ser quimiadsorvidos pela trietanolamina (TEA) como ião NO_3^- . A concentração do ião nitrato após a exposição é analisada por cromatografia iónica ou por espectrofotometria, na banda do espectro visível (FSM, 2007).

Os locais onde são efectuadas as medições podem dividir-se, dependendo do tipo de ambiente, em: rurais, suburbanos e urbanos; e, do tipo de fontes de influência, em: fundo, tráfego e industrial (CE, 2001). A combinação destes dois tipos de classificações permite definir o tipo de localização, como por exemplo: rural de fundo, urbano de tráfego, ou suburbano industrial.

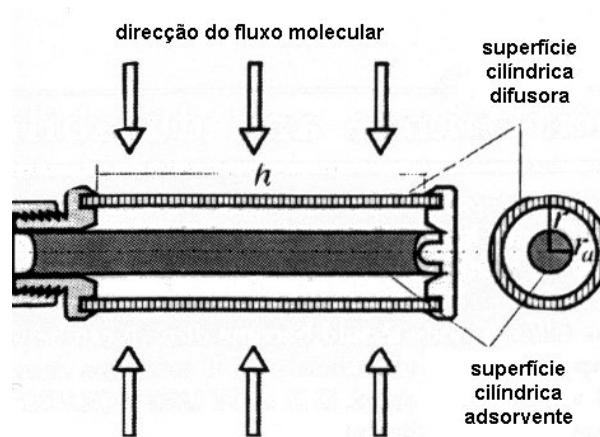


Figura 3. Método de difusão passiva (FSM, 2007).

3.3.2 Modelação

Os modelos atmosféricos correspondem, de um modo geral, a qualquer procedimento matemático que resulta numa estimativa de dados de qualidade de ar ambiente (i.e. concentrações, excedências, deposição). A modelação, nomeadamente, da qualidade do ar, em termos do método pode dividir-se em dois grupos:

- Determinística;
- Estatística ou estocástica.

Os modelos determinísticos são baseados na descrição dos processos físicos e químicos de dispersão dos poluentes; considerando as emissões, advecção e turbulência, são calculadas a transformação e deposição químicas. Estes tipos de modelos são capazes de dar uma descrição das relações causa - efeito (Greig, A., et al., 2001).

Os modelos estatísticos são uma valiosa ferramenta no diagnóstico e predição de qualidade do ar por interpolação e extrapolação do dado monitorizado. O diagnóstico constitui o fundamento estatístico das predições dos modelos (Greig, A., et al., 2001). Os modelos estatísticos fundamentam-se em relações, geralmente baseadas em conceitos físicos, entre observações e outros parâmetros relevantes. Existe um conjunto variado de técnicas estatísticas que podem ser usadas incluindo a regressão, redes neuronais, entre outros. Uma vez que as relações se baseiam nas observações, os modelos estatísticos devem ser estabelecidos em todos os tipos de ambientes a que se pretendem aplicar.

Na **Tabela 3**, apresenta-se uma lista do tipo de modelo e processos requeridos, normalmente utilizados para modelar a qualidade do ar a várias escalas, e mais especificamente, o NO_2 (Denby et al., 2008).

Local/hotspot	Urbano/aglomeração	Regional/continental
Modelo		
Modelos parametrizados Gaussianos e não Gaussianos Modelos estatísticos Modelos dinâmicos de fluidos Modelos lagrangianos (<i>ensemble</i>)	Modelos parametrizados Gaussianos e não Gaussianos Modelos de transporte químicos eulerianos Modelos lagrangianos (<i>ensemble</i>)	Modelos de transporte químicos eulerianos Modelos lagrangianos químicos (<i>ensemble</i>)
Processo NO ₂		
Foto-oxidante químico simples Relações estatísticas/empíricas	Foto-oxidante químico limitado Esquema foto-estacionário Relações estatísticas/empíricas	Deposição Foto-oxidante químico completo

Tabela 3. Tipo de modelo e processos requeridos normalmente utilizados para cada escala e para o NO₂ (Denby et al., 2008).

Na aplicação da modelação determinística a várias escalas em Portugal, salienta-se o grupo de investigação GEMAC da Universidade de Aveiro que desenvolveu, entre vários outros trabalhos, uma aplicação do modelo CHIMERE para a avaliação e previsão da qualidade do ar à escala nacional (resolução 9x9 km) actualmente operacional para vários poluentes e produzindo previsões diárias (Monteiro et al., 2005, Monteiro et al., 2007a, Monteiro et al., 2007b). Especificamente para o nível da cidade de Lisboa, não existe actualmente um modelo determinístico operacional com resoluções mais elevadas.

Para a escala local, existe em funcionamento um sistema de modelos estatísticos para previsão do dia seguinte que abrange a cidade de Lisboa desenvolvido pelo Departamento de Ciências e Engenharia do Ambiente da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa (DCEA/FCT/UNL) (Marques et al., 2006). Neste sistema é desenvolvido um modelo estatístico usando regressão linear múltipla para cada local onde existe uma estação de monitorização, não existindo uma continuidade espacial.

Em nenhum dos casos são disponibilizados mapas para a cidade de Lisboa com resolução suficiente para estimar áreas e população exposta a excedências provocadas pelo tráfego rodoviário. No primeiro caso, apesar de se apresentar uma modelação espacial e temporal, a resolução espacial usada não é suficiente para a escala urbana e no segundo caso a modelação é apenas temporal, não sendo produzidos mapas de distribuição espacial das concentrações.

3.3.3 Interpolação

As técnicas de interpolação envolvem a estimativa do valor de um parâmetro para locais não amostrados baseada no cálculo de funções que utilizam os vários valores medidos num dado número de pontos de amostragem conhecidos. Elas estão sujeitas a incerteza, tanto a nível dos valores estimados como da sua distribuição espacial (Lindley et al., 2005).

Na base das técnicas de interpolação mais frequentemente utilizadas está a primeira lei da Geografia ou Lei de Tobler (1970) que afirma: "*Everything is related to everything else, but near things are more related than distant things*". O que para efeitos de aplicação à interpolação se pode traduzir como "pontos vizinhos no espaço tem valores mais correlacionados que pontos distantes".

Normalmente, distinguem-se dois tipos de técnicas de interpolação: determinística e geoestatística. Ambas assentam na semelhança de pontos próximos para criar uma superfície. As técnicas determinísticas usam funções matemáticas para criar as superfícies. As técnicas geoestatísticas para além das funções matemáticas recorrem a métodos estatísticos podendo ser usadas para criar superfícies, e também estimar a incerteza das previsões (Johnston et al., 2001).

A utilização de ferramentas de interpolação para mapear a qualidade do ar a partir de dados de estações ou de campanhas de monitorização tem sido usada em Portugal com algum sucesso quando se pretende modelar espacialmente as concentrações de fundo e/ou concentrações de âmbito regional de um poluente. No entanto, quando o objectivo é obter mapas de maior resolução para a escala urbana, que modelem tanto locais de tráfego como de fundo, a escassez de pontos de amostragem pode inviabilizar a utilização deste tipo de métodos, apenas baseados na distância, sendo que outras soluções poderão ser mais adequadas. Uma possibilidade é recorrer a métodos que utilizem, para além dos dados de monitorização da variável em análise, variáveis auxiliares com elevada correlação com esta, mas com maior número de pontos de medição. Este é o caso do *ordinary cokriging* e dos modelos de regressão linear que modelam a distribuição espacial de uma variável integrando informação auxiliar com alta resolução.

3.3.3.1 *Inverso da Distância Ponderada (IDW)*

De entre as técnicas determinísticas de interpolação encontram-se por exemplo: o *IDW*, o *spline*, o polinomial global, o polinomial local e a função de base radial. Dada a sua simplicidade, o *IDW* é um método frequentemente usado para interpolar dados de monitorização da qualidade do ar (Lindley et al., 2005; Hoek et al., 2001; Fiala et al., 2000; Horálek et al., 2005; Marshal et al., 2008). Deste modo, neste trabalho foi usado este método como exemplo de interpolação com métodos determinísticos.

O método *IDW* considera que os dados pontuais são ponderados durante a interpolação, de tal forma que a influência de um determinado dado pontual em relação a outro diminui com o aumento da distância (Davis, 1986). A interpolação usando *IDW* implementa explicitamente o pressuposto que pontos mais próximos serão provavelmente mais parecidos do que pontos mais distantes. Para prever um valor para uma localização sem

medição, o *IDW* usará os valores medidos na vizinhança desse local. Os valores medidos em locais mais próximos terão maior influência para a previsão do que os medidos mais longe. Assim, o *IDW* assume que cada ponto medido tem uma influência local que diminui com a distância.

O *IDW* é um interpolador determinístico exacto, onde os valores máximo e mínimo na superfície de interpolação não podem ocorrer fora dos pontos de amostragem, e que assume que a superfície é determinada pela variação local que pode ser determinada a partir da vizinhança. A aplicação do método é bastante rápida, havendo poucas decisões a tomar relativamente aos parâmetros do modelo. A superfície calculada usando o *IDW* depende da selecção do expoente (*Power*) e da estratégia de pesquisa da vizinhança. Este método pode ser uma boa forma de rapidamente obter uma primeira superfície. No entanto, não permite uma avaliação dos erros da previsão, e podem ser produzidos círculos em volta de alguns pontos de amostragem, ou seja, a superfície produzida é sensível a *clusters* e à presença de *outliers* (Johnston et al., 2001).

O *IDW* pode ser definido pela seguinte equação:

$$f(x, y) = \left[\frac{\sum_{i=1}^n w(d_i) z_i}{\sum_{i=1}^n w(d_i)} \right], i = 1, 2, \dots, n$$

Em que: z_i - é o valor da observação i ; d_i - é a distância entre o ponto a estimar e a observação i ; $w(d_i) = 1/(d_i)^p$ - é a ponderação atribuída à observação i (função do inverso da distância); p - é a função expoente (*Power*)

No *IDW* os pesos são proporcionais ao inverso da distância elevada ao valor de p (expoente ou *power*). Como resultado, à medida que a distância aumenta os pesos decrescem, sendo o decréscimo tanto mais acentuado, quanto maior for o p . O valor da função expoente (*power*) (p) óptimo é determinado minimizando o *RMSE* (*Root Mean Squared Error*), que é calculado a partir da validação cruzada (**Figura 4**).

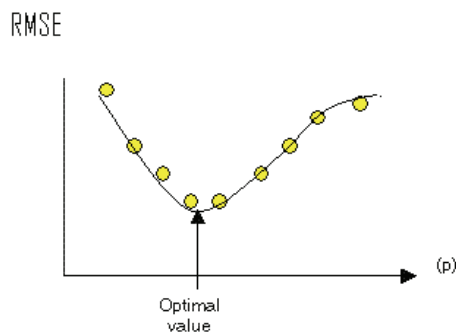


Figura 4. Determinação do valor da função expoente (*Power*) ótimo (Johnston et al., 2001).

Porque os pontos próximos são mais semelhantes do que os pontos mais distantes, à medida que as localizações se afastam, os valores medidos têm pouca relação com o valor da localização a prever. Para agilizar os cálculos podem-se considerar inexistentes os pontos a partir de determinada distância. Como resultado, é prática comum limitar o número de pontos medidos que são usados na previsão de um valor desconhecido numa localização através de uma vizinhança de pesquisa específica. A forma específica da vizinhança restringe até onde, e onde, se devem pesquisar os valores medidos para usar na previsão.

3.3.3.2 Ordinary Kriging

Os métodos de *kriging*, de entre os vários métodos de interpolação geoestatística são os mais conhecidos. Este método tem sido frequentemente aplicado para o mapeamento da qualidade do ar (Oliver and Webster, 1990; Myers, 1994; Liu et al., 1994; Ferreira et al., 2001, 2003; Mesquita et al., 2004; Horálek et al., 2005; Lindley et al., 2005, Briggs et al., 2005), nomeadamente, na construção de mapas de concentrações de fundo pelo *EMEP* (*European Monitoring and Evaluation Programme*) (Hjellbrekke, 2000).

O *Kriging* é um interpolador relativamente rápido que pode ser exacto ou aproximado dependendo do modelo associado à análise do erro relativo à medição. É muito flexível e permite a intervenção do utilizador na construção dos gráficos de autocorrelação espacial. O *Kriging* usa modelos estatísticos que permitem a construção de uma variedade de mapas incluindo previsões, erros das previsões e probabilidades. A flexibilidade do *kriging* pode requerer uma série de decisões sobre transformações nos dados, tendências, modelos, parâmetros, e vizinhanças. O *kriging* assume que os dados provêm de um processo estocástico estacionário, e alguns métodos assumem que os dados seguem uma distribuição normal (Johnston et al., 2001). Uma grande vantagem do *kriging* relativamente ao *IDW* e a outros métodos determinísticos é a facilidade na produção de mapas de erro e superfícies de probabilidade (Lindley et al., 2005).

O *kriging* assume que a distância ou direcção entre pontos amostrados reflecte a correlação espacial que pode ser usada para explicar a variação numa superfície. No *kriging* é ajustada uma função matemática a um número de pontos específico, ou todos os pontos dentro de um raio especificado, para determinar o valor para cada localização. O *kriging* é um processo com vários passos. Inclui a análise estatística exploratória dos dados, modelação do variograma, criação da superfície e, opcionalmente, a exploração da variância da superfície. É o método mais apropriado quando se sabe que existe uma correlação espacial com a distância ou um viés direccional nos dados (Johnston et al., 2001).

A fórmula da metodologia de *kriging* é semelhante à do *IDW* uma vez que pondera os valores medidos na vizinhança para derivar uma previsão para um local não medido. A fórmula geral para ambos os interpoladores baseia-se na soma ponderada dos dados (Johnston et al., 2001):

$$Z(S_0) = \sum_{i=1}^N \lambda_i Z(S_i)$$

Em que: $Z(S_i)$ é o valor medido na i -ésima localização; λ_i é um peso desconhecido para o valor medido na i -ésima localização; S_0 é a localização a prever; N é o número de valores medidos.

No *kriging* o peso é baseado não apenas na distância entre localizações medidas e previstas (como no *IDW*), mas também no arranjo espacial entre pontos de medição. Para usar o arranjo espacial dos pontos de medição nos pesos, a autocorrelação espacial tem que ser quantificada. Então, no *ordinary kriging*, o peso, depende do modelo ajustado aos pontos de medição, da distância ao local da previsão, e das relações espaciais entre os valores medidos em volta do local da previsão (Johnston et al., 2001).

Para criar uma superfície de previsão com o método de interpolação *kriging*, é necessário:

- Descobrir as regras de dependência entre as observações criando um variograma e funções de covariância para estimar a dependência estatística (chamada autocorrelação espacial) entre os valores que dependem do modelo de autocorrelação (ajustamento do modelo);
- Realizar as previsões (previsão dos pontos com valor desconhecido).

Portanto, no *kriging* os dados são usados em dois passos distintos: primeiro para estimar a autocorrelação espacial (definir o variograma) e segundo para fazer as previsões (Johnston et al., 2001).

O gráfico que representa a semivariância em função das distâncias entre pares de pontos amostrados, designa-se por variograma experimental. A semivariância é uma medida positiva do grau de dependência espacial entre os pontos amostrados calculada através de:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_i)]^2$$

Em que: $\gamma(h)$ é a semivariância para o intervalo h ; $N(h)$ é o número de pares de dados que distam h entre eles; $Z(x_i + h) - Z(x_i)$ é o valor da diferença de um par de dados que distam h entre si.

Na construção do variograma, todos os possíveis pares são examinados. Quando a distância entre pares de pontos é zero, o valor em cada ponto é comparado com ele próprio, e logo as diferenças são zero e o valor da semivariância também zero. Se a distância h é pequena, os pontos a serem comparados são muito semelhantes e estão bastante relacionados entre si, pelo que o valor da semivariância é reduzido. O efeito de pepita (*nugget*) é o valor da semivariância quando a distância tende para zero. À medida que aumenta a distância entre os pontos a serem comparados, maior é a semivariância, até que, a partir de determinada distância (*range*), a semivariância estabiliza num valor (*sill*) que é igual à variância dos dados da população considerada, significando que já não existe qualquer relação entre os pares de dados considerados a esta distância (**Figura 5**).

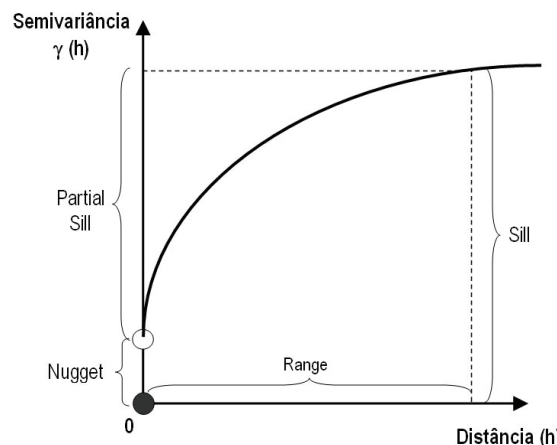


Figura 5. Forma característica de um semivariograma (Johnston et al., 2001).

É habitual encontrar dois tipos de variogramas elementares, variogramas com e sem patamares. Na **Figura 6**, identificam-se formas e equações de modelos matemáticos que podem descrever a semivariância, sendo o modelo esférico um dos mais frequentemente aplicados.

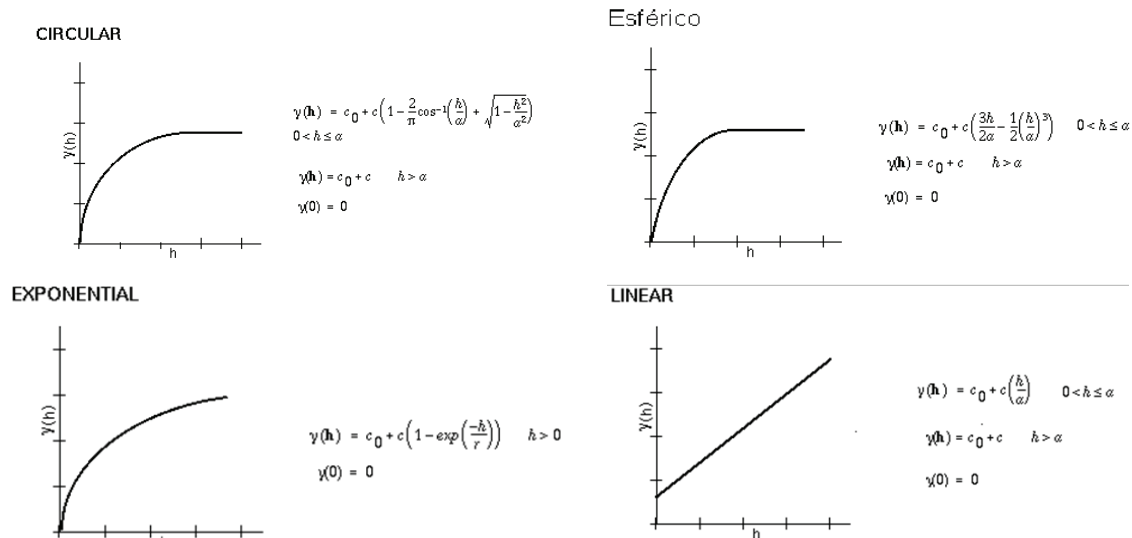


Figura 6. Exemplos de modelos matemáticos aplicados a variogramas (Johnston et al., 2001).

A dependência espacial de uma dada variável pode depender da direcção, pelo que se o variograma apresentar uma forma semelhante em todas as direcções do espaço, só dependendo de h , diz-se que a estrutura do fenómeno é isotrópica, ou seja, sem direcções privilegiadas de variabilidade espacial.

3.3.3.3 Ordinary CoKriging

O *ordinary cokriging* é um interpolador que permite integrar informação adicional no mapeamento da qualidade do ar (Roth and Bournel-Bosson, 2001; Jeannée, 2004). Este método é uma extensão do *ordinary kriging* a situações em que duas ou mais variáveis são espacialmente dependentes e a variável a estimar não está amostrada com a intensidade com que estão as outras variáveis dependentes, utilizando-se os valores destas e as suas dependências para estimar a variável requerida. Este método corresponde a uma alteração no princípio de base dos métodos de interpolação acrescentando ao factor da semelhança devida à proximidade, a possibilidade de usar a co-variância existente entre a variável alvo e outras variáveis correlacionadas mas com maior número de pontos de amostragem.

O *cokriging* usa informação sobre vários tipos de variáveis. Se a principal variável é Z_1 , são usadas para fazer melhores previsões, para além da autocorrelação para o Z_1 , a correlação cruzada entre Z_1 e todas as outras variáveis. É bastante tentador usar informação de outras variáveis para auxiliar nas previsões, mas existe um custo associado. O *cokriging* exige mais estimação, incluindo a estimativa da autocorrelação de cada variável assim como as correlações cruzadas. Teoricamente não é possível piorar os resultados do *kriging* porque se não houver correlação cruzada, pode-se voltar a recorrer à autocorrelação de Z_1 . No entanto, cada vez que é feita uma estimativa dos parâmetros de auto-correlação, é

introduzida mais variabilidade. Assim, os ganhos na precisão das previsões podem não justificar o esforço extra (Johnston et al., 2001).

3.3.3.4 *Regressão Linear Múltipla*

A utilização de modelos de regressão linear é outro dos métodos possíveis para modelar a distribuição espacial de uma variável integrando informação auxiliar com alta resolução, sendo para tal necessário encontrar uma ou mais variáveis preditoras capazes de explicar o fenómeno em análise. A regressão linear baseia-se nos princípios de que:

- As condições ambientais para a variável de interesse podem ser estimadas a partir de um número reduzido de variáveis preditoras disponíveis;
- A relação entre a variável alvo e as preditoras pode ser estabelecida com confiança, baseando-se apenas num pequeno número de amostras ou área de treino (Briggs et al., 2000).

O maior uso desta técnica tem sido na interpretação de dados de detecção remota (Fuller et al., 1998; Li et al. 1998), onde os métodos de regressão são usados para determinar a relação entre a reflectância medida e a ocupação do solo ou outros atributos derivados de levantamentos ao nível do solo, nomeadamente a poluição atmosférica (Grosso et al., 2006). No entanto, os métodos de regressão têm sido usados para um amplo leque de outras utilizações, incluindo: mapeamento de qualidade da paisagem (Briggs and France, 1980), condições do solo (Knotters et al., 1995), contaminação salina (Mattson and Godfrey, 1994) e poluição do ar (Wagner, 1994). Na área da poluição do ar salienta-se o projecto SAVIAH (*Small Area Variations in Air quality and Health*) financiado pela União Europeia, uma vez que foi o primeiro projecto com alguma dimensão (quatro cidades europeias), onde a regressão foi usada para estimar e mapear as concentrações médias anuais da poluição relacionada com o tráfego rodoviário como base para o estudo de variações em áreas pequenas de qualidade do ar e de doenças respiratórias crónicas (Briggs et al., 2000).

Mais recentemente, o método de regressão linear múltipla tem vindo a ser usado frequentemente para o mapeamento da qualidade do ar, sendo designado correntemente por modelos de regressão de ocupação do solo (*Land Use Regression models - LUR*) ou mapeamento usando regressão (*Regression Mapping*). Numa revisão dos modelos de regressão linear múltipla para avaliar a variação espacial da poluição no ar ambiente publicada por Hoek, G., et al. (2008), referenciam-se 25 estudos que recorrem a este tipo de metodologia. Nesta revisão, os autores verificaram que estes modelos combinam tipicamente 20-100 localizações de monitorização, espalhadas pela área em estudo, e o desenvolvimento de modelos estatísticos usando variáveis preditoras usualmente obtidas usando SIG. Frequentemente, a monitorização é de curta duração: uma a quatro

campanhas, com uma a duas semanas de duração. As variáveis preditoras utilizadas incluem várias representações de tráfego, densidade populacional, ocupação do solo, características geográficas como a altitude e o clima. Estes modelos têm sido geralmente aplicados com sucesso à média anual de NO₂, NO_x, PM_{2,5}, COV, parte de fuligem das partículas PM_{2,5} em cidades Norte Americanas e Europeias. O desempenho destes modelos tem sido de uma forma geral superior ou equivalente a métodos como o *kriging* e modelos determinísticos.

A regressão linear múltipla é uma extensão da regressão linear simples onde existem duas ou mais variáveis independentes que explicam uma variável dependente em vez de uma. Um modelo de regressão linear múltipla pode ser formulado da seguinte forma:

$$Y = A + B_1X_1 + B_2X_2 + \dots$$

onde: Y = variável dependente; X₁, X₂, ... = variáveis independentes ou preditoras ; B₁, B₂, ... = os coeficientes das variáveis independentes; A = ordenada na origem (Tabachnick e Fidell, 1996).

A avaliação do ajustamento do modelo é feita usando o coeficiente de correlação de Pearson que se calcula com a seguinte expressão:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Em que: x₁, x₂, ..., x_n e y₁, y₂, ..., y_n são os valores medidos de ambas as variáveis. Para além disso $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ e $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ são as médias aritméticas de ambas as variáveis.

A análise correlacional indica a relação entre 2 variáveis lineares e os valores estarão sempre compreendidos entre +1 e -1. O sinal indica a direcção da correlação (positiva ou negativa), a magnitude do coeficiente de correlação indica a força da correlação.

O coeficiente de determinação (R²), o quadrado do coeficiente de correlação de Pearson, representa a variância explicada pelo modelo e varia entre [0,1].

As hipóteses subjacentes a este modelo são semelhantes às consideradas no caso do modelo de regressão linear simples:

- Os valores das variáveis preditoras são considerados predeterminados, sem erro;

- Os erros são independentes, têm valor esperado nulo, variância constante e são normalmente distribuídos (Guimarães e Cabral, 1997).

Quando o número de variáveis preditoras potencialmente interessantes para incluir no modelo é muito elevado podem ser usados vários métodos automáticos para seleccionar as variáveis. Guimarães e Cabral (1997), apontam 4 métodos para a selecção das variáveis:

- Método exaustivo: consiste em criar todos os modelos que combinem as variáveis existentes de todas as maneiras possíveis, usar um critério de qualidade para ordenar os modelos e seleccionar um conjunto que valha a pena analisar em pormenor;
- Método progressivo (*forward*): consiste em fazer um modelo com a variável independente que tiver maior correlação com a variável dependente que se pretende calcular e ir acrescentando as variáveis cujos coeficientes de correlação sejam sucessivamente mais baixos;
- Método regressivo (*backward*): fazer um modelo com todas as variáveis como independentes em relação à variável dependente que se pretende calcular e ir excluindo as variáveis cujos coeficientes não sejam significativamente diferentes de zero ou que menos alterem a variância explicada;
- Método passo a passo (*stepwise*): é uma versão dos métodos regressivos ou progressivos nos quais as variáveis excluídas ou incluídas em passos anteriores voltam a ser analisadas.

O maior perigo associado aos métodos automáticos de selecção de variáveis independentes (preditoras) consiste na sua utilização acrítica. É fundamental não esquecer que, ao construir um modelo de regressão, a inclusão de preditores tem de fazer sentido. Isto é, entre preditores e a variável dependente têm que existir relações plausíveis causa-efeito. Assim, um modelo sugerido por um método automático não deve ser considerado definitivo e pode ser alterado pelo analista (Guimarães e Cabral, 1997).

Ao estabelecer modelos de regressão linear múltipla é importante verificar se existem problemas de multicolineariedade, dois ou mais preditores que expliquem o mesmo processo ou fenómeno, e em que apenas um preditor deve ser mantido. A multicolinearidade pode ser avaliada pelo factor VIF (*Variance Inflation Factor*), pela tolerância (*tolerance* = 1/VIF) ou outros parâmetros. O VIF é obtido usando a seguinte expressão:

$$VIF_i = \frac{1}{1 - R_i^2}$$

Considera-se que existe multicolinearidade quando os valores de VIF são elevados e quando os valores de tolerância são muito baixos. No entanto, não estão definidos valores consensuais para se considerar se existe ou não multicolinearidade. Nalguns casos apontase VIF superiores a 2, 5 ou 10 para cada uma das variáveis. O valor mais frequentemente indicado foi o 10 para valores individuais e 6 para a média das variáveis de regressão. Os efeitos da multicolinearidade são os seguintes:

- A imprecisão na estimativa dos coeficientes da regressão;
- Ligeiras flutuações na correlação conduzem a grandes diferenças nos coeficientes de regressão;
- Acrescentar ou retirar casos pode levar a grandes diferenças nos coeficientes de regressão;
- Aumenta o erro padrão dos coeficientes, conduzindo à redução da significância dos testes (Webb, 2006).

3.4 Estimativa das áreas e população expostas à poluição atmosférica

A Comissão Europeia, no guia para resposta ao questionário anual relativo à aplicação da legislação em matéria de qualidade do ar (CE, 2002a) e quanto à exposição da população, refere a exposição como sendo a exposição potencial, que apenas tem em consideração as concentrações verificadas no exterior. A exposição da população acima de um valor limite é calculada como o total de pessoas vivendo em casas na área onde os níveis no ar ambiente estão acima do valor limite. Para exposição ao longo das ruas, são tidas em conta as casas que tiverem pelo menos um lado exposto ao nível de concentração acima do valor limite.

Na prática, o número de pessoas calculado vai depender da resolução espacial dos dados da população e dos mapas das concentrações. Idealmente, uma resolução da ordem dos 10 m seria necessária, mas na prática resoluções mais baixas podem também ser usadas. Como o principal objectivo deste parâmetro é dar uma indicação da exposição, não necessariamente um número preciso, podem ser dadas estimativas, em particular quando as excedências ocorrem apenas em alguns *hotspots* da zona.

No caso de cálculos detalhados em que são distinguidas as casas individuais, é recomendado contabilizar os residentes de uma casa como expostos a níveis acima do valor limite se a concentração se pelo menos um lado da casa estiver exposta a esses níveis.

Se se desejar realizar cálculos mais avançados pode ser tida em consideração a presença de não residentes. Por exemplo, em zonas turísticas, o número de pessoas presente em alturas de férias é consideravelmente mais alto que o número de habitantes permanentes, o mesmo sucedendo em áreas de serviços das grandes cidades, durante as horas de trabalho. Por razões de consistência, será necessário compensar a ausência dessas pessoas noutra local, para evitar uma dupla contagem (CE, 2002a).

3.5 Avaliação da incerteza

Nenhuma avaliação deve ser feita sem alguma estimativa de incerteza dos seus resultados, mesmo que de uma forma descritiva. Isto pode ser feito através de uma análise de sensibilidade, de níveis de confiança ou de estimativas directas da incerteza nos dados à medida que se propagam através da aplicação de um modelo (Rossiter, 1994).

A incerteza emergiu na comunidade científica como o termo preferido para tudo o que uma base de dados não conseguiu captar acerca do mundo real, ou a diferença entre o que a base de dados indica e o que existe na realidade. A qualidade de uma base de dados SIG é avaliada pelo seu nível de incerteza. Podem-se distinguir duas classes de incerteza (Eastman et al., 1993):

- Incerteza inerente aos dados tem a ver com a observação da natureza ou da sociedade: não se tem a certeza do que se observa ou mede. Pode estar associada à medição e/ou à amostragem;
- Incerteza inerente à regra ou modelação, tem a ver com o modo como se interpretam essas observações: existe incerteza acerca das conclusões a retirar dos dados (mesmo que perfeitos).

Quanto à incerteza associada aos dados, em SIG as fontes óbvias incluem a idade dos dados, cobertura da área (parcial ou completa), escala do mapa, densidade das observações, relevância e formato dos dados. Deve também ser considerada a incerteza resultante da variabilidade natural das medições originais, que pode estar associada à exactidão posicional, exactidão de conteúdo (qualitativo e quantitativo), fontes de variação nos dados, falhas na entrada ou saída dos dados, enviesamento do observador ou variação natural.

