

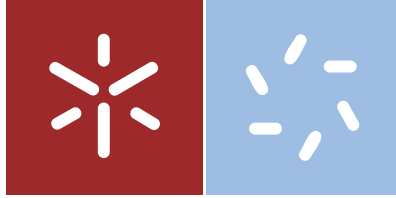


Universidade do Minho
Escola de Ciências

Gustavo Domingos da Costa Coelho Soutinho

Estimação do uso e ocupação do
solo ao nível das NUTS 2 e NUTS 3

Gustavo Soutinho
Estimação do uso e ocupação do
solo ao nível das NUTS 2 e NUTS 3



Universidade do Minho
Escola de Ciências

Gustavo Domingos da Costa Coelho Soutinho

Estimação do uso e ocupação do
solo ao nível das NUTS 2 e NUTS 3

Tese de Mestrado
Estatística

Trabalho efectuado sob a orientação de
Professora Doutora Susana Margarida Ferreira Sá
Faria
Professora Doutora Raquel Menezes da Mota Leite

e co-orientação de
Professor Doutor Pedro José Ramos Moreira Campos
Mestre Maria da Conceição da Silva Ferreira

Anexo 3

DECLARAÇÃO

Nome: **Gustavo Domingos da Costa Coelho Soutinho**

Endereço eletrónico: **gustavosoutinho@sapo.pt**

Telemóvel: **916944517**

Número do Bilhete de Identidade: **10879478**

Título da Tese: **“Estimação do uso e ocupação do solo ao nível das NUTS 2 e NUTS 3”**

Orientadores:

Professora Doutora Raquel Menezes da Mota Leite

Professora Doutora Susana Margarida Ferreira Sá Faria

Professor Doutor Pedro José Ramos Moreira Campos

Mestre Maria da Conceição da Silva Ferreira

Ano de conclusão: **2014**

Designação do Mestrado ou do Ramo de Conhecimento do Doutoramento: **Mestrado em Estatística**

Nos exemplares das teses de doutoramento ou de mestrado ou de outros trabalhos entregues para prestação de provas públicas nas universidades ou outros estabelecimentos de ensino, e dos quais é obrigatoriamente enviado um exemplar para depósito legal na Biblioteca Nacional e, pelo menos outro para a biblioteca da universidade respectiva, deve constar uma das seguintes declarações:

1. É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;
2. É AUTORIZADA A REPRODUÇÃO PARCIAL DESTA TESE (indicar, caso tal seja necessário, nº máximo de páginas, ilustrações, gráficos, etc.), APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;
3. DE ACORDO COM A LEGISLAÇÃO EM VIGOR, NÃO É PERMITIDA A REPRODUÇÃO DE QUALQUER PARTE DESTA TESE.

Universidade do Minho, ___/___/_____

Assinatura: _____

Agradecimentos

Dedico este trabalho ao meu filho Luís Afonso e à minha família pelo carinho, apoio e incentivo que sempre me deram ao longo destes meses.

Um agradecimento muito especial às professoras Susana Faria e Raquel Menezes e aos meus orientadores do Instituto Nacional de Estatística, Pedro Campos e Conceição Ferreira, pelos conhecimentos científicos, disponibilidade e o humanismo revelados na realização deste trabalho.

Agradeço igualmente ao Instituto Nacional de Estatística pela possibilidade que me deu em realizar um estágio pedagógico que considero ter sido muito importante para o meu enriquecimento em termos estatísticos e profissionais.

Por fim, gostava também de relembrar os meus colegas de mestrado com quem partilhei momentos de amizade, confraternização e de aprendizagem que jamais esquecerei.

A todos, o meu mais profundo agradecimento.

Resumo

Land Use and Cover Area Frame Statistical Survey (LUCAS) é um projeto de amostragem areolar em duas fases (*doubling area frame sampling*) que tem sido implementado pelo Eurostat com o intuito de uniformizar a multiplicidade de sistemas de classificação existentes nos países membros da União Europeia.

De modo a aferir a viabilidade do LUCAS se tornar o método de amostragem espacial preferencial ao nível do espaço europeu para a obtenção de estatísticas, os países participantes no LUCAS (incluindo Portugal) desenvolveram o estudo LUCAS piloto (“sinergias entre o LUCAS e Sistemas de classificação nacionais”) no âmbito do qual foi realizado este trabalho. Para além da obtenção de estimativas para as áreas de uso e ocupação do solo ao nível das NUTS 1, NUTS 2 e NUTS 3, foram quantificadas as diferenças entre as estimativas obtidas através do LUCAS e as áreas provenientes de sistemas de classificação nacionais.

Pelos resultados obtidos, ao nível 1 da nomenclatura de classificação do LUCAS por NUTS 2, observa-se que para a generalidade das classes de uso e ocupação do solo as estimativas das áreas apresentam precisões abaixo dos valores máximos estipulados pelo Eurostat. Para os níveis 2 e 3 da nomenclatura de classificação, observou-se uma diminuição da precisão das estimativas devido ao menor número de pontos pertencentes a cada classe. No que se refere às NUTS 3, excetuando algumas classes, o esquema de amostragem usado no LUCAS evidenciou fornecer estimativas com precisões baixas.

Relativamente às diferenças entre os diferentes sistemas de classificação, os resultados obtidos confirmaram, ao nível das NUTS 2 e NUTS 3, diferenças significativas. As diferenças observadas podem dever-se a alguns dos seguintes factores: dificuldades na uniformização das nomenclaturas usadas nos sistemas de classificação nacionais comparativamente com a proposta no projeto LUCAS; os sistemas de classificação basearem-se em abordagens de amostragem espacial muito diferentes; o facto de haver um iato temporal entre as fontes de dados utilizadas.

Abstract

Land Use and Cover Area Frame Statistical Survey (LUCAS) is a doubling area frame sampling project conducted by Eurostat in order to uniform a large number of national data sources with different land cover and land use classification systems existing through European countries.

With the propose to assess whether LUCAS may become in European zone the preferable spatial sampling method for obtaining estimates, countries which belong to the project carried out a pilot study (“synergies among LUCAS and National data sources”) under which this work was conducted. In addition to obtaining land use and land cover areas estimates at NUTS 1, NUTS 2 and NUTS3, the differences between LUCAS estimation and NDS (National Data Sources) estimation were computed.

By computing the results at level 1 of LUCAS nomenclature and at NUTS 2, it is observed that the major part of land cover and land use classes reveal levels of precision which stand under the maximum values defined by Eurostat at NUTS 2. As it was expected, at NUTS level 2 and 3 when compared to level 1, precisions decrease due to the small number of points belonging to each class. With regards to NUTS 3, apart from some classes, the number of points selected in the sampling scheme used in LUCAS seems not to be appropriated due to high coefficients of variation.

Concerning the differences between LUCAS and NDS estimates, at level NUTS 2 and NUTS 3, considerable differences exist which may due to: the difficultty to uniform the nomenclature used in NDS when compared to LUCAS nomenclature; classification systems based on different spatial sampling approaches; the temporal gap among the data sources used.

Abreviaturas

AFN – Autoridade Florestal Nacional
COS – Carta de Ocupação do Solo
CV – Coeficiente de Variação
DG AGRI - Diretório Geral para o Desenvolvimento Agrícola e Rural
DGT – Direção Geral do Território
EC – Comissão Europeia
EEA - Agência Ambiental Europeia
UE – União Europeia
EUNIS - *European University Information Systems*
FAO – *Food and Agriculture Organization*
GEE – Gases com Efeito Estufa
GRTS - *Generalized Random-Tessellation Stratified*
HRL - *High Resolution Layers*
IFN – Inventário de Florestas Nacionais
INE – Instituto Nacional de Estatística
LC – *Land Cover* (ocupação do solo)
LU – *Land Use* (uso do solo)
LUCAS- *Land Use/Cover Area Frame Sampling*
LUCC – *Land Use and land Cover Classes*
NDS – *National Data Sources*
NRCS - *Natural Resources Conservation Service*
NRI - *National Resources Inventory*
NUTS – Nomenclatura das Unidades Territoriais para Fins Estatísticos
 p_{eso_f2} - Peso da segunda fase de estratificação
PSU – Unidades de Amostragem Principais
SEI - *Shannon Evenness Index*
SIG – Sistemas de Informação Geográfica
SSU – Unidades de Amostragem Secundárias
TIG – Tecnologias de Informação Geográfica
 w_h – Peso da primeira fase de estratificação

Índice

Índice	ix
Lista de figuras	xiii
Lista de tabelas.....	xv
Capítulo 1: Introdução.....	1
1.1. Enquadramento	1
1.2. Motivação	1
1.3. Objetivos	2
1.4. Métodos utilizados	3
1.5. Estrutura do trabalho	3
Capítulo 2: Generalidades sobre amostragem e estimação espacial	5
2.1. Unidades amostrais	5
2.2. Correlação espacial	6
2.3. Heterogeneidade espacial	7
2.4. Métodos de estimação.....	8
2.4.1. Métodos <i>design-based</i>	8
2.4.2. Métodos <i>model-based</i>	9
2.5. Esquemas de amostragem	10
2.5.1. Amostragem aleatória simples.....	11
2.5.2. Amostragem aleatória estratificada.....	11
2.5.3. Amostragem aleatória por <i>clusters</i> (grupos)	12
2.5.4. Amostragem aleatória sistemática	13
2.5.5. Amostragem aleatória multietápica.....	14
2.5.6. Amostragem aleatória multifásica	15
2.5.7. Amostragem espacialmente balanceada (<i>spatially-balanced sampling</i>).....	16
2.6. Estimação.....	17
2.6.1. Conceitos fundamentais	17
2.6.2. Métodos de estimação.....	19
2.7. Estimadores relacionados com os esquemas de amostragem	20
2.7.1. Amostragem aleatória simples.....	20
2.7.2. Amostragem aleatória estratificada.....	21
2.7.3. Amostragem aleatória por <i>clusters</i> (grupos)	23

2.7.4.	Amostragem aleatória sistemática	24
2.7.5.	Amostragem aleatória em duas etapas.....	24
2.7.6.	Amostragem aleatória em duas fases (por estratificação).....	25
2.8.	Projetos de amostragem areolar sobre uso e ocupação do Solo	26
2.8.1.	Inventário de Florestas Nacional (IFN).....	26
2.8.2.	Projeto Landyn	28
2.8.3.	Projeto TERUTI	28
2.8.4.	Projeto <i>Countryside Survey</i>	29
2.8.5.	Projeto National Resources Inventory (NRI)	29
Capítulo 3: Land Use and Cover Area Frame Statistical Survey (LUCAS)		31
3.1.	Objetivos do projeto LUCAS	32
3.2.	Sistema de classificação do projeto LUCAS	33
3.2.1.	Classes de ocupação do solo	33
3.2.2.	Classes de uso do solo.....	34
3.2.3.	Comparabilidade da nomenclatura de classificação do projeto LUCAS com outros sistemas de classificação.....	35
3.3.	Esquema de amostragem implementado no LUCAS	36
3.3.1.	Primeira fase do esquema de amostragem	37
3.3.2.	Segunda fase do esquema de amostragem	38
3.4.	Metodologia de estimação utilizada	40
3.4.1.	Descrição dos tipos de pesos referentes à metodologia de estimação	41
3.4.2.	Cálculo dos pesos <i>wh</i> e <i>peso_f2</i> , por NUTS 1 e NUTS 3	47
3.4.3.	Estimadores para a proporção e área total	48
3.4.4.	Estimador para a variância da proporção e área total.....	49
Capítulo 4: Resultados obtidos		51
4.1.	Análise exploratória da base de dados LUCAS <i>survey</i>	51
4.1.1.	Tipo de estrato (variável STRATA)	53
4.1.1.1.	Distribuição do tipo de estrato na amostra da primeira fase	53
4.1.1.2.	Distribuição do tipo de estrato na amostra da segunda fase.....	55
4.1.2.	Classes de ocupação do solo (variáveis LC1 e LC2)	57
4.1.2.1.	Distribuição do tipo de classe de ocupação do solo por NUTS 1, 2 e 3.....	57
4.1.2.2.	Distribuição de pontos com duplos registos para as variáveis LC1 e LC2	61
4.1.2.3.	Alteração do tipo de classe de ocupação do solo entre 2009 e 2012.....	62
4.1.3.	Classes de uso do solo (variáveis LU1 e LU2).....	63
4.1.3.1.	Distribuição do tipo de classe de uso do solo por NUTS 1, NUTS 2 e NUTS 3.....	63

4.1.4.	Raio atribuído ao ponto (variável <i>OBS_RADIUS</i>)	64
4.1.4.1.	Distribuição dos pontos por tipo de raio observado (NUTS 2 e NUTS 3)	65
4.1.4.2.	Distribuição dos pontos por tipo de raio (classes de ocupação do solonível 1)... 66	
4.1.5.	Variável <i>AREA_SIZE</i> (dimensão da área em torno do ponto).....	67
4.1.5.1.	Distribuição dos pontos por tipo de <i>area size</i> (ao nível das NUTS 2).....	67
4.1.5.2.	Distribuição dos pontos por tipo de <i>area size</i> (Nível 1 das classes de ocupação) 68	
4.2.	Estimativas para as classes de ocupação do solo	69
4.2.1.	Avaliação das estimativas para a NUTS 1	69
4.2.1.1.	Nível 1 (da nomenclatura de classificação LUCAS).....	69
4.2.1.2.	Nível 2 (da nomenclatura de classificação LUCAS).....	70
4.2.1.3.	Nível 3 (da nomenclatura de classificação LUCAS).....	70
4.2.2.	Avaliação das estimativas para a NUTS 2	72
4.2.2.1.	Nível 1 (da nomenclatura de classificação LUCAS).....	72
4.2.2.2.	Nível 2 e 3 (da nomenclatura de classificação LUCAS)	73
4.2.3.	Avaliação das estimativas para a NUTS 3	74
4.2.3.1.	Nível 1 (da nomenclatura de classificação LUCAS).....	74
4.2.3.2.	Para o nível 2 e 3 (da nomenclatura de classificação LUCAS)	74
4.3.	Comparação de estimativas para as classes de ocupação do solo	76
4.3.1.	Nível 1 da nomenclatura de classificação LUCAS	76
4.3.1.1.	Comparação de resultados por NUTS 1 (Portugal Continental).....	76
4.3.1.2.	Comparação de resultados por NUTS 2.....	77
4.3.1.3.	Comparação de resultados por NUTS 3.....	77
4.3.2.	Nível 2 e 3 da nomenclatura de classificação LUCAS	79
4.3.2.1.	Comparação de resultados por NUTS 1 (Portugal Continental).....	79
4.3.2.2.	Comparação de resultados por NUTS 2.....	80
4.3.2.3.	Comparação de resultados por NUTS 3.....	81
4.4.	Estimativas para as classes de uso do solo.....	82
4.5.	Comparação de estimativas para as áreas de uso do solo	85
Capítulo 5:	Conclusões	89
5.1.	Avaliação das Estimativas obtidas através da metodologia de estimação	90
5.2.	Comparação de resultados com outras fontes nacionais de classificação de dados (NDS)	92
5.3.	Sugestões para trabalho futuro.....	93
Capítulo 6:	Bibliografia.....	95

Anexos.....	99
Anexo 1: Descrição das NUTS 1, NUTS 2 e NUTS 3 ao nível do território de Portugal (utilizadas no projeto LUCAS).....	99
Anexo 2: Nomenclatura do IFN6 (Fonte: AFN, 2012).....	100
Anexo 3: Nomenclatura utilizada no projeto <i>Landyn</i> (Fonte: DGT, 2013).....	101
Anexo 4: Nomenclatura do sistema de classificação do projeto LUCAS 2012	102
Anexo 5: Precisão (Limites máximos de erro esperados por classe de ocupação do solo por NUTS definidos pelo Eurostat)	108
Anexo 6: Estimativas da área (km ²) e respetivos CV, ao nível 2 e 3 das classes de ocupação do solo, por NUTS 2, para os dados recolhidos em 2009	109
Anexo 7: Estimativas da área (km ²) e respetivos CV, ao nível 2 e 3 das classes de ocupação do solo, por NUTS 2, para os dados recolhidos em 2012	110
Anexo 8: Estimativas da área (km ²) e respetivos CV, ao nível 1 das classes de ocupação do solo, por NUTS 3, para os dados recolhidos no LUCAS de 2009.....	111
Anexo 9: Estimativas da área (km ²) e respetivos CV, ao nível 1 das classes de ocupação do solo, por NUTS 3, para os dados recolhidos no LUCAS de 2012.....	112
Anexo 10: Estimativas da área (km ²) para o nível 1 das classes de ocupação do solo, por NUTS 2, para os dados recolhidos no LUCAS de 2009 e 2012 e resultados obtidos de “outras fontes” e resultados exclusivos da COS 2007.....	113
Anexo 11: Diferenças entre os resultados estimados com base nos dados LUCAS de 2009 e 2012 com os resultados obtidos de outras fontes de classificação (NDS) e resultados exclusivos da COS 2007 (por NUTS 2 ao nível 1 da nomenclatura de classificação LUCAS)	113
Anexo 12: Diferenças em percentagem entre as estimativas LUCAS 2009 e os resultados provenientes de NDS e COS 2007	114
Anexo 13: Diferenças em percentagem entre as estimativas LUCAS 2012 e os resultados provenientes de NDS e COS 2007	115
Anexo 14: Comparação das Estimativas para as classes de uso do solo U110, U120, U130, U340 e U612, dadas através dos dados LUCAS 2009, ao nível das NUTS 3, com os resultados obtidos de outras fontes de dados	116
Anexo 15: Comparação das Estimativas para as classes de uso do solo U110, U120, U130, U340 e U612, dadas através dos dados LUCAS 2012, ao nível das NUTS 3, com os resultados obtidos de outras fontes de dados	117

Lista de figuras

2.1: Semelhanças e diferenças entre esquemas de amostragem (adaptado de Gruijter <i>et al.</i> , 2006).....	10
2.2: Exemplo da implementação de um esquema de amostragem em duas fases. Os símbolos sombreados representam a amostra da segunda fase (retirado de Gruijter <i>et al.</i> , 2006).....	15
2.3: Esquema de amostragem espacialmente balanceado segundo o algoritmo GRTS (retirado de Stevens e Olsen, 2004).....	17
2.4: Estrutura da grelha nacional de amostragem (ICNF, 2012).....	27
2.5: Esquema simplificado dos processos de recolha e processamento de dado do IFN de 2010 (retirado de IFN, 2012).....	27
3.1: Esquema de amostragem em duas fases para 2012	37
3.2: Subgrelhas obtidas a partir da grelha da primeira fase	38
3.3: Indicação das localizações de acordo com as distâncias a um ponto localizado na extremidade à esquerda e no meio do lado esquerdo (adaptado de Brus <i>et al.</i> , 2011)	39
3.4: Probabilidades de inclusão estimadas usando o algoritmo implementado	40

Lista de tabelas

3.1: Classes de ocupação do solo (níveis 1, 2 e 3 do sistema de classificação LUCAS)	34
3.2: Classes de Uso do Solo (níveis 1 e 2 do sistema de classificação LUCAS).....	35
3.3: Sete estratos da primeira fase do esquema de amostragem LUCAS.....	37
3.4: Pesos da segunda fase, $peso_f2$, atribuído ao ponto, de acordo o subconjunto do esquema de amostragem LUCAS a que pertença (Martino e Palmieri, 2009).....	44
3.5: Número de pontos para cada subconjunto $Sh, 1$ e $Sh, 2$ (para Portugal em 2009)....	45
3.6: Número de pontos para cada subconjunto $Sh, 1$ e $Sh, 2$ (para Portugal em 2012)	46
3.7: Número de pontos $Nh, 21$ e $Nh, 22$ para cada domínio (para Portugal, em 2012)....	46
3.8: Pesos w_h e $peso_f2$ ao nível das NUTS 3 (2009).....	48
3.9: Pesos w_h e $peso_f2$ ao nível das NUTS 3 (2012).....	48
4.1: Descrição das variáveis da base de dados LUCASsurvey envolvidas no processo de estimação e exemplo dos valores para o ponto com ID=26601762	52
4.2: Distribuição de pontos por domínio (estrato/NUTS2) na primeira fase do esquema do esquema de amostragem LUCAS para Portugal.....	53
4.3: Distribuição de pontos por domínio (estrato/NUTS 3) na primeira fase do esquema do esquema de amostragem LUCAS para Portugal continental (em 2009 e 2012).54	
4.4: Distribuição de pontos por domínio (estrato/NUTS 2) na segunda fase do esquema do esquema de amostragem LUCAS para Portugal continental (em 2009).....	55
4.5: Distribuição de pontos por domínio (estrato/NUTS 2) na segunda fase do esquema do esquema de amostragem LUCAS para Portugal continental (em 2012).....	55
4.6: Distribuição de pontos por domínio (estrato/NUTS 3) na segunda fase do esquema do esquema de amostragem LUCAS para Portugal continental.....	56
4.7: Distribuição de pontos, por tipo de classe de ocupação do solo ao nível 1 da nomenclatura de classificação LUCAS por NUTS 1 e NUTS 2	58
4.8: Distribuição dos pontos da amostra da segunda fase para o nível 3 das classes de ocupação do solo por NUTS 1 e NUTS 2.....	59
4.9: Distribuição dos pontos da amostra da segunda fase para o nível 3 das classes de ocupação do solo por NUTS 3 (em 2012)	60
4.10: Distribuição de pontos em 2009 com duplos registos na segunda fase do esquema de amostragem para as classes de ocupação do solo (ao nível da nomenclatura LUCAS)	61

4.11: Distribuição de pontos em 2012 com duplos registos na segunda fase do esquema de amostragem para as classes de ocupação do solo (ao nível da nomenclatura LUCAS)	61
4.12: Descrição do número de pontos cuja classe de ocupação do solo em 2012 foi diferente da atribuída em 2009 (ao nível 1 da nomenclatura do sistema de classificação LUCAS).....	62
4.13: Distribuição dos pontos da amostra da segunda fase para as classes U110, U120, U130, U340 e U612 por NUTS 1 e 2 (em 2009 e 2012)	64
4.14: Distribuição dos pontos da amostra da segunda fase para as classes U110, U120, U130, U340 e U612 por NUTS 3 (em 2009 e 2012).....	64
4.15: Descrição das categorias da variável <i>AREA_SIZE</i>	67
4.16: Estimativas da área (km ²) e respetivos CV, ao nível 1 das classes de ocupação do solo, para os dados LUCAS recolhidos para Portugal continental (PT1).....	70
4.17: Estimativas da área (km ²) e respetivos CV, ao nível 2 das classes de ocupação do solo, para os dados LUCAS recolhidos para Portugal continental (PT1).....	71
4.18: Estimativas da área (km ²) e respetivos CV, ao nível 3 das classes de ocupação do solo, para os dados LUCAS recolhidos para Portugal continental (PT1).....	71
4.19: Estimativas da área (km ²) e respetivos CV, ao nível 1 das classes de ocupação do solo (NUTS 2)	72
4.20: Comparação das medidas de localização para os valores CV ao nível 1 das classes de ocupação do solo, por NUTS 2, para os dados recolhidos nos LUCAS de 2009 e 2012.....	73
4.21: Comparação das medidas de localização para os valores CV ao nível 1 das classes de ocupação do solo, por NUTS 3, para os dados recolhidos nos LUCAS de 2009 e 2012.....	74
4.22: Estimativas da área (km ²) e respetivos CV, para os níveis 2 e 3 das classes de ocupação do solo, por NUTS 3, cujos CV estão dentro dos limites máximos de precisão (em 2009)	75
4.23: Estimativas da área (km ²) e respetivos CV, para os níveis 2 e 3 das classes de ocupação do solo, por NUTS 3, cujos CV estão dentro dos limites máximos de precisão (em 2012).....	75
4.24: Diferenças entre os resultados estimados com base nos dados LUCAS de 2009 e 2012 com os resultados obtidos de outras fontes de classificação (NDS) e resultados exclusivos da COS 2007 (por NUTS 1 ao nível 1 da nomenclatura de classificação LUCAS).....	76
4.25: Diferenças entre os resultados estimados com base nos dados LUCAS de 2009 com os resultados obtidos de outras fontes de classificação (NDS), por NUTS 3 ao	

nível 1 da nomenclatura de classificação LUCAS (com diferenças inferiores a 10 % e dentro do limites de qualidade de precisão)	78
4.26:Diferenças entre os resultados estimados com base nos dados LUCAS de 2012 com os resultados obtidos de outras fontes de classificação (NDS), por NUTS 3 ao nível 1 da nomenclatura de classificação LUCAS (com diferenças inferiores a 10 % e dentro do limites de qualidade de precisão)	78
4.27:Diferenças entre os resultados estimados com base nos dados LUCAS de 2012 com os resultados obtidos de outras fontes de classificação (NDS), por NUTS 3 ao nível 1 da nomenclatura de classificação LUCAS	79
4.28:Medidas de localização das diferenças (em %) entre dados LUCAS, por nível 2 e 3 da nomenclatura do projeto LUCAS, com os resultados obtidos através de outras fontes de classificação nacionais para Portugal Continental (PT1)	80
4.29:Diferenças entre os resultados estimados com base nos dados LUCAS de 2012 com os resultados obtidos de outras fontes de classificação (NDS), por NUTS 2 para os níveis 2 e 3 da nomenclatura de classificação LUCAS (com diferenças inferiores a 20 % e dentro do limites de qualidade de precisão).....	80
4.30:Medidas de localização das diferenças (em %) entre dados LUCAS, por nível 2 e 3 das classes de ocupação do solo, para NUTS 2, com os resultados obtidos de outras fontes de classificação nacionais	81
4.31:Diferenças entre os resultados estimados com base nos dados LUCAS de 2012 com os resultados obtidos de outras fontes de classificação (NDS), por NUTS 3 para os níveis 2 e 3 da nomenclatura de classificação LUCAS (com diferenças inferiores a 20 % e dentro do limites de qualidade de precisão).....	81
4.32:Medidas de Localização das diferenças (em %) entre dados LUCAS, por nível 2 e 3 das classes de ocupação do solo, para NUTS 3, com os resultados obtidos de outras fontes de classificação nacionais	82
4.33: Estimativas da área (km ²) e respetivos CV, para as classes de uso do solo para os dados LUCAS recolhidos em 2009 e 2012 para Portugal continental (PT1).....	83
4.34: Estimativas da área (km ²) e respetivos CV, para as classes de uso do solo para os dados LUCAS recolhidos ao nível das NUTS 2 em 2009 e 2012	83
4.35:Medidas de localização para os valores CV das classes de uso do solo, por NUTS 2, para os dados recolhidos nos LUCAS de 2009.....	84
4.36: Medidas de localização para os valores CV das classes de uso do solo, por NUTS 2, para os dados recolhidos nos LUCAS de 2012	84
4.37:Medidas de localização para os valores dos CV das estimativas LUCAS 2009 para uso do solo, ao nível das NUTS 3.....	85