Quanto à incerteza associada à regra ou modelação, esta surge do processamento dos dados (e.g. arredondamentos, conversões do tipo de dados), limites da representação numérica (precisão, armazenamento, aspectos da geometria e dos atributos não

capturáveis), erros da digitalização e geocodificação, problemas associados à sobreposição de mapas, problemas de classificação e generalização, metodologia, definição dos intervalos de classes e interpolação (Foote e Huebner, 1996).

Dependendo da área do conhecimento em avaliação os indicadores de incerteza e as suas designações são muito variáveis. Por vezes, a mesma designação usada em contextos diferentes tem significados ligeiramente diferentes. É o caso de precisão, que num contexto SIG tem um significado idêntico à resolução, num contexto de representação numérica significa número de casas decimais, e no contexto de métodos de medição significa a proximidade (desvio padrão) entre várias medições (da mesma situação) usando o mesmo método.

Outro aspecto a considerar na avaliação da incerteza tem a ver com a propagação do erro ao longo de uma cadeia de procedimentos. Uma das mais fortes capacidades de um SIG, particularmente na área das ciências da terra e do ambiente, é permitir derivar novos atributos a partir de outros atributos já inseridos na base de dados SIG, usando as várias funções ou operações de álgebra de mapas e de interpolação. Quando mapas armazenados numa base de dados SIG são usados como entrada de uma operação SIG, os seus erros são propagados para o produto resultante. A propagação do erro continua quando o resultado de uma operação é usado como dado de entrada de outra operação. Consequentemente, quando não é guardada informação sobre a incerteza dos mapas intermédios, é extremamente difícil avaliar a exactidão dos resultados.

Deste modo, o desenvolvimento de estratégias para gerir a incerteza é fundamental num projecto SIG e deverá ter em consideração os seguintes factores:

- Desenvolvimento de modelos de incerteza rigorosos e formais;
- Entendimento de como a incerteza se propaga através do processamento e tomada de decisão;
- Desenho de técnicas para obter informação geográfica adequada para uso e reduzir a incerteza para níveis geríveis;
- Comunicação da incerteza em formas mais claras/ explícitas (esclarecedoras) tendo em consideração os diferentes níveis de utilizadores;
- Aprender a tomar decisões quando a incerteza é apresentada em informação geográfica, isto é, sendo capaz de absorver a incerteza e conviver com ela diariamente (Gary J. Hunter, 1998).

3.6 Utilização de ferramentas de SIG na avaliação da qualidade do ar e estimativa da população exposta

Os SIG podem ser definidos como “...um conjunto poderoso de ferramentas para colectar, armazenar, recuperar, transformar e visualizar dados sobre o mundo real com um determinado objectivo...” (Burrough, 1986).

A utilização de ferramentas SIG para avaliação da qualidade do ar, dado o carácter eminentemente espacial dos fenómenos em estudo, é de grande utilidade uma vez que apresentam boas potencialidades de integração e análise de informação relativa a qualidade do ar e a variáveis relacionadas, sendo uma ferramenta em constante desenvolvimento. Técnicas avançadas em SIG têm vindo a ser desenvolvidas para fornecer capacidades de análise espacial para objectivos de modelação.

Tal como já referido anteriormente, uma das necessidades em qualidade do ar relaciona-se com a obtenção de concentrações onde não existem estações de medição através de técnicas de interpolação que transformam valores pontuais em superfícies contidas no limite da malha de pontos considerada. Os SIG integram esta técnica e possibilitam procedimentos de análise, visualização, integração, combinação flexível, armazenamento dos conjuntos de dados e obtenção de cartografia georreferenciada, contribuindo desta forma com uma mais-valia para o mapeamento da qualidade do ar (Silva e Amorim, 2002).

A visualização é outra capacidade dos SIG que pode ser utilizada em modelação da qualidade do ar. As ferramentas comerciais dos SIG detêm formas intuitivas de visualização e manipulação dos dados. Os resultados numéricos dos modelos podem ser extremamente volumosos e a sua visualização através de mapas pode sistematizar a informação e facilitar o processo de tomada de decisão. Tem sido argumentado que as tecnologias de visualização, significando geração computacional de gráficos e áudio, têm uma importância sem paralelo para lidar com estes resultados. Existem evidências de que a visualização tornou importantes introspecções científicas possíveis, e afirma-se que, quanto mais interactiva é uma visualização, e maior a quantidade de dados que possa apresentar de uma só vez sem serem ambíguos, mais útil esta pode ser (Bishop e Karadaglis, 1996).

A modelação baseada em dados espaciais dentro de um SIG pode ser incorporada como uma operação *stand-alone* (independente) com os dados a serem transferidos entre o SIG e o software de modelação com recurso a importações e exportações. Esta abordagem tem a vantagem do processo de modelação poder ser muito rápido e eficiente correndo num código compilado e optimizado. Enquanto que o processo de modelação pode ser rápido, o processo de transferência de dados pode ser lento e inadequado. O processo pode ser

dividido nos seguintes módulos: Especificação dos objectivos; Operação dos modelos; Análise e mapeamento dos resultados (Bishop e Karadaglis, 1996).

Em síntese, nas situações em que os dados de monitorização são escassos, os SIG disponibilizam um conjunto de ferramentas para estimar os níveis de poluição em locais sem monitorização e para construir mapas de poluição, recorrendo à interpolação usando métodos como o *Kriging*, o *IDW* ou o mapeamento por regressão. Disponibilizam também os meios para ligar estas estimativas de poluição à distribuição da população, e assim avaliar as exposições por indivíduo ou grupos de população. À data, a maioria das aplicações destes métodos têm sido relativamente simples: a maioria usa a proximidade ou distância, e técnicas como a análise de ponto-para-polígono, intersecção ou zonas tampão (*buffers*). Na maioria dos casos, os métodos têm-se focado nas exposições crónicas (longo prazo), onde os factores de variação temporal podem ser ignorados. Assim sendo, pouca atenção tem vindo a ser dada às dinâmicas populacionais, tendo em vez disso a exposição sido apenas definida em termos do local de residência da população. Estas simplificações são normalmente justificáveis, por exemplo, devido à natureza geral da hipótese a ser testada, incertezas acerca de um tóxico específico envolvido, o período de exposição de interesse ser longo ou a pobre qualidade dos dados de saúde. No entanto, não é necessário fazer estas simplificações. Com a ajuda dos SIG é agora possível desenvolver modelos de exposição mais realísticos, incluindo os padrões das actividades temporais das pessoas. Neste caso, será necessário usar modelos dinâmicos, para avaliar quer a dispersão da poluição quer a exposição da população (Briggs, 2005).

4 Metodologia

No presente capítulo apresenta-se a metodologia seguida para a obtenção dos mapas de distribuição das concentrações dos poluentes atmosféricos e estimativa da área e população expostas a excedências aos valores limite de NO_2 . No diagrama da **Figura 7** apresentam-se as várias fases da metodologia seguida.

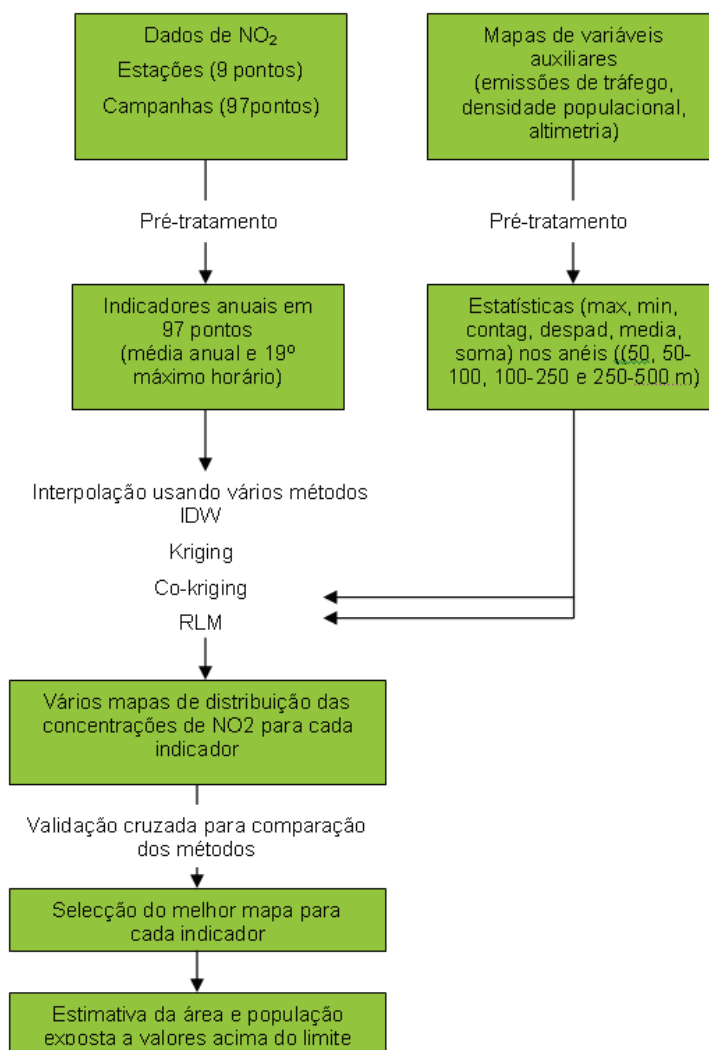


Figura 7. Diagrama geral da metodologia.

4.1 Pré-tratamento de dados de NO_2

Os dados das concentrações de NO_2 , no ar ambiente, para a cidade de Lisboa em 2001 e 2002, usados neste estudo, foram obtidos junto da Comissão de Coordenação do Desenvolvimento Regional de Lisboa e Vale do Tejo (CCDR-LVT). Dizem respeito a um reduzido número de estações de monitorização fixa (nove) e a duas campanhas de curta

duração, usando o método de difusão passiva, com um número elevado de pontos (cerca de 100). O pré-tratamento efectuado consistiu na estimativa dos indicadores anuais de NO₂ para as estações de monitorização fixas e a sua extrapolação para as 100 localizações das campanhas.

4.1.1 Estações de monitorização fixa

Em 2001 e 2002, existiam na cidade de Lisboa, para além das 7 estações de monitorização da qualidade do ar actuais, mais 2 estações em funcionamento, Av. Casal Ribeiro (2001 e 2002) e Rua da Prata (2001). Estas estações pertencem à rede de estações de monitorização em contínuo gerida pela CCDR-LVT e estão equipadas com o método de referência previsto na legislação actual.

O pré-tratamento dos dados das estações de monitorização fixa implicou o cálculo das concentrações de NO₂, expressas nos indicadores estatísticos anuais, média anual e 19º máximo horário, e da eficiência anual (percentagem de valores horários disponíveis relativamente ao número de horas de um ano), que permitem avaliar o cumprimento dos valores limite para o NO₂ em cada estação para os anos em estudo. Foram ainda calculadas as concentrações médias de NO₂, verificadas nas várias estações, durante o período das campanhas de monitorização.

4.1.2 Campanhas de monitorização

Dado o número de estações existentes dentro da cidade de Lisboa ser reduzido tendo em conta o presente objectivo de modelar a distribuição espacial dos poluentes usando vários métodos de interpolação, foi necessário recorrer a duas campanhas de monitorização de curta duração promovidas pela CCDR-LVT com a colaboração, entre outras entidades, da FCT/UNL (Ferreira, 2003). As campanhas foram realizadas, utilizando tubos de difusão *Radiello* (Método de Difusão Passiva - MDP) (**Figura 8**).



Figura 8. Tubos de difusão standard da Radiello.

Em cada campanha, e durante uma semana, foram expostos tubos de difusão para monitorizar vários poluentes, entre eles, o NO_2 . Cerca de 100 pontos de amostragem foram localizados dentro da cidade de Lisboa e tendo sido classificados de acordo com o nível de tráfego do local em fundo, médio e *hotspot* (**Figura 9**). As análises dos tubos, após a sua exposição de 7 dias, foram realizadas na Fondazione Salvatore Maugeri, sediada em Pádova (Itália) – por serem os fabricantes desta marca e por possuírem bastante experiência em medições deste tipo e assegurando uma série de procedimentos de controlo e garantia de qualidade das análises.

Estas campanhas foram realizadas durante uma semana no Verão de 2001 (de 3 a 10 de Julho) e outra semana no Inverno de 2002 (de 29 de Janeiro a 5 de Fevereiro) com o objectivo de avaliar a distribuição espacial da qualidade do ar na região de Lisboa sob duas condições meteorológicas diferentes, complementando assim as medições obtidas nas estações fixas de monitorização em contínuo.



Figura 9. Localização dos pontos de monitorização, estações fixas e tubos de difusão em Lisboa.

O mapeamento dos resultados das campanhas permite verificar quais os pontos e tipos de pontos com concentrações mais elevadas, assim como perceber se existem diferenças significativas entre os dois períodos meteorológicos amostrados. A média das duas campanhas poderá ser indicativa da média anual para 2001 e 2002.

Para efeitos de controlo e garantia de qualidade, dado que o método associado aos tubos de difusão, método de difusão passiva (MDP), não é um método de referência (MR) para a avaliação de NO_2 , foram colocados três tubos de difusão (MDP) junto a cada uma das nove

estações fixas equipadas com o MR, para posterior cálculo da incerteza global do MDP relativamente ao MR (**Figura 10**).



Figura 10. Colocação de tubos de difusão em triplicado junto a cada estação de monitorização da qualidade do ar fixa para controlo de qualidade.

4.1.3 Obtenção de indicadores anuais extrapolando os dados das campanhas usando regressão linear simples

Uma vez que os resultados obtidos nas campanhas descritas no ponto anterior, representam concentrações médias semanais e os dois tipos de valores limite especificados na legislação para o NO₂ (ver **Tabela 1**) dizem respeito a um ano, não era possível saber, usando apenas os dados das campanhas, se os níveis registados estavam abaixo ou acima dos valores definidos na legislação.

No entanto, a utilização conjunta dos dados das estações fixas e dos três tubos colocados junto a elas (para efeitos de garantia e controlo de qualidade) permitiu fazer uma estimativa dos indicadores anuais nos vários pontos. Deste modo, para as estações que obtiveram uma eficiência anual superior a 85% ¹, em 2001 e 2002, foram estabelecidas regressões lineares simples entre as concentrações de NO₂, expressas nos indicadores média anual e 19º máximo horário, obtidas pelo método de referência, e as concentrações médias semanais dadas pelas médias dos 3 tubos de difusão (método de difusão passiva).

¹ De acordo com o Anexo X do Decreto-Lei n.º 111/2002 a recolha de dados anual deve ser no mínimo de 90% para que se considere que existiu medição em contínuo, no entanto, segundo o Guidance on the Annexes to Decision 97/101/EC on *Exchange of Information as revised by Decision 2001/752/EC for the European Commission* (Comissão Europeia, 2002) tendo em conta que a manutenção e calibração dos aparelhos pode representar uma perda de 5% dos dados, a eficiência mínima anual requerida para se considerar que há monitorização em contínuo é reduzida para 85%

Estas relações lineares foram testadas usando a média dos tubos para cada uma das campanhas individualmente e para a média das duas. Relativamente aos dados das estações, foi testado cada ano individualmente e também a média dos dois anos (2001 e 2002).

4.2 Pré-tratamento de variáveis auxiliares

Dois dos métodos de interpolação a utilizar no presente trabalho, como se verá adiante, o *ordinary cokriging* e a regressão linear múltipla permitem a utilização de variáveis auxiliares para melhorar a estimativa das concentrações em locais sem medição.

O aumento da disponibilidade de dados espaciais (mapas) relativos a diversas variáveis com crescente qualidade, permitiu a selecção, aquisição e utilização de vários tipos de variáveis que, por estarem relacionadas com a emissão e dispersão dos poluentes em ambientes urbanos, podem auxiliar na modelação da distribuição espacial da qualidade do ar dentro da cidade de Lisboa.

Os mapas de dados espaciais possíveis de reunir para a cidade de Lisboa, e que foram testados como variáveis auxiliares estão discriminados na **Tabela 4** de acordo com o modelo de dados, tipo de entidades, resolução, e a fonte/autoria dos mapas.

Variável auxiliar	Tipo de dados	Modelo de dados	Tipo de entidades	Resolução	Fonte
Emissões de tráfego	Quantitativos	Vectorial	Linhas		Gois et al., 2006
Densidade populacional por subsecções estatísticas	Quantitativos	Vectorial	Polígonos		Calculado a partir do mapa de população residente do INE Censos 2001
Ocupação do solo	Categóricos	Vectorial	Polígonos		IGP: Corine Land Cover 2000
Altimetria	Quantitativos	Raster	Polígonos	90 m	NASA

Tabela 4. Variáveis auxiliares usadas.

O modelo de tráfego rodoviário detalhado e o modelo de distribuição espacial das emissões que foram usados no presente trabalho estão incluídos no inventário das emissões atmosféricas para a cidade de Lisboa que foi recentemente sujeito a melhorias consideráveis (Gois et al., 2006). O mapa das vias de tráfego, que serviu de base à desagregação espacial das emissões, foi disponibilizado pela Câmara Municipal de Lisboa. O desenvolvimento destes modelos baseou-se nas seguintes actividades:

- Identificação de veículos em circulação usando fotografia aérea;
- Velocidade através de monitorização por GPS no veículo;
- Vendas de combustível e movimentos pendulares;

- Campanhas de amostragem para a caracterização da frota em circulação.

O pré-tratamento efectuado a algumas variáveis auxiliares, baseado em anéis de influência, resultou da verificação de que a relação entre as concentrações do poluente num determinado local estão mais correlacionadas com as características das fontes emissoras e condições de dispersão na sua envolvente do que com as fontes emissoras e condições de dispersão nesse local, e ainda que essa correlação varia com a distância da envolvente considerada. Deste modo as variáveis utilizadas necessitaram de um pré-tratamento que consistiu no cálculo de várias estatísticas da variável auxiliar (como a soma, máximo, média, mínimo, desvio padrão e contagem) na envolvente do ponto amostrado que se pretende modelar. Para tal foram construídos anéis a várias distâncias relativamente aos pontos de monitorização. O objectivo é caracterizar a estrutura espacial da variável auxiliar a várias distâncias do ponto de amostragem. Na **Figura 11** é possível ver o exemplo dos anéis construídos para modelar as concentrações de NO₂ na cidade de Lisboa partindo dos resultados das campanhas de tubos de difusão.

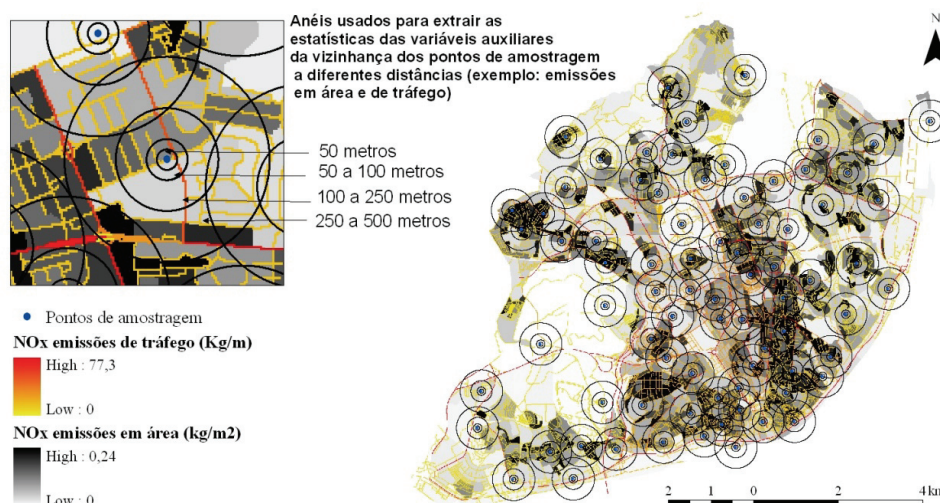


Figura 11. Anéis usados para calcular as estatísticas das variáveis auxiliares na envolvente dos locais amostrados na cidade de Lisboa durante a campanha com tubos de difusão.

O primeiro passo efectuado foi a transformação, no caso de algumas das variáveis, do modelo de dados vectorial para *raster* com uma resolução de 10 m. Seguidamente foram calculadas para as variáveis quantitativas, como a densidade populacional e as emissões de tráfego, as estatísticas (soma, máximo, média, mínimo, desvio padrão, contagem) para as várias distâncias definidas (50 m, 50-100 m, 100-250 m, 250-500 m) que resultou num conjunto de 24 estatísticas para cada variável auxiliar. Para a variável categórica, ocupação do solo, foi usada apenas a estatística contagem. A variável altitude não se revelou útil para a modelação da distribuição espacial das concentrações de NO₂, dado não se terem verificado relações significativas entre si quer no local dos pontos de amostragem, quer na sua envolvente a várias distâncias.

Usando os dados em análise (concentrações de NO₂, expressas na média anual e 19º máximo horário, para cada ponto de amostragem e estatísticas de cada variável auxiliar dentro de cada anel concêntrico), foi ainda produzida uma matriz de correlações entre o indicador do poluente que se pretende modelar e as estatísticas de cada variável auxiliar. O objectivo deste procedimento foi identificar as variáveis, estatísticas e anéis que mais influenciavam as concentrações de NO₂.

4.3 Modelação da distribuição espacial

Na modelação da distribuição espacial ou mapeamento do NO₂ partindo dos resultados dos 97 pontos das campanhas, extrapoladas para os dois indicadores, média anual e 19º máximo horário (ver capítulo 4.1.3), foram testados vários métodos de interpolação alternativos: método determinístico, o inverso da distância ponderada (*IDW*), métodos geoestatísticos, o *ordinary kriging* e o *ordinary cokriging* e método de regressão linear múltipla. Para a implementação destes métodos determinísticos e geoestatísticos recorreu-se à extensão *Geostatistical Analyst* do *ArcGIS* 9.2 e para o método estatístico usou-se o *SPSS 15.0* para *Windows*.

4.3.1 Inverso da Distância Ponderada (IDW)

Para a obtenção da superfície de concentrações de NO₂, relativas à média anual e 19º máximo horário, utilizando o método *IDW* foi necessário definir os seguintes parâmetros: o expoente (*power*), o número de vizinhos a incluir e o formato da figura para pesquisa da vizinhança do ponto.

4.3.2 Ordinary Kriging

Para obter as superfícies de concentrações de NO₂, relativas à média anual e 19º máximo horário, usando o método de interpolação *Ordinary Kriging* foi necessário, em primeiro lugar ajustar o variograma testando vários valores para cada um dos parâmetros que o definem (tipo de modelo matemático; *range*; *sill*; *nugget*; *lag*), e seguidamente foi definida a vizinhança de pesquisa.

4.3.3 Ordinary CoKriging

A obtenção das superfícies de interpolação usando o método *ordinary cokriging* é muito semelhante à seguida para o método *ordinary kriging*. A principal diferença é a possibilidade de introdução de variáveis auxiliares. É necessário definir o variograma, que entra não só com a autocorrelação da variável a interpolar (concentrações de NO₂), como com as correlações cruzadas entre esta e as variáveis auxiliares, usadas para melhorar a previsão,

sendo ainda, tal como no *Ordinary Kriging*, necessário definir a vizinhança de pesquisa. Foram testadas algumas das estatísticas de variáveis auxiliares (ver capítulo 4.2) seleccionadas tendo em consideração as que obtiveram melhor correlação com o indicador a interpolar.

4.3.4 Regressão linear múltipla

Após as fases de obtenção de dados de NO₂ relativos a 97 pontos de amostragem, expressos nos indicadores, média anual e 19º máximo horário (ver capítulo 4.1.3) e de selecção e preparação das variáveis auxiliares (produção de estatísticas das variáveis tal como descrito no capítulo 4.2), foi possível a utilização de modelos de regressão linear múltipla como método de interpolação fazendo:

- O ajustamento de modelos de RLM;
- A aplicação dos modelos a uma grelha de pontos.

O ajustamento de vários modelos de regressão linear múltipla foi feito usando como variáveis dependentes a média anual e 19º máximo horário de dióxido de azoto (NO₂) nos 97 pontos das campanhas, e como variáveis independentes i) as variáveis estatísticas, soma, máximo, média, mínimo, desvio padrão e contagem de células dentro dos vários anéis (50 m, 50-100 m, 100-250 m e 250-500 m) criados na envolvente dos 97 pontos de amostragem, ii) das emissões em área de NO_x (directamente proporcional à densidade populacional) iii) emissões de tráfego de NO_x ao nível da via, iv) da estatística contagem dentro dos mesmos anéis para as várias classes de ocupação do solo (urbano contínuo, urbano descontínuo, aeroporto, parque urbanos, agricultura e Porto de Lisboa) (**Tabela 5**).

Anéis/Estatísticas	Mínimo	Máximo	Média	Desvio Padrão	Soma	Contagem
Anel 50 m						
Anel 50-100 m	EANOX;	EANOX;	EANOX;	EANOX;	EANOX;	EANOX;
Anel 100-250 m	ETNOx	ETNOx	ETNOx	ETNOx	ETNOx	ETNOx;
Anel 250-500 m						COS

Nota: EANOX - Emissões em área de NO_x; ETNOx – Emissões de tráfego de NO_x; COS – classes de ocupação do solo

Tabela 5. Estatísticas das variáveis auxiliares usadas como variáveis independentes.

Foram construídos diversos modelos usando várias combinações de variáveis auxiliares, anéis e estatísticas. Nomeadamente, modelos incluindo apenas as várias estatísticas de cada uma das variáveis, as várias estatísticas para cada um dos anéis e todas as estatísticas para todas as variáveis. Para a selecção das variáveis a considerar nos modelos, foram testados os métodos automáticos de *Forward* e *Backward*, tendo sido imposto que as variáveis de entrada deveriam ter um nível de significância inferior a 0,05 e as de saída um nível de significância superior a 0,1. Em qualquer dos casos, foi avaliada a

multicolinearidade entre as variáveis seleccionadas. Foi ainda analisada a eventual presença de pontos considerados *outliers* e posterior exclusão desses casos.

Na selecção dos melhores modelos foram tidos em consideração os seguintes factores:

- As variáveis seleccionadas pelos métodos automáticos como o ponto de partida para a definição do melhor modelo.
- A existência ou não de multicolinearidade. Não deve haver colinearidade entre variáveis a incluir no modelo;
- O número de variáveis preditoras incluídas em cada modelo. O número deve ser o mais baixo possível;
- A avaliação da incerteza: ajustamento dos modelos e análise de erro com base em critérios objectivos, que se apresentam adiante (ver capítulo 4.4.3).

Finalmente, com o objectivo de obter uma superfície de concentrações de NO_2 , foi construída uma grelha de pontos distando 100 m entre si (**Figura 12**).

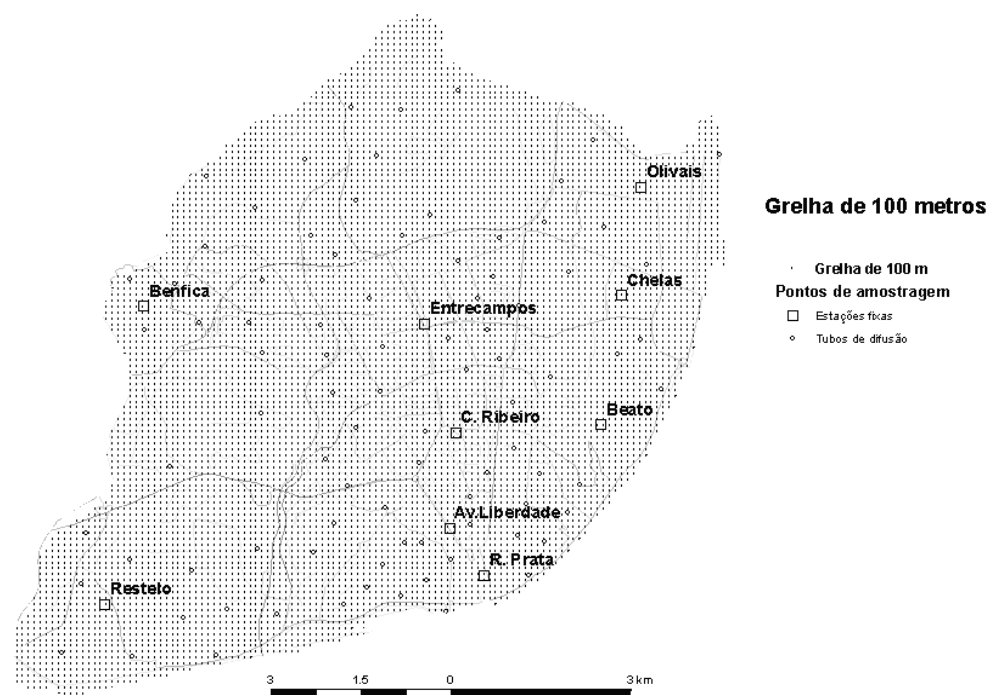


Figura 12. Grelha de 100 m aplicada à cidade de Lisboa.

Foram depois criados, para cada um dos 8500 pontos resultantes, anéis de 50 m, 50-100 m, 100-250 m e 250-500 m e calculadas estatísticas (soma, máximo, média, mínimo, desvio padrão e contagem) para as variáveis auxiliares mais adequadas, para posterior aplicação dos modelos de regressão seleccionados. Finalmente, a grelha de 8500 pontos foi convertida em polígonos, para se obter uma superfície contínua.

A escolha da grelha de 100x100 m teve em consideração os seguintes factores:

- Possuir resolução suficiente para distinguir o efeito das estradas;
- Minimizar o tempo de cálculo dos indicadores dentro de cada anel;
- Estar de acordo com a resolução indicada na Directiva 2008/50/CE para pontos de tráfego;
- O raio mínimo dos anéis (50 m) usado na modelação.

4.4 Avaliação da incerteza

Associadas a todos os passos da metodologia exposta anteriormente existem fontes de incerteza. O erro pode ser calculado para cada procedimento da metodologia ou, pode ser avaliado o erro global da metodologia, comparando os valores reais com valores modelados resultantes da sua aplicação.

As fontes de incerteza relacionadas com a metodologia exposta são várias e podem salientar-se os erros associados aos dados, onde se incluem os de entrada, nomeadamente:

- Medições, tanto as medições de concentrações de NO₂ obtidas com o método de referência (assumidos como valores reais) como com o método de difusão passiva;
- Variáveis auxiliares (estimativas de emissões e de densidade populacional).

Para além destes, existem também erros associados aos modelos e métodos usados, onde se incluem os seguintes procedimentos:

- Pré-tratamento das variáveis auxiliares;
- Pré-tratamento dos dados (extrapolação dos resultados da campanhas com MDP para o MR, do período da campanha para o período anual);
- Interpolação determinística, geoestatística e usando regressão múltipla;
- Estimativa da população.

Nem sempre foi possível fazer uma estimativa quantitativa dos erros associados a cada operação/procedimento efectuado. De facto, a obtenção de valores reais necessários à validação de cada procedimento, para muitos dos passos da metodologia, revelou-se muito difícil. Este é o caso da incerteza associada às variáveis auxiliares (em termos de correcção posicional e de conteúdo), nomeadamente às emissões de tráfego, e ao seu pré-tratamento, e também às estimativas de áreas e população expostas.

Dado serem conhecidos os valores “reais” das concentrações de NO₂ para os locais das estações fixas existentes na cidade de Lisboa, foi possível obter a incerteza global associada às campanhas de monitorização, à extrapolação destes dados para os indicadores anuais e à aplicação de cada método de mapeamento/interpolação escolhido.

Não é possível saber com rigor, para os diferentes métodos em que se recorreu a variáveis auxiliares (sujeitos a pré-tratamentos), a contribuição destas para a incerteza final, nem como ocorreu a propagação dessa incerteza ao longo da metodologia.