4.38:Medidas de localização para os valores dos CV das estimativas LUCAS 2012 parauso do Solo, ao nível das NUTS 3	85
4.39: Diferenças entre os resultados estimados através dos dados LUCAS 2009 e 2012, para as classes LU, para Portugal continental, com os resultados obtidos de outras fontes de dados nacionais (NDS).....	86
4.40: Comparação de resultados ao nível das NUTS 2 entre os dados LUCAS 2009 e os dados provenientes de outras fontes de dados nacionais (NDS)	87
4.41:Comparação de resultados ao nível das NUTS 2 entre os dados LUCAS 2012 e os dados provenientes de outras fontes de dados nacionais (NDS)	87
4.42: Medidas de localização das diferenças (em %) entre dados LUCAS 2009, por Uso do Solo, ao nível das NUTS 2, com os resultados obtidos de outras fontes de classificação nacionais.....	88
4.43: Medidas de localização das diferenças (em %) entre dados LUCAS 2012, por Uso do Solo, ao nível das NUTS 2, com os resultados obtidos de outras fontes de Classificação nacionais.....	88

Capítulo 1

Introdução

1.1. Enquadramento

A utilização de dados estatísticos associados ao uso e ocupação do solo é um tema de enorme atualidade ao nível do espaço europeu influenciando decisivamente as políticas comunitárias em áreas como a proteção da natureza; a gestão de recursos hídricos e florestais; o planeamento urbano e de transportes; as políticas agrícolas; a prevenção de riscos naturais; a biodiversidade; a monitorização de alterações climáticas, entre outras.

Devido à importância do tema e de forma a responder às solicitações por parte de diferentes organismos europeus, tais como a Agência Ambiental Europeia (EEA) eo Diretório Geral para o Desenvolvimento Agrícola e Rural (DG AGRI), o Eurostat tem empreendido uma estratégia de recolha, produção e disseminação uniformizada de estatísticas sobre o uso e a ocupação do solo em cooperação com os diferentes estados membros e organismos internacionais.

De entre as medidas já realizadas por parte do Eurostat, há a destacar o projeto LUCAS (*Land Use/Cover Area Frame Sampling*), que tem como principal objetivo a produção regular de informação estatística harmonizada ao nível dos países da União Europeia (UE), tendo por base uma clara distinção entre os conceitos de ocupação e uso do solo.

Com este projeto, pretende-se uniformizar a multiplicidade de sistemas de classificação existentes nos países membros da UE, através da nomenclatura definida no projeto LUCAS de modo a melhorar a comparabilidade entre as estimativas produzidas ao nível do espaço europeu.

1.2. Motivação

Após a realização dos inquéritos do projeto LUCAS em 2009 e 2012, o Eurostat empreendeu o desenvolvimento de um estudo LUCAS piloto pelos estados

membros participantes, tendo por finalidade o estabelecimento de sinergias entre o LUCAS e outros sistemas de classificação nacionais (NDS - *National Data Sources*), de acordo com alguns dos seguintes propósitos:

- Examinar e comparar resultados provenientes de outros sistemas de classificação nacionais com as estimativas obtidas através dos dados recolhidos no LUCAS;
- Estabelecer metodologias e soluções técnicas para estimar áreas totais para as classes de uso e ocupação do solo da nomenclatura de classificação do LUCAS;
- Obter as áreas totais para os diferentes níveis da nomenclatura LUCAS para NUTS 1, NUTS 2 e NUTS 3;
- Avaliar a qualidade das estimativas em termos de precisão;
- Avaliar os benefícios e os obstáculos da integração de esquemas de amostragem areolares nacionais com o projeto LUCAS.

Este trabalho de investigação decorreu no contexto de um estágio pedagógico, realizado no Instituto Nacional Estatística (INE), entre janeiro e julho de 2014, integrado na fase IV do referido estudo piloto para a obtenção de estatísticas harmonizadas sobre o uso e ocupação do solo. As atividades realizadas são apresentadas no subcapítulo seguinte.

1.3. Objetivos

Os objetivos propostos para o trabalho de investigação foram os seguintes:

- Analisar a metodologia de amostragem usada no projeto LUCAS;
- Implementar uma metodologia de estimação envolvendo amostragem areolar através do desenvolvimento de programas escritos numa linguagem de programação adequada (R);
- Proceder ao cálculo das estimativas para classes de uso e ocupação do solo
- Obter as áreas das classes de uso e ocupação do solo para os níveis 1, 2 e 3 da nomenclatura do sistema de classificação do projeto LUCAS para o território português (por NUTS 1, NUTS 2 e NUTS 3) a partir dos dados obtidos nos inquéritos de 2009 e 2012;

- Avaliar a qualidade das estimativas produzidas, tendo por base os valores dos coeficientes de variação obtidos e os limites definidos pelo Eurostat por NUTS 2 apresentados no Anexo 5;
- Calcular desvios em percentagem entre as estimativas obtidas usando a metodologia de estimação implementada no LUCAS e os resultados provenientes de sistemas de classificação nacionais, denominados neste estudo também por fontes nacionais (NDS – National Data Sources);
- Avaliar as classes de uso e ocupação do solo da nomenclatura do projeto LUCAS que apresentam maior concordância com as classes provenientes de outras fontes de classificação nacionais.

1.4. Métodos utilizados

Para a obtenção das estimativas LUCAS foi utilizado o estimador proposto por Martino e Palmieiri (2009) para esquemas de amostragem areolar composto por duas fases (com estratificação), que tem sido implementado, desde 2006, pelo Eurostat, com intuito de reduzir a correlação espacial dos dados amostrais.

Para a obtenção dos pesos associados aos estimadores para as classes de uso e ocupação do solo para os dados LUCAS recolhidos em Portugal, em 2009 e 2012, foi utilizada a metodologia de estimação descrita por Martino e Palmieri (2009) que se caracteriza por uma divisão dos pontos selecionados na amostra em subconjuntos como é apresentado no subcapítulo 3.4.

Os resultados provenientes de outros sistemas de classificação nacionais foram recolhidos durante as primeiras três fases do estudo piloto, tendo por base dados recolhidos a partir de fontes nacionais como o inventário de florestas nacionais (IFN) ou a carta de ocupação do solo de 2007.

1.5. Estrutura do trabalho

Este trabalho de investigação está estruturado em 5 capítulos e 15 anexos.

Neste capítulo é apresentado um enquadramento do tema de investigação, a importância do mesmo ao nível da União Europeia, as estratégias

implementadas pelo Eurostat e os objetivos do estudo piloto “Sinergias entre o LUCAS e sistemas de classificação nacionais”.

No capítulo 2 é efetuado um enquadramento teórico sobre os conceitos relacionadas com amostragem espacial e em particular com as abordagens *design-based* e *model-based*. Tendo em consideração os objetivos do trabalho, no âmbito do projeto LUCAS, procedeu-se a uma análise sobre diferentes tipos esquemas de amostragem areolar com implicações diretas na metodologia LUCAS. Por fim, são apresentados projetos nacionais e internacionais de amostragem areolar para uma melhor contextualização do tema.

No capítulo 3 é introduzido o projeto LUCAS, em termos de evolução histórica, objetivos e a nomenclatura de classificação proposta. É igualmente apresentada a metodologia de amostragem e a metodologia de estimação usada neste projeto de investigação.

No capítulo 4, depois de efetuada uma análise exploratória dos dados LUCAS recolhidos nos inquéritos realizados em 2009 e 2012 em Portugal Continental, nas duas fases do esquema de amostragem, são apresentadas as estimativas obtidas pela aplicação da metodologia de estimação no âmbito do estudo LUCAS piloto de produção de estimativas e comparação das mesmas com outras fontes de dados nacionais.

No capítulo 5 são apresentadas as principais conclusões do trabalho e indicadas algumas sugestões para futuras investigações.

Capítulo 2

Generalidades sobre amostragem e estimação espacial

Amostragem Espacial é um tipo de amostragem probabilística na qual a escolha da amostra depende da posição geográfica das localizações.

Segundo Wang *et al.* (2012), o processo de amostragem espacial, envolve as seguintes etapas: (1) clarificação dos objetivos da amostragem: se se pretende estimar um parâmetro populacional ou modelar uma tendência média do parâmetro; (2) definição da população e das unidades amostrais que são constituídas por localizações geográficas que não se sobrepõem; (3) escolha do método de amostragem e quantificação da dimensão da amostra, tendo em consideração o orçamento disponível e o valor desejado para as precisões das estimativas; (4) definição do esquema de amostragem que descreva *onde*, *quando* e *como* são recolhidas as amostras; (5) recolha da amostra e (6) análise quantitativa dos dados espaciais e espaço-temporais, relativamente à tendência global das estimativas, aleatoriedade, erros ou dependência espacial, usando estatística espacial.

Entre os campos de aplicação, segundo Ruiz-Medina (2012), a amostragem espacial divide-se em duas áreas principais: uma relacionada com *domínios físicos* (agricultura, geologia, hidrologia ou ecologia, por exemplo) e outra, envolvendo *domínios socioeconómicos* (epidemiologia ou econometria espacial, por exemplo).

2.1. Unidades amostrais

Em amostragem espacial, as unidades amostrais podem ser de três tipos: pontos, linhas ou segmentos. Pontos e segmentos são os mais utilizados em inquéritos para a obtenção de estatísticas sobre uso e ocupação do solo, enquanto a utilização de linhas é menos frequente.

Embora, teoricamente, um ponto seja uma representação geométrica sem uma área associada, em termos operacionais, é possível atribuir-lhe uma área que está associada a uma classe de uso ou ocupação do solo, como por exemplo florestas ou matos, habitualmente atribuída a uma classe dominante (Schoenmakers, 2005).

Segundo o mesmo autor, a utilização de pontos como unidades amostrais, apesar dos erros que podem ocorrer na sua localização no terreno, permite reduzir os custos de observação comparativamente com o que acontece com os segmentos, em que é necessário delinear e digitalizar os limites das classes de uso ou ocupação do solo.

Os segmentos podem ser definidos de três formas (FAO, 1996): (1) através de barreiras físicas permanentes e perfeitamente identificáveis (como rios ou estradas); (2) por polígonos regulares (representados por grelhas); (3) através de propriedades agrícolas ou divisões administrativas, tais como o sistema de divisão territorial europeu representado pela nomenclatura das unidades territoriais para fins estatísticos (NUTS). A descrição das NUTS para Portugal, por nível 1, 2 e 3, é apresentada no Anexo 1.

Nas unidades amostrais dadas por linhas, procede-se à contagem do número de ocorrências (pontos) pertencentes às classes de uso ou ocupação do solo em estudo. A utilização de linhas é muito frequente nos estudos ambientais e florestais, podendo a direção espacial das mesmas ser predefinida ou obtida aleatoriamente (Gruijter *et al*, 2006).

2.2. Correlação espacial

A existência de dependência espacial entre as unidades amostrais é um aspeto inerente à amostragem espacial em que segundo princípios geográficos, observações mais próximas encontram-se mais correlacionadas do que observações mais afastadas (Sahoo *et al.*, 2006).

A correlação entre os dados amostrais implica, necessariamente, que a amostragem espacial se diferencie, em vários aspetos, da amostragem clássica, que assenta nos pressupostos de que as observações são realizações independentes entre si e que provêm de variáveis aleatórias identicamente

distribuídas, usualmente, de acordo com uma distribuição normal (Chun e Griffith, 2013).

O impacto mais comum da existência de correlação espacial é o da diminuição da eficiência dos estimadores (pelo aumento da variância). Por isso, segundo Sahoo *et al.* (2006) é fundamental na escolha do esquema de amostragem, que se tenha em conta a estrutura espacial dos dados de modo a capturar a correlação espacial e limitar problemas relacionados com propriedades dos estimadores tais como eficiência, enviesamento e consistência.

Para o caso de amostragens espaciais, a modelação da autocorrelação, pode ser feita através de uma função designada por variograma, a partir do qual é possível calcular a variância considerando apenas os pares de localizações (pontos) que se encontram a uma determinada distância e obter padrões espaciais para os dados em estudo, como por exemplo, regiões de variação de temperaturas ou poluição do solo (Barnett, 2004).

Assim, segundo o mesmo autor, a partir do variograma é possível calcular três focos de variação:

- Variância total do processo (*sill*), que corresponde ao valor para o qual o variograma tende à medida que a distância entre as localizações aumenta;
- Distância a partir da qual a correlação entre variáveis é nula (*range*);
- Efeito pepita ou ruído branco (*nugget*).

2.3. Heterogeneidade espacial

Heterogeneidade espacial refere-se a uma situação em que não existe uma variabilidade uniforme associado aos valores da variável em estudo de cada localização da área populacional.

O pressuposto estabelecido em estatística clássica de que os dados têm de ser independentes e identicamente distribuídos, em contextos espaciais, é equivalente a assumir a homogeneidade espacial dos dados. Por outras palavras, significa que a média e a variância da variável são constantes ao longo da área da região em estudo e que os seus valores têm uma distribuição eventualmente representada por uma curva normal (Storch e Zwiers, 1999).

Associada à heterogeneidade espacial podem-se distinguir dois tipos de variabilidade dos dados: a *variância global* (relacionada com a população de dados) e a estrutura espacial dessa variância (autocorrelação espacial da população) (Chun e Griffith, 2013).

De acordo com os mesmos autores, tendo em consideração estes factos, estes dois elementos da variabilidade do processo devem ser reconhecidos e avaliados no esquema de amostragem e na escolha do estimador.

2.4. Métodos de estimação

Segundo Brus e Knotters (2008) há a considerar duas abordagens para a inferência de parâmetros populacionais: métodos baseados no desenho amostral (*design-based*) e métodos baseados num modelo estatístico (*model-based*).

2.4.1. Métodos *design-based*

A abordagem *design-based* utiliza a teoria clássica de amostragem, na qual a população é constituída por um conjunto de pontos (ou localizações) numa região e em que o processo de amostragem consiste na seleção de subconjuntos de unidades amostrais de acordo com uma probabilidade de ser selecionada na amostra, probabilidade de inclusão, que se deseja que seja idêntica para cada elemento amostral (Brus, 2010).

Os métodos de amostragem espacial obtidos através de *design-based* dividem-se em dois grupos: os métodos baseados em listas (*list frame*) e os métodos baseados em áreas (*area frame sampling*) (FAO, 1996).

Nos métodos baseados em listas, as bases de amostragem são compostas por unidades, como por exemplo explorações agrícolas, a partir das quais as unidades de amostragem são selecionadas. O Inventário de Florestas Nacional (IFN) é um exemplo deste tipo de método implementado em Portugal (IFN, 2012).

Nos métodos baseados em áreas ou areolares, o desenho de amostragem é definido habitualmente por uma representação cartográfica do território e por um esquema de amostragem que define como este se divide nas unidades amostrais espaciais que são observadas no local. Portanto, é a partir do

esquema de amostragem utilizado que é definida a probabilidade de inclusão do conjunto de unidades na amostra o que permite procedimentos de inferência válidos sem a necessidade de introduzir pressupostos adicionais (Gallego, 1995).

De acordo com Bettio *et al.* (2002), a utilização de métodos de amostragem areolar, comparativamente com os baseados em listas, permite reduzir o número de erros não amostrais resultante de entrevistas; fornece uma informação mais detalhada e atualizada sobre o tipo de classe de uso e ocupação do solo, assim como facilita o processo de monitorização das alterações ocorridas no solo, através de recolhas sazonais dos dados.

Segundo Gallego (2007), apesar das melhorias verificadas nos últimos anos na qualidade de resolução ao nível de sistemas de deteção remota, os métodos areolares apresentam estimativas mais precisas do que estes, em particular para situações de heterogeneidade do solo. Por isso, segundo o mesmo autor, os métodos de deteção remota devem utilizar-se, principalmente, como fontes auxiliares de informação.

Ainda de acordo com Martino e Fritz (2008), uma vez que os métodos areolares utilizam procedimentos de estatística espacial através da observação direta apenas de uma parte da população de pontos, é possível obter estimativas que são usadas como generalizações válidas para toda a área em estudo. Deste modo, ao contrário do que acontece em abordagens de cartografia (*mapping*), como por exemplo, no projeto *corine land cover*, não é necessário realizar um levantamento exaustivo da área em estudo.

2.4.2. Métodos *model-based*

Os métodos *model-based* constituem uma alternativa à abordagem *design-based*, em particular, em situações em que a dimensão da amostra impossibilita a obtenção de estimativas com precisão aceitáveis.

Como o próprio nome indica, nos métodos *model-based*, o conjunto de valores associados aos pontos são considerados como realizações de um processo aleatório que pode ser formalizado num modelo geoestatístico, desde que pelo menos algumas características do processo se assumem como conhecidas, como por exemplo, a normalidade dos dados (Gruijter e Braak, 1990).

Um exemplo da implementação de um método *model-based*, em situações em que a dimensão da amostra é reduzida, é o da estimação em pequenos domínios, cujos resultados são baseados em modelos de regressão entre variáveis dependentes e variáveis auxiliares, que estão associadas a diferentes fontes disponíveis (Rao, 2003).

2.5. Esquemas de amostragem

Os métodos de amostragem probabilísticos tradicionais aplicados a localizações espaciais georreferenciadas podem ser denominados como métodos de amostragem espaciais, a partir dos quais é possível modelar a covariância entre os pontos.

De acordo com Gruijter *et al.* (2006) a forma como estes métodos se relacionam entre si, é esquematizada na Figura 2.1.

Segundo os mesmos autores é frequente em amostragem espacial que o esquema de amostragem seja obtido pela combinação de mais do que um dos referidos anteriormente, como por exemplo ocorre no caso do projeto LUCAS, em que é utilizado um esquema de amostragem em duas fases, cuja descrição será apresentada no capítulo 3 deste trabalho.

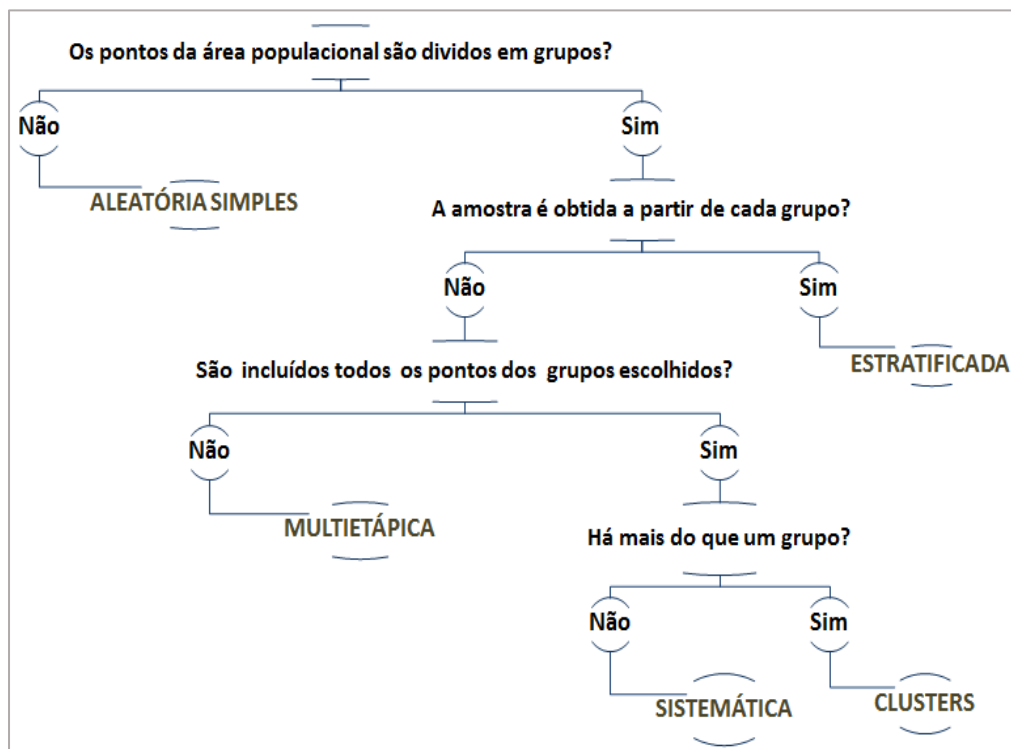


Figura 2.1: Semelhanças e diferenças entre esquemas de amostragem (adaptado de Gruijter *et al.*, 2006)

2.5.1. Amostragem aleatória simples

A amostragem aleatória simples é a forma mais simples de selecionar uma amostra. Neste tipo de amostragem todas as localizações (pontos ou segmentos) têm a mesma probabilidade de ser selecionados e são independentes entre si.

Este tipo de amostragem, por si só, não é habitualmente utilizado em amostragem espacial, excetuando os casos em que faz parte de esquemas de amostragem mais complexos (Cadima *et al.*, 2005).

De entre as razões para a sua não utilização, destacam-se: (1) apresenta uma cobertura espacial pobre, com a tendência para os pontos selecionados se agruparem (Chun e Griffith, 2013); (2) a distribuição espacial é irregular o que origina custos elevados e mais tempo despendido na recolha da amostra (Gruitjer *et al.*, 2006) e (3) a estimação em pequenos domínios pode ser impossível por causa dos largos espaços vazios que podem ocorrer entre os pontos (Brus, 2011).

De acordo com Gruijter *et al.* (2006), a vantagem deste tipo de método reside na simplicidade de implementação e no facto da estimação dos parâmetros e da variância ser relativamente simples de calcular.

2.5.2. Amostragem aleatória estratificada

O processo de estratificação consiste na divisão de uma área populacional em uma ou várias subáreas, designadas por estratos, em que para cada um deles, um número predeterminado de localizações amostrais é selecionado utilizando amostragem aleatória simples (Manly, 2001).

Um estrato homogéneo é um estrato no qual as características de ocupação do solo variam apenas ligeiramente comparativamente com a variação espacial dentro de toda a área populacional.

Para os casos em que após a estratificação a amostra final escolhida é formada por estratos homogéneos, observa-se uma diminuição da variância espacial e consequentemente um ganho na eficiência das estimativas comparativamente com as obtidas através de amostragem aleatória simples (Manly, 2001).

Segundo Gruijter *et al.* (2006), uma correta estratificação pode ser conseguida de duas formas:

- Usando variáveis auxiliares correlacionadas com a variável aleatória em estudo (a partir de um conhecimento prévio sobre o tipo de classe de uso ou ocupação do solo como ocorre no projeto LUCAS);
- Para variáveis aleatórias quantitativas em que os estratos podem ser obtidos por análise de *clusters*.

Nos casos em que a estratificação é incorretamente realizada, habitualmente ocorre não um ganho, mas sim uma perda na precisão do estimador. No trabalho realizado por Brus (2011) é apresentado um exemplo de estratificação em que se teve o cuidado de não considerar um número elevado de estratos com o intuito de evitar um número de localizações por estrato passasse a ser fortemente desproporcional relativamente ao que se verifica para a superfície da área total.

De acordo com Schoenmakers (2005), em estudos de estimação de ocupação do solo, no processo de estratificação não se deve utilizar limites administrativos pela habitual falta de representatividade destes relativamente à cobertura do solo o que poderá originar estratos menos homogêneos e estimativas com variância alta.

2.5.3. Amostragem aleatória por *clusters* (grupos)

A amostragem aleatória por *clusters* é um tipo de esquema de amostragem em que não é desejável obter uma amostra aleatória a partir da área total. Para tal, começa-se por aleatoriamente se selecionar alguns grupos (*clusters*) a partir dos quais são observadas as localizações que os compõem. Cada grupo contém necessariamente um ou mais elementos, realçando-se que são os grupos e não os elementos que o compõem que são as unidades amostrais (Barnett, 2004).