Nos pontos seguintes são apresentados os parâmetros estatísticos escolhidos para avaliar a incerteza associada a cada um destes três passos principais da metodologia.

4.4.1 Avaliação da incerteza das campanhas de monitorização

Na avaliação da incerteza associada às campanhas de monitorização com tubos de difusão (entre o MDP e o MR) foram usados os parâmetros estatísticos: exactidão, precisão e incerteza global, de acordo com o especificado na EN 13528-1:2002 (CEN, 2002). Para obter dos parâmetros estatísticos associados a cada campanha foi calculada a média destes parâmetros em todas as estações para as quais existiam dados disponíveis.

Os parâmetros estatísticos de incerteza para cada estação foram calculados de acordo com a norma referida usando as seguintes formulas:

$$\begin{aligned} incerteza (\%) &= \frac{|\bar{x} - x_{ref}| + 2\sigma}{x_{ref}} \times 100 & exactidão (\%) &= \frac{\bar{x} - x_{ref}}{x_{ref}} \times 100 \\ precisão (\%) &= \frac{2\sigma}{x_{ref}} \times 100 \end{aligned}$$

Em que: $\bar{x}$ – é o valor da média dos resultados das repetições das medições usando tubos de difusão; x_{ref} – é o valor de concentração verdadeiro ou aceite como tal, neste caso, o valor medido pelo equipamento de referência da estação fixa; σ – é o desvio padrão das medições com tubos de difusão.

4.4.2 Avaliação da incerteza na obtenção de indicadores anuais

A avaliação da incerteza associada à obtenção dos dois indicadores anuais (média anual 2001-2002 e do 19º máximo horário 2001-2002) para todos os locais amostrados durante as campanhas, recorrendo a relações lineares estabelecidas entre valores dos equipamentos

fixos com MR (reais) e valores dos tubos de difusão (MDP) para o local das estações (ver capítulo 4.1.3), resultou dos seguintes procedimentos:

- Ajustamento dos modelos, em termos de coeficiente de determinação (R^2) da regressão linear entre as concentrações observadas e as modeladas;
- Validação cruzada (VC), que consiste em deixar de fora do ajustamento do modelo um ponto de cada vez, aplicar a esse ponto o modelo construído e calcular o erro entre a concentração medida e modelada. Este procedimento é repetido para todos os pontos. Foi considerado o seguinte conjunto de parâmetros estatísticos de erro resultantes da validação cruzada efectuada:

Erro médio	$\bar{E} = \frac{\sum_{i=1}^n M_i - O_i}{n}$
Erro absoluto médio	$\bar{E}_{abs} = \frac{\sum_{i=1}^n M_i - O_i }{n}$
RMSE (<i>Root Mean Square Error</i>)	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (M_i - O_i)^2}$
RMSE normalizado (NRMSE)	$NRMSE = \frac{RMSE}{\bar{O}_i}$

Em que: M – valor modelado; O – valor observado; n – número de casos; $\bar{M}_i$ – média dos valores observados.

Os modelos são tanto melhores quanto mais próxima de zero seja a média do erro absoluto e mais baixo seja o RMSE.

4.4.3 Avaliação da incerteza da modelação da distribuição espacial

Para avaliar os vários métodos testados foi necessário estipular um conjunto de indicadores objectivos. Como já foi referido no capítulo 3.2.2, a Directiva 2008/50/CE define a incerteza associada ao modelo como o desvio máximo entre as concentrações medidas e calculadas para 90% das estações de medição individuais, durante o período considerado, em relação ao valor limite, independentemente da cronologia das ocorrências. A incerteza associada ao modelo será interpretada como sendo aplicável na região do valor limite pertinente. As medições fixas que forem seleccionadas para comparação com os resultados da modelação devem ser representativas da escala abrangida pelo modelo. O cálculo da incerteza associada ao modelo pode assim ser formulado como (Denby et al, 2008):

$$RDE = \frac{|O - M|}{VL}$$

Em que,

RDE – erro relativo da directiva;

VL – valor limite;

O – valor observado;

M – valor modelado.

Para além deste indicador, que permite averiguar se o mapa é adequado para avaliação do cumprimento da legislação, na avaliação e comparação da qualidade de cada mapa, quer seja obtido usando *IDW*, *ordinary kriging*, *ordinary co-kriging* ou regressão linear múltipla recorreu-se às mesmas técnicas e parâmetros estatísticos de erro apresentados no ponto anterior (Erro médio, Erro absoluto médio, RMSE e NRMSE).

4.4.4 Selecção do melhor mapa e produção de mapas de erro

A selecção do melhor mapa para cada caso é feita utilizando a validação cruzada como teste comparativo. Como parâmetros estatísticos do erro são usados, o erro médio e o RMSE da validação cruzada. Na avaliação do erro foi tido em consideração que a média do erro deve ser o mais próximo de zero possível e o RMSE o mais baixo possível.

Para seleccionar o melhor mapa para cada método, para além da análise visual, é prática comum criar um número elevado de mapas, fazendo variar os parâmetros do método, antes de um deles ser identificado como o melhor e venha a ser escolhido como o final. O procedimento deve ser, ir sistematicamente comparando dois mapas produzidos, em termos visuais e dos parâmetros estatísticos de erro, e ir retirando o pior, até se obter o melhor mapa possível com a técnica em análise.

Este procedimento foi efectuado para seleccionar o melhor mapa produzido com os vários métodos de interpolação, *IDW*, *Ordinary Kriging* e *Ordinary Cokriging* e regressão linear múltipla (RLM). No caso da regressão, quando o número de observações é elevado, é esperado que o *RMSE* total e o *RMSE* da validação cruzada sejam equivalentes. No entanto, no caso de existirem poucas observações, o efeito da validação cruzada pode ser significativo (Denby, 2007).

Após a selecção do melhor mapa para cada método interpolação usado, foram produzidos os mapas de erro associados com o objectivo de permitir a comparação entre os vários métodos. Estes mapas foram obtidos pela multiplicação do NRMSE resultante da validação cruzada (usando apenas os resultados do método de referencia, ou seja, apenas os dados da estações de monitorização fixa) pela superfície de concentrações obtida.

4.5 Estimativa das áreas e da população exposta a poluição atmosférica

Para a estimativa da população exposta, em primeiro lugar, foi necessário fazer uma estimativa das áreas em excedência do valor máximo legislado. Para tal foram usadas varias ferramentas de SIG para calcular as áreas para as quais existe elevada probabilidade de excedência ao valor limite para:

- Seleccionar as áreas;
- Isolar as áreas;
- Calcular as áreas – contagem do número de grelhas e multiplicação pela área de cada grelha (100 m²).

Depois de obtidas as áreas em excedência, foi estimada a população residente nessas mesmas áreas. Para tal foi feita a sobreposição das grelhas que ultrapassam os valores legais com o mapa da densidade populacional em modelo de dados *raster* (100 m²) os dados do Instituto Nacional de Estatística (INE) de população residente para a subsecção estatística definida pelo INE (INE, 2001).

4.6 Diagrama geral: operações em ambiente SIG

Como resumo da metodologia exposta nos pontos anteriores deste capítulo, foram desenhados dois diagramas. Na **Figura 13**, é possível visualizar o diagrama geral para obtenção e selecção dos mapas de distribuição das concentrações de NO₂ e na **Figura 14**, encontra-se o esquema geral para a obtenção das área e população expostas a determinados níveis de poluição, partindo dos mapas de distribuição das concentrações seleccionados. Os diagramas produzidos salientam em caixas, os dados de entrada e os vários produtos obtidos em cada fase da metodologia. Entre cada caixa, a cinzento, estão especificadas as operações implementadas em ambiente de SIG para obter o produto discriminado dentro de cada caixa.

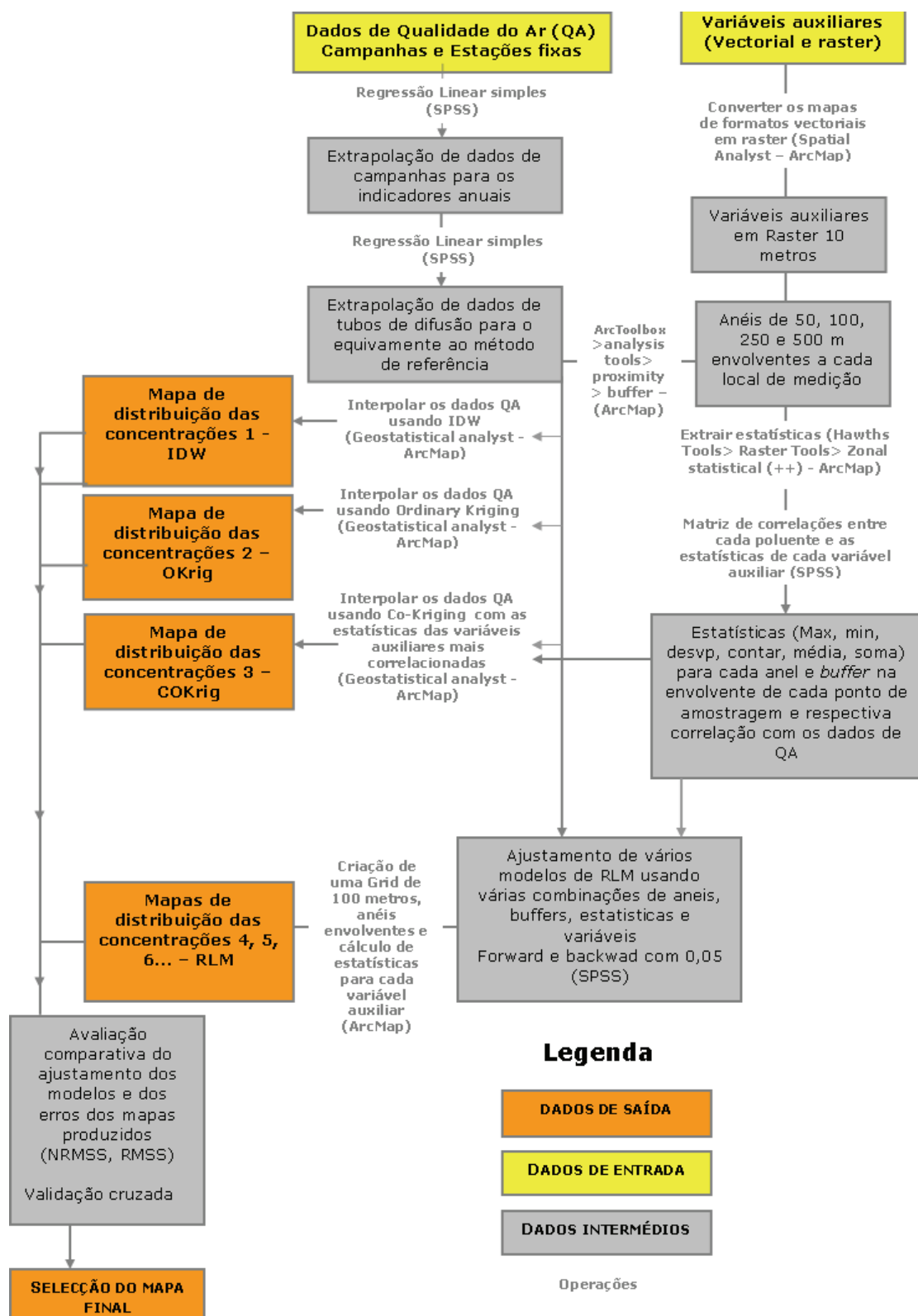


Figura 13. Esquema geral para a obtenção e selecção dos mapas de distribuição das concentrações.

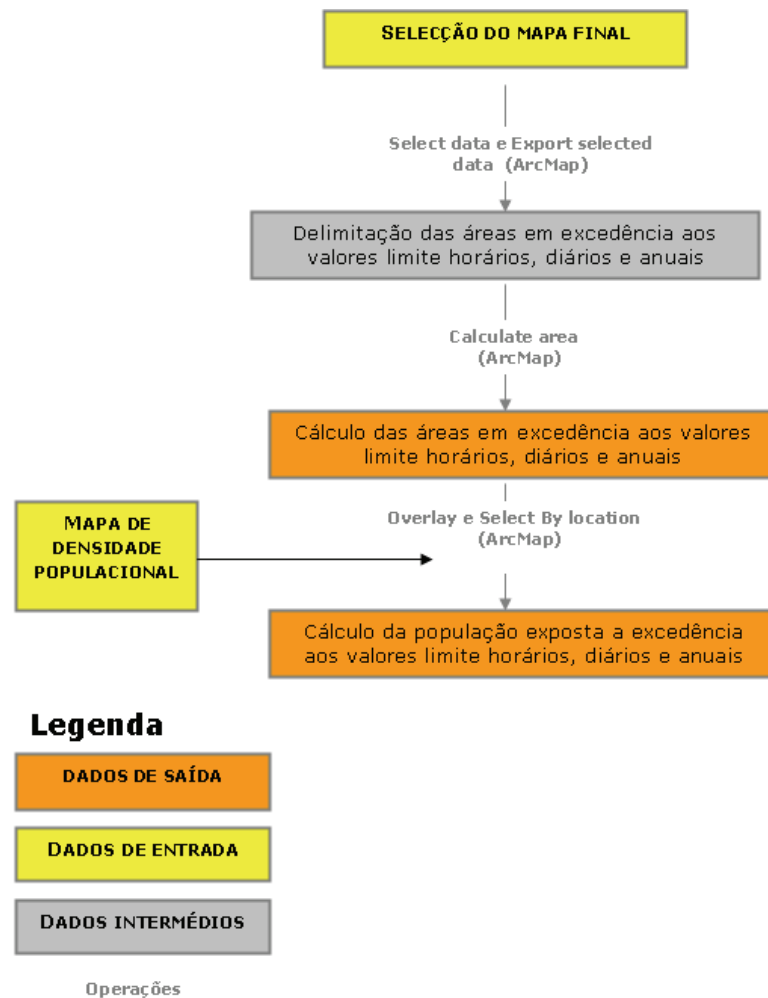


Figura 14. Esquema geral para a obtenção das áreas e população expostas a determinados níveis de poluição partindo dos mapas de distribuição das concentrações seleccionados.

5 Resultados e discussão

Neste capítulo são apresentados os resultados da aplicação da metodologia exposta anteriormente, utilizando, como exemplo, as concentrações de NO₂ medidas na cidade de Lisboa durante os anos de 2001 e 2002.

5.1 Pré-tratamento de dados de NO₂

Neste ponto apresenta-se o pré-tratamento realizado aos dados das estações de monitorização fixa, referentes a 2001, e das campanhas de monitorização com duração semanal, realizadas no Verão de 2001 e Inverno de 2002, com tubos de difusão colocados em 97 locais na cidade de Lisboa, com o objectivo de estimar as concentrações de NO₂ expressas nos indicadores média anual e 19º máximo horário, para análise do cumprimento da legislação.

5.1.1 Estações de monitorização fixa

Os resultados da avaliação do cumprimento da legislação actual relativamente ao NO₂, para as nove estações de monitorização existentes na cidade de Lisboa, durante os anos de 2001 e 2002, apresentam-se na **Tabela 6** e, em mapa, na **Figura 15**. Da análise dos resultados pode concluir-se que a média anual apresentou ultrapassagens ao valor limite (VL) em ambos os anos, em todas as estações de tráfego, à excepção de Entrecampos em 2001. Na Avenida da Liberdade em 2002 foi mesmo ultrapassado o valor limite acrescido da margem de tolerância (VL+Mt). Para o indicador horário à escala anual, o 19º máximo horário, foi apenas ultrapassado o valor limite na estação de tráfego de Benfica. É ainda de salientar que nas estações do Restelo e Entrecampos, em 2002, e na estação da Rua da Prata, em 2001, não foi cumprida a eficiência exigida na legislação para estações fixas, pelo que os resultados nestes casos devem ser analisados com reservas.

A avaliação global dos resultados indica que as concentrações de NO₂ são mais elevadas junto às vias de maiores fluxos de tráfego. Este facto é consistente com o inventário de emissões relativo a 2001 e 2002 produzido pela CCDRLVT e INVENTAR em 2006 (Gois et al., 2006), que indica que o sector dos transportes rodoviários é responsável por mais de 75% das emissões de NO_x na cidade de Lisboa, sendo o sector residencial, comércio e serviços responsável por 10%, a indústria por 8% e os restantes cerca de 5% são atribuídos à aviação (**Figura 16**).

Nome Estação	Tipo Estação	Tipo Zona Estação	Ano	Eficiência (%)	Valor máximo anual ($\mu\text{g}/\text{m}^3$)	Média anual ($\mu\text{g}/\text{m}^3$)	19º Máximo horário ($\mu\text{g}/\text{m}^3$)
Beato	Fundo	Urbana	2001	100	122	20	90
			2002	95	164	23	118
Chelas	Fundo	Urbana	2001	99	210	26	149
			2002	94	177	23	108
Olivais	Fundo	Urbana	2001	95	135	22	104
			2002	99	148	23	103
Restelo	Fundo	Urbana	2002	83	216	18	140
Avenida da Liberdade	Tráfego	Urbana	2001	96	309	54	178
			2002	96	212	62	169
Benfica	Tráfego	Urbana	2001	96	263	48	216
			2002	91	199	45	154
Casal Ribeiro	Tráfego	Urbana	2001	92	229	49	160
			2002	96	226	44	180
Entrecampos	Tráfego	Urbana	2001	85	125	27	103
			2002	75	151	43	127
Rua da Prata	Tráfego	Urbana	2001	66	171	44	136

Nota: Incumprimento do mínimo de eficiência da estação

Ultrapassagem do VL

Ultrapassagem do VL+MT

Tabela 6. Avaliação do cumprimento do DL n.º 111/2002 para o NO_2 relativa à protecção da saúde humana em 2001 e 2002

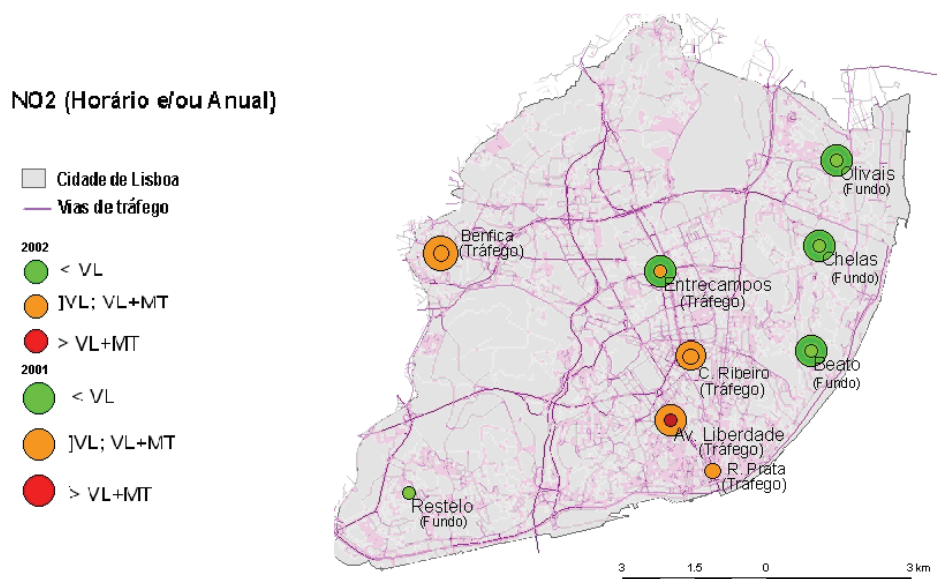


Figura 15. Verificação do cumprimento da legislação relativa ao NO_2 dentro da cidade de Lisboa.

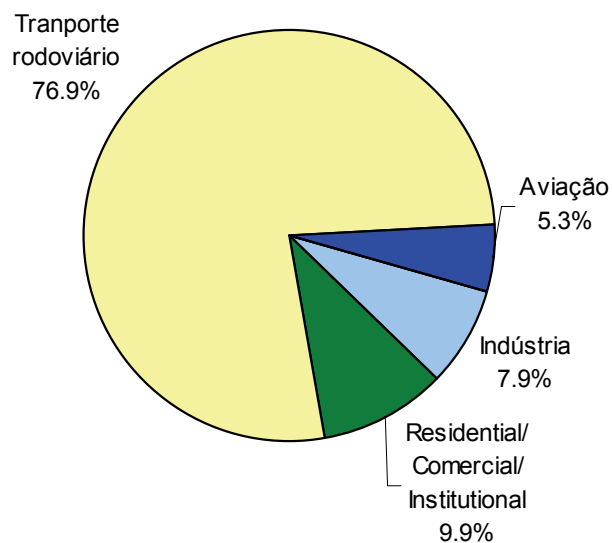


Figura 16. Distribuição das emissões de NO₂ (média 2001-2002) por sector de actividade na cidade de Lisboa (Gois et al., 2006).

5.1.2 Campanhas de monitorização

Os resultados das duas campanhas realizadas no Verão de 2001 e Inverno de 2002 pela CCDR-LVT, relativos às concentrações médias de uma semana de medição com tubos de difusão (método de difusão passiva - MDP) realizada em 100 locais da cidade de Lisboa apresentam-se na **Figura 17** e na **Figura 18**. Nas figuras são realçados os pontos que correspondem a locais das estações de monitorização fixa, onde foram colocados 3 tubos de difusão para efeitos de controlo de qualidade, e é apresentada uma classificação relativamente ao tipo de influência de cada ponto. A observação das duas imagens permite verificar que as concentrações mais elevadas verificaram na campanha de Inverno. Em ambas as campanhas, os pontos de tráfego de uma forma geral, são os que apresentam concentrações mais elevadas e os de fundo (pontos sem influência directa de tráfego) estão associados a níveis de concentração mais baixos.

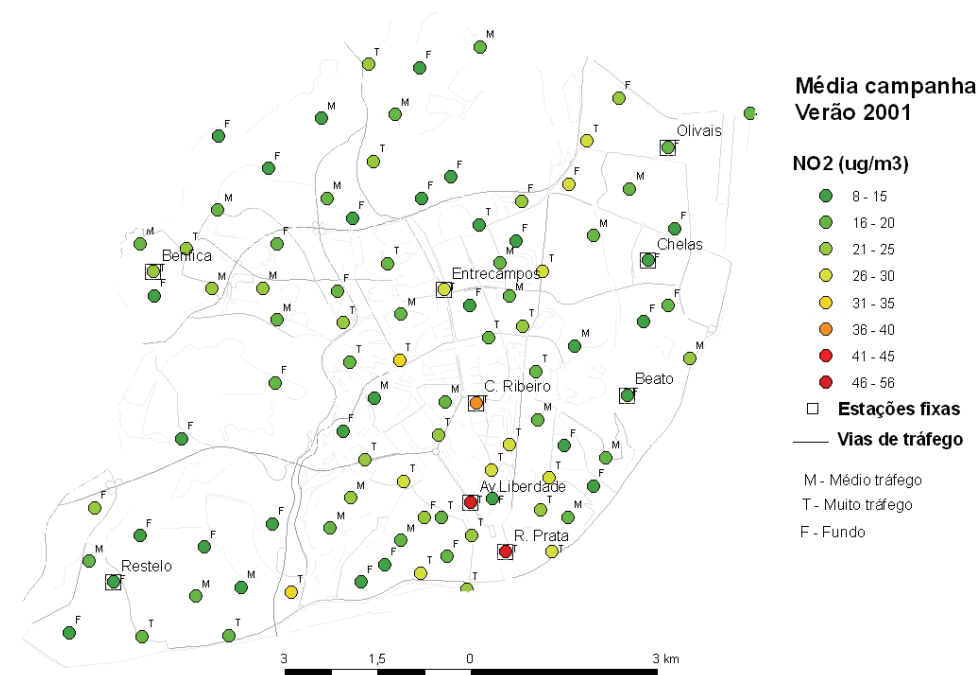


Figura 17. Concentrações médias da semana de 03/07 a 10/07 (campanha Verão 2001).

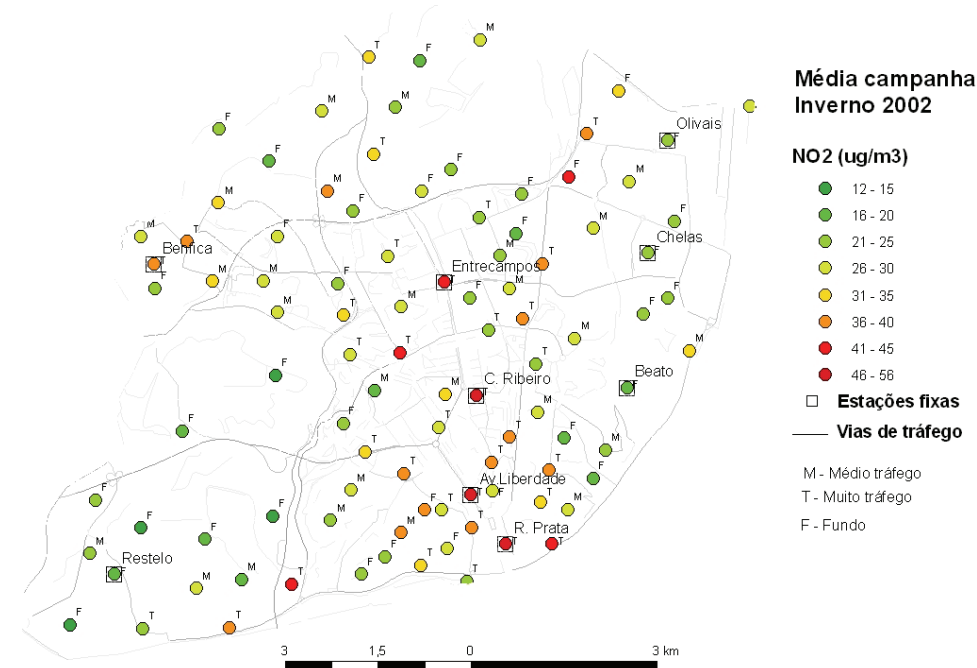


Figura 18. Concentrações médias da semana de 29/01 a 05/02 (campanha Inverno 2002).

Como referido na metodologia, foram colocados 3 tubos de difusão (MDP) para controlo e garantia de qualidade do processo junto às estações fixas. Os resultados dos triplicados permitiram avaliar a incerteza das campanhas (**Tabela 7** e **Tabela 8**) para cada estação

com dados durante o período de cada campanha, e em média para cada uma das campanhas. É de salientar, que não foi possível avaliar a incerteza, por falta de dados suficientes nas estações Olivais, Av. Casal Ribeiro durante o período da primeira campanha e nas estações Rua da Prata, Entrecampos e Restelo durante a segunda campanha. Relativamente à incerteza global, que inclui as duas componentes: precisão e exactidão (ver capítulo 4.4.1), a campanha do Verão de 2001, teve uma incerteza média de 51% (variando entre 25 e 107%) e para o Inverno de 2002, teve uma incerteza média de 35% (entre 21 e 75%).

Os resultados apresentados mostram uma grande variabilidade de qualidade, tanto a nível da precisão como da exactidão, de estação para estação, que poderá estar relacionada com diferentes condições meteorológicas entre locais, nomeadamente em parâmetros como a temperatura, direcção e intensidade do vento e humidade, que podem afectar a absorvância do tubos de difusão. Olhando para os resultados da exactidão, nota-se também para o MDP uma tendência para sobrestimar concentrações mais baixas e subestimar as mais elevadas.

De acordo com o Anexo VII do Decreto-Lei n.º 111/2002, de 16 de Abril, os objectivos de qualidade dos dados para medições indicativas de NO₂ apontam para uma incerteza máxima de 25% e para uma cobertura mínima de tempo de 14% do ano, pelo que se pode concluir que estas campanhas de apenas uma semana e com incertezas médias de 51% e de 35%, não cumpriram o mínimo necessário para serem consideradas medições indicativas e permitirem directamente a verificação do cumprimento da legislação.

Apesar da incerteza global ser algo elevada, é de destacar o baixo desvio padrão apresentado pelo método dos tubos de difusão (MDP), para a maioria dos locais, o que indicia que, apesar das concentrações obtidas por este método apresentarem uma incerteza relativamente ao método de referencia elevada, o padrão de distribuição espacial das concentrações dados pelos tubos apresenta um erro baixo. Deste modo, apesar dos resultados das campanhas não poderem ser usados directamente para verificação do cumprimento da legislação, pode ser testada a sua utilização para extrapolar os indicadores anuais para cada ponto amostrado nas campanhas recorrendo a regressões lineares entre os resultados dos equipamentos fixos e os tubos de difusão.

Estação (tipo de estação)	Equipamento Fixo (MR) ($\mu\text{g}/\text{m}^3$)	Tubos difusão (MDP)				Incerteza da campanha		
		Ponto	Resultados ($\mu\text{g}/\text{m}^3$)	Média ($\mu\text{g}/\text{m}^3$)	Desvio Padrão ($\mu\text{g}/\text{m}^3$)	Precisão (%)	Exactidão (%)	Incerteza global (%)
Chelas (fundo)	13	1	12	15	3	40	15	56
		2	17					
		3	13					
		4	17					
Beato (fundo)	11	1	9	11	3	48	0	48
		2	10					
		3	14					
Olivais (fundo)	22	1	17	17	1	9	-23	32
		2	17					
		3	16					
Av. C.Ribeiro (tráfego)	31	1	34	40	5	34	29	63
		2	42					
		3	44					
R.Prata (tráfego)	35	1	46	49	3	15	40	55
		2	50					
		3	51					
Entrecampos (tráfego)	14	1	27	27	1	14	93	107
		2	28					
		3	26					
Av. Liberdade (tráfego)	65	1	57	54	3	8	-17	25
		2	54					
		3	52					
Benfica (tráfego)	29	1	23	23	1	4	-21	25
		2	22					
		3	23					
Média	28		29	34	3	25	17	51

Tabela 7. Avaliação da incerteza associada aos dados de NO₂ obtidos na campanha de Verão 2001 usando os indicadores precisão, exactidão e incerteza global.

Estação (tipo de estação)	Equipamento Fixo (MR) ($\mu\text{g}/\text{m}^3$)	Tubos difusão (MDP)				Incerteza da campanha		
		Ponto	Resultados ($\mu\text{g}/\text{m}^3$)	Média ($\mu\text{g}/\text{m}^3$)	Desvio Padrão ($\mu\text{g}/\text{m}^3$)	Imprecisão (%)	Inexactidão (%)	Incerteza global (%)
Olivais (fundo)	30	1	36	36	1	7	20	27
		2	37					
		3	35					
Chelas (fundo)	25	1	40	38	3	23	52	75
		2	40					
		3	35					
Beato (fundo)	40	1	37	35	3	16	-13	29
		2	36					
		3	31					
Av. C. Ribeiro (tráfego)	55	1	53	54	6	20	-2	22
		2	60					
		3	49					
Av. Liberdade (tráfego)	82	1	52	55	2	6	-33	39
		2	56					
		3	56					
Benfica (tráfego)	64	1	56	54	2	5	-16	21
		2	53					
		3	53					
Média	49		45	45	3	13	22	35

Tabela 8. Avaliação da incerteza associada aos dados de NO_2 obtidos na campanha de Inverno 2002 usando os indicadores precisão, exactidão e incerteza global.

5.1.3 Obtenção de indicadores anuais extrapolando os dados das campanhas usando regressão linear simples

O pré-tratamento dos dados de NO₂, usando o método da regressão linear simples teve como objectivo o cálculo dos indicadores, média anual e 19º máximo horário, para os locais de amostragem das campanhas, partindo dos resultados obtidos junto às estações fixas.

Na **Tabela 9** encontram-se os resultados para o NO₂ verificados na localização das estações fixas durante o período das duas campanhas e na média destas, usando tubos de difusão (método da amostragem por difusão passiva - MDP), e os equipamentos automáticos das estações (método de referência - MR), e os resultados para os períodos anuais dos equipamentos automáticos das estações (método de referência - MR).