Uma vez que apenas se extraem pontos pertencentes a alguns grupos, a área a analisar é obviamente menor do que se considerasse a área populacional. Consequentemente, usando este esquema de amostragem, é possível reduzir os custos operacionais e os tempos de viagem entre as diferentes localizações, principalmente quando são utilizadas amostragens com regularidades. Esta

vantagem é tanto maior quanto maior for a dimensão da área total a observar (Gruijter *et al.*, 2006).

Para os casos em que o número de pontos a observar em cada grupo é muito grande, pode-se utilizar um método de amostragem em duas etapas de modo a obter amostragens mais eficientes.

Apesar das vantagens da utilização deste esquema de amostragem, ele assenta num pressuposto que nem sempre é possível de ocorrer: que uma dada população compreende subgrupos heterogéneos que funcionam como microcosmos da população inteira e que é possível, seleccionando apenas alguns pontos destes grupos, obter subgrupos que são representativos da área total.

Este método areolar é comum em amostragem espacial, como por exemplo, o apresentado em Cadima *et al.* (2005) no âmbito de um estudo sobre pesca. Numa primeira fase como estudo piloto, entre 2001 e 2005, esta metodologia chegou a ser implementada no projeto LUCAS, tendo sido alterada com o intuito melhorar a precisão das estimativas de uso e ocupação do solo (Gallego e Bamps, 2008).

2.5.4. Amostragem aleatória sistemática

A conveniência com que se processa a recolha da amostra espacial e a necessidade de a amostra, composta por localizações geoespaciais a analisar *in situ*¹, se espalhe por toda a área em estudo são aspetos relevantes na definição do esquema de amostragem areolar.

Num processo de amostragem sistemática, a primeira unidade amostral começa por ser seleccionada aleatoriamente e as outras localizações (pontos), dentro da área total, são escolhidas de acordo com um padrão regular. Este padrão é usualmente definido através de grelhas que se espalham por toda a área em estudo o que permite obter uma cobertura espacial dos pontos mais ampla que compõem a população (Scheaffer *et al.*, 1996).

Segundo os mesmos autores, devido a esta otimização espacial é possível obter estimativas mais precisas, o que habitualmente não acontece em amostragem aleatória simples. Mesmo para o caso da amostragem estratificada em que a

¹Locução latina para descrever a observação no local exato

seleção das amostras para cada estrato é realizada separadamente, tal facto pode originar uma diminuição na precisão das estimativas.

A amostragem aleatória sistemática é um caso concreto de utilização de uma amostragem por *clusters* em que o esquema de amostragem é formado por um único grupo.

Na primeira fase do esquema de amostragem do projeto LUCAS, que será desenvolvido no capítulo 3 deste trabalho, tem sido implementada uma amostragem sistemática constituída por uma grelha formada por pontos espaçados $2 \times 2 \text{ km}^2$ entre si.

2.5.5. Amostragem aleatória multietápica

O método de amostragem multietápica é uma combinação de métodos referidos anteriormente. Para cada etapa, procede-se a uma seleção aleatória de unidades amostrais que são obtidas a partir da amostra da etapa anterior (Cadima *et al.*, 2005).

Para o caso de uma amostragem em duas etapas, a área total começa por ser dividida em subáreas, designadas por unidades de amostragem principais (PSU). Contrariamente ao que acontece na amostragem estratificada, em que a amostra final recolhida é formada por elementos que pertencem a todos os estratos, no esquema de amostragem em duas etapas, apenas um grupo das PSU são aleatoriamente selecionadas. Na segunda etapa, as unidades de amostragem secundárias (SSU) são extraídas aleatoriamente a partir das PSU selecionadas (Gruijter *et al.*, 2006)

Segundo Gruijter *et al.* (2006), através deste tipo de método é possível aceder mais eficientemente aos pontos por parte dos observadores (comparando com os casos em que é selecionada, indiferenciadamente, uma amostra a partir da área total). Tem igualmente a vantagem de reduzir os tempos de viagem entre pontos.

De acordo com os mesmos autores, como desvantagem na utilização de um esquema em duas etapas há a realçar que as estimativas produzidas podem ser menos precisas do que o esperado para situações em que a variação dentro das unidades primárias é mais pequena do que o previsto aquando da decisão do número de localizações a atribuir a cada uma delas.

2.5.6. Amostragem aleatória multifásica

A amostragem aleatória multifásica é um esquema de amostragem amplamente utilizado em contextos de estimação espacial, principalmente nos casos em que existe pouca informação sobre a população que se pretendem estimar.

Assim, neste tipo de amostragem, a amostra recolhida na fase imediatamente anterior fornece informação auxiliar para a realização da amostragem da fase seguinte do esquema de amostragem (Gruidjer *et al.*, 2006).

Por exemplo, segundo os mesmos autores, no caso da amostragem em duas fases (*double sampling*), ao contrário do que acontece na amostragem em duas etapas, os elementos amostrais que compõem as PSU e as SSU são medidos de acordo com duas variáveis que estão correlacionadas entre si. Consequentemente, o sucesso da amostragem em duas fases depende consideravelmente do grau de associação entre a variável primária e secundária que quantificam as unidades amostrais.

Ainda de acordo com Gruidjer *et al.* (2006), um exemplo prático deste tipo de amostragem em estudos espaciais, ocorre quando se pretende estratificar a área total, mas o número e a dimensão dos estratos populacionais é desconhecida. Para tal, os autores propõem a definição de classes (representadas, no caso, por uma variável categórica) que são selecionadas aleatoriamente na primeira fase e usadas como estratos na segunda fase, como está representado na Figura 2.2.

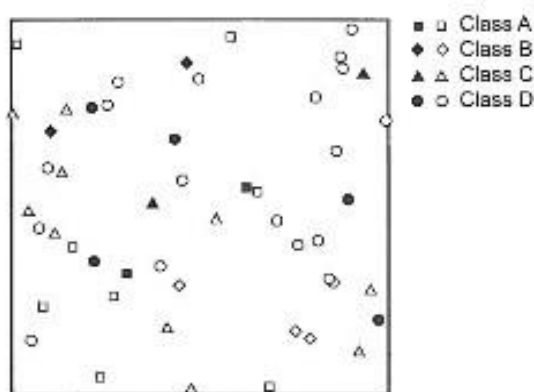


Figura 2.2: Exemplo da implementação de um esquema de amostragem em duas fases. Os símbolos sombreados representam a amostra da segunda fase (retirado de Gruijter *et al.*, 2006)

Kim e Lee (2013) realizaram um estudo em que sugerem que antes de se adotar um esquema de amostragem em duas fases é importante ter-se a

certeza de que as variáveis em estudo sejam altamente correlacionadas caso contrário pode ser preferível escolher outros tipos de esquemas de amostragem para melhorar a eficiência das estimativas.

Os principais inconvenientes na implementação deste método são a complexidade da computação das estimativas e a determinação da dimensão da amostra.

Como será apresentado no subcapítulo 3.3, no projeto LUCAS é utilizado um esquema de amostragem em duas fases (com estratificação) cujos pontos em cada uma das fases são classificados de acordo com estratos e classes de uso e ocupação do solo correlacionadas entre si.

2.5.7. Amostragem espacialmente balanceada (*spatially-balanced sampling*)

Uma amostra diz-se espacialmente balanceada sempre que esta apresenta uma distribuição uniforme ao longo da área em estudo obtida através da maximização da independência espacial entre as localizações geográficas que compõem a amostra (Theobald *et al.*, 2007).

De acordo com os mesmos autores, a utilização deste tipo de amostragem permite melhorar a eficiência das estimativas comparativamente com outros casos de amostragens areolares, como por exemplo a amostragem aleatória sistemática, principalmente em situações em que as unidades amostrais a extrair dentro da área em estudo são não contíguas (como ocorre em zonas com lagos ou zonas húmidas).

Há várias técnicas para obter amostras espacialmente balanceadas entre as quais se destaca a designada por *generalized random-tessellation stratified* (GRTS), apresentada por Stevens e Olsen (2004), cujo algoritmo envolve um processo aleatório hierárquico recursivo que preserva as relações espaciais das unidades amostrais.

Neste algoritmo, a área da população é dividida em segmentos quadrados com dimensões iguais, que por sua vez se dividem em novos quadrados com a mesma ordem de numeração dos anteriores. Este processo recursivo continua até que cada quadrado corresponda a uma localização que será numerada com o número atribuído ao quadrado a que pertence. Por fim, as localizações são

ordenadas de acordo com os números que lhe foram atribuídos, como é representado na Figura 2.3.

O *package spsurvey* desenvolvido por Kincaid *et al.* (2014), apresenta um conjunto de funções na linguagem R para a implementação do esquema de amostragem GRTS.

Este tipo de esquema de amostragem areolar apresenta as seguintes vantagens: fornece um bom balanceamento espacial das localizações; utiliza seleção de amostras através de probabilidades desiguais; pode ser utilizado em bases amostrais dadas por áreas, assim como em linhas e zonas não contíguas.

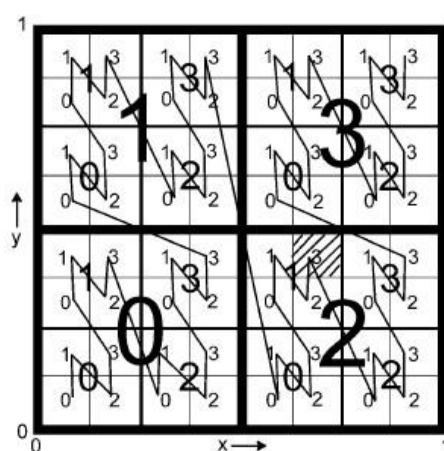


Figura 2.3: Esquema de amostragem espacialmente balanceado segundo o algoritmo GRTS (retirado de Stevens e Olsen, 2004)

Por outro lado, este tipo de esquema de amostragem apresenta como desvantagem o facto de utilizar técnicas relativamente difíceis de implementar e compreender. Além disso, para os casos em que se pretenda quantificar a correlação espacial dos parâmetros este esquema de amostragem é ineficiente.

2.6. Estimação

2.6.1. Conceitos fundamentais

Os parâmetros da população são medidas que resumiriam a forma como os valores de uma variável aleatória se distribuem na população. A cada variável aleatória corresponde uma função X , que a cada elemento pertencente ao espaço de resultados Ω , $\omega \in \Omega$, associa um valor numérico real $X(\omega) = x \in IR$.

Para estimar os parâmetros são utilizadas funções que se designam por estatísticas, cujos valores obtidos (estimativas) se pretende que sejam próximos dos valores reais dos parâmetros.

Neste processo de inferência estatística, a precisão da estimativa depende, entre outros fatores, da representatividade da amostra e da dimensão da amostra (Sanders,1995).

Considere-se uma população com um determinado número de elementos, para uma dada característica representada pela variável aleatória Y , em que Y_i é o valor da variável aleatória na unidades estatística i .

As populações podem ser finitas ou infinitas. O número total de elementos de uma população finita é dado por N . Sempre que o número de elementos é muito grande, N pode considerar-se como infinito.

Se a população é finita o valor total das observações para a característica em estudo é dado por

$$Y = \sum_{i=1}^N Y_i . \quad (2.1)$$

A média populacional é dada por

$$\bar{Y} = \frac{\sum_{i=1}^N Y_i}{N} . \quad (2.2)$$

O parâmetro total é obtido multiplicando o número de elementos da população pela média

$$Y_T = N\bar{Y} . \quad (2.3)$$

De entre os parâmetros utilizado como medidas de dispersão para a populações finitas, a variância populacional σ^2 é dada por

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2 . \quad (2.4)$$

O coeficiente de variação, CV , é uma medida relativa de dispersão que permite comparar a dispersão de duas populações mesmo com diferentes unidades de medida. Para o caso da média populacional o CV é dado por

$$CV = \frac{\sigma}{\bar{Y}} . \quad (2.5)$$

Para variáveis categóricas, o parâmetro de interesse corresponde à proporção populacional de elementos com uma determinada característica. Por exemplo, para o caso de amostragens espaciais de classificação de uso ou ocupação do solo, a variável Y pode ser dicotómica

$$Y_i = \begin{cases} 1, & \text{ponto pertence à classe de uso ou ocupação do solo} \\ 0, & \text{ponto não pertence à classe de uso ou ocupação do solo} \end{cases}$$

A proporção de elementos pertencentes à categoria de interesse, P , é igual à média de Y

$$\bar{Y} = \frac{\sum_{i=1}^N Y_i}{N} = P . \quad (2.6)$$

A variância da população em termos de proporções é dada por

$$\sigma^2 = P(1 - P) . \quad (2.7)$$

2.6.2. Métodos de estimação

Excetuando quando é efetuado um censo, os valores dos parâmetros populacionais não são conhecidos. Para os estimar podem ser utilizados dois métodos:

- Método de estimação pontual: é obtido um único valor (estimativa) através de um estimador associado ao tipo esquema de amostragem. De entre os métodos de estimação pontual há a considerar o método dos momentos e o método de máxima verosimilhança;
- Método de estimação intervalar: é obtido um intervalo de valores de acordo com um modelo paramétrico associado e um nível de confiança predefinido. O nível de confiança refere-se à probabilidade de corretamente incluir o parâmetro no intervalo estimado.