Estações	Ano	Período das campanhas		Período anual	
		Tubos de difusão (MDP)		Equipamento Fixo (MR)	
		Média (µg/m ³ de NO ₂)	Média (µg/m ³ de NO ₂)	19º máximo horário (µg/m ³ de NO ₂)	
Olivais (fundo)	2001	17	22	104	
	2002	36	23	103	
	Média	26	22	103	
Beato (fundo)	2001	11	20	90	
	2002	35	23	118	
	Média	23	21	104	
Chelas (fundo)	2001	15	26	149	
	2002	38	23	108	
	Média	27	24	128	
Av. Casal Ribeiro (tráfego)	2001	40	49	160	
	2002	54	44	180	
	Média	47	47	170	
R. Prata (tráfego)	2001	49	44	136	
	2002	49	*	*	
	Média	49	*	*	
Entrecampos (tráfego)	2001	27	27	103	
	2002	58	43	127	
	Média	43	35	115	
Av. Liberdade (tráfego)	2001	54	54	178	
	2002	55	62	169	
	Média	55	58	173	
Benfica (tráfego)	2001	23	48	216	
	2002	54	45	154	
	Média	38	46	185	
Restelo (tráfego)	2001	10	*	*	
	2002	28	18	140	
	Média	19	*	*	

Nota: * a estação não estava a funcionar; não tem eficiência superior a 85% durante o período em análise.

Tabela 9. Resultados de NO₂ nas 9 estações fixas durante o período das campanhas (Verão 2001, Inverno 2002 e média de ambas) obtidos usando tubos de difusão e os equipamentos das estações, e o período anual (2001, 2002 e média) usando os equipamentos das estações.

O estabelecimento de regressões lineares entre dois indicadores correlacionados, permite fazer estimativas e inferir um indicador para a partir de outro com menor quantidade de informação. Para tal, é necessário que o ajustamento da regressão linear, avaliado pelo coeficiente de determinação (R^2), e significância do coeficiente de correlação, atinja determinados níveis de qualidade.

No presente caso o objectivo foi avaliar a adequabilidade das regressões lineares estabelecidas para os locais das estações fixas entre os dois métodos de análise, MR (variável dependente) e MDP (variável independente), para estimar os indicadores anuais, média anual e 19º máximo horário, para as cerca de 100 localizações para as quais existem dados obtidos com o MDP. Os resultados do MR, expressos na média anual e no 19º máximo horário, são a variável dependente porque correspondem aos dados “reais” que pretendemos modelar a partir dos resultados obtidos com o método de difusão passiva e estimar para outros locais.

Na **Tabela 10**, podemos encontrar os coeficientes de determinação do ajustamento das regressões lineares entre os resultados de NO_2 nas estações existentes usando tubos de difusão (MDP) durante o período das campanhas (Verão 2001, Inverno 2002 e média dos dois períodos de campanha) e os equipamentos das estações (MR) durante o período anual (2001, 2002 e média dos dois anos). As estações de Entrecampos, Rua da Prata e Restelo foram excluídas da análise por não terem eficiência para algum dos períodos em análise. A análise da tabela revela que:

- Os ajustamentos dos dois indicadores anuais de NO_2 , obtida com o MR, (variável dependente) às médias semanais dos tubos (variável independente), para cada um dos anos individualmente e para a média dos anos de 2001 e 2002, têm genericamente R^2 acima de 0,7 podendo considerar-se bons.
- Os melhores ajustamentos para os dois indicadores anuais de NO_2 , para o ano de 2001, 2002 e média dos anos 2001 e 2002, foram conseguidos recorrendo, como variável independente, à campanha de Inverno de 2002 ou fazendo a média das duas campanhas. A campanha de verão de 2001 obteve os resultados piores para estimar os indicadores anuais.

R ²		Tubos de difusão (MDP) Variável independente (y)			
		Média dos tubos			
		Verão 2001	Inverno 2002	Verão 2001-Inverno 2002	
Equipamento estações fixas (MR) Variável dependente (x)	Média anual	2001	0,75	<u>0,99</u>	0,90
		2002	0,85	0,86	<u>0,93</u>
		Média 2001-2002	0,83	0,95	<u>0,95</u>
	19º máximo horário	2001	0,27	<u>0,73</u>	0,44
		2002	0,71	<u>0,89</u>	0,84
		Média 2001-2002	0,50	<u>0,95</u>	0,70

Nota: Melhor modelo para estimar a variável dependente

Modelos seleccionados para prosseguir a metodologia

Todas as regressões lineares são aceitáveis para um nível de significância superior a 0,05;

Tabela 10. Coeficientes de determinação do ajustamento das regressões lineares, para estações existentes, entre os resultados de NO₂ usando tubos de difusão (MDP) durante o período das campanhas (Verão 2001, Inverno 2002 e média das duas) e os equipamentos das estações fixas (MR) no período anual (para 2001, 2002 e média dos dois).

Apesar de também ser possível estimar os resultados das campanhas para a média anual e 19º máximo horário dos anos 2001 e 2002 individualmente, optou-se por considerar apenas a média dos dois anos 2001 e 2002, obtida através da média das duas campanhas, para ambos os indicadores de NO₂.

A regressão linear entre a média dos dois períodos das campanhas usando tubos de difusão e a média dos dois anos 2001-2002 tem 95% da variância explicada (ver **Figura 19**). Já a relação linear entre a média dos dois períodos das campanhas usando tubos de difusão e o 19º máximo horário médio dos dois anos 2001-2002, explica 70% da variância (ver **Figura 20**).

A aplicação de cada uma destas relações à média dos dois períodos das campanhas usando tubos de difusão, obtido em cada local amostrado durante as campanhas apresenta-se nas **Figura 21** e **Figura 22**, respectivamente. Na escala de cores escolhida para os mapas de pontos optou-se por pôr a vermelho os pontos para os quais se estimaram concentrações superiores aos valores limite anuais e horários. Os resultados mostram que para a média anual se estima ocorrerem ultrapassagens aos valores limite em 8 pontos da cidade, localizados no centro da cidade e junto a grandes vias de tráfego e cruzamentos enquanto que para o 19º máximo horário nenhum ponto está em ultrapassagem.

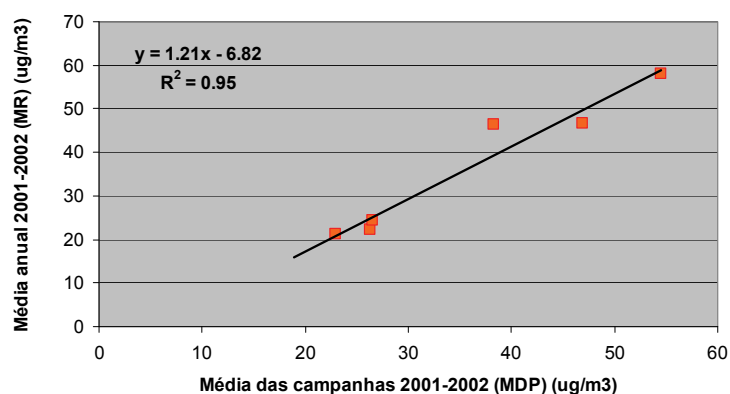


Figura 19. Regressão linear estabelecida para o local das estações fixas entre a média das campanhas usando tubos de difusão (MDP) e a média anual 2001-2002 medida usando o método de referência (MR).

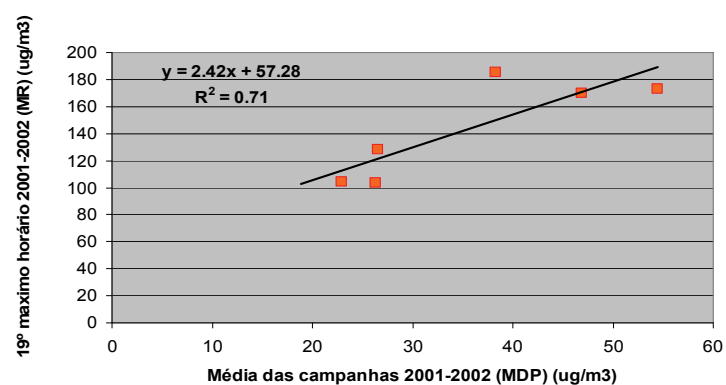


Figura 20. Regressão linear estabelecida, para o local das estações fixas, entre a média das campanhas usando tubos de difusão (MDP) e o 19º máximo horário usando o método de referência (MR).

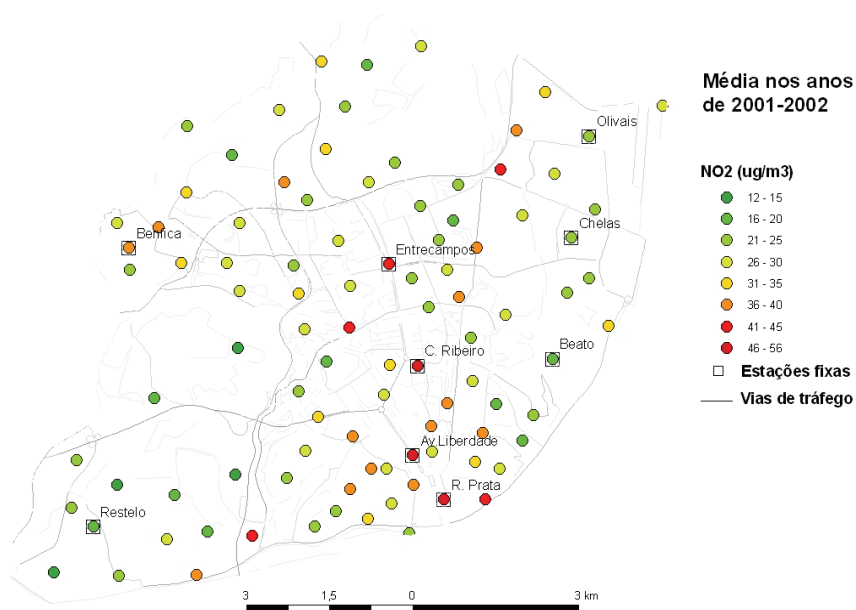


Figura 21. Estimativa da média das médias anuais de 2001-2002 para os pontos de amostragem da campanha MDP usando a relação linear estabelecida na Figura 19.

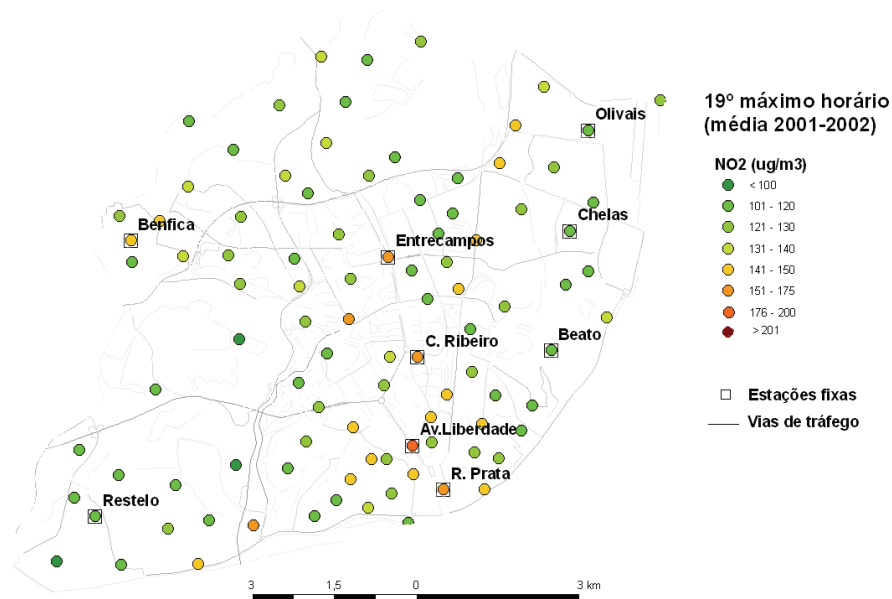


Figura 22. Estimativa do 19º máximo horário (média dos anos 2001-2002) para os pontos de amostragem da campanha MDP usando a relação linear estabelecida Figura 20.

O erro associado à estimativa dos resultados de NO₂ obtidos nas campanhas com tubos de difusão (MDP) para a média anual e 19º máximo horário médios dos anos 2001 e 2002, foi avaliado fazendo a validação cruzada para os locais das seis estações e calculando os parâmetros de erro definidos na metodologia (ver capítulo 4.4.2). De acordo com os resultados desta avaliação expressos na **Tabela 11** e na **Tabela 12**, o erro global da estimativa expresso em NRMSE, em ambas as situações, encontra-se abaixo de 25%, sendo, como expectável, os resultados para a média anual melhores do que para o 19º máximo horário. Estes resultados foram considerados aceitáveis podendo-se prosseguir a aplicação dos modelos de interpolação com estas estimativas de ambos os indicadores. É também de notar, que a estação de Benfica condiciona o ajustamento da relação e tem um erro considerável.

Estações	Tubos de Difusão (MDP)	Valor real Equipamento fixo (MR)	Parâmetros da regressão linear (sem a estação)			Valor modelado aplicando a regressão linear VC	Erro entre o valor modelado e o valor real ($\mu\text{g}/\text{m}^3$)
	Média das campanhas ($\mu\text{g}/\text{m}^3$)	Média anual ($\mu\text{g}/\text{m}^3$)	m*	b*	R ²	MA ₂₀₀₁₋₂₀₀₂ ($\mu\text{g}/\text{m}^3$)	
Olivais (fundo)	26	22	1,16	-4,85	0,94	25	3
Beato (fundo)	23	21	1,20	-7,10	0,93	21	-1
Chelas (fundo)	27	24	1,18	-5,65	0,94	26	2
Casal Ribeiro (tráfego)	47	47	1,25	-7,90	0,95	51	4
Av. Liberdade (tráfego)	55	58	1,26	-8,60	0,91	61	3
Benfica (tráfego)	38	46	1,18	-7,52	1,00	37	-9
Erro médio							0
Erro médio absoluto							4
RMSE							4
NRMSE (%)							12

* $y = mx + b$; MA₂₀₀₁₋₂₀₀₂ – Média da média anual dos anos 2001 e 2002

Tabela 11. Cálculo do erro da estimativa dos resultados da campanha para média anual média dos anos 2001 e 2002 usando a validação cruzada

Estações	Tubos de difusão (MDP)	Valor real Equipamento fixo (MR)	Parâmetros da regressão linear (sem a estação)			Valor modelado aplicando a regressão linear VC	Erro entre o valor modelado e o valor real ($\mu\text{g}/\text{m}^3$)
	Média das campanhas ($\mu\text{g}/\text{m}^3$)	Média anual ($\mu\text{g}/\text{m}^3$)	m*	b*	R ²	19ºMH ₂₀₀₁₋₂₀₀₂ ($\mu\text{g}/\text{m}^3$)	
Olivais (fundo)	26	103	1,9	72,0	0,44	122	19
Beato (fundo)	23	104	2,0	67,6	0,42	114	10
Chelas (fundo)	27	128	2,3	53,5	0,53	116	-12
Casal Ribeiro (tráfego)	47	170	2,0	63,6	0,47	159	-11
Entrecampos (tráfego)	43	115	2,5	55,4	0,73	162	47
Av. Liberdade (tráfego)	55	177	2,2	57,9	0,40	181	4
Benfica (tráfego)	38	185	2,1	55,5	0,72	135	-50
Erro médio							1
Erro médio absoluto							22
RMSE							28
NRMSE (%)							20

* $y = mx + b$; 19º MH₂₀₀₁₋₂₀₀₂ – Média do 19º máximo horário dos anos 2001 e 2002

Tabela 12. Cálculo do erro da estimativa dos resultados da campanha para o 19º máximo horário médio dos anos 2001 e 2002 usando a validação cruzada

5.2 Pré-tratamento das variáveis auxiliares

Para a modelação da distribuição espacial da qualidade do ar na cidade de Lisboa foi testada a utilização de variáveis auxiliares que se pressupõe estarem relacionadas com as emissões e dispersão dos poluentes produzidos pelo tráfego rodoviário. Para o NO_2 assim como para outros poluentes gerados por esta fonte emissora, estas variáveis auxiliares poderão ser por exemplo os níveis de tráfego automóvel, a rede de tráfego, a ocupação do solo, a densidade populacional, ou o inventário de emissões desagregado espacialmente.

Como já referido no capítulo 4.2, no pré-tratamento das variáveis seleccionadas foi inicialmente feita a conversão do modelo de dados vectorial para grelha (*raster*) com uma resolução de 10 metros, e depois foi extraído o valor da variável no local de amostragem ou calculados estimadores estatísticos básicos (média, máximo, mínimo, soma, desvio padrão e contagem de células) para cada variável usando os anéis de vizinhança de cada ponto de amostragem, exemplificados na **Figura 23**.

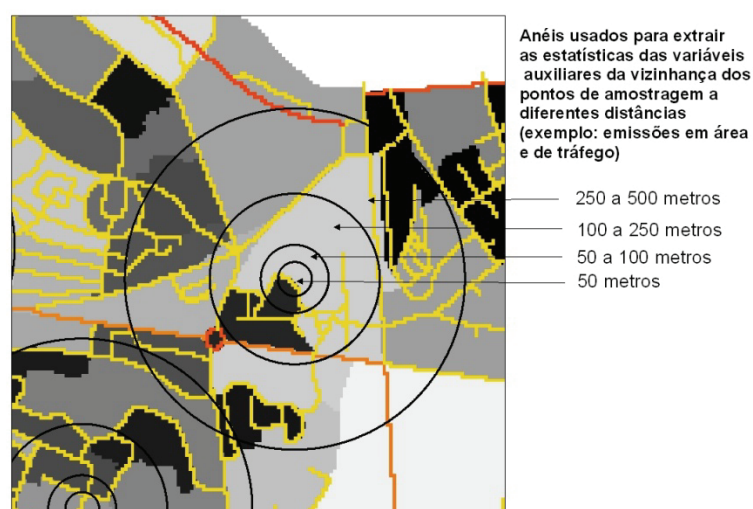


Figura 23. Anéis usados para extrair as estatísticas das variáveis auxiliares da vizinhança dos pontos de amostragem a diferentes distâncias (exemplo: emissões em área e de tráfego).

5.2.1 Ocupação do solo

Na **Figura 24**, apresenta-se a ocupação do solo obtida para a cidade de Lisboa usando uma adaptação do nível 3 de informação da série cartográfica de ocupação do solo do Corine LandCover de 2000 à escala 1:100.000 disponibilizada pelo Instituto Geográfico Português. Dado esta ser uma variável categórica apenas se avaliou o número de quadrículas de cada classe existentes dentro de cada anel de influência. Uma breve análise visual do mapa desta figura mostra que as concentrações mais elevadas de NO_2 se encontram junto às

principais vias de tráfego e predominantemente dentro da área urbana contínua. Dentro ou junto a áreas agrícolas e parques urbanos as concentrações são mais baixas.

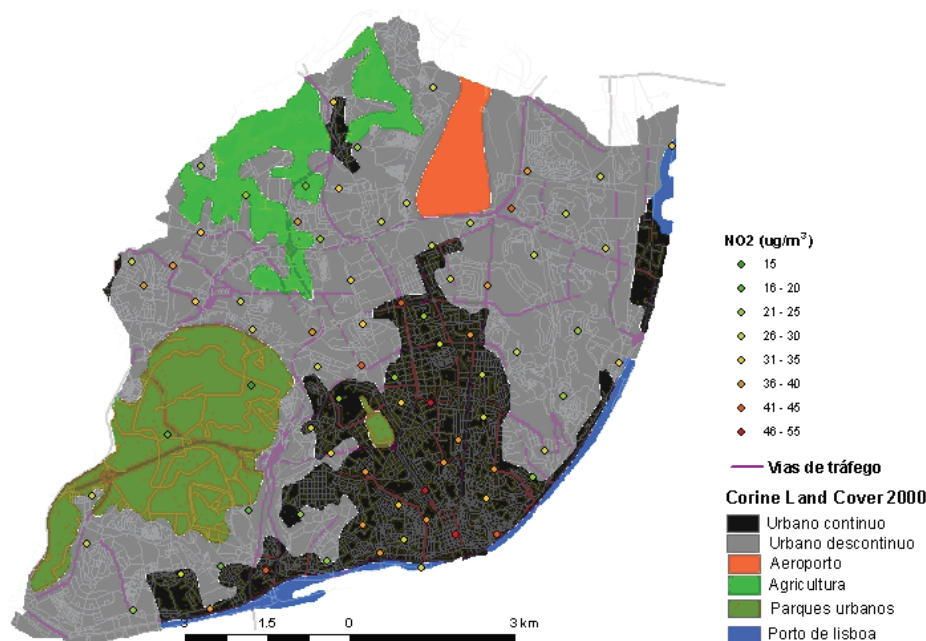


Figura 24. Ocupação do solo da cidade de Lisboa (CORINE Land cover 2000), média anual de NO₂ e vias de tráfego

O coeficiente de determinação da relação linear simples entre a concentração de NO₂ e o número de células de cada classe para cada anel apresentam-se na **Tabela 13** e em gráfico na **Figura 25**.

R	Urbano contínuo	Urbano descontínuo	Parque urbano	Agricultura	Porto de Lisboa	Aeroporto
50 m	0,41	0,20	0,29	0,05	0,01	
50-100 m	0,41	0,22	0,29	0,08	0,01	
100-250 m	0,38	0,29	0,35	0,14	0,02	0,07
250-500 m	0,40	0,28	0,33	0,12	0,01	0,06

Nota: A relação não é significativa (nível de significância de $(\alpha) \geq 0,05$)

Tabela 13. Coeficiente de correlação (R) da relação linear entre as concentrações do NO₂ e da contagem do número de células dentro de cada anel relativamente à ocupação do solo

Relativamente às correlações com as concentrações de NO₂ apresentadas, verificou-se que estas são de uma forma geral baixas, só se obtendo relações significativas (nível de significância $(\alpha) < 0,05$) com as classes: urbano contínuo, urbano descontínuo e parque urbano. Verificou-se ainda que os coeficientes de correlação não têm uma variação muito marcada em função da distância (raio dos anéis de influência). A variância explicada para a classe urbano contínuo tende a diminuir com a distância enquanto que para o parque urbano, urbano descontínuo e agricultura existe um aumento até aos 100-250 m, seguindo-se uma diminuição nos 250-500 m. Pelo facto de esta variável ser qualitativa e pouco

pormenorizada, a maioria dos anéis dos pontos abrangem apenas uma classe na sua área de influência o que dificulta a sua utilização para modelar a distribuição das concentrações a uma resolução tão baixa (100 m).

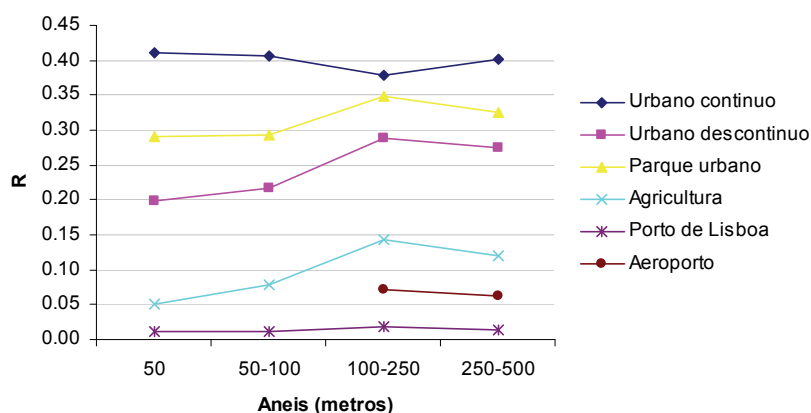


Figura 25. Variação do coeficiente de correlação da relação linear, entre as concentrações do NO_2 e da contagem do número de células de cada tipo de ocupação, dentro de cada anel de influência.

5.2.2 Emissões de tráfego rodoviário

O resultado da alocação das emissões anuais de NO_x a cada via de tráfego, tendo em consideração os níveis de tráfego e as velocidades, a partir dos modelos desenvolvidos para os anos de 2001 e 2002 (Gois et al., 2006) apresenta-se na **Figura 26**.

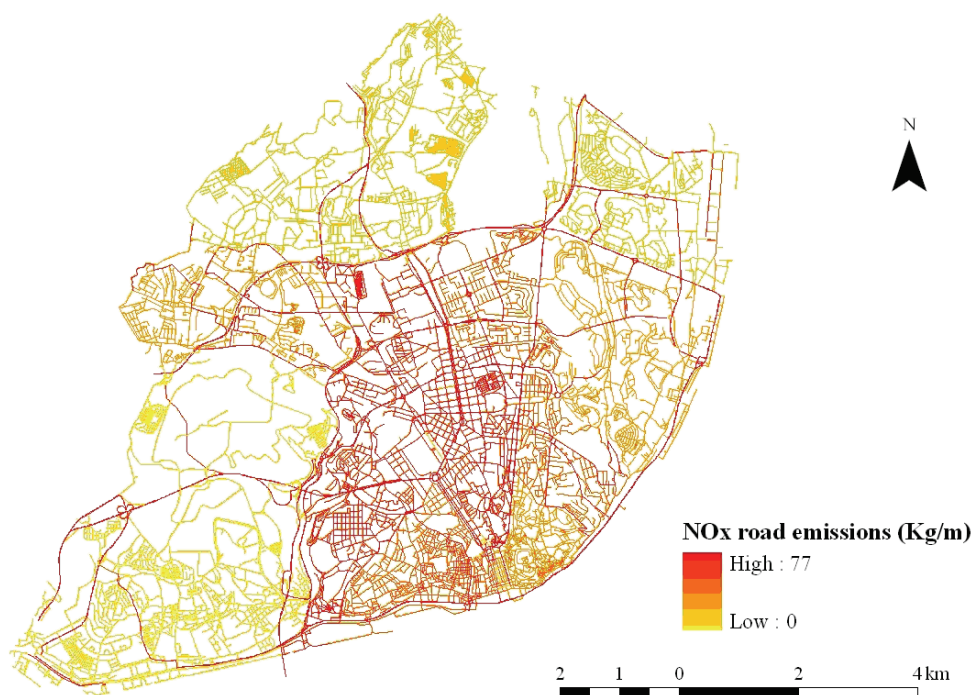


Figura 26. Emissões anuais de NO_x resultantes do tráfego rodoviário em Lisboa desagregadas ao nível da via (Gois et al., 2006).

Os coeficientes de determinação das regressões lineares simples entre a concentração de NO_2 e cada um dos estimadores estatísticos (média, máximo, mínimo, soma, desvio padrão

e contagem de células) calculados para cada anel relativamente às emissões de NO_x proveniente do tráfego rodoviário, apresentam-se na **Tabela 14** e no gráfico da **Figura 27**. Através deste gráfico é possível verificar que:

- A estatística mais correlacionada com as concentrações de NO_2 é a soma das emissões para o anel de 50 m;
- Para a generalidade das estatísticas a correlação vai descendo com o afastamento dos anéis ao ponto de amostragem;
- Para a estatística “contagem” a correlação sobe com o aumento da distância.

Apesar de algumas estatísticas apresentarem correlações (relativamente às concentrações de NO_2) individuais baixas e em alguns casos sem significado estatístico, todas as estatísticas vão ser testadas na regressão linear múltipla, uma vez que, apesar de individualmente não serem significativas em conjunto com outras variáveis podem resultar em modelos significativos.

R	0-50 m	50-100 m	100-250 m	250-500 m
Média	0,44	0,33	0,39	0,26
Contagem	0,17	0,17	0,32	0,35
Máximo	0,44	0,48	0,35	0,28
Mínimo	0,28	0,20	0,14	0,10
Desvio Padrão	0,39	0,37	0,28	0,17
Soma	0,55	0,54	0,52	0,40

Nota: A relação não é significativa (nível de significância de $(\alpha) \geq 0,05$)

Tabela 14. Coeficiente de correlação (R) da relação linear entre as concentrações do NO_2 e as estatísticas para cada anel de influência relativamente às emissões do tráfego rodoviário

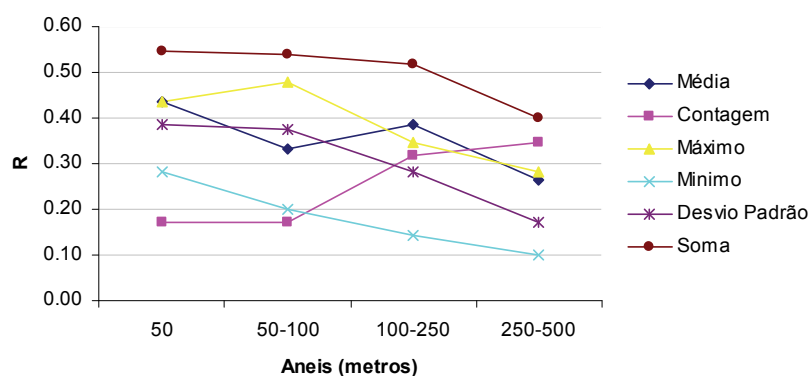


Figura 27. Variação do coeficiente de determinação da relação linear, entre as concentrações do NO_2 e das estatísticas das emissões de tráfego dentro de cada anel de influência.

5.2.3 Emissões em área

As emissões em área que englobam as emissões residenciais, comerciais e de serviços, foram desagregadas espacialmente (**Figura 28**) tendo em conta somente a densidade

populacional calculada com base nos dados do Instituto Nacional de Estatística (INE) de população residente para a subsecção estatística definida pelo INE (INE, 2001). Por este facto a distribuição espacial obtida é equivalente à da densidade populacional.

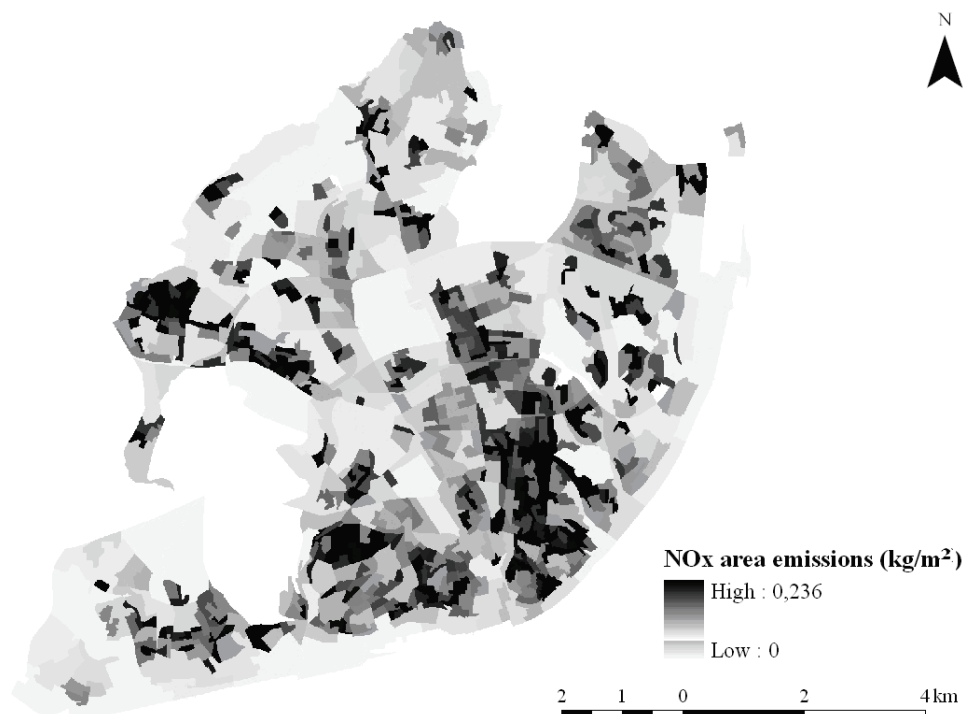


Figura 28. Emissões de NO₂ em área em Lisboa desagregadas para o nível da subsecção estatística (Gois et al., 2006; INE, 2001).

A análise dos resultados permite verificar que as correlações individuais são baixas e na maioria dos casos não têm significado. A única estatística das emissões em área com correlação moderada é o mínimo no anel de 250 a 500 m. Ainda assim, tal como no caso das emissões de tráfego, todas as estatísticas vão ser testadas na regressão linear múltipla uma vez que apesar de individualmente não serem significativas em conjunto com outras variáveis podem produzir modelos significativos.

R	50 m	50-100 m	100-250 m	250-500 m
Média	0,19	0,18	0,19	0,29
Contagem	0,11	0,02	0,01	0,06
Máximo	0,14	0,15	0,19	0,19
Mínimo	0,14	0,14	0,21	0,33
Desvio Padrão	0,05	0,07	0,09	0,16
Soma	0,19	0,18	0,19	0,28

Nota: A relação não é significativa (nível de significância de $(\alpha) \geq 0.05$)

Tabela 15. Coeficiente de correlação (R) da relação linear entre as concentrações do NO₂ e as estatísticas para cada anel de influência relativamente às emissões em área

O gráfico da **Figura 29**, é possível verificar que para a generalidade das estatísticas a correlação aumenta com a distância do anel ao ponto de amostragem. É de notar, que

neste caso será relevante o facto do aumento da distância corresponder também a um aumento da área dos anéis.

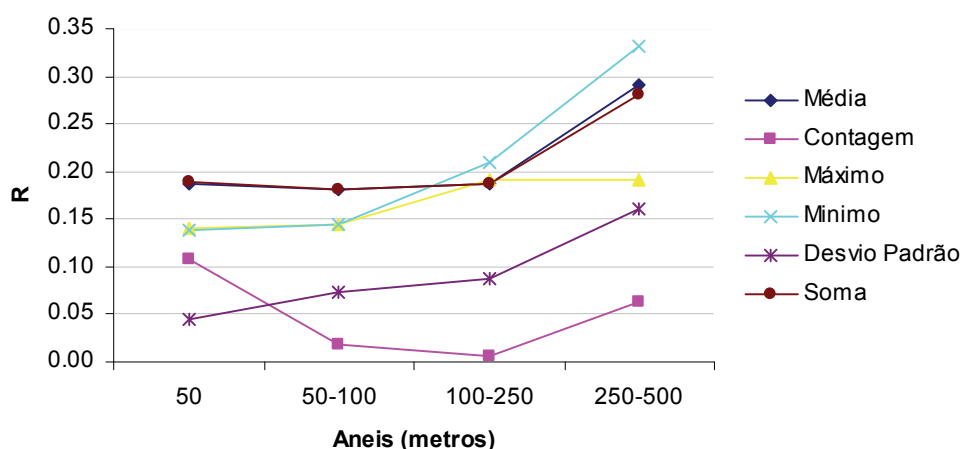


Figura 29. Variação do coeficiente de correlação da relação linear entre as concentrações do NO₂ e as estatísticas das emissões em área para cada anel de influência.

5.3 Modelação da distribuição espacial

Neste capítulo apresentam-se os mapas com menor erro associado que foi possível obter, usando cada um dos métodos de interpolação descritos no enquadramento (ver capítulo 3.3.3) e implementados de acordo com o exposto na metodologia (ver capítulo 4.3). Para a interpolação foram usados os resultados dos 97 pontos com medições em ambas as campanhas com tubos de difusão, para os dois indicadores de dióxido de azoto (NO₂), médios de 2001 e 2002, média anual e 19º máximo horário, estimados de acordo com o exposto no capítulo 5.1.3. No último ponto apresenta-se a comparação dos mapas de cada método usando medidas de erro obtidas a partir da validação cruzada para os 97 pontos de amostragem (ver capítulo 4.4.3).

5.3.1 Inverso da Distância Ponderada (IDW)

A utilização da técnica de interpolação *IDW* permitiu produzir vários mapas, testando vários valores para os parâmetros de interpolação do *IDW*, expoente (*power*) e vizinhança de pesquisa. A selecção do melhor mapa de distribuição das concentrações de NO₂, para média anual e 19º máximo horário, obtido com este método, foi efectuada avaliando a estatística de erro. Os melhores mapas obtidos para cada um dos indicadores, e os respectivos parâmetros de interpolação, apresentam-se na **Figura 30** e na **Figura 31**. É possível encontrar informação mais detalhada sobre os parâmetros de interpolação no **Anexo 1**.

A observação destes mapas permite identificar em ambos uma área de concentrações mais elevadas ($> 30 \mu\text{g}/\text{m}^3$ e $> 130 \mu\text{g}/\text{m}^3$) na baixa de Lisboa num triângulo com vértices em Entrecampos, Santa Apolónia e Santos, com um pico entre a Avenida da Liberdade e Rua da Prata, e uma zona de concentrações mais baixas ($< 25 \mu\text{g}/\text{m}^3$ e $< 120 \mu\text{g}/\text{m}^3$) na zona do parque de Monsanto, Restelo e Belém. Uma vez que o *IDW* é um interpolador exacto, o ajustamento dos modelos de interpolação é de 100%, são também identificados outros locais pontuais de concentrações mais elevadas, normalmente junto a grandes vias de tráfego, ou mais baixas, em locais menos densamente povoados. Em termos de avaliação do cumprimento dos valores limite, usando este método foram identificadas apenas três pequenas áreas, junto às estações da Avenida da Liberdade, Avenida Casal Ribeiro e a um ponto de medição em Alcântara, em excedência ao valor limite para a média anual em 2001-2002.

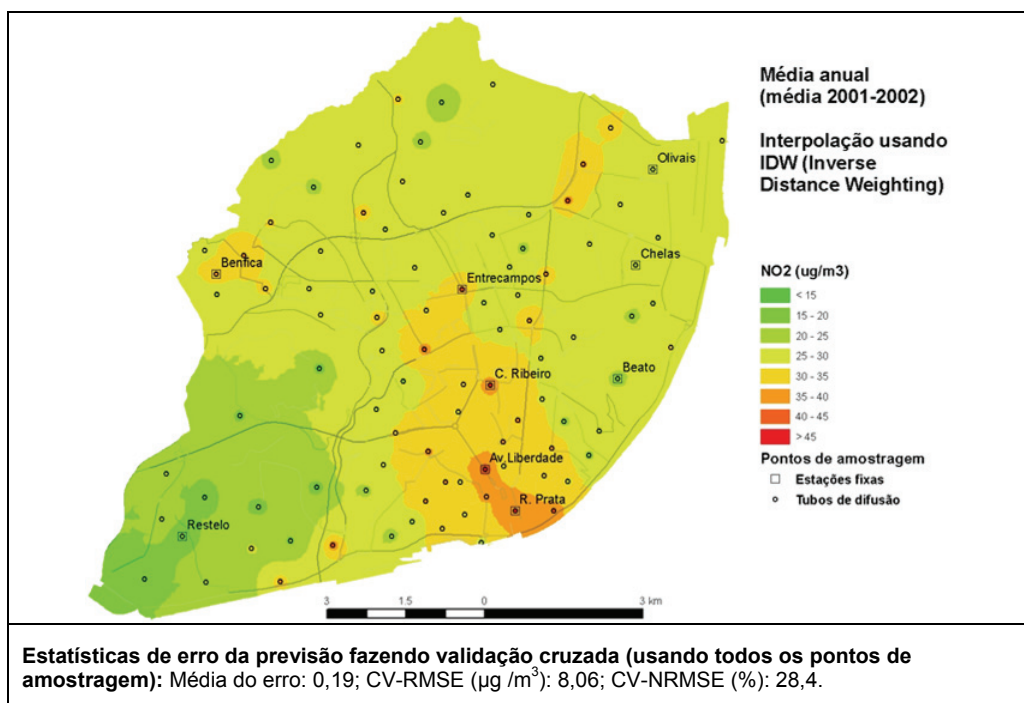


Figura 30. Interpolação da média de NO_2 durante de 2001 e 2002 usando *IDW* e estatísticas de erro da previsão fazendo validação cruzada.

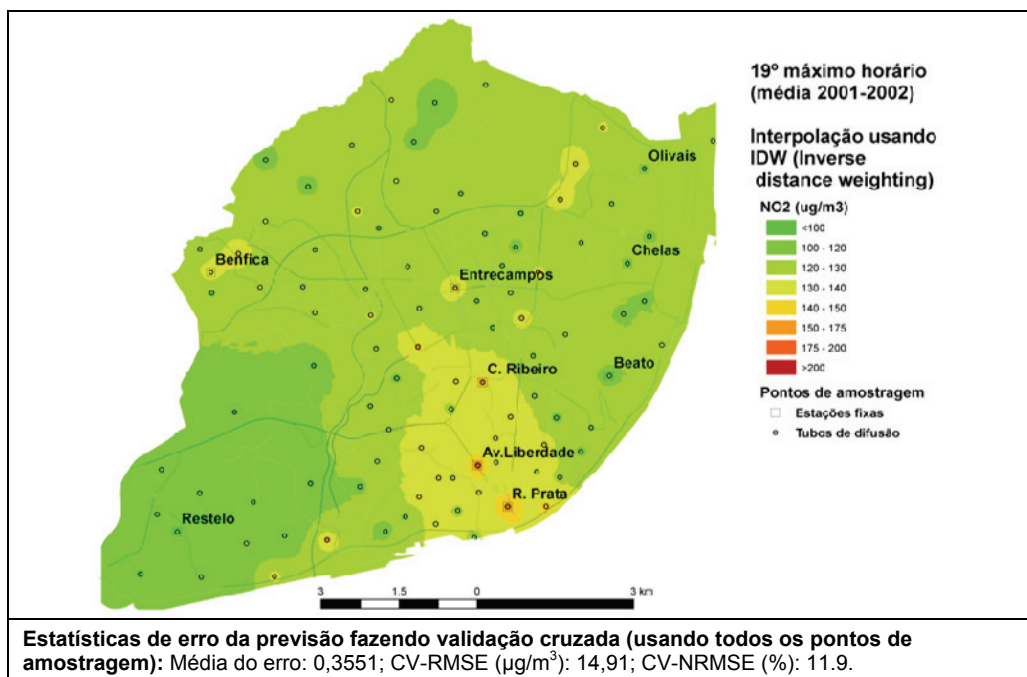


Figura 31. Interpolação do 19º máximo horário de NO₂ médio de 2001 e 2002 usando o *IDW* e estatísticas de erro da previsão fazendo validação cruzada.

5.3.2 Ordinary Kriging

Tal como para o *IDW* foram produzidos vários mapas usando o método *Ordinary Kriging*, testando vários valores para o variograma (nomeadamente, o *range*, *nugget*, *lag*, forma do modelo) e vizinhança de pesquisa. A selecção do melhor mapa de distribuição das concentrações de dióxido de azoto (NO₂), para média anual e 19º máximo horário, obtido com este método, foi efectuada avaliando a estatística de erro. O mapa do erro padrão, as opções dos parâmetros usadas para definir o variograma e a vizinhança de pesquisa e as estatísticas de erro fazendo a validação cruzada dos resultados são apresentados no **Anexo 1**. Na **Figura 33** e na **Figura 32**, apresentam-se os melhores mapas de interpolação da média anual e do 19º máximo horário de NO₂ médios para os anos de 2001 e 2002 obtidos usando *ordinary kriging*.

Os mapas produzidos usando *ordinary kriging* não identificam picos como o *IDW* mas apenas tendências gerais do mapa. Tal como nos resultados usando *IDW*, os mapas de *kriging* permitem identificar em ambos os mapas uma área de concentrações mais elevadas (> 30 µg/m³ e > 130 µg/m³) na baixa de Lisboa num triângulo entre Entrecampos, Santa Apolónia e Santos, com um pico entre a Avenida da Liberdade e Rua da Prata, e uma zona de concentrações mais baixas (< 25 µg/m³ e < 120 µg/m³) na zona do parque de Monsanto, Restelo e Belém. Em termos de avaliação do cumprimento dos valores limite, não são

identificadas áreas em excedência ao valor limite para a média anual ou 19º máximo horário no período 2001-02.

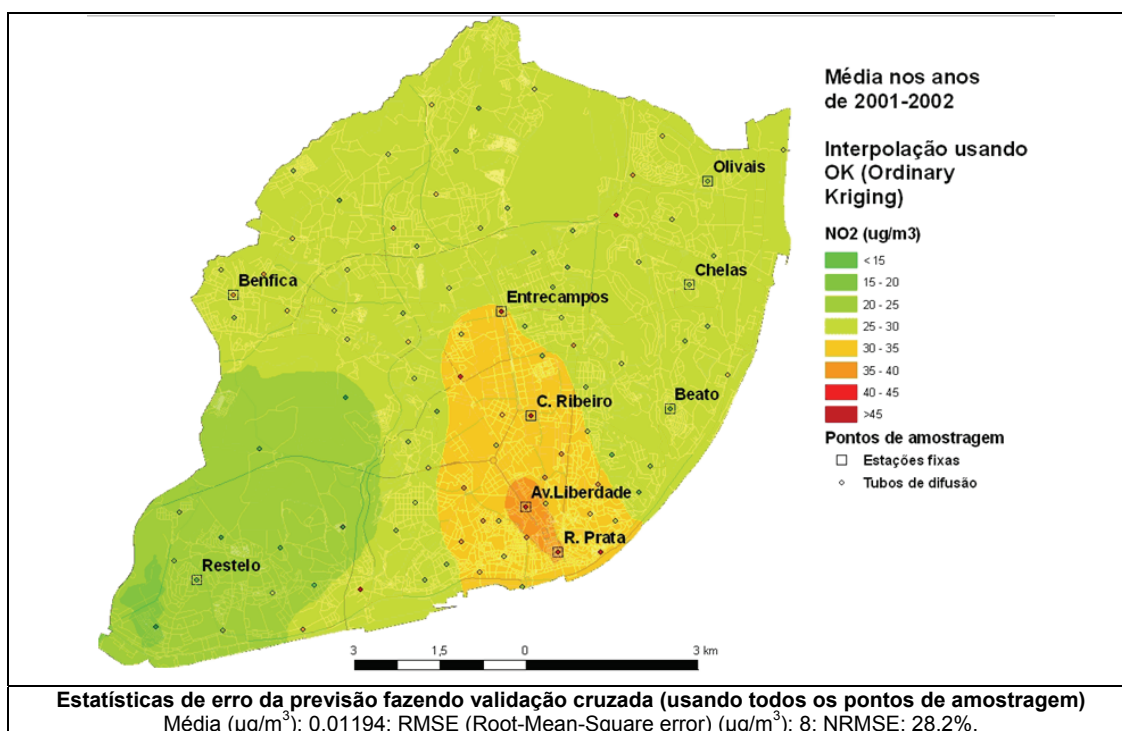


Figura 32. Interpolação da média de NO₂ durante os anos de 2001 e 2002 usando o método ordinary kriging e estatísticas de erro da previsão fazendo validação cruzada

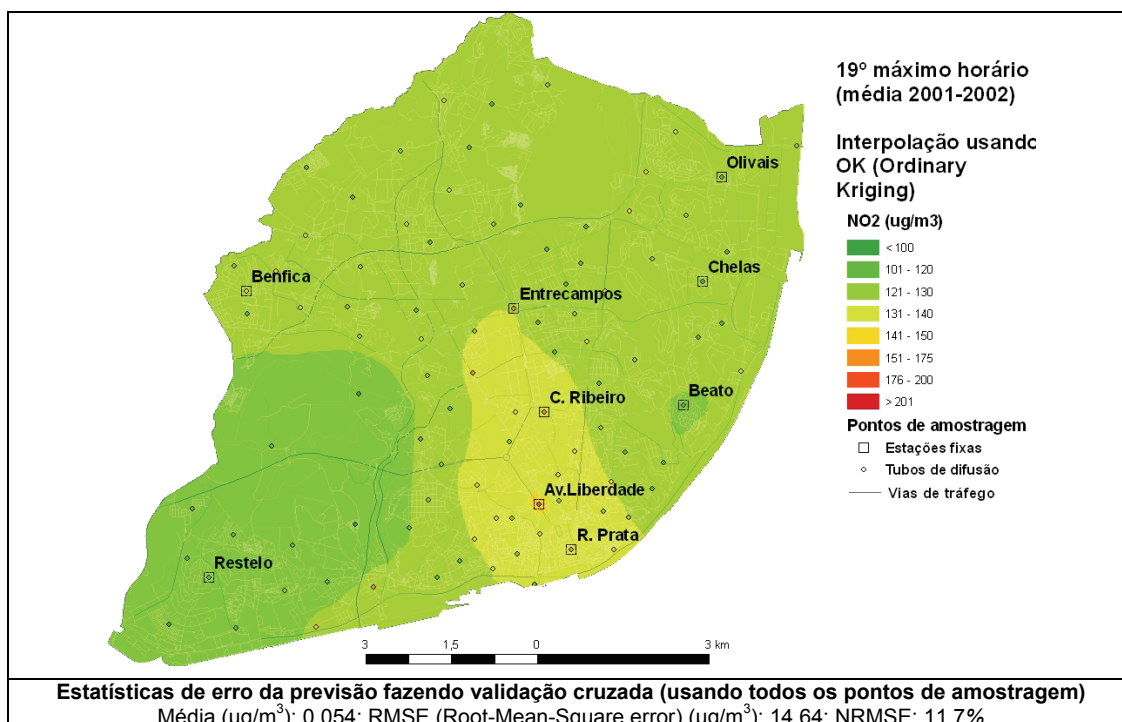


Figura 33. Interpolação do 19º máximo horário de NO₂ médio de 2001 e 2002 usando *ordinary kriging* e estatísticas de erro da previsão fazendo validação cruzada.

5.3.3 Ordinary CoKriging

Na **Figura 34** e na **Figura 35** apresentam-se os mapas obtidos por interpolação usando *ordinary cokriging*, para a média anual e 19º máximo horário de NO₂ para o período de 2001 e 2002, respectivamente. Foram seleccionadas como as melhores variáveis auxiliares a introduzir na interpolação, as três variáveis com maior correlação com os níveis de NO₂, ou seja: a soma das emissões NO_x devidas ao tráfego rodoviário na envolvente de 50 m, entre 50 e 100 m e entre 100 e 250 m. Tal como para o *ordinary kriging* apresentam-se o variograma, as estatísticas de erro e o mapa do erro padrão relativas aos mapas obtidos com o *ordinary cokriging* no **Anexo 1**.

Os mapas obtidos com este método de interpolação são semelhantes aos obtidos usando *ordinary kriging* mas permitem melhorar os resultados identificando concentrações elevadas em algumas áreas sem pontos de amostragem mas que as variáveis auxiliares identificaram como sendo de emissão de NO_x elevada. Este é o caso da rotunda do Marquês de Pombal que foi incluída na área de concentração máxima e de algumas áreas envolventes a grandes vias de tráfego e cruzamentos.

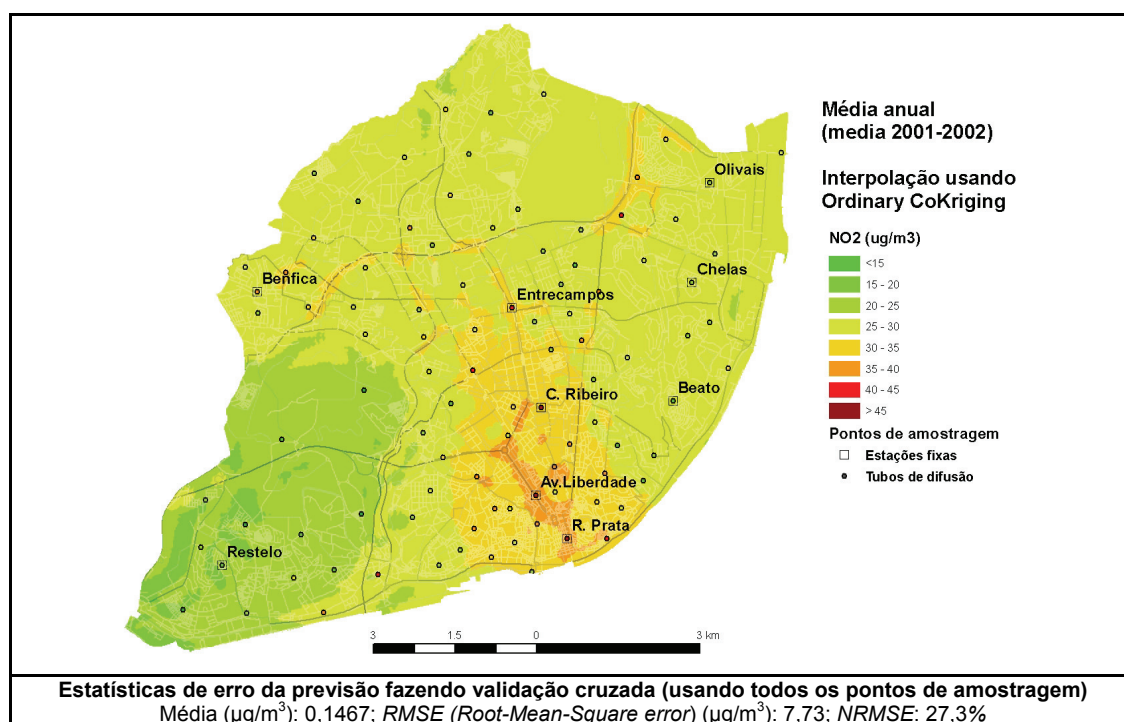


Figura 34. Interpolação da média anual de NO₂ para a média de 2001 e 2002 usando *cokriging* com as variáveis auxiliares: soma das emissões NO_x devidas ao tráfego rodoviário na envolvente de 50m, entre 50 e 100m e entre 100 e 250m e estatísticas de erro da previsão fazendo validação cruzada.

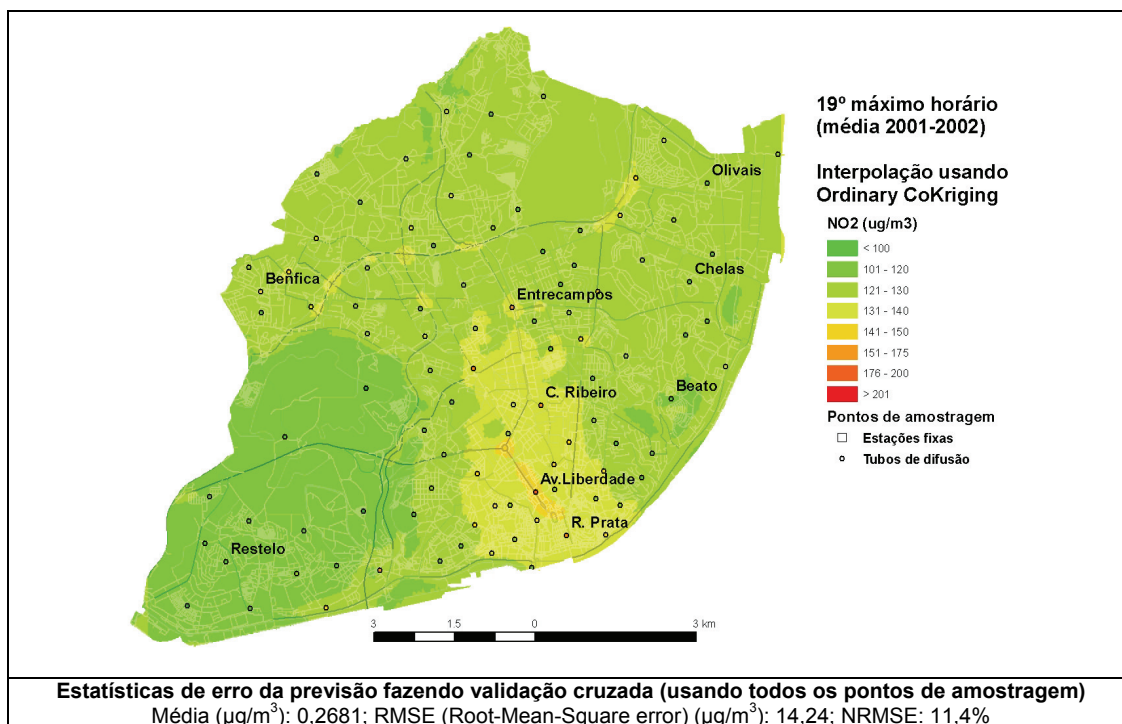


Figura 35. Interpolação do 19º máximo horário de NO_2 médio de 2001 e 2002 usando *ordinary cokriging* com as variáveis auxiliares: soma das emissões derivadas do NO_x do tráfego na envolvente de 50m, entre 50 e 100m e entre 100 e 250m e estatísticas de erro da previsão fazendo validação cruzada

5.3.4 Regressão linear múltipla

Na interpolação das concentrações de NO_2 (variável dependente) obtidas no capítulo 5.1.3, usando a regressão linear múltipla foram testadas como variáveis auxiliares as emissões de tráfego desagregadas para cada via (ET), as emissões em área desagregadas usando a densidade populacional (EA) e a ocupação do solo (OS). Estas variáveis foram tratadas de acordo com o exposto no capítulo 5.2 e os estimadores estatísticos produzidos foram usados como variáveis independentes (preditores) das regressões lineares.

Vários modelos foram produzidos usando diferentes conjuntos de estatísticas como variáveis de entrada e usando como métodos de selecção/exclusão de variáveis: os métodos de progressivo (*forward*) e de regressivo (*backward*). Em ambos os casos foram seleccionadas como variáveis de entrada estatísticas das emissões de tráfego desagregadas ao nível da via e da densidade populacional ao nível da subsecção estatística. O melhor modelo produzido, tendo em consideração o ajustamento do modelo, foi o obtido usando o método regressivo (coeficiente de determinação de 0,63) e as estatísticas das variáveis das ET e EA. No entanto, este modelo incluía um número elevado de variáveis (doze variáveis), que tornava difícil a sua compreensão, e tinha problemas de multicolinearidade (*variance inflation factor* (VIF) superior a 5). Assim, foi seleccionado o modelo obtido pelo método *forward*. Este modelo foi ainda ligeiramente melhorado fazendo

algumas alterações nas variáveis, nomeadamente foi considerado o mínimo não do anel 250-500 m, mas sim, do anel de 0-500 m, e a soma do anel de 50-250 m em vez de 100-250 m. Este modelo, adaptado do obtido pelo método *forward*, teve um ajustamento de 52% (coeficiente de determinação de 0,52).

Os resultados dos modelos obtidos para a média anual e 19º máximo horário em termos de coeficientes e estatísticas de colinearidade apresentam-se na **Tabela 16**, e o sumário dos ajustamentos na **Tabela 17**. Uma vez que os dados de NO₂, expressos na média anual e 19º máximo horário, são directamente proporcionais a única diferença entre os dois modelos está nos coeficientes da regressão, sendo as variáveis independentes e o ajustamento iguais. A aplicação dos modelos, expressos nas equações lineares que se apresentam de seguida, a uma grelha com uma resolução de 100 m resultou nos mapas que se apresentam na **Figura 36**, para a média anual, e na **Figura 37**, para o 19º máximo horário.

Modelo	Coeficientes não padronizados Média anual		Coeficientes não padronizados 19º Máximo horário		t	Sig.	Estatísticas de colinearidade	
	B	Erro padrão	B	Erro padrão			Tol.	VIF
Constante	23,510	1,736	111,670	3,660	13,541	<0,05		
Emissões NO _x de tráfego (ET) Soma no anel de 50m	0,041	0,008	0,086	0,018	4,842	<0,05	0,715	1,398
Emissões NO _x de tráfego (ET) Soma no anel de 50-250m	0,004	0,001	0,009	0,002	4,747	<0,05	0,429	2,33
Emissões NO _x de tráfego (ET) Média no anel de 250-500m	-2,126	0,578	-4,478	1,218	-3,675	<0,05	0,213	4,705
Emissões NO _x de tráfego (ET) Desvio padrão no anel de 250-500m	0,889	0,320	1,871	0,674	2,776	<0,05	0,298	3,361
Emissões NO _x em área (EA) Mínimo no anel de 0-500m	458,458	148,994	965,650	313,826	3,077	<0,05	0,844	1,185

Tabela 16. Coeficientes do modelo de regressão linear múltipla obtido para a média anual e 19º máximo horário (método *forward*)

R	0,722
R ²	0,521
R ² ajustado	0,491
Erro padrão da estimativa	6,439

Tabela 17. Sumário do ajustamento do modelo de regressão linear múltipla obtido para a média anual e para o 19º máximo horário (método *forward*)

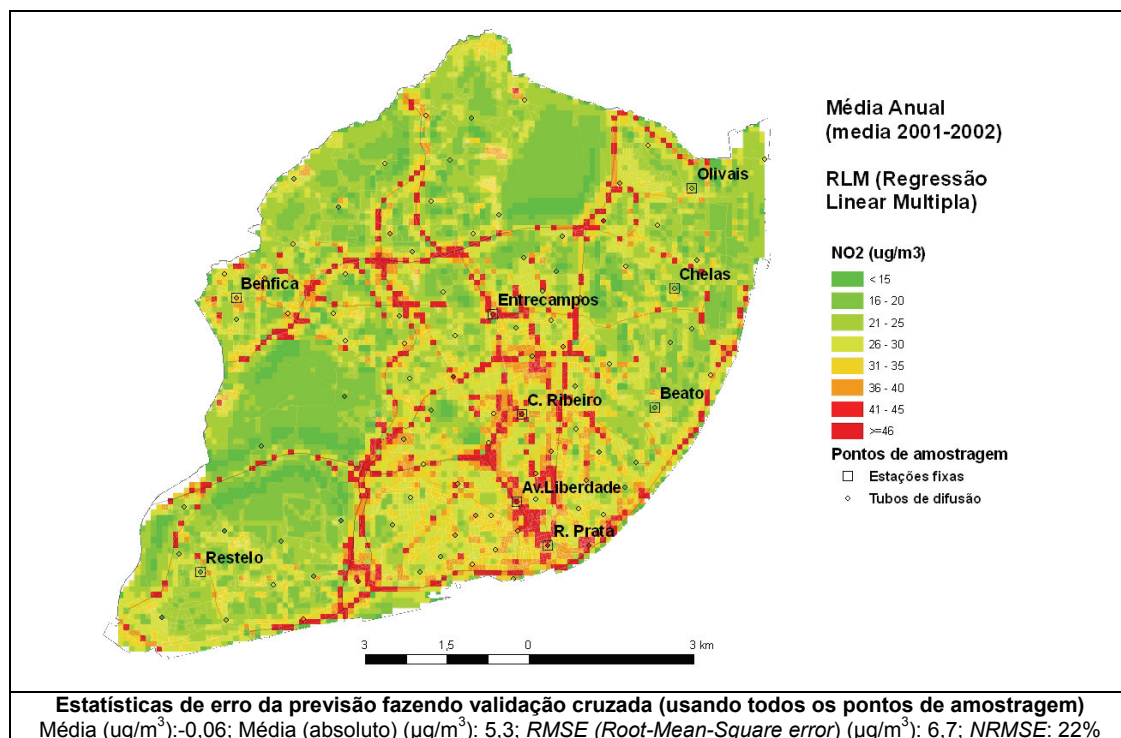


Figura 36. Interpolação da média de NO₂ em 2001 e 2002 usando RLM com as variáveis auxiliares e estatísticas de erro da validação cruzada.