2.7. Estimadores relacionados com os esquemas de amostragem

Efetuada uma revisão de literatura sobre os conceitos fundamentais relacionados com estimação, neste capítulo são apresentados os estimadores associados a cada tipo de esquema de amostragem referidos no subcapítulo 2.5.

2.7.1. Amostragem aleatória simples

Estimador para a média

Um estimador, \bar{y} , não enviesado para a média da área espacial (para tipo de cobertura do solo), aplicando uma amostragem aleatória simples, é intuitivamente obtido por

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad (2.8)$$

em que n é a dimensão da amostra e y_i corresponde à área associada ao ponto (relativa ao tipo de solo, por exemplo).

Estimador para a variância da média

De acordo com Barnett (2004), a variância da média deste tipo de amostragem é dada por

$$\text{var}(\bar{y}) = \frac{s^2}{n} = \frac{1}{n(n-1)} \sum_{i=1}^n (y_i - \bar{y})^2, \quad (2.9)$$

em que s^2 é a estimativa da variância das medidas obtidas para a área total em estudo.

Segundo Barnett (2004), para o caso de amostragens sem reposição para populações finitas, o valor da variância é reduzido multiplicando pelo fator de correção da população finita, $1 - f$, em que f é a taxa de amostragem dada por n/N , ou seja

$$\text{var}(\bar{y}) = (1 - f) \frac{s^2}{n}. \quad (2.10)$$

No caso em que f é menor ou igual a 0.05, este fator pode ser ignorado.

Estimador para a proporção

O estimador da proporção dos elementos pertencerem a uma categoria é dado por

$$p = \frac{1}{n} \sum_{i=1}^n y_i \quad (2.11)$$

em que

$$y_i = \begin{cases} 1, & \text{ponto que pertence à categoria} \\ 0, & \text{ponto que não pertence à categoria} \end{cases}$$

Segundo Barnett (2004), a variância da proporção é dada por

$$\text{var}(p) = \frac{np(1-p)}{n-1}. \quad (2.12)$$

Para populações finitas, aplicando o fator de correção,

$$\text{var}(p) = (1 - f) \frac{np(1-p)}{n-1}. \quad (2.13)$$

2.7.2. Amostragem aleatória estratificada

Estimador para a média

De acordo com Cochran (1977), aplicando um esquema de amostragem estratificada a área média é estimada por

$$\bar{y}_{st} = \frac{1}{N} \sum_{h=1}^L N_h \bar{y}_h = \sum_{h=1}^L w_h \bar{y}_h, \quad (2.14)$$

onde L é o número de estratos; N é o número de unidades estatísticas na população de pontos; N_h é o número de unidades estatísticas no h -ésimo estrato e \bar{y}_h é a média estimada da área no h -ésimo estrato.

Estimador para a variância da média

Segundo Cochran (1977), um estimador não enviesado para a variância estratificada de \bar{y}_{st} é dado por

$$\text{var}(\bar{y}_{st}) = \sum_{h=1}^L w_h^2 \frac{s_h^2}{n_h}, \quad (2.15)$$

em que, n_h é a dimensão da amostra no h -ésimo estrato e s_h^2 é variância dentro do estrato dada por

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2, \quad (2.16)$$

em que \bar{y}_h é a média estimada da área no h -ésimo estrato.

Para o caso de populações finitas, o fator de correção seria dado por $1 - f_h$, onde $f_h = \frac{n_h}{N_h}$, e o estimador da variância para média é dado por

$$\text{var}(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \frac{s_h^2}{n_h}. \quad (2.17)$$

Estimador para a proporção

De acordo com Cochran (1977), no caso de uma amostragem aleatória estratificada, a estimativa da proporção de unidades amostrais pertencentes a uma categoria, é obtida por

$$p_{st} = \sum_{h=1}^L w_h p_h, \quad (2.18)$$

em que p_h é a proporção estimada da categoria no h -ésimo estrato.

Estimador para a variância da proporção

Para uma amostragem aleatória estratificada, a variância de p_{st} é dada por

$$\text{var}(p_{st}) = \sum_{h=1}^L w_h^2 \times \frac{p_h(1-p_h)}{n_h}. \quad (2.19)$$

Para populações finitas, em que o fator de correção deve ser considerado, a variância é dada por

$$\text{var}(p_{st}) = \frac{1}{N^2} \sum_{i=1}^L N_h^2 (N_h - n_h) \frac{p_h(1-p_h)}{n_h(N_h-1)}. \quad (2.20)$$

2.7.3. Amostragem aleatória por *clusters* (grupos)

Estimador para a média

Segundo Gruijter *et al.* (2006), a área média é estimada por

$$\bar{y}_{cl} = \frac{1}{m} \sum_{i=1}^m \bar{y}_i, \quad (2.21)$$

em que m é o número de *clusters* e \bar{y}_i é a média amostral do i -ésimo *cluster*.

Estimador para a variância da média

Segundo Gruijter *et al.* (2006) o estimador para a variância da média é dado por

$$\text{var}(\bar{y}_{cl}) = \frac{1}{m(m-1)} \sum_{i=1}^m (\bar{y}_i - \bar{y}_{cl})^2. \quad (2.22)$$

Para o caso dos estimadores para a proporção, estes são obtidos substituindo \bar{y}_i , por p_i e \bar{y}_{cl} por p_{cl} .

2.7.4. Amostragem aleatória sistemática

Estimador para a média

A amostragem aleatória sistemática é um caso particular da amostragem por *cluster*, sendo constituída apenas por um único grupo. Assim, pelo estimador da média da amostragem por *clusters*, o estimador da média para uma amostra aleatória sistemática reduz ao exposto para a amostragem aleatória simples.

Estimador para a variância da média

Segundo Cochran (1977), a variância da média é difícil de estimar uma vez que esta requer algumas considerações sobre a ordem com que os pontos se posicionam. Para os casos em que se considera que a média da amostra é aleatória, então a variância é dada por

$$\text{var}(\bar{y}_{sy}) = (1 - f) \frac{s^2}{n}. \quad (2.23)$$

Para o caso do estimador da variação da proporção, é obtido por

$$\text{var}(p_{sy}) = (1 - f) \frac{np(1-p)}{n-1}. \quad (2.24)$$

2.7.5. Amostragem aleatória em duas etapas

Estimador para a média

Segundo Gruidjer *et al.* (2006), a área média para um esquema de amostragem em duas etapas é obtido pelo estimador

$$\bar{y}_{TS} = \frac{1}{n} \sum_{i=1}^n \bar{y}_i, \quad (2.25)$$

em que n é o número unidades de amostragem principais (PSU) seleccionadas e \bar{y}_i é a média amostral da i -ésima PSU seleccionada.

Estimador para a variância da média

Segundo Gruidjer *et al.*, 2006, o estimador da variância da média é dado por

$$\text{var}(\bar{y}_{TS}) = \frac{1}{n(n-1)} \sum_{i=1}^n (\bar{y}_i - \bar{y}_{TS})^2 . \quad (2.26)$$

Para o caso dos estimadores para a proporção, estes são obtidos substituindo \bar{y}_i , por p_i e \bar{y}_{TS} por p_{TS} .

2.7.6. Amostragem aleatória em duas fases (por estratificação)

Estimador para a média

De acordo com Cochran (1977), o estimador da média para uma amostragem em duas fases (por estratificação) é dado por

$$\bar{y}_{TP} = \frac{1}{N} \sum_{h=1}^H N_h \bar{y}_h = \sum_{h=1}^H w_h \bar{y}_h, \quad (2.27)$$

em que N é número de pontos selecionados na primeira fase, N_h é o número de pontos da primeira fase pertencentes ao h -ésimo estrato e \bar{y}_h é a média da amostra na segunda fase do esquema de amostragem.

Estimador para a variância da média

Segundo Cochran (1977), o estimador da variância da média é dado por

$$\text{var}(\bar{y}_{TP}) = \sum_{h=1}^H \frac{w_h^2 \times s_h^2}{n_h} - \sum_{h=1}^H \frac{w_h \times s_h^2}{N} + \frac{g'}{n'} \sum_{h=1}^H w_h (\bar{y}_h - \bar{y}_{TP}), \quad (2.28)$$

em que, $g' = \frac{(N'-N)}{(N'-1)}$, N' é o número de pontos na população, N é o número de pontos da primeira fase do esquema de amostragem e n_h é o número de pontos pertencentes ao h -ésimo estrato na segunda fase do esquema de amostragem.

No subcapítulo 3.4 sobre a metodologia de estimação implementada no projeto LUCAS será apresentada a forma como se procedeu à utilização deste tipo de estimadores na obtenção das proporções e respetivas áreas associadas ao uso e ocupação do solo.

2.8. Projetos de amostragem areolar sobre uso e ocupação do Solo

De seguida são apresentados exemplos de sistemas de classificação nacionais e internacionais baseados em amostragem areolar.

Dada a relevância do projeto LUCAS neste trabalho, a descrição do mesmo será efetuada no capítulo 3.

2.8.1. Inventário de Florestas Nacional (IFN)

O Inventário de Florestas Nacional abrange todos os tipos de solos com uso florestal a nível de Portugal Continental. Este tem vindo a ser implementado desde 1965, com uma periodicidade aproximada de 10 anos, tendo o último IFN6 sido realizado em 2010 (AFN, 2012)

O IFN utiliza uma metodologia de amostragem areolar com um esquema de amostragem em duas fases. A primeira fase é basicamente composta por trabalho de gabinete em que se procede à avaliação das áreas a partir da amostra constituída por fotopontos. Os dados da fotografia aérea são, igualmente, analisados espacialmente, através de um sistema de informação geográfica para produção de cartografia genérica relativa a uso e ocupação do solo do território continental (Ferreira e Campos, 2013).

Na segunda fase são efetuadas medições nas parcelas do inventário (para avaliar as características dos povoamento florestais e matos) e recolhas de amostra de solo (em solos agrícolas e florestais) (AFN, 2012).

O esquema de amostragem é realizado através de uma grelha com orientação N-S e E-O em que na primeira fase os pontos estão espaçados 500x500 m² (a qual dá origem a aproximadamente 360 000 pontos fotointerpretados distribuídos de forma regular por Portugal continental).

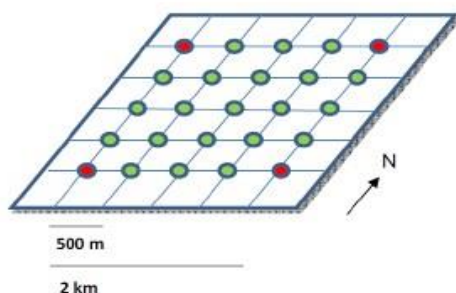


Figura 2.4: Estrutura da grelha nacional de amostragem (AFN, 2012)

Para a sua implementação, o IFN envolve quatro passos principais: no primeiro procede-se à fotointerpretação para avaliação de áreas sobre uso e ocupação de solos florestais e agrícolas; seguidamente são recolhidos dados de campo para caracterização quantitativa dos espaços florestais e para quantificação do carbono dos solos florestais e agrícolas. Por fim, ocorre o processamento de dados e publicação de informação.

A forma como as quatro passos são estruturadas é apresentada na figura 2.5

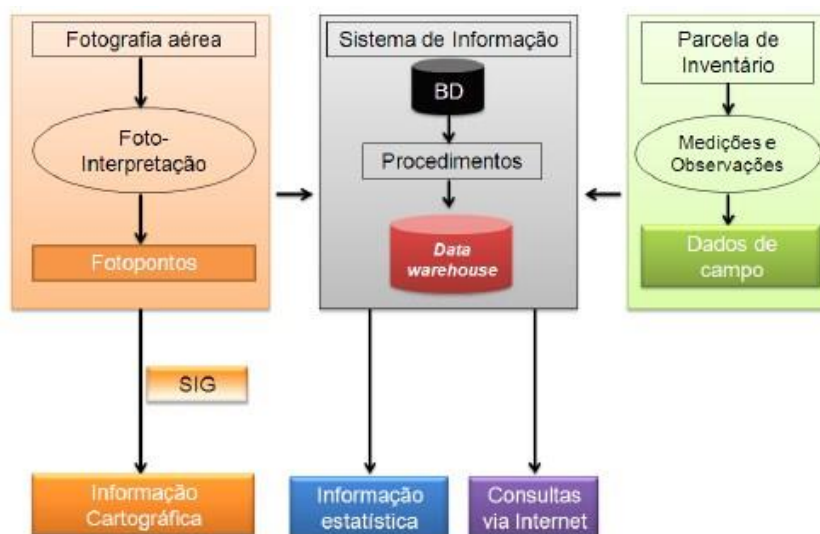


Figura 2.5: Esquema simplificado dos processos de recolha e processamento de dados do IFN de 2010 (retirado de AFN, 2012)

Na obtenção de estatísticas de uso e ocupação do solo a área de cada classe tem sido estimada a partir da proporção entre o número de pontos e o número de pontos na área em que unidade estatística é uma porção de terra de forma circular.

A nomenclatura de uso e ocupação do solo utilizada no IFN6 encontra-se estruturada em 9 níveis de informação, organizados em três grandes temas:

uso do solo, ocupação do solo e análise da paisagem, cuja descrição pode ser consultada no Anexo 2.

Para a definição das classes de uso e ocupação do solo os valores mínimos estabelecidos para dimensão e forma da mancha homogénea a classificar são (AFN, 2012):

- Área maior ou igual a 5000 m² (i.e., 0,5 ha);
- Largura não inferior a 20 m.

2.8.2. Projeto Landyn

Landyn é um projeto nacional de classificação de uso e ocupação do solo cujos objetivos principais são: (1) analisar as alterações que ocorreram ao nível do uso e da ocupação do solo desde a década de 1980 até à atualidade; (2) identificar as principais causas para as alterações ocorridas; (3) formular cenários possíveis para as classes de uso e ocupação do solo até 2040, usando para tal, modelos espaciais; (4) usar a informação das classes de uso e ocupação do solo disponível para que seja utilizada em estudos sobre emissões e remoções de gases com efeito de estufa (GEE) (Landyn, 2013).

A nomenclatura de classificação do projeto Landyn é composta por 32 classes, agrupadas em sete classes principais em função da generalização do uso e tipo de ocupação do solo (artificializados, agrícolas, agro-florestais, corpos de água, florestas, incultos e zonas húmidas) (Anexo 3).

O projeto Landyn utiliza uma metodologia areolar em que na definição da base de amostragem foi utilizada uma grelha de referência de 1x1 Km² da Agência Europeia do Ambiente (EEA). A partir desta grelha, de forma a garantir a consistência espacial com os dados recolhidos no projeto LUCAS em 2009, utilizou-se uma amostragem por *clusters* (constituída por 1279 unidades amostrais distribuídas pelo território de Portugal Continental espaçadas 2x2 km²) (DGT, 2013).

2.8.3. Projeto TERUTI

A nível europeu o TERUTI é um projeto francês que tem sido implementado desde a década de 1970, cujo esquema de amostragem areolar em duas etapas

sem estratificação e que tem servido de referência para outros projetos, como o BANCÍK. A metodologia de amostragem baseia-se na observação regular de um conjunto de pontos, as unidades secundárias, situadas numa seleção de segmentos quadrados de 2x2 Km², as unidades primárias (Betio *et al*, 2002).

Segundo Schoenmakers (2005), na primeira etapa, por cada segmento de 12x12 km², são selecionados aleatoriamente quatro segmentos quadrados (2x2 km²), representando 11% do universo original. Na segunda etapa são selecionados 36 pontos, escolhidos sistematicamente, com uma distância de 300 metros entre cada ponto.

Para analisar a ocupação do solo de cada ponto efetua-se trabalho de campo, excetuando nos casos dos pontos com má acessibilidade utilizando-se, nestes casos, interpretação de fotografias aéreas. Cada ponto representa uma área de 3x3 metros, observando-se nesta área o tipo de ocupação.

2.8.4. Projeto *Countryside Survey*

Countryside Survey é um projeto que tem sido implementado no Reino Unido com o intuito de recolher informação relacionada com cobertura do solo, tipo de *habitats* naturais e espécies, utilizando um esquema de amostragem areolar em duas etapas estratificada com segmentos quadrados de 1 km².

O primeiro inquérito ocorreu em 1978 tendo-se procedido à seleção de 256 segmentos quadrados (correspondente a 8 por cada um dos 32 tipos de classes de ocupação do solo). Os inquéritos seguintes ocorreram em 1984, 1990, 1998 e 2007, tendo sido selecionados, respetivamente, 384, 508, 569 e aproximadamente 620 segmentos. Destes segmentos, 244 segmentos que têm sido sempre analisados com o objetivo de avaliar as alterações ocorridas ao longo do tempo (Brus *et al.*, 2011).

Em termos de estimadores, neste projeto tem sido utilizada uma abordagem *design-based*, dada por estimadores *ratio* e uma *model based* proposta por Scott (2008).

2.8.5. Projeto National Resources Inventory (NRI)

O NRI é um inquérito para a deteção de condições e tendências relacionadas com a ocupação e uso do solo, água e outros recursos naturais que tem sido

implementado pelo *US Department of Agriculture's Natural Resources Conservation Service* (NRCS). Entre 1977 e 1997 realizou-se com uma periodicidade de 5 anos. Desde 2000 tem sido realizado anualmente (Opsomer *et al.*, 2003)

Os dados recolhidos pertencem a uma amostra estratificada correspondendo a áreas não-federais dos Estados Unidos e Porto Rico. O esquema de amostragem é realizado em duas etapas com probabilidades desiguais de seleção. As unidades primárias de amostragem são representadas por segmentos. As unidades secundárias de amostragem são representadas por pontos que são selecionados aleatoriamente dentro de cada segmento. A amostra atual é composta por 300 000 segmentos e cerca de 844 000 pontos. Para uma descrição mais completa acerca da metodologia de amostragem e características dos solos estudadas no NRI, consultar Fuller (1999).

No processo de estimação do NRI há a referir a combinação de informação proveniente de diversas fontes para produzir os dados finais que são compostos por registos obtidos a partir de todos os inquéritos. Para mais informações sobre o processo de estimação podem ser obtidas em Fuller (1999).

Capítulo 3

Land Use and Cover Area Frame Statistical Survey (LUCAS)

Land Use and Cover Area Frame Statistical Survey (LUCAS) é um projeto de amostragem espacial que tem sido conduzido pelo Eurostat, de acordo com a decisão do Conselho Europeu de serem implementadas técnicas de amostragem areolar para a produção de estimativas relacionadas com o uso e ocupação do solo ao nível do espaço europeu (Bettio *et al.*, 2002).

Desde 2006 que o projeto LUCAS tem sido implementado com uma periodicidade de três anos. Já foram produzidos resultados em 2006, 2009 e 2012. O próximo inquérito está previsto para 2015 (Eurostat, 2013a).

O LUCAS caracteriza-se por um esquema de amostragem em duas fases (*double sampling*) cuja primeira base de amostragem é composta por aproximadamente um milhão de pontos sobrepostos numa grelha quadrangular com um espaçamento de 2x2 km² entre os pontos (Gallego, 2013).

Os pontos observados na segunda fase de amostragem são selecionados a partir da amostra inicial e classificados *in situ*, segundo a nomenclatura de classificação de uso e ocupação definida no projeto LUCAS (Eurostat, 2013b).

Uma vez que o LUCAS utiliza um processo de inferência estatística na produção de estimativas, não há a necessidade de se observar toda a área populacional, tal como acontece em abordagens de deteção remota. Além disso, como se procede a uma observação direta das localizações geográficas (pontos), através dos inquéritos do LUCAS é possível recolher uma descrição mais completa da composição do solo e, conseqüentemente, aumentar a eficiência das estimativas em situações de heterogeneidade do solo.

Ainda não é possível obter uma tão detalhada classificação usando processos de fotointerpretação ou recolha de imagens via satélite, como os obtidos pela *Corine Land Cover* (Martino *et al.*, 2009). Por isso, segundo os mesmos autores, os métodos espaciais de deteção remota devem-se utilizar idealmente como ferramenta auxiliar para melhorar o esquema de amostragem.

No caso do LUCAS, as imagens de satélite foram utilizadas para atribuição de classes em casos de pontos inacessíveis (exemplo de florestas impenetráveis) e a deteção remota foi utilizada para uma localização mais precisa dos pontos no processo de observação no terreno.

3.1. Objetivos do projeto LUCAS

Apesar da importância crescente que é atribuída aos sistemas de classificação na definição e avaliação de políticas comuns no espaço europeu, ao longo dos anos tem-se verificado na maioria dos casos, que as estimativas produzidas são realizadas independentemente em cada estado membro.

Assim, com o intuito de uniformizar procedimentos e aumentar a comparabilidade entre os resultados, a partir da decisão N°1445/2000/EC do Parlamento Europeu e do Conselho Europeu de 22.05.2000 “*On the application of area-frame survey and remote-sensing techniques to the agricultural statistics for 1999 to 2003*”, continuada até 2007 pela decisão 2066/2003/EC, de 10 Novembro de 2003 e prolongada pela decisão 786/2004/EC, de 21 Abril de 2004, o projeto LUCAS começou a ser implementado em 2001 em 13 países europeus (Bettio *et al.*, 2002).

De um modo geral, conforme referido em Eurostat (2013a), os objetivos do LUCAS são:

- Desenvolver um sistema de classificação com uma nomenclatura hierárquica que permita uma clara distinção entre as classes de ocupação e uso do solo;
- Implementar um método de amostragem espacial, ao nível do espaço europeu, que simultaneamente combine técnicas de amostragem areolar e de deteção remota (tais como, *Corine Land Cover*);
- Obter estatísticas não-enviesadas e harmonizadas que possam ser comparadas e analisadas ao longo do tempo a uma escala europeia e que sejam utilizadas na definição de estratégias comuns conduzidas pela Agência Ambiental Europeia;
- Fornecer informação estatística destinada, não apenas, a utilizadores de estatísticas sobre agricultura, mas igualmente, para especialistas

ambientais que possam ter ao seu dispor bases de dados homogéneas sobre erosão, paisagísticos, riscos naturais, ruído, por exemplo;

- Monitorizar a evolução das condições ambientais pelos países europeus ao longo do tempo.

3.2. Sistema de classificação do projeto LUCAS

Enquanto o conceito de ocupação do solo diz respeito à cobertura física da superfície que compõe o solo, o uso do solo está relacionado com as funções socioeconómicas (por exemplo, um edifício pode ter um uso residencial, industrial ou comercial) (Jacques e Gallego, 2005).

O sistema de classificação LUCAS caracteriza-se por uma clara distinção entre estes dois conceitos o que não acontece por exemplo com a *corine land cover* ou no caso do sistema de classificação nacional da carta e ocupação do solo COS 2007.

Para tal, a nomenclatura definida no LUCAS é composta por classes mutuamente exclusivas, definidas hierarquicamente em três níveis, em que o nível 1 corresponde a uma descrição mais abrangente e que se subdivide em classes mais detalhadas sobre o uso e ocupação do solo (níveis 2 e 3) como podem ser consultadas em Eurostat (2013a).

A classificação dos pontos de acordo com a nomenclatura do projeto LUCAS é realizada na segunda fase do esquema de amostragem no momento em que os pontos são observados *in situ* por parte dos observadores.

3.2.1. Classes de ocupação do solo

Para o nível 1 do sistema de classificação LUCAS foram definidas 8 classes principais como se encontram apresentadas na Tabela 3.1, designadas por territórios artificializados (A), terrenos agrícolas (B), floresta (C), matos (D), prados (E), solo nu e líquenes (F), corpos de água (G) e zonas húmidas (H).

No total, para o inquérito realizado em 2012, os pontos foram classificados de acordo com 83 classes de ocupação do solo cuja descrição é apresentada no documento de referência técnica sobre a classificação dos pontos selecionados (Eurostat, 2013c).

Para este trabalho de investigação foram obtidas estimativas das áreas das classes de ocupação do solo para os níveis 1, 2 e 3 da nomenclatura do sistema de classificação LUCAS apresentadas na Tabela 3.1.