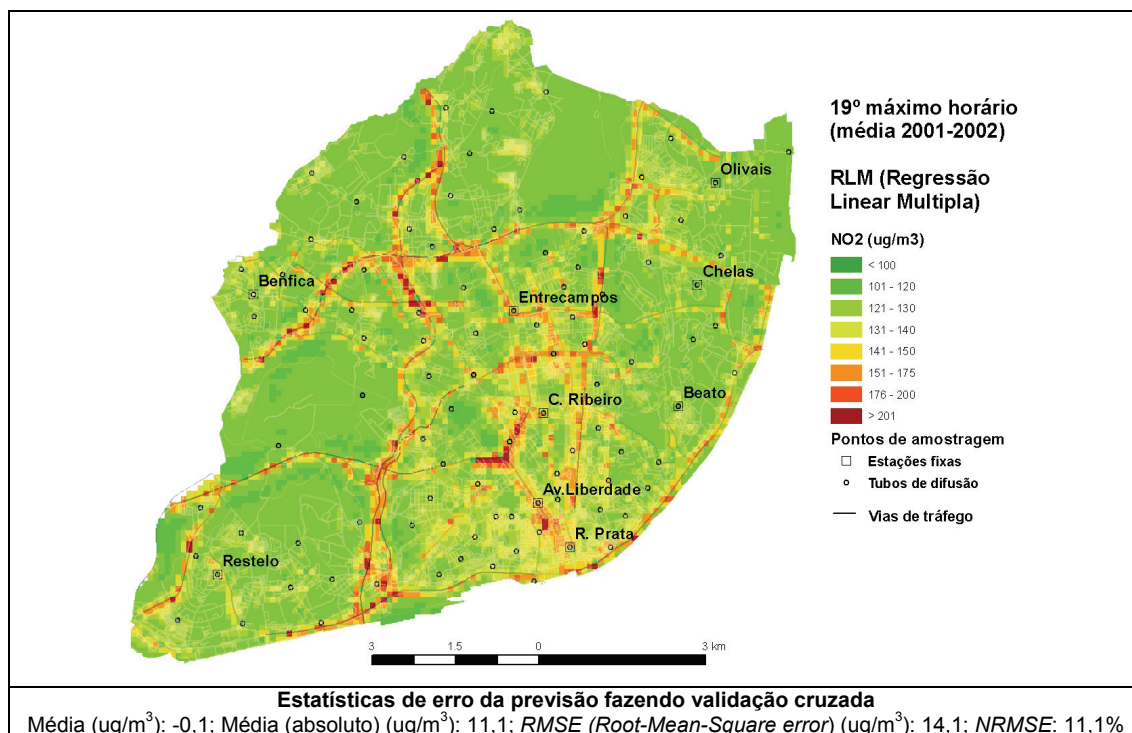


Figura 37. Interpolação do 19º máximo horário de NO₂ médio em 2001 e 2002 usando RLM com as variáveis auxiliares e estatísticas de erro da validação cruzada.

A análise dos mapas produzidos permite verificar que usando o método de interpolação RLM se obtiveram mapas bastante diferentes dos obtidos com os métodos anteriores, que tinham em consideração apenas, ou principalmente (*ordinary cokriging*), as concentrações

nos pontos de amostragem. Com este método foi possível identificar várias áreas, junto às principais vias de tráfego, para as quais se estima ocorrerem ultrapassagens ao valor limite estabelecido para a média anual de NO₂ e um pequeno conjunto de quadrículas onde se estima ocorrerem excedências ao valor limite definido para o 19º máximo horário de NO₂. Focalizando a atenção nos locais de fundo (sem influência directa de tráfego significativo), tal como para os mapas anteriores, é possível notar que no centro de Lisboa as concentrações de NO₂ são mais elevadas. É também de notar que para a zona do Aeroporto de Lisboa, onde não existem medições, se estimam concentrações muito baixas, o que também estará relacionado com o facto de neste estudo não terem sido consideradas as suas emissões.

Os modelos obtidos para a média anual e para o 19º máximo horário, são os seguintes:

$$\text{Média anual}_{2001-2002} = 23.51 + 0.04 \times \text{ET}_{\text{Soma 50 m}} + 0.004 \times \text{ET}_{\text{Soma 50-250 m}} - 2.13 \times \text{ET}_{\text{Média 250-500 m}} + 0.89 \times \text{ET}_{\text{DesPad 250-500 m}} + 458.5 \times \text{EA}_{\text{Mín 0-500 m}}$$

$$\text{19º Max}_{2001-2002} = 111.67 + 0.08 \times \text{ET}_{\text{Soma 50 m}} + 0.009 \times \text{ET}_{\text{Soma 50-250 m}} - 24.48 \times \text{ET}_{\text{Média 250-500 m}} + 1.87 \times \text{ET}_{\text{DesPad 250-500 m}} + 965.6 \times \text{EA}_{\text{Mín 0-500 m}}$$

Sendo: ET = Emissões NOx de tráfego e EA = Emissões NOx em área.

5.3.5 Comparação e avaliação da incerteza dos métodos de interpolação usados

Após terem sido testados e aplicados os vários métodos para modelar a distribuição espacial de NO₂ (sob a forma da sua média anual e do 19º máximo horário) na cidade de Lisboa, é importante proceder-se à sua avaliação usando para tal critérios objectivos. Para avaliar e comparar os resultados dos vários métodos recorreu-se ao:

- Ajustamento dos modelos, em termos de coeficiente de determinação (R^2) da regressão linear entre o observado e o modelado;
- Usando a técnica da validação cruzada, para cada estação fixa, ao seguinte conjunto de parâmetros estatísticos de erro (em concentração e em percentagem em relação à média das estações): Erro médio; Erro absoluto médio; *RMSE* (*Root Mean Square Error* - raiz da média do quadrado do erro).
- Análise visual dos mapas produzidos e dos mapas de erro associados, fazendo a multiplicação do VC-RMSE pela superfície de concentrações.

5.3.5.1 Ajustamentos dos modelos

Na **Figura 38**, para a média anual, e na **Figura 39**, para o 19º máximo horário, pode visualizar-se o ajustamento de cada um dos modelos de interpolação aos dados resultantes da estimativa dos resultados das campanhas, a partir dos dados das estações fixas. Verifica-se que o ajustamento para o *IDW* é de aproximadamente 100%, em ambos os casos, o que decorre do facto de este ser um interpolador exacto. Para a média anual verifica-se que a interpolação obtida pela RLM e pelos dois métodos de *kriging* obtiveram um ajustamento semelhante. Enquanto que, para o 19º máximo horários dois métodos de *kriging* obtiveram um ajustamento melhor do que a média anual.

Considera-se que o ajustamento dos valores modelados, relativamente aos valores reais, não é um critério adequado para a comparação da qualidade dos modelos de interpolação uma vez que alguns destes métodos são exactos ou quase exactos, ou seja, fazem passar a superfície pelos valores dos pontos de amostragem. A comparação do ajustamento só pode ser feita dentro de cada método de interpolação ou usando valores modelados fazendo validação cruzada, ou seja, os resultados sem pontos terem sido incluídos na modelação.

A validação cruzada foi feita usando os dados das seis estações fixas (equipamentos usando o método de referência) com eficiência superior a 85% em 2001 e 2002. As estações do Restelo, Entrecampos e Rua da Prata foram excluídas desta análise uma vez que não obtiveram 85% de eficiência para algum dos anos em análise. Os ajustamentos dos modelos, para as seis estações, fazendo a validação cruzada, apresentam-se nas **Figura 40** e **Figura 41**. Estes resultados mostram que o melhor modelo de interpolação é a RLM, seguida do método *coKriging*, *kriging* e por último o *IDW*. Verificou-se ainda que, o ajustamento para média anual é melhor que para o 19º máximo horário.

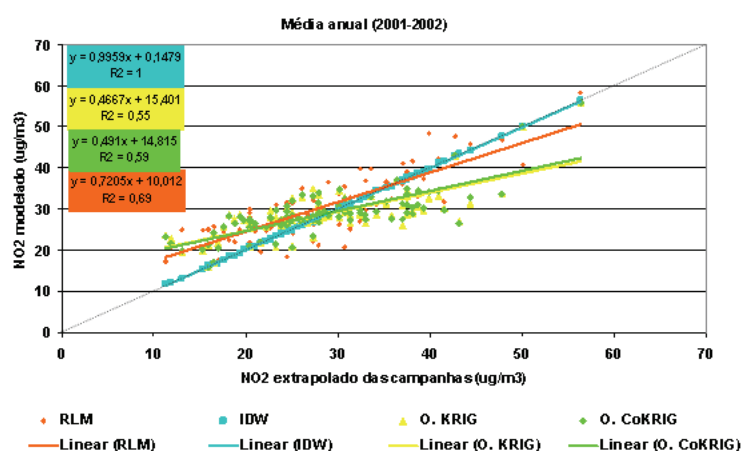


Figura 38. Ajustamento dos modelos de interpolação aos dados obtidos da estimativa dos resultados das campanhas para a média anual, usando os resultados das estações fixas.

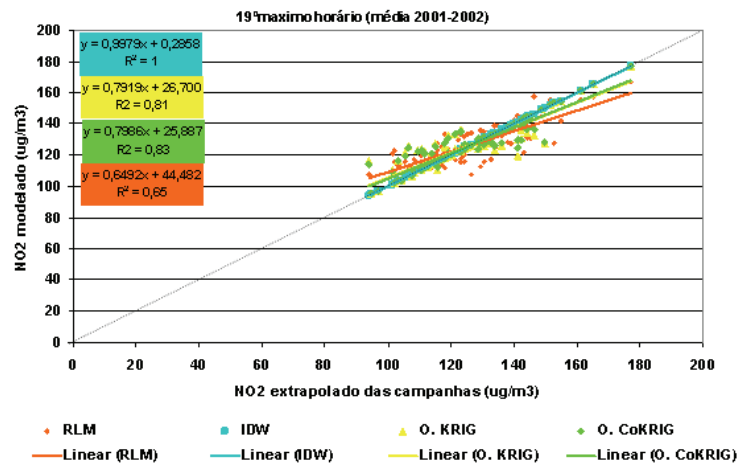


Figura 39. Ajustamento dos modelos de interpolação aos dados obtidos da estimativa dos resultados das campanhas para o 19º máximo horário, usando os resultados das estações fixas.

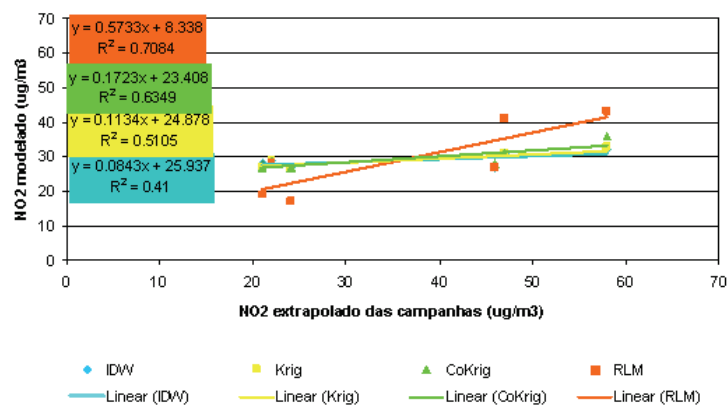


Figura 40. Ajustamento dos resultados dos modelos obtidos através da validação cruzada com os valores reais de cada uma das estações (média anual).

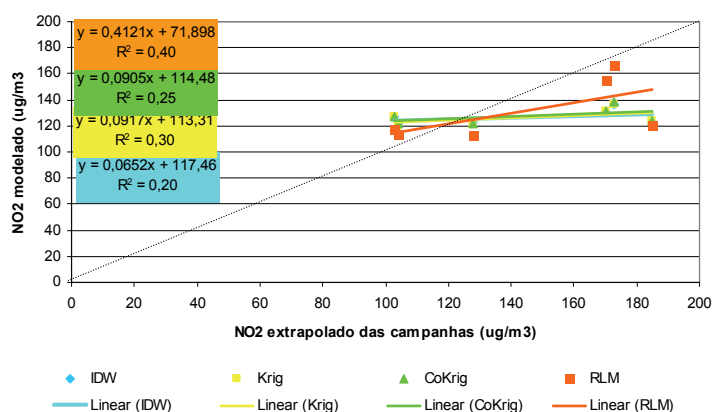


Figura 41. Ajustamento dos resultados dos modelos obtidos através da validação cruzada com os valores reais de cada uma das estações (19º máximo horário).

5.3.5.2 Análise de erro

O cálculo do erro, usando a validação cruzada, para o local das seis estações fixas, é feito relativamente a dois valores: o resultado do equipamento de referência e o resultado dos tubos de difusão. Deste modo obtém-se o erro individual de cada método de interpolação usando, *IDW*, *Ordinary Kriging*, *Ordinary Cokriging* e *RLM*, e o erro global da metodologia de mapeamento usada (método de estimativa da média anual e 19^a máximo diário + método de interpolação).

Nas **Tabela 18** e **Tabela 19** apresentam-se os resultados da validação cruzada por estação, para cada um dos métodos isoladamente e o erro acumulado da estimativa dos resultados das campanhas, com cada um dos métodos de interpolação, designado por erro global. Apresentam-se ainda, para cada um dos casos, as seguintes estatísticas: erro médio, erro médio absoluto e RMSE.

Na **Figura 42** e na **Figura 43**, apresentam-se em gráfico os erros globais por estação para a média anual e 19^o máximo horário, respectivamente, usando cada um dos métodos de interpolação. Finalmente, na **Figura 44** e na **Figura 45**, é feita uma comparação do valor real da média anual (2001-2002) de NO₂ nas 6 estações fixas com os resultados da validação cruzada para cada um dos métodos de interpolação, para a média anual e 19^o máximo horário, respectivamente.

A análise dos resultados da validação cruzada, através das tabelas e figuras apresentadas, mostra que:

- Comparando as estatísticas de erro para a média anual e para o 19^o máximo horário, verifica-se que, para a generalidade das estações e métodos de interpolação, as percentagens de erro são superiores para a média anual. Este facto era expectável uma vez que os valores para o indicador horário têm uma ordem de grandeza superior.
- O erro global é inferior à soma do erro da estimativa dos resultados das campanhas e do erro de cada um dos métodos de interpolação.
- Ordenando os métodos de interpolação do menor para o maior nível de erro obtemos: *RLM*, *CoKrig*, *Krig* e *IDW*. Ou seja, quanto mais complexo é o método melhores são os resultados.
- Todos os métodos de interpolação testados globalmente subestimam as concentrações de NO₂. De uma forma geral, as localizações com concentrações mais baixas são sobrestimadas e as localizações com concentrações mais elevadas são subestimadas.

- A RLM modela melhor a variabilidade espacial das concentrações principalmente junto a locais de tráfego. Os métodos *IDW*, *Krig* e *CoKrig* tendem a homogeneizar as superfícies de concentrações.
- Apesar do erro do método *CoKrig* ser para generalidade dos locais mais baixo que para os métodos *Krig* e *IDW*, as diferenças são muito pouco significativas. O acréscimo de tempo e custos despendidos para adquirir e tratar informação relativa a variáveis auxiliares não parece justificar a pequena melhoria verificada.
- Olhando para o erro médio absoluto verifica-se que a RLM tem um erro de 21% para a média anual e de 14% para o 19º máximo horário, ou seja, encontra-se em ambos os casos globalmente abaixo dos 30%.

De acordo com a Directiva 2008/50/CE, para o NO_2 , para efeitos da adequação dos mapas para verificação do cumprimento da legislação, o desvio entre o valor real e o valor modelado em relação ao valor limite (o erro relativo da directiva apresentado no ponto 4.4.3.) não pode ultrapassar, em 90% das estações fixas, 30% para o caso da média anual, e 50% para o 19º máximo horário. Deste modo, é necessário comparar os resultados da validação cruzada, para cada uma das estações, com 30% do valor limite para a média anual, $12 \mu\text{g}/\text{m}^3$ e 50% do valor limite para o 19º máximo horário, $100 \mu\text{g}/\text{m}^3$. Se estes valores de erro forem ultrapassados em mais do que uma das 6 estações ($6 \times 0.1 = 0.6 \approx 1$), o mapa não é adequado para verificação do cumprimento da legislação. Os resultados de erro obtidos permitem concluir que, para o valor limite horário qualquer dos métodos é adequado para verificação do cumprimento legal, enquanto que, para o valor limite anual apenas o mapa obtido por RLM é adequado.

Tendo ainda, em consideração o objectivo de estimar a população exposta a concentrações em excedência aos valores legislados verifica-se que, apesar de também subestimar as concentrações, o método RLM é o único que permite obter áreas em excedência.

Estações		Olivais (Fundo) (µg/m ³)	Beato (Fundo) (µg/m ³)	Chelas (Fundo) (µg/m ³)	Av. Casal Ribeiro (Tráfego) (µg/m ³)	Av. Liberdade (Tráfego) (µg/m ³)	Benfica (Tráfego) (µg/m ³)	Erro médio (µg/m ³)	Erro médio Absolute (µg/m ³)	RMSE (µg/m ³)
Concentração Real		22	21	24	47	58	46			
Estimativa campanhas	Concentração	24	20	25	47	57	37			
	Concentração	29	28	27	31	32	27			
	Erro	5	8	2	-16	-25	-10	-6 -17%	11 31%	13 38%
	Erro global	7	7	3	-16	-26	-19	-7 -20%	13 36%	15 42%
IDW	Concentração	29	27	27	31	33	27			
	Erro	5	7	2	-16	-24	-10	-6 -17%	11 30%	13 37%
	Erro global	7	6	3	-16	-25	-19	-7 -19%	13 35%	15 41%
	Concentração	29	27	27	31	36	28			
Ordinary Kriging	Concentração	29	27	27	31	36	28			
	Erro	5	7	2	-16	-21	-9	-5 -15%	10 29%	12 34%
	Erro global	7	6	3	-16	-22	-18	-7 -18%	12 33%	14 38%
	Concentração	26	27	24	38	49	28			
Ordinary CoKriging	Concentração	26	27	24	38	49	28			
	Erro	5	7	2	-16	-21	-9	-5 -15%	10 29%	12 34%
	Erro global	7	6	3	-16	-22	-18	-7 -18%	12 33%	14 38%
	Concentração	26	27	24	38	49	28			
Regressão Linear Múltipla	Concentração	26	27	24	38	49	28			
	Erro	2	7	-1	-9	-8	-9	-3 -9%	6 17%	7 20%
	Erro global	4	6	0	-9	-9	-18	-4 -12%	8 21%	9 26%
	Concentração	26	27	24	38	49	28			

Nota: casos em que o erro é superior a +/-12 µg/m³ (30% do VL para a média anual)

Tabela 18. Resultados da validação cruzada por estação para a média anual. Avaliação de cada um dos métodos de interpolação isoladamente (erro) e em conjunto com a estimativa dos resultados da campanha (erro global).

Estações		Olivais (Fundo) (µg/m ³)	Beato (Fundo) (µg/m ³)	Chelas (Fundo) (µg/m ³)	Av. Casal Ribeiro (Tráfego) (µg/m ³)	Av. Liberdade (Tráfego) (µg/m ³)	Benfica (Tráfego) (µg/m ³)	Erro médio Absoluto (µg/m ³)	RMSE (µg/m ³)
Concentração Real		103	104	128	170	173	185		
Estimativa campanhas	Concentração	118	111	118	161	177	143		
	Concentração	127	122	123	130	136	123		
	Erro	9	11	5	-31	-41	-20	-11 -8%	23 -16%
	Erro global	24	18	-5	-40	-37	-62	-17 -12%	36 -25%
<i>IDW</i>	Concentração	127	120	121	131	136	124		
	Erro	9	9	3	-30	-41	-19	-12 -8%	23 -16%
	Erro global	24	16	-7	-39	-37	-61	-17 -12%	35 25%
	Concentração	127	122	122	131	139	124		
<i>Ordinary Kriging</i>	Erro	9	11	4	-30	-38	-19	-11 -7%	22 15%
	Erro global	24	18	-6	-39	-34	-61	-16 -11%	35 24%
	Concentração	118	114	112	155	167	121		
	Erro	0	3	-6	-6	-10	-22	-7 -5%	11 -7%
<i>Ordinary CoKriging</i>	Erro global	15	10	-16	-15	-6	-64	-13 -9%	29 20%
	Concentração	127	122	122	131	139	124		
<i>Regressão Linear Múltipla</i>	Erro	9	11	4	-30	-38	-19	-11 -7%	22 15%
	Erro global	24	18	-6	-39	-34	-61	-16 -11%	35 24%
	Concentração	118	114	112	155	167	121		
	Erro	0	3	-6	-6	-10	-22	-7 -5%	11 -7%
	Erro global	15	10	-16	-15	-6	-64	-13 -9%	29 20%

Nota: casos em que o erro é superior a +/-100 µg/m³ (50% do VL horário)

Tabela 19. Resultados da validação cruzada por estação para o 19º máximo horário. Avaliação de cada um dos métodos de interpolação isoladamente (erro) e em conjunto com a estimativa dos resultados das campanhas (erro global).

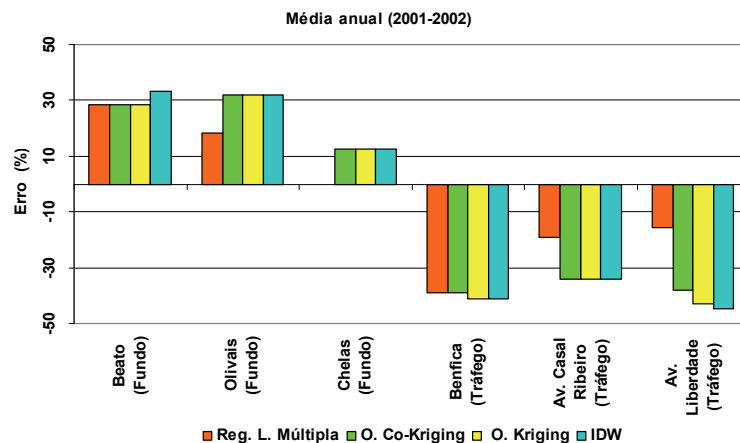


Figura 42. Erro global obtido para as estações fixas usando cada um dos métodos de interpolação (média anual).

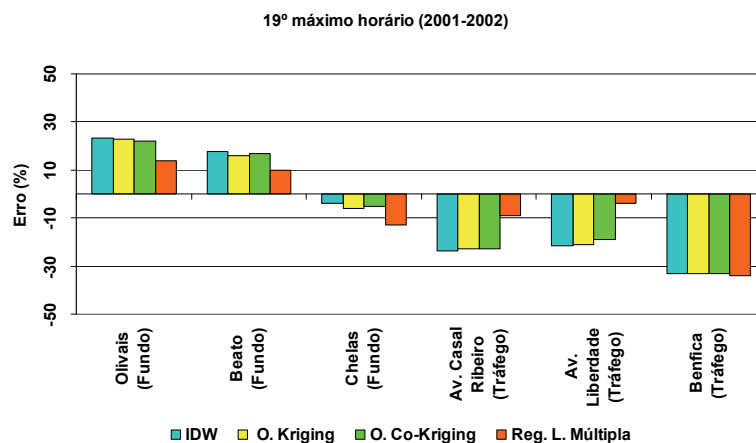


Figura 43. Erro global obtido para as estações fixas usando cada um dos métodos de interpolação (19º máximo horário).

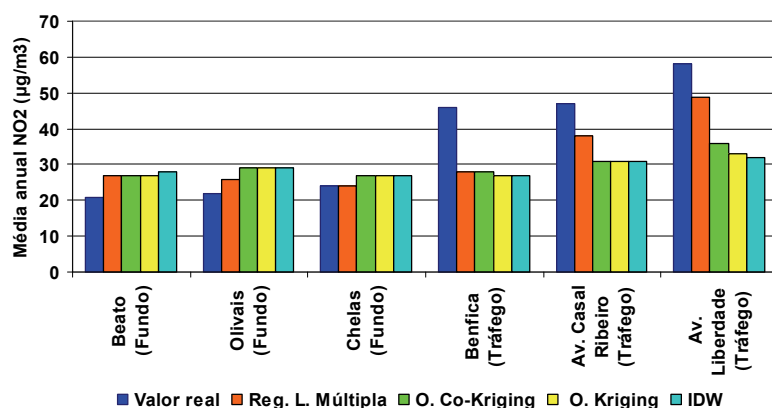


Figura 44. Valor real da média anual (2001-2002) de NO₂ nas 6 estações fixas comparado com os resultados da validação cruzada para cada um dos métodos de interpolação.

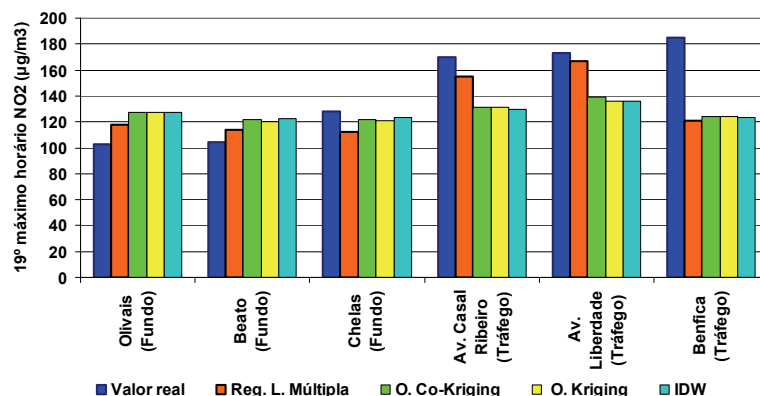


Figura 45. Valor real do 19º máximo horário (2001-2002) de NO₂ nas 6 estações fixas comparado com os resultados da validação cruzada para cada um dos métodos de interpolação.

5.3.5.3 Análise visual dos mapas produzidos e dos mapas de erro associados

Finalmente foi feita uma comparação visual dos mapas produzidos e dos mapas de erro associados. Na **Figura 46** e na **Figura 47** apresentam-se os mapas das concentrações de NO₂ (expressas pela média anual e 19º máximo horário, respectivamente) e respectivos mapas de erro resultantes da multiplicação das superfícies de concentração pelos CV-NRMSE obtidos fazendo validação cruzada para os 4 métodos de interpolação em análise.

Da avaliação dos mapas de distribuição das concentrações obtidos pode constatar-se que a RLM é o único dos métodos de interpolação avaliados, que permite obter um mapa de alta resolução capaz de modelar a variabilidade das concentrações junto às vias de tráfego, sendo por isso o mais indicado para modelar a distribuição espacial em áreas urbanas e assim estimar as áreas em exposição a excedências dos VL e a população exposta.

Os mapas obtidos pelos métodos de interpolação mais tradicionais, que se baseiam principalmente (*ordinary cokriging*) ou exclusivamente (*IDW* e *ordinary kriging*) nos dados de qualidade do ar, parecem ser mais adequados para identificar a poluição de fundo. Ou seja, permitem identificar de uma forma mais generalizada as áreas onde a poluição de fundo é mais alta, zona da baixa, ou mais baixa, zona do Parque Florestal de Monsanto, não detectando no entanto, os picos na envolvente das maiores vias de tráfego. Para que estes métodos pudessem atingir um nível de detalhe semelhante ao da RLM, seria necessário realizar campanhas de monitorização com um número de pontos de amostragem tão elevado que as tornaria impraticáveis.

A análise visual dos mapas de erro permitiu verificar que o erro na RLM em valor absoluto é maior junto às vias de tráfego enquanto que, para os restantes métodos os valores de erro mais elevados estão mais dispersos por toda a área da cidade de Lisboa em particular na baixa da cidade.

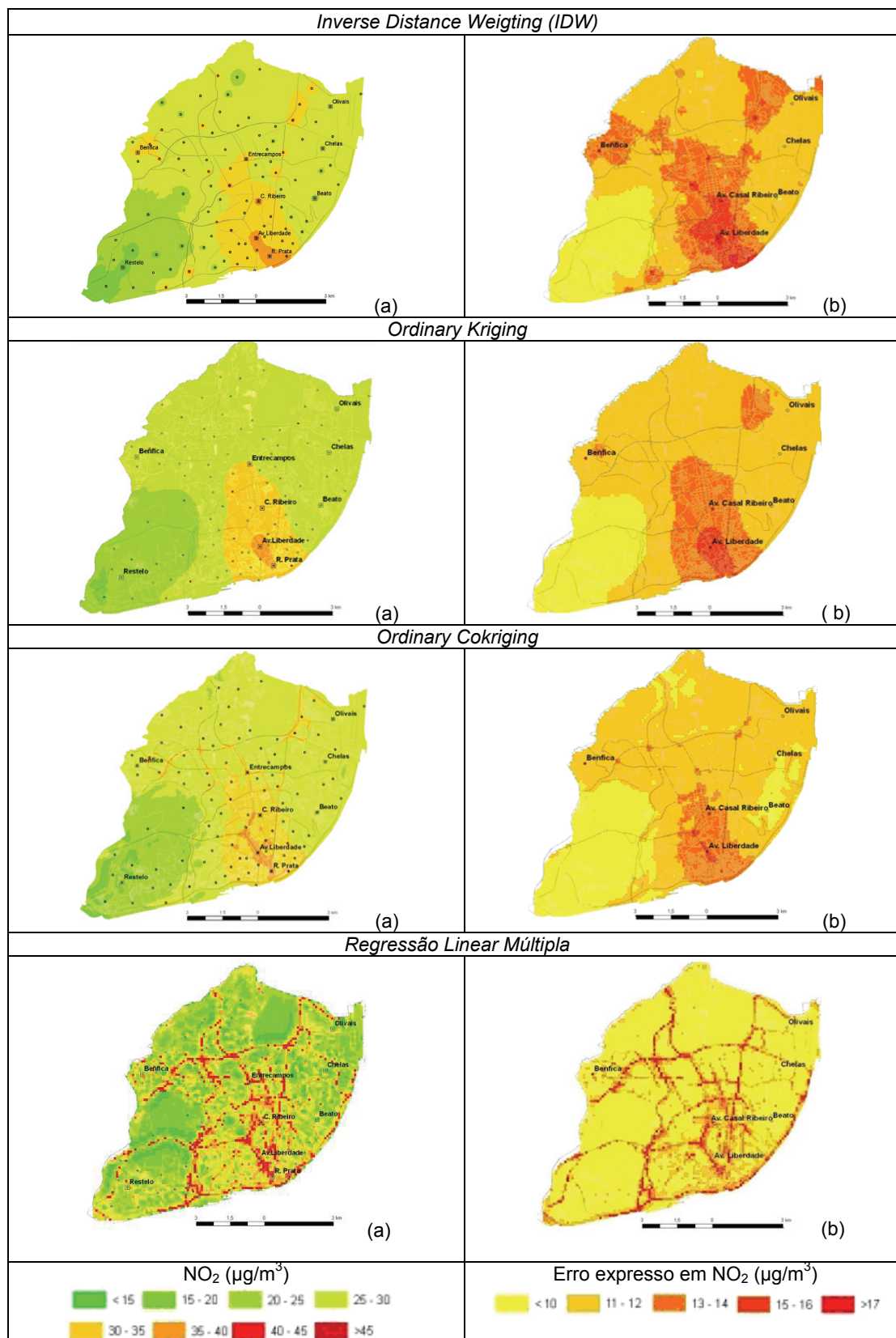


Figura 46. Concentrações de NO₂, expressas pela média anual, (a) e respectivos mapas de erro (b) fazendo validação cruzada para os quatro métodos de interpolação em análise.

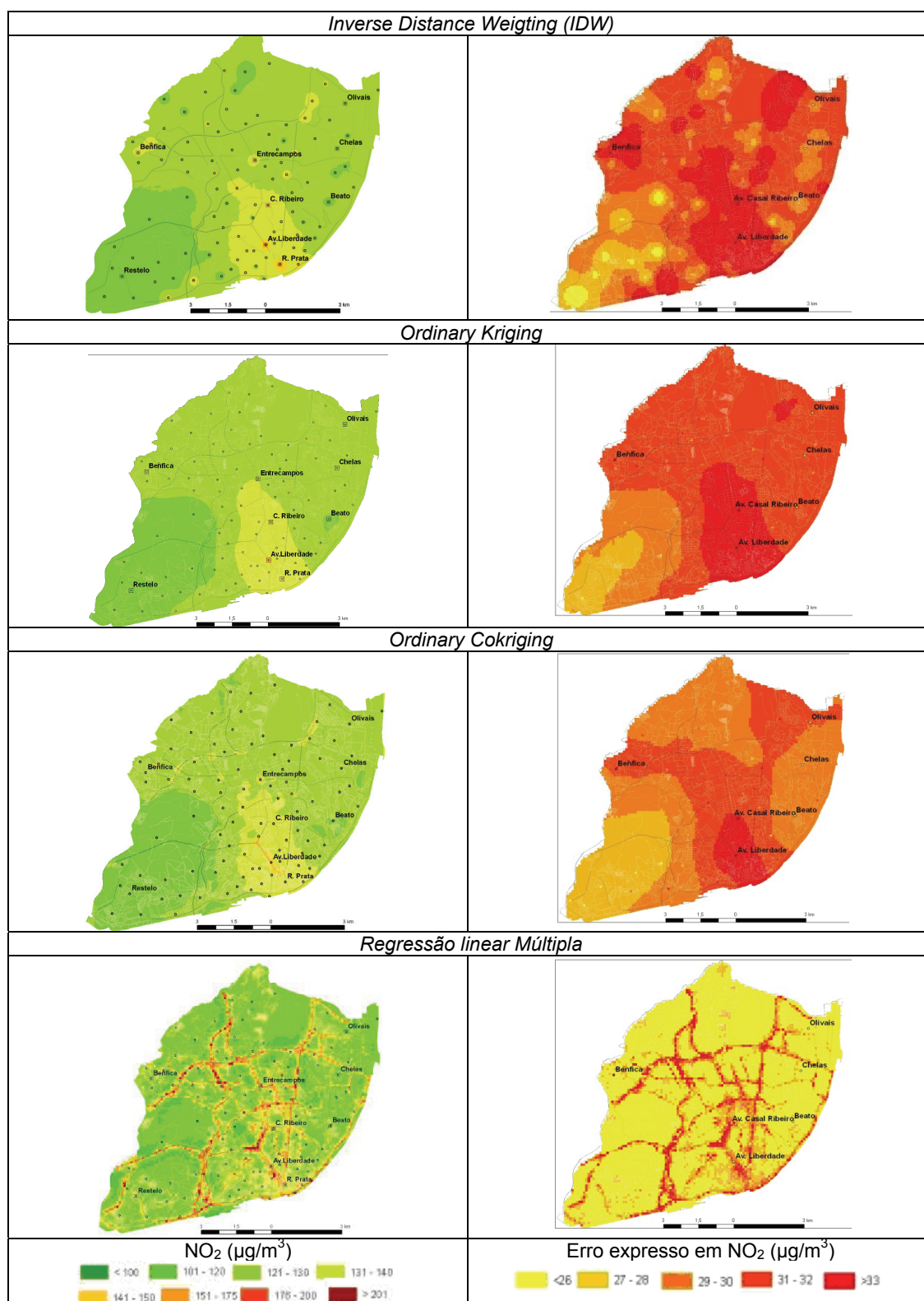


Figura 47. Concentrações de NO₂, expressas pelo 19º máximo horário, e respectivos mapas de erro fazendo validação cruzada para os quatro métodos de interpolação em análise.

5.4 Estimativa das áreas e da população exposta a poluição atmosférica

A avaliação/comparação dos mapas produzidos para representar a distribuição das concentrações de NO₂ na cidade de Lisboa, tendo em vista a estimativa da população exposta, indicou que o método mais adequado, é a regressão linear múltipla (RLM). Isto porque permite obter mapas de alta resolução e com menor erro associado. O melhor mapa obtido por este método correspondeu a um modelo recorrendo às emissões de tráfego e a densidade populacional como variáveis auxiliares.

Após a selecção dos mapas mais adequados, foi feita a extracção das áreas/quadrículas em excedência ao valor limite (VL), e valor limite adicionado da margem de tolerância (VL+MT) para a média anual (**Figura 48**) e das áreas em excedência ao valor limite horário (**Figura 49**). Dado que o VL+Mt para a média anual é diferente para 2001 e 2002 usou-se o valor mais baixo ou seja o valor de 2002 (56 µg/m³).

Seguidamente foi feito o cálculo da área em excedência e a estimativa da população exposta a níveis de NO₂ superiores aos permitidos por lei: A população exposta, foi obtida sobrepondo o mapa das áreas em excedência com o mapa da densidade populacional ao nível da subsecção estatística para 2001 (INE, 2001) (**Figura 50**).

Na **Tabela 20** apresentam-se a área em excedência e a estimativa da população exposta e as respectivas percentagens relativamente ao total de área do concelho e da população que habitava a cidade de Lisboa em 2001.

A área em excedência ao valor limite anual corresponde a 6,6% da área da cidade de Lisboa e nela habitam cerca de 44 mil habitantes. Relativamente ao valor limite anual adicionado da margem de tolerância, a área em excedência corresponde a 1,7% e a população residente afectada ronda os 7 mil habitantes. Para o valor limite horário a área afectada por concentrações acima do permitido por lei é reduzida afectando apenas cerca de 1% da área do concelho de Lisboa e cerca de 2000 habitantes.

É de recordar, que a cidade de Lisboa aumenta substancialmente a sua população durante o dia, pelo que estas estimativas que consideram apenas a população residente, estão muito subestimadas.

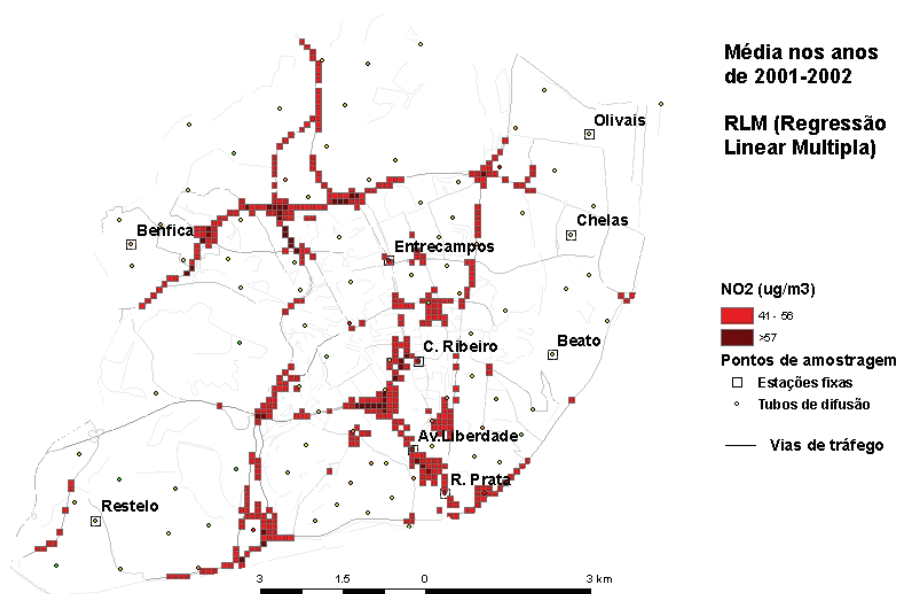


Figura 48. Extracção das grelhas de 100m que o modelo de RLM estimou estarem acima do VL e ao VL+MT para a média anual.

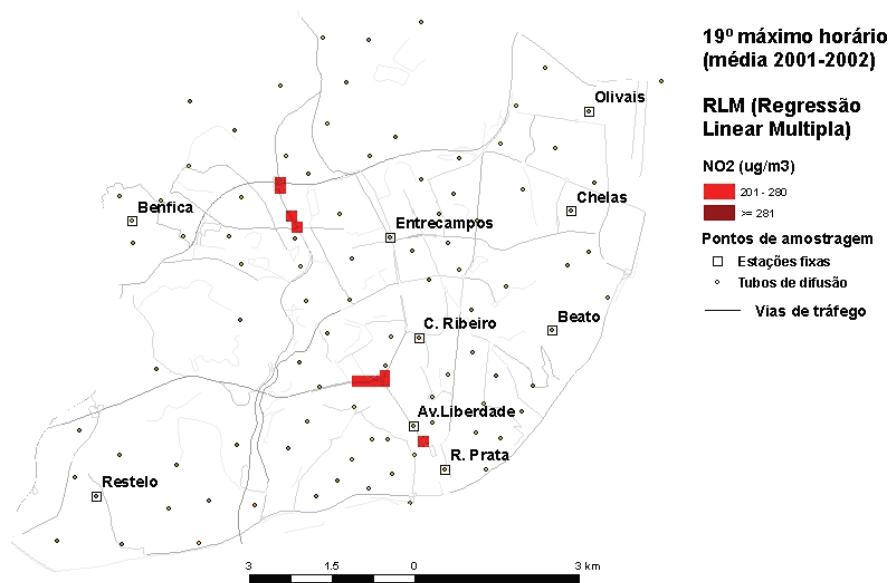


Figura 49. Extracção das grelhas de 100m que o modelo de RLM estimou estarem acima do VL para o 19º máximo horário.

Valores legais	Área (km ²)	Percentagem de área total (%)	População residente (hab.)	Percentagem da população residente total (%)
Anual	VL	7,8	44 197	5,5
	VL+Mt	1,3	7 142	0,5
Horário	VL	0,4	2 228	0,1

Tabela 20. Áreas e população em excelência ao VL e VL+MT anual e para o VL horário

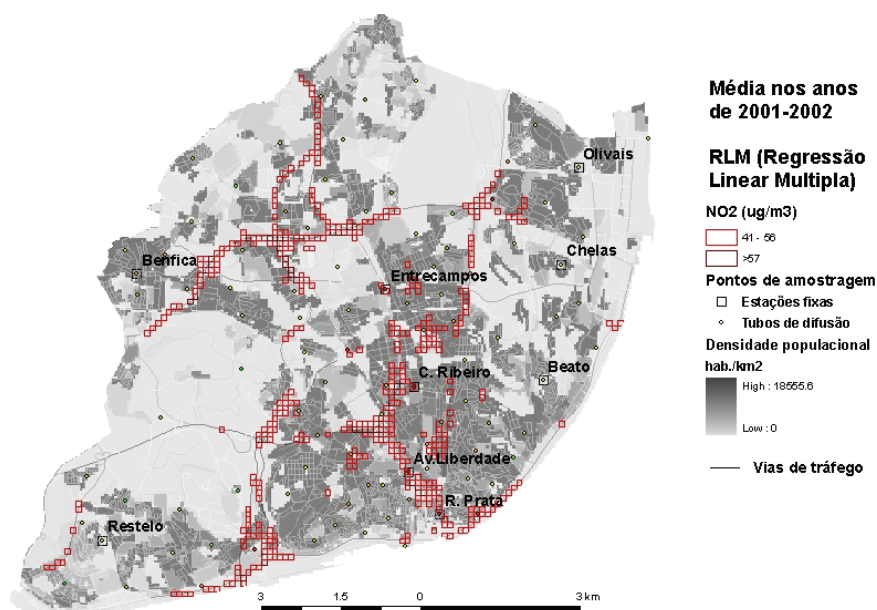


Figura 50. Sobreposição das grelhas de 100m, que o modelo de RLM estimou estarem acima do VL ($40 \mu\text{g}/\text{m}^3$) e ao VL+Mt ($56 \mu\text{g}/\text{m}^3$) para a média anual, com o mapa da densidade populacional.

5.5 Utilização de ferramentas de SIG

Após a realização do presente trabalho verificou-se que na obtenção de mapas de distribuição da qualidade do ar e da estimativa das áreas e população expostas a excedências ao valor limite, a utilização de operações implementadas numa ferramenta de SIG é fundamental. A análise global da metodologia, permitiu verificar que esta assentou em operações maioritariamente efectuadas em SIG, salientando-se as seguintes:

- Visualização e produção de mapas digitais com alteração de escalas e legendas;
- Sobreposição de mapas;
- Conversão de formatos de dados vectoriais em quadrícula (*raster*);
- Geração de grelhas regulares de pontos;
- Conversão de pontos em polígonos;
- Selecção e extracção de atributos de variáveis e criação de novos mapas;
- Operações de distância: Criação de áreas de influencia (*buffer*) em círculos e anéis;
- Operações entre camadas;
- Cálculo de estatísticas na vizinhança (*zonal statistics*);
- Cálculo de áreas;
- Métodos de interpolação determinísticos e geoestatísticos.

6 Conclusões e desenvolvimentos futuros

A obtenção de mapas da distribuição das concentrações dos poluentes atmosféricos em todo o território é uma das exigências das directivas comunitárias, uma vez que é fundamental para obter estimativas da área e da população exposta a excedências aos valores limite. Em áreas urbanas, a tarefa de mapeamento das concentrações dos poluentes, resultantes principalmente das emissões do tráfego rodoviário, tem algumas particularidades. Junto a vias de tráfego existem flutuações abruptas dos níveis de poluentes que não são fáceis de modelar.

Nesta tese, foram testadas e comparadas várias metodologias, implementadas em ambiente SIG, para o mapeamento da qualidade do ar, na cidade de Lisboa, usando métodos de interpolação determinísticos (*Inverse Distance Weighting - IDW*), geoestatísticos (*Ordinary Kriging* e *Ordinary CoKriging*) e estatísticos (regressão linear múltipla). Recorreu-se a dados de qualidade do ar e de variáveis auxiliares como os emissões de tráfego, densidade populacional, ocupação do solo e orografia.

A utilização de resultados de campanhas de curta duração com tubos de difusão, para aumentar o número de pontos espaciais a interpolar, revelou-se exequível e muito útil. Foi possível estimar indicadores anuais de longo e de curto prazo para as localizações de campanhas estabelecendo para tal regressões lineares simples entre os resultados anuais das estações de monitorização fixa e das médias semanais obtidas com os tubos de difusão colocadas junto a estas. O erro global associado a esta estimativa foi estimado usando a técnica de validação cruzada, tendo sido o VC-NRMSE de 15% para a média anual e de 20% para o 19º máximo horário, o que se considerou aceitável, face às exigências expostas na legislação em vigor para métodos de monitorização indicativos. Deste modo, os métodos de interpolação seleccionados foram testados para os cerca de 97 pontos das campanhas usando para tal os resultados da média anual e para o 19º máximo horário obtidos destas estimativas.

Tendo em vista a sua utilização na interpolação com regressão linear múltipla e *ordinary coKriging*, os mapas das variáveis auxiliares seleccionadas foram tratados fazendo a conversão do formato vectorial para *raster* com uma resolução de 10 m. Seguidamente, foi extraído o valor da variável no local de amostragem ou calculados os estimadores estatísticos básicos (média, máximo, mínimo, soma, desvio padrão e contagem de células) para cada variável, usando os anéis de vizinhança de cada ponto de amostragem com várias distâncias até um máximo de 500 m.

O melhor mapa obtido pelo método da regressão linear múltipla, avaliado pelo ajustamento do modelo ($R^2=0,52$), recorreu, como variáveis de entrada, às emissões de tráfego

desagregadas ao nível da via e à densidade populacional ao nível da subsecção estatística. As variáveis de entrada seleccionadas usando o método progressivo foram: soma das emissões de tráfego no anel de 50 m, soma das emissões de tráfego no anel de 50-250 m, média das emissões de tráfego no anel de 250-500 m, desvio padrão das emissões de tráfego no anel de 250-500 m e mínimo das emissões em área no anel de 0-500 m.

A análise da incerteza comparativa dos quatro métodos de interpolação foi feita recorrendo à técnica de validação cruzada, para os pontos das estações de monitorização (6 estações), usando como indicador de ajustamento dos modelos o coeficiente de determinação (R^2) e como estatísticas de erro: o erro médio, o erro médio absoluto e o RMSE. Os resultados obtidos para a média anual apresentam-se na **Tabela 21** e revelam que o método de regressão linear múltipla é o mais adequado para interpolar os resultados das campanhas extrapolados. Com este método foi possível obter um mapa com erro abaixo de 30% que permitiu estimar a população e áreas expostas a níveis elevados de poluição devidos ao tráfego rodoviário. Para o 19º máximo horário (**Tabela 22**) as conclusões são idênticas sendo, no entanto, o ajustamento dos modelos mais baixo, o que era expectável uma vez que o erro da extrapolação dos valores das campanhas para o indicador horário é superior. As estatísticas de erro são mais baixas para o 19º máximo horário, uma vez que os valores são de uma ordem de grandeza superior. Em ambas as tabelas os resultados apresentados incluem o erro associado à extrapolação dos resultados dos tubos de difusão para a média anual e 19º máximo horário.

Método	R^2	Erro médio (%)	Erro médio absoluto (%)	RMSE (%)
IDW	0,41	20%	36%	42%
Ordinary Kriging	0,50	-19%	35%	41%
Ordinary Cokriging	0,63	-18%	33%	38%
Regressão Linear Múltipla	0,73	-12%	21%	26%

Tabela 21. Ajustamento dos resultados, para a média anual, dos métodos de interpolação e estatísticas de erro usando a técnica de validação cruzada relativamente aos valores reais obtidos nas seis estações de monitorização

Método	R^2	Erro médio (%)	Erro médio absoluto (%)	RMSE (%)
IDW	0,20	-12	20	25
Ordinary Kriging	0,30	-12	20	25
Ordinary Cokriging	0,25	-11	20	24
Regressão Linear Múltipla	0,40	-9	14	20

Tabela 22. Ajustamento dos resultados, para o 19º Máximo horário, dos métodos de interpolação e estatísticas de erro usando a técnica de validação cruzada relativamente aos valores reais obtidos nas seis estações de monitorização

Os mapas obtidos com RLM permitiram o cálculo da área em excedência ao valor limite anual que corresponde a 6,6% da área da cidade de Lisboa e a uma população de 44 197 habitantes (7,8% da população de Lisboa). Desses 7 142 habitantes (1,3% da população de

Lisboa) habitam locais expostos a excedências ao valor limite anual adicionado da margem de tolerância e 2.228 habitam locais com excedências ao valor limite horário.

Para além da avaliação da incerteza na comparação dos vários métodos, considerando aspectos como: os objectivos a que se destinam os mapas desenvolvidos, custo e tempo despendido no desenvolvimento do método, nomeadamente na aquisição das variáveis auxiliares, resolução temporal e espacial possíveis e requeridas, aplicabilidade a outros períodos de tempo e locais, pode concluir-se que o método da RLM é bastante mais exigente do que os restantes em termos de informação requerida e tempo de processamento despendido. No entanto, para o presente caso, dado que o objectivo era modelar a distribuição espacial de um poluente originado pelo tráfego numa área urbana, a RLM foi o único dos métodos de interpolação avaliados que permitiu obter um mapa de alta resolução capaz de modelar a variabilidade das concentrações junto às vias de tráfego, sendo por isso o mais indicado para modelar a distribuição espacial em áreas urbanas.

Apesar dos resultados da metodologia de regressão linear múltipla desenvolvida, no presente trabalho, serem positivos, a metodologia desenvolvida tem algumas limitações. As principais limitações estão relacionadas com a não consideração de alguns factores, como variáveis auxiliares, que podem ser determinantes para as concentrações verificadas em cada ponto. Este é o caso das emissões de fontes pontuais e de condições meteorológicas diferenciadas para cada local monitorizado. Apesar da cidade de Lisboa ter uma área relativamente reduzida (85 km²), dado tratar-se de uma área urbana, é uma generalização considerar que em todos os pontos as condições meteorológicas são semelhantes. Isto porque, num ambiente urbano, aliadas às condições meteorológicas, as características físicas de cada rua, relativamente ao seu comprimento, largura, orientação, altura e densidade dos edifícios, condicionam a capacidade de dispersão dos poluentes. Dos factores referidos, o único que é considerado na metodologia ainda que de uma forma indirecta, é a densidade e altura dos edifícios através do uso da variável densidade populacional ao nível da subsecção estatística (INE, 2001). No entanto, o facto de serem mapeados indicadores anuais, e não de uma dada hora ou dia, pode atenuar consideravelmente o facto de não serem consideradas as condições meteorológicas desagregadas espacial e temporalmente.

Relativamente à utilização de dados de campanhas de curta duração, assumir que, as concentrações obtidas são representativas das concentrações dessa localização relativamente à média anual, e que a distribuição espacial das emissões verificada durante cada uma das campanhas se manteve durante os dois anos em análise, tem obviamente erros associados.

Como limitações na determinação da população exposta à poluição atmosférica é de referir que apenas foi tida em consideração a população residente, o que dentro da cidade de Lisboa e principalmente nas áreas de maior densidade de comércio e serviços que correspondem a zonas de menor densidade de população residente.

O presente trabalho permitiu concluir que ao nível local a regressão linear múltipla é o método de interpolação mais adequado para obter mapas de alta resolução relativos a poluentes relacionados com o tráfego rodoviário. Dos métodos de interpolação testados foi o único que permitiu identificar a população e áreas expostas a níveis de poluição superiores ao permitido legalmente. Pode concluir-se que, o mapeamento usando interpoladores determinísticos e geoestatísticos, que assumem como princípio fundamental que pontos vizinhos no espaço tem valores mais correlacionados que pontos distantes, não é o mais adequado para a obtenção de superfícies de distribuição das concentrações em áreas urbanas. Isto acontece porque normalmente na envolvente de cada fonte emissora (principalmente estradas) não existem pontos de amostragem suficientes para se conseguir distinguir/definir um padrão. No caso de se pretender modelar um poluente como o ozono, mais de influência regional, ou concentrações de fundo de outros poluentes, os métodos de interpolação apenas baseados na distância entre pontos os resultados deverão ser mais encorajadores.

A disponibilidade de mapas de variáveis auxiliares cada vez mais exactos e precisos conjugada com a evolução da ferramentas SIG, que realizam operações cada vez mais complexas em menos tempo, permitiu realizar as tarefas objecto desta tese, que até há algum tempo seriam muito difíceis, nomeadamente a aplicação de RLM como método de interpolação, e no futuro permitirá melhorar a qualidade dos mapas produzidos.

Em trabalhos futuros, os objectivos deverão ser a aplicação desta metodologia a outros poluentes como o PM_{10} e O_3 , a outras áreas urbanas dentro da Região de Lisboa e Vale do Tejo e à escala nacional para uma grelha de $1 \times 1 \text{ km}^2$. Recomenda-se testar a inclusão de outras variáveis auxiliares como variáveis meteorológicas e variáveis relacionadas com a densidade de construção (altura dos edifícios e distância da via aos edifícios), assim como a produção de mapas de probabilidade de excedência para complementar a avaliação efectuada.

Referências bibliográficas

Agência Portuguesa do Ambiente (APA), 2008. *QualAr - Base de dados on-line sobre qualidade do ar em Portugal*. (URL: <http://www.qualar.org>, consulta em Dezembro de 2008).

Bishop, I.D., Karadaglis, C., 1996. *Combining GIS Based Environmental Modelling and Visualisation: Another Window on the Modelling Process*. Centre for Geographic Information Systems and Modelling. University of Melbourne .

Briggs, D.J., France J., 1980. Landscape evaluation: a comparative study. *Journal of Environmental Management*, 10, 263-275.

Briggs, D.J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebre, E., Pryl, K., Van Der Veen, A., 1997. Mapping urban air pollution using GIS: A regression-based approach. *International Journal of Geographical Information Science*, 11 (7), 699-718.

Briggs, D.J., De Hoogh, C., Gulliver, J., Wills, J., Elliott, P., Kingham, S., Smallbone, K. 2000. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Science of the Total Environment*, 253 (1-3), 151-167.

Briggs, D., 2005. The role of GIS: Coping with space (and time) in air pollution exposure assessment. *Journal of Toxicology and Environmental Health, Part A*, 68 (13-14), 1243-1261.

Briggs, D.J., Aaheim, A., Dore, C., Hoek, G., Petrakis, M., Shaddick, G., 2005. *Air pollution modelling for support to policy on health and environmental risks in Europe*. Final Report Section 6. Imperial College, London. EVK2-2002-00176. (URL: <http://www.apmosphere.org/>, consulta em Dezembro de 2008).

Burrough, P. A., 1986. *Principles of Geographical Information Systems for Land Resources Assessment*. Monographs on soil and resources survey, n. 12 (Oxford: Clarendon Press).

CEN, 2002. Ambient air quality - Diffusive samplers for the determination of concentrations of gases and vapours - Requirements and test methods. Part 2: Specific requirements and test methods. *European Standards* EN 13528-1:2002, 11 October 2002.

Comissão Europeia (CE), 1997. *Position paper on air quality: Nitrogen Dioxide*. Working Group on Nitrogen Dioxide. Novembro de 1997.

Comissão Europeia (CE), 2002a. *Guideline referring to Commission Decision 2001/839/EC laying down a questionnaire to be used for annual reporting on ambient air quality assessment under Council Directives 96/62/EC and 1999/30/EC*. January 28th 2002

Comissão Europeia (CE), 2002b. *Guidance on Assessment under the EU Air Quality Directives*. Working Group for Ambient Air Quality Directives. Final draft. (URL: <http://europa.eu.int/comm/environment/air/pdf/guidanceunderairquality.pdf>, consultado em 15 de Março de 2006)

Comissão Europeia (CE), 2005. *Thematic Strategy on Air Pollution. Communication from the Commission to the council and the European Parliament*. (COM(2005) 446) Brussels, 21.9.2005. (URL: <http://europa.eu.int/comm/environment/air/cafe/index.htm>, consultado em 15 de Março de 2006).

Davis, J. C., 1986. *Statistics and data analysis in geology*. Second Edition, John Wiley & Sons 646 p. (URL: <http://www.cpatc.embrapa.br/labgeo/srgsr2/pdfs/poster35.pdf>, consulta em Dezembro de 2008).

Decisão da Comissão 2001/752/CE de 17 Outubro de 2001, relativa à revisão dos anexos da Decisão 97/101/CE do Conselho que estabelece um intercâmbio recíproco de informações e de dados provenientes das redes e estações individuais que medem a poluição atmosférica nos Estados-Membros. *Jornal Oficial das Comunidades Europeias*. L 282, 26.10.2001, 69-76.

Decisão da Comissão 2004/461/EC, de 29 Abril de 2004. Laying down a questionnaire to be used for annual reporting on ambient air quality assessment under Council Directives 96/62/EC and 1999/30/EC and under Directives 2000/69/EC and 2002/3/EC of the European Parliament and of the Council. *Jornal Oficial das Comunidades Europeias* L 156, 30.4.2004, 78–125.

De Saeger, E., Gerbolès, M., Ballesta, P. P., Amantini, L. , Payrissat, M., 1995. *Air Quality Measurements in Brussels (1993-1994) - NO₂ and BTX Monitoring Campaigns by Diffusive Samplers*. EUR Report 16310 EN, 1995.

Denby, B., Walker, S. E., Eben, K., Fiala, J., 2005. *Interpolation and assimilation methods for European scale air quality assessment and mapping Part I: Review and recommendations*. Final Draft. December 2005. ETC/ACC Technical Paper 2005/7.

Denby, B., 2007. *Individual case study report 6: Basic data assimilation: application to the urban scale*. March 2007. Air4EU-Air Quality Assessment for Europe: from local to continental scale- 6th Framework Programme- Policy oriented Research Priority 8.1 Topic 1.5 Task 2.

Denby, B., Larssen, S., Guerreiro, C., Douros, J., Moussiopoulos N., Fragkou, L., Gauss, M., Olesen, H., Miranda, A. I., 2008. *Guidance on the use of models for the European air quality directive*. FAIRMODE - A working document of the Forum for Air Quality Modelling in Europe. ETC/ACC report. Version 2.6.

Directiva 96/62/CE do Conselho, de 27 de Setembro relativa à avaliação e gestão da qualidade do ar ambiente. *Jornal Oficial das Comunidades Europeias* L 296, 21/11/1996, 55-63.

Directiva 1999/30/CE do Conselho, de 22 de Abril, relativa a valores limite para o dióxido de enxofre, dióxido de azoto e óxidos de azoto, partículas em suspensão e chumbo no ar ambiente, *Jornal Oficial das Comunidades Europeias* L 163. 29/06/1999. 41-60.

Directiva 2008/50/CE do parlamento Europeu e do Conselho, de 21 de Maio de 2008 relativa à qualidade do ar ambiente e a um ar mais limpo na Europa. *Jornal Oficial da União Europeia* L 151, 11/06/2008, 1-44.

Eastman, J. R., Kyem, P. A. K., Toledano, J., and Jin, W., 1993. Explorations in Geographic Information Systems. In *Volume 4: GIS and decision making*. Geneva (Switzerland). United Nations Institute for Training and Research (UNITAR).

Ferreira, F., Mesquita, S., Tente, H., Torres, P., Santos, E., Carreira, P., Jardim, D., Esgalhado, F., 2002. *Avaliação preliminar da Qualidade do Ar em Portugal, no âmbito da Directiva 1999/30/CE - SO₂, NO₂, NO_x, PM₁₀ e Pb*, Julho de 2002. (URL: <http://www.iambiente.pt>, consulta em Dezembro de 2008).

Ferreira, F., Nogueira, L., Meireles, P., Mesquita, S., Almeida, C., Tente, H. Torres, P., 2003. Diffusive sampling and monitoring stations to evaluate air quality in urban areas, In *Proceedings of the UAQ4 conference*, Charles University Prague, Czech Republic, 25-28 March 2003.

Ferreira, F.; Tente, H.; Torres, P.; Mesquita, S.; Santos, E.; Jardim, D.; Esgalhado, F., 2001. Background levels of Sulphur Dioxide, Nitrogen Dioxide and Ozone in Portugal, In *Proceedings of the International Conference Measuring Air Pollutants by Diffusive Sampling* – Montpellier, 26-28 Setembro.

Fiala, J., Srněnský, R., Livorová, H., Novotný, P., Riedl, P., Beňo, P. 2000. *Methods of assimilation of measured and modelled data in the production of fields of air quality characteristics*. In Environmental project VaV/740/2/00. CHMI, Praha.

Field, R. A., Ballesta, P. P., Caracena, A. B., Nikolova, I., Connolly, R., Cao, N., Gerboles, M., Buzica, D., Amantini, L., Lagler, F., Borowiak, A., Marelli, L., De Santi, G., De Saeger, E., 2005. *Population Exposure to Air Pollutants in Europe (PEOPLE) Methodological Strategy and Basic Results*. EUR 21810 EN European Commission, Joint Research Centre, Institute for Environment and Sustainability, Emissions and Health Unit, TP 441, 21020 Ispra (VA), Italy.

FSM (Fondazione Salvatore Maugeri), 2007. *Radiello- the radial diffusive sampler*. (URL: http://www.radiello.com/english/index_en.html, consulta em Dezembro de 2008).

Foot, K. E., Huebner, D. J., 1996. *Managing Error*. Department of Geography, University of Texas at Austin, (URL: http://www.colorado.edu/geography/gcraft/notes/manerror/manerror_f.html, consulta em Dezembro de 2008).

Fuller, R.M., Wyatt, B.K., Barr, C.J. 1998. Countryside survey from ground and space: different perspectives, complementary results. *Journal of Environment Management* 1998;54:101-126.

Gois, V., Torres, P., Nogueira, L., Maciel, H., Almeida, C., 2006. *Inventário de Emissões Atmosféricas da Região de Lisboa e Vale do Tejo 2000-2001*. Comissão de Coordenação e Desenvolvimento Regional de Lisboa e Vale do Tejo. 1ª edição, Lisboa, Dezembro de 2006. ISBN: 972-8872-11-9.

Greig, A., Pohjola, M., Kukkonen, J., 2001. *Literature review on carbon monoxide modelling*. Project funded by the European Community under the “Information Society Technology” Programme (1998-2002).

Grosso, N., 2005. Retrieving aerosol optical thickness from a MODIS dataset for the Lisbon Metropolitan Area using a contrast reduction method. In *Proceedings of the 25th EARSeL Symposium - Global Developments in Environmental Earth Observation from Space*, June 6 - 11, Porto, Portugal.

Grosso, N., Ferreira, F., Mesquita, S. 2006. Using remote sensing data to improve PM10 mapping and population exposure. In *Proceedings of the ITM 2006. 28th NATO/CCMS International Technical Meeting on Air Pollution Modelling and its Application*. 15-19 May 2006 Leipzig, Germany.

Guimarães, R. C., Cabral, J.A.S. 1997. *Estatística*. Edição revista. MacGraw-Hill.

Hewitt, C.N. 1991. Spatial variation in nitrogen dioxide concentrations in an urban area. *Atmospheric Environment*. 1991:25B:429-434.

Hjellbrekke, A.G., 2000. *Ozone measurements 1998*. Kjeller, Norwegian Institute for Air Research. EMEP/CCC-Report 5/2000.

Hoek, G., Beelen, R., Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, Volume 42, Issue 33, October 2008, 7561-7578.

Hoek, G., Fischer, P., van den Brandt, P., Goldbohm, S., and Brunekreef, B., 2001. Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. *Journal of Exposure Science and Environmental Epidemiology*. 11, 459–469.

Horálek, J., Kurfürst, P., Denby, B., Smet, P., Leeuw, F., Brabec, M., Fiala, J., 2005. *Interpolation and assimilation methods for European scale air quality assessment and mapping Part II: Development and testing new methodologies*. Final draft. December 2005. ETC/ACC Technical Paper 2005/8.

Hunter, G.J., 1998. "Managing Uncertainty in GIS", *NCGIA Core Curriculum in GIScience*, (URL: <http://www.ncgia.ucsb.edu/giscc/units/u187/u1871.html>, posted in February 03 1998, consulta em Dezembro de 2008).

Instituto Geográfico Português, 2000. Série cartográfica de ocupação do solo à escala 1:100000 "*Corine Land Cover de 2000*".

Instituto Nacional de Estatística (INE), 2001. Dados de população ao nível da subsecção estatística. *Censos 2001*.

Instituto Português de Qualidade (IPQ), 1992. Norma Portuguesa NP 4172:1992 Qualidade do ar - Determinação da concentração em massa de óxidos de azoto no ar ambiente.

Jeannée, N., Nedellec, V., Bouallala, S., Deraisme, J., Desqueyroux, H., 2004. Geostatistical assessment of long term human exposure to air pollution – Geostatistics for Environmental Applications. In *Proceedings of the GEOENV2004*. Springer. 161-172.

Johnston, K., Ver Hoef, J.M., Krivoruchko, k., Lucas, N., 2001. *Using ArcGIS Geostatistical Analyst*. GIS by ESRI.

Knotters M, Brus DJ, Voshaar JHO. 1995. A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. *Geoderma* 1995; 67: 227-246.

Laxen EAH, Noordally E., 1987. Nitrogen dioxide distribution in street canyons. *Atmospheric Environment* 1987; 21: 1899-1903.

Li, J.L., Liang, T.G., Chen, Q.G., 1998. Estimating grassland yields using remote sensing and GIS technologies in China. *New Zealand Journal of Agricultural Research*, 1998;41:31-38.

Lindley, S.J., Walsh T. 2005. Inter-comparison of interpolated background nitrogen dioxide concentrations across Greater Manchester, UK. *Atmospheric Environment* 39 (2005) 2709–2724.

Liu, L. J. S., Rossini, A. J. 1996. Use of kriging models to predict 12-hour mean ozone concentrations in Metropolitan Toronto—A pilot study. *Environment International*. 22:677–692.

Marques F., Cassmassi, J., Ferreira, F. E Mesquita, S. 2006. Air quality forecast for Lisbon, Portugal: from data analysis to outreach. In *Proceedings of the U.S. EPA's 2006 National Air Quality Conferences*. Fevereiro 5-8, 2006, em San Antonio, Texas, USA. (URL: http://www.epa.gov/airnow//2006conference/posters/Air_quality_forecast_for_Lisbon,_Portugal_Filipa_Marques.PDF, consulta em Dezembro de 2008).

Marshall, J. D., Nethery, E. Brauer, M., 2008. Within-urban variability in ambient air pollution: Comparison of estimation methods. *Atmospheric Environment* 42 (2008) 1359–1369.

Mattson MD, Godfrey PJ. Identification of road salt contamination using multiple regression and GIS. *Environmental Management* 1994;18:767-773.

Mesquita, S, Ferreira, F.; Tente, H.; Torres, P., 2004. Using systematic diffusive sampling campaigns and geostatistics to map air pollution in Portugal. In *Proceedings of the GEOENV2004*. Centre for Hydrogeology, University of Neuchâtel, Switzerland, on October 13-15.

Monteiro, A., Vautard, R., Borrego C., Miranda, A.I., 2005. Long-term simulations of photo oxidant pollution over Portugal using the CHIMERE model, *Atmospheric Environment* 39, 3089–3101.

Monteiro, A., Miranda, A.I., Borrego, C., Vautard, R., 2007a. Air quality assessment for Portugal, *Science of the Total Environment*. 373 (2007), 22–31.

Monteiro, A., Miranda, A.I., Borrego, C., Vautard, R., Ferreira, J., Perez, A.T., 2007b. Long-term assessment of particulate matter using CHIMERE model, *Atmospheric Environment* 41 (2007) 7726–7738.

Myers, D. E., 1994, Spatial interpolation: an overview. *Geoderma*, 62, 17- 28.

Nevers, N., 1995. *Air Pollution Control Engineering*. McGraw-Hill Series in water resources and environmental engineering.

Oliver, M. A., and Webster, R., 1990, Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information Systems*, 4, 313- 32.

Decreto-Lei n.º 276/99, de 23 de Julho. *Diário da República*. I SÉRIE-A. Nº 170. 4599-4604

Decreto-Lei n.º 111/2002, de 16 de Abril. Nº 89, *Diário da República*. I SÉRIE-A. 3711-3722.

Rossiter, G. D., 1994. "Uncertainty, Part 5", Course Soil, Crop & Atmospheric Sciences, Cornell University. (URL: http://wwwscas.cit.cornell.edu/landeval/le_notes/s494ch5p.html, consultado em Dezembro de 2008).

Roth, C., Bournel-Bosson, C., 2001. Mapping diffusive sampling results: including uncertainty and indirect information. In *Proceedings of the International Conference Measuring Air Pollutants by Diffusive Sampling, Montpellier, France*, 26-28 September 2001.

Silva, Â., Amorim, A., 2002. Estimação da Temperatura Média do Ar em Portugal Continental Utilização, Comparação e Avaliação de alguns Métodos de Interpolação em SIG. ESIG, 2002. (URL: http://www.igeo.pt/servicos/cdi/biblioteca/Publicacoes/IGP/esig_2002/papers/p046.pdf, consulta em Dezembro de 2008).

Stedman, J.R., Vincent, K.J., Campbell, G.W., Goodwin, J.W.L., Downing, C.E.H., 1997. New high resolution maps of estimated background ambient NO_x and NO₂ concentrations in the UK. *Atmospheric Environment* 31,3591 –3602.

Stedman, J.R., Bush, T.J., Vincent, K.J., 2002. *UK Air Quality Modelling for Annual Reporting 2001 on Ambient Air Quality Assessment under Council Directives 96/62/EC and 1999/30/EC*. AEA Technology, AEAT/ENV/R/1221 Culham.

Tabachnick, B. G., Fidell, L.S., 1996. *Using multivariate statistics*. Third edition. HarperCollins College Publishers.

Tobler, W. R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography*. 46(2): 234-240.

Van Aalst, R., Lynne E., Pulles, T., Emile, S., Tombrou, M., Tonnesen, D., 1998. *Guidance report on preliminary assessment under EC air quality directives. Technical report No 11*. (URL: <http://europa.eu.int/comm/environment/air/cafe/index.htm>, consulta em Dezembro de 2008).

Wagner, E. 1994. Impacts on air pollution in urban areas. *Environmental Management*. 1994;18:759-765.

Webb, N. 2006. *Education 230B Linear Statistical Models*. University of California. (URL: <http://www.gseis.ucla.edu/courses/ed231c/notes2/collin.html>, consulta em Dezembro de 2008).

World Health Organization (WHO), 2004. *Health aspects of air pollution results from the WHO project stematic review of health aspects of air pollution in Europe*. (URL: <http://europa.eu.int/comm/environment/air/cafe/activities/pdf/e83080.pdf>, consulta em Dezembro de 2008).

ANEXOS

1 – Parâmetros de Modelação da Distribuição Espacial

No presente anexo apresentam-se os parâmetros de interpolação relativos à vizinhança de pesquisa para os métodos de interpolação determinística e geoestatística. No caso do *ordinary Kriging* e *ordinary Cokriging* apresenta-se também o mapa do erro padrão fazendo a validação cruzada (neste caso usando todos pontos de amostragem, ou seja, usando os resultados dos tubos de difusão). O erro padrão é uma medida da precisão da média da amostra calculada e obtém-se dividindo o desvio padrão (σ) pela raiz quadrada do tamanho da amostra (n), ou seja:

$$Erro.Padr\tilde{a}o = \frac{\sigma}{\sqrt{n}}$$

O mapa do erro padrão apresentado nas duas figuras, permite visualizar que a incerteza dos mapas produzidos é superior nos limites da área de interpolação e onde há menos pontos de monitorização. No centro dos mapas existe uma maior concentração de pontos de amostragem sendo por isso a incerteza mais baixa.

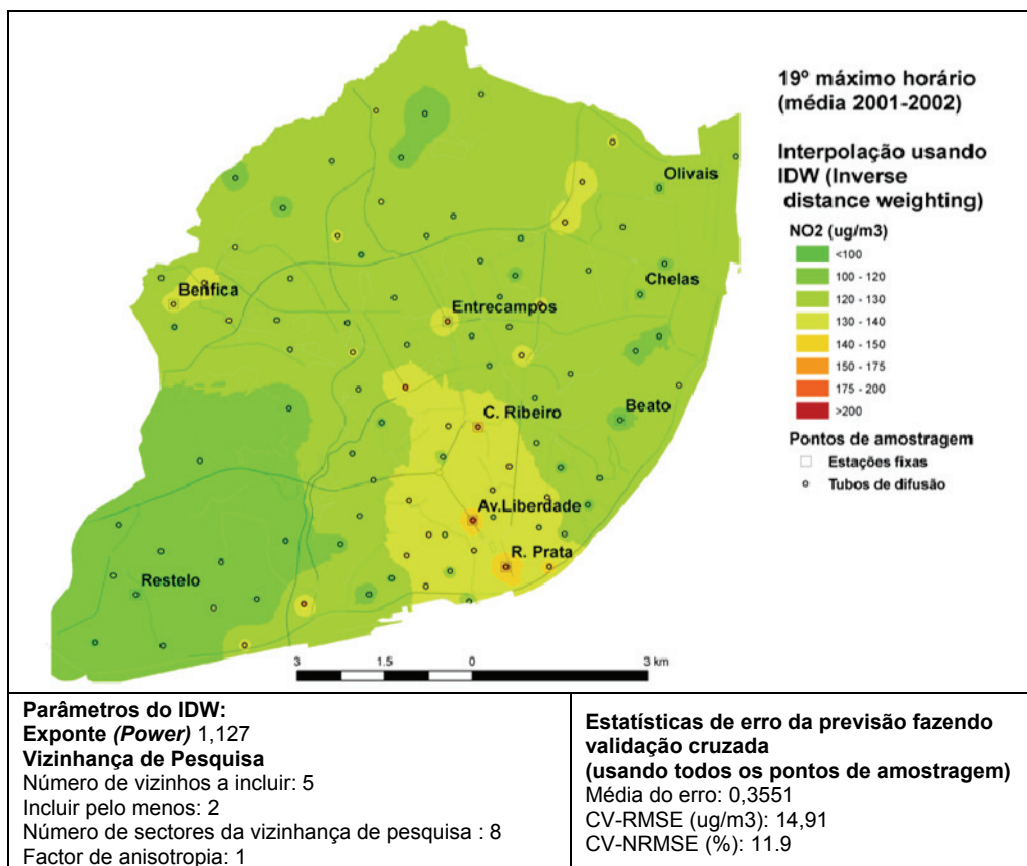


Figura A.1. 1 Interpolação do 19º máximo horário de NO₂ médio de 2001 e 2002 usando o IDW . Parâmetros do método e resultados da validação cruzada incluindo todos os pontos

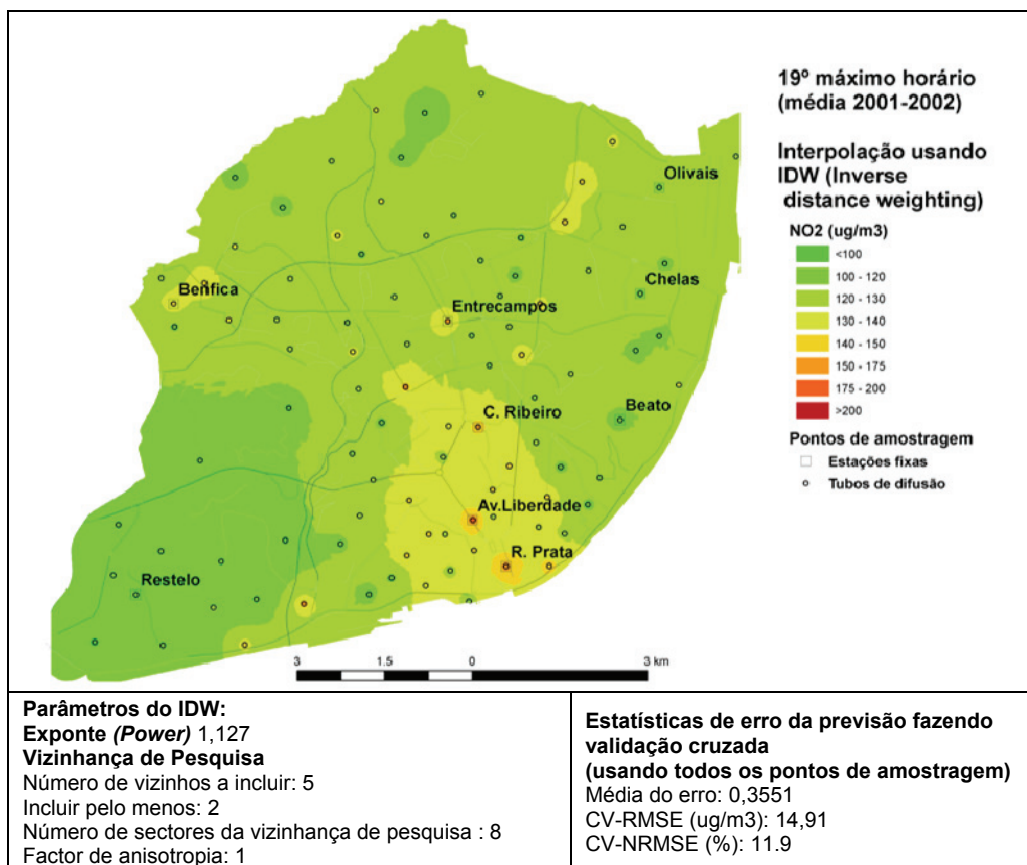


Figura A.1. 2 Interpolação do 19º máximo horário de NO₂ médio de 2001 e 2002 usando o IDW . Parâmetros do método e resultados da validação cruzada incluindo todos os pontos

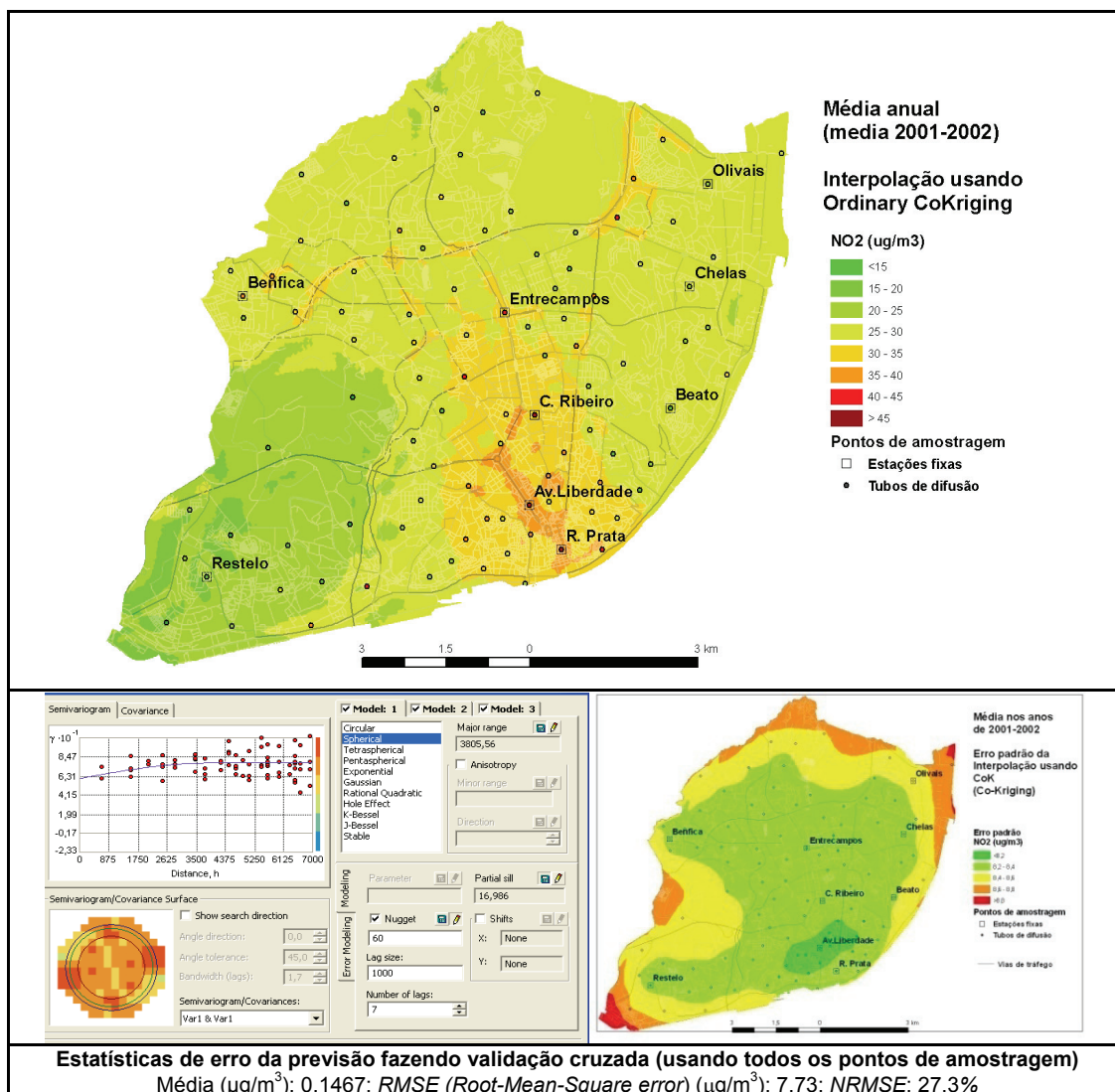


Figura A.1. 3 Interpolação da média anual de NO₂ para a média de 2001 e 2002 usando *cokriging* com as variáveis auxiliares: soma das emissões NO_x devidas ao tráfego rodoviário na envolvente de 50m, entre 50 e 100m e entre 100 e 250m

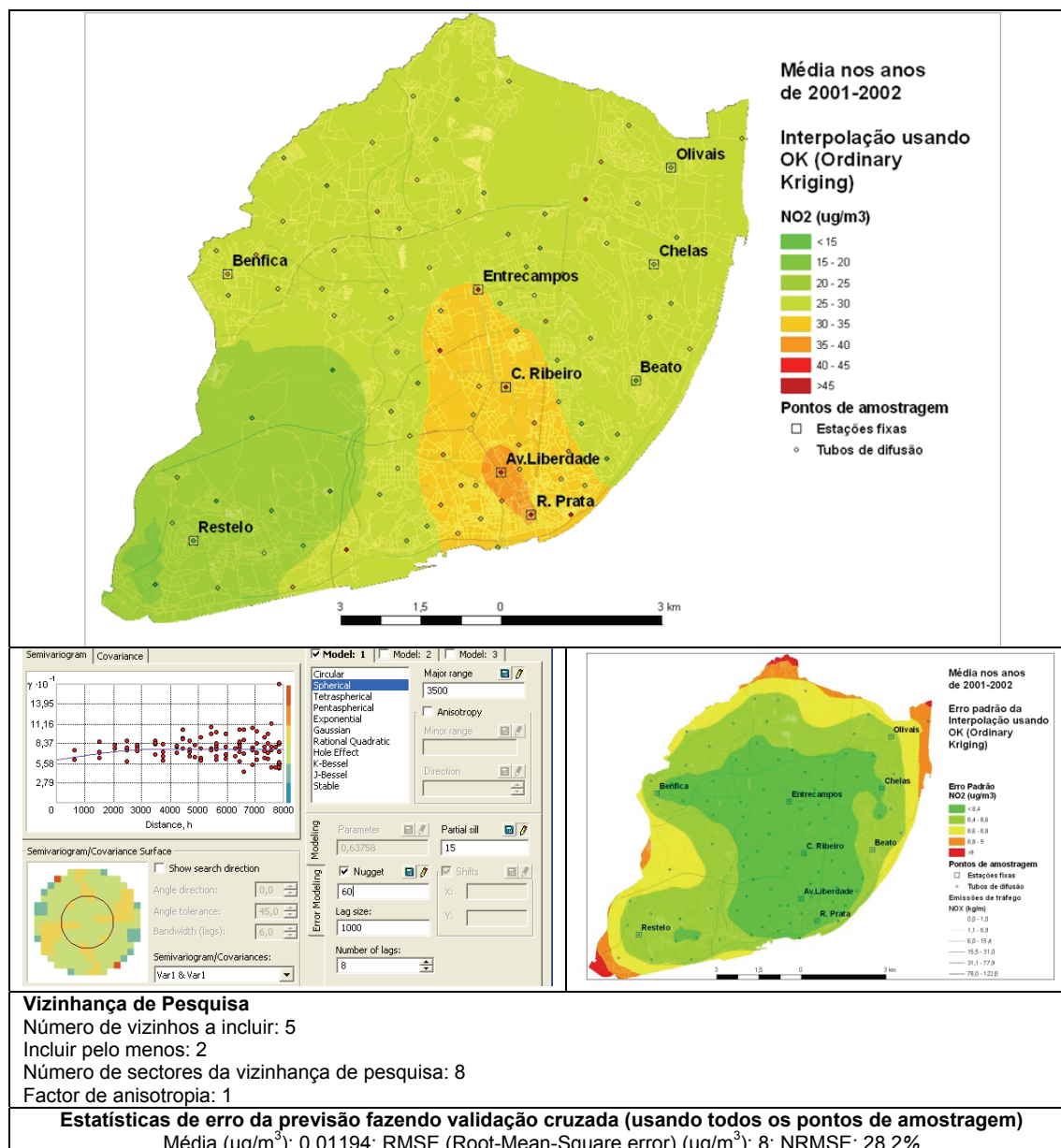


Figura A.1. 4. Interpolação da média de NO₂ durante os anos de 2001 e 2002 usando o método *ordinary kriging*. Variograma, mapa do erro padrão e estatísticas de erro da previsão fazendo validação cruzada.

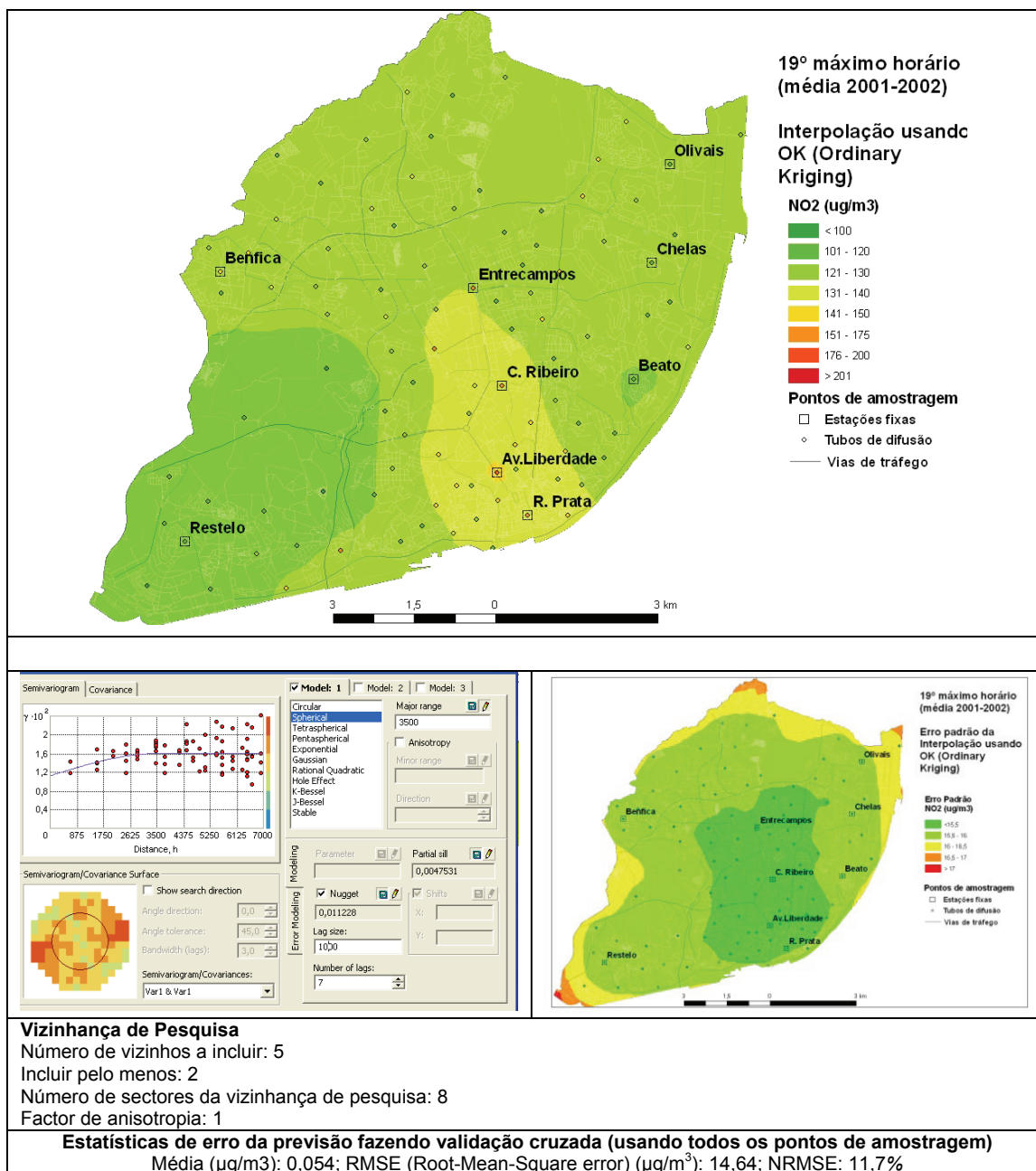


Figura A.1. 5 Interpolação do 19º máximo horário de NO₂ médio de 2001 e 2002 usando *ordinary kriging*. Variograma, mapa do erro padrão e estatísticas de erro da previsão fazendo validação cruzada

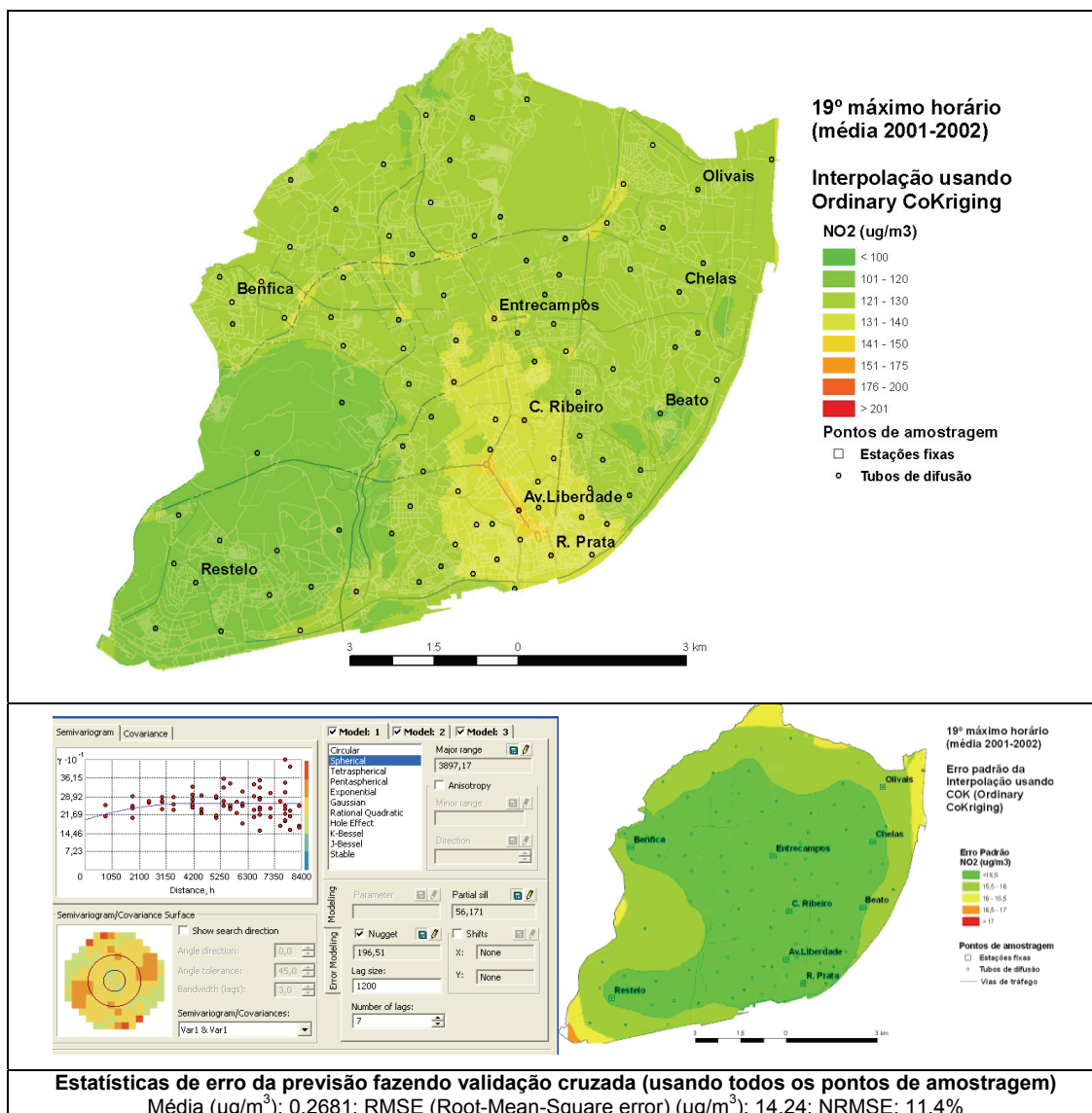


Figura A.1. 6 Interpolação do 19º máximo horário de NO₂ médio de 2001 e 2002 usando *ordinary cokriging* com as variáveis auxiliares: soma das emissões derivadas do NO_x do tráfego na envolvente de 50m, entre 50 e 100m e entre 100 e 250m