Tabela 3.1: Classes de ocupação do solo (para os níveis 1, 2 e 3 do sistema de classificação LUCAS)

Nível 1	Nível 2	Nível 3
A – Territórios artificializados	A10 (Áreas edificadas), A20 (Áreas não construídas artificiais)	
B – Terrenos agrícolas	B10 (Cereais), B20 (Culturas de raízes comestíveis), B30 (Culturas temporárias industriais), B40 (Leguminosas secas, hortícolas e flores), B50 (Culturas forrageiras), B70 (Frutos secos, frutos de casca rijá, frutos pequenos de baga, citrinos e frutos sub- tropicais), B80 (Vinha, olival, viveiros e culturas permanentes industriais).	
C – Floresta	C10 (Floresta (de acordo com a FAO)), C20 (Outras superfícies arborizadas (FAO)), C30 (Outras superfícies arborizadas (não-FAO))	C11 (Floresta de folhosas caducifólias (75%)), C12 (Floresta de coníferas (75%)), C13 (Floresta mista)
D - Matos	D10 (Matos com árvores esparsas), D20 (Matos sem árvores)	
E - Prados	E10 (Prados com árvores esparsas), E20 (Prados sem árvores), E30 (Vegetação espontânea)	
F – Solo nu e líquenes	F10 (Afloramentos rochosos e zonas pedregosas), F20 (Areias), F30 (Líquenes), F40 (Outras áreas de solo nu)	
G – Corpos de água	G10 (Planos de água interiores), G20 (Águas interiores correntes), G30 (Corpos de água costeiros), G50 (Glaciares e neves permanentes)	
H – Zonas húmidas	H10 (Zonas húmidas interiores), H20 (Zonas húmidas costeiras)	

3.2.2. Classes de uso do solo

No que se refere às classes de uso de solo os pontos podem ser classificados por 33 classes cuja descrição se encontra no Anexo 4. Na Tabela 3.2 são apresentadas as classes de uso do solo para os níveis 1 e 2 da nomenclatura de classificação LUCAS. Para este trabalho de investigação, foram obtidas estimativas das áreas das classes agricultura (U110), floresta (U120), indústria

extrativa (U130), serviços culturais, entretenimento e recreativos (U340) e áreas naturais ou com outros usos socioeconómicos (U612).

Tabela 3.2: Classes de Uso do Solo (para os níveis 1 e 2 do sistema de classificação LUCAS)

Nível 1	Nível 2
U100 – Produção primária	U110 (Agricultura), U120 (Floresta), U130 (Indústria extrativa), U140 (Aqüicultura e pesca), U150 (Outra produção primária)
U200 – Produção secundária	U210 (Indústria de matérias primas), U220 (Indústria pesada), U230 (Indústria ligeira), U240 (Produção de energia)
U300 – Produção terciária	U310 (Serviços comerciais), U320 (Serviços financeiros, profissionais e informação), U330 (Serviços de comunidade), U340 (Serviços culturais, entretenimento e recreativos).
U400 – Redes de transporte e serviços de utilidade pública	U410 (Redes de transportes), U420 (Serviços de logística e armazenagem), U430 (Serviços de utilidade pública)
U500 – Utilização residencial	U510 (Utilização residencial permanente), U520 (Utilização residencial temporária)
U600 – Outras utilizações ou utilizações não sócio-económicas	U610 (Utilização temporária), U611 (áreas abandonadas), U612 (Áreas naturais ou com outros usos socioeconómicos)

3.2.3. Comparabilidade da nomenclatura de classificação do projeto LUCAS com outros sistemas de classificação

O sistema de classificação LUCAS foi concebido de forma a permitir a maior comparabilidade possível relativamente a outros sistemas de classificação, seguindo para tal definições usadas internacionalmente sobre uso e ocupação do solo (Martino *et al.*, 2009).

Por exemplo, a classificação de florestas (C10) e outras superfícies arborizadas (C20), estas estão de acordo com a legislação europeia, tendo sido, igualmente, adotada a classificação sobre florestas da FAO: Reg (EC) N.º 2152/2003 de

17/11/2003 relativamente à monitorização de florestas e interações ambientais na *Community - Forest Focus*. (Eurostat, 2013a)

Ainda no que se refere à comparabilidade entre os sistemas de classificação destaca-se o facto de o tipo de florestas definido na nomenclatura do projeto LUCAS ter sido implementado segundo as indicações da EUNIS (*European University Information Systems*).

Os resultados obtidos nos inquéritos implementados no LUCAS têm sido utilizados como indicadores agro-ambientais e de eficiência energética, assim como tem sido utilizados na produção, verificação e validação de processos ao nível de iniciativas *mapping* de deteção remota, tais como, *Corine Land Cover* e outros projetos *mapping* como Copernicus HRL (*High Resolution Layers*) (Eurostat, 2013a).

3.3. Esquema de amostragem implementado no LUCAS

Desde 2006 que no projeto LUCAS tem sido implementado um esquema de amostragem areolar em duas fases (por estratificação) (*double sampling*), depois de um período experimental, entre 2001 e 2005, em que foi utilizado um esquema de amostragem em duas etapas (Martino *et al.*, 2009).

Segundo Gallego (2007), com este novo esquema de amostragem foi possível melhorar a eficiência das estimativas, reduzir a variância e extrair unidades primárias de amostragem (PSU) mais representativas.

De facto, com a utilização de uma amostragem sistemática, na primeira fase do LUCAS, é possível obter uma distribuição geográfica mais homogénea dos pontos por todo o espaço europeu o que conjugado com a estratificação da população (na segunda fase) torna a estimativas com melhor precisão (pelo aumento da representatividade das amostras dentro dos estratos homogéneos).

A forma como se processa o esquema de amostragem LUCAS é descrita na Figura 3.1.

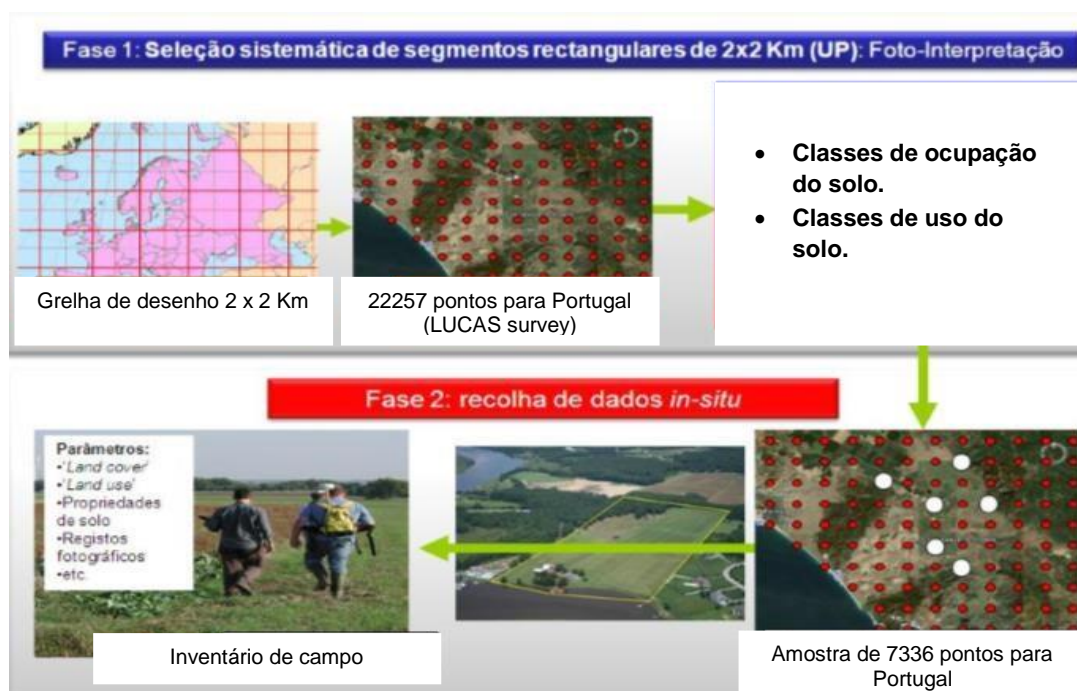


Figura 3.1: Esquema de amostragem em duas fases para 2012 (adaptado de Eurostat, 2012)

3.3.1. Primeira fase do esquema de amostragem

A amostra da primeira fase é obtida a partir de uma amostragem sistemática de pontos sobrepostos numa grelha quadrangular, com pontos espaçados por 2x2 km² segundo os 4 pontos cardiais, cobrindo 27 países da União Europeia participantes no projeto LUCAS. O número total de pontos em 2012 foi 1077247 (Eurostat, 2013a).

A base de amostragem da primeira fase é implementada com intuito de estratificar a população, usando para tal informação auxiliar obtida por fotointerpretação, em que a cada ponto é atribuído um de 7 estratos de cobertura do solo apresentados na Tabela 3.3 (Jacques e Gallego, 2005).

Tabela 3.3: Sete estratos da primeira fase do esquema de amostragem LUCAS

Estrato	Designação
1	Terras cultivadas (<i>Arable Lands</i>)
2	Culturas Permanentes (<i>Permanent Crops</i>)
3	Pastagens (<i>Grassland</i>)
4	Áreas de Floresta e Matos (<i>Wooded areas and Shrubland</i>)
5	Solo nu, vegetação esparsa (<i>Bare land, Rare Vegetation</i>)
6	Territórios artificializados (<i>Artificial lands</i>)
7	Corpos de água (<i>Water</i>)

3.3.2. Segunda fase do esquema de amostragem

A partir dos pontos que compõem a amostra da primeira fase, um novo conjunto de pontos é extraído e classificado *in situ*, pelos observadores, de acordo com as classes de ocupação e uso do solo, apresentadas no Anexo 4.

A precisão da estratificação da primeira fase é avaliada nesta segunda fase pela comparação do estrato atribuído por fotointerpretação e as classes de ocupação do solo que foi atribuída ao ponto. Informação mais detalhada sobre os procedimentos que envolvem este processo de validação pode ser consultada em Martino *et al.* (2009).

Segundo os mesmos autores, com o intuito de reduzir a correlação espacial e obter amostras mais homogêneas, a seleção dos pontos da segunda fase é realizada através da maximização da distância dos pontos *intra* e *inter* estratos, evitando-se, assim, a recolha de informação redundante que influenciaria negativamente a precisão das estimativas.

Para alcançar este objetivo, a grelha de pontos que compõe a amostra da primeira fase do processo de amostragem começa por ser dividida em subgrelhas com 9x9 pontos formando segmentos quadrados de lado com 18 km (Figura 3.2).

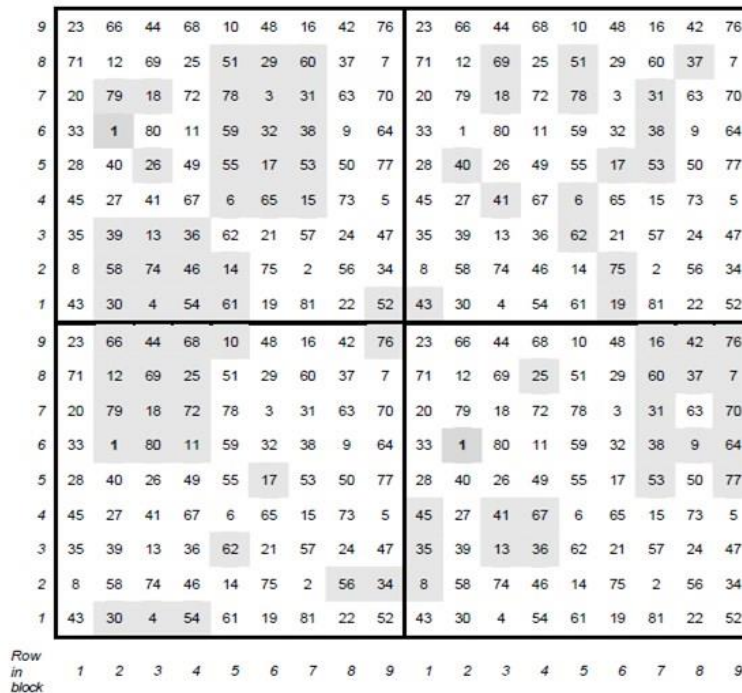


Figura 3.2: Subgrelhas obtidas a partir da grelha da primeira fase (em que os números representam réplicas de pontos) (retirado de Martino *et al.*, 2009).

A alocação da numeração dos pontos na grelha decorre pela implementação de um esquema de amostragem com balanceamento espacial, de acordo com seguinte algoritmo:

1. O primeiro ponto é selecionado aleatoriamente sendo-lhe atribuído o número 1;
2. O segundo ponto é selecionado aleatoriamente a partir do grupo de pontos que maximizam a distância relativamente ao primeiro ponto, sendo-lhe atribuído o número 2;
3. A seleção do terceiro ponto até ao ponto com o número 81, efetua-se sempre aleatoriamente a partir do grupo que maximiza a distância mínima relativamente aos pontos já selecionados (Figura 3.2).

A distância entre dois pontos usada neste algoritmo não é a euclideana, uma vez que, segundo Brus *et al.* (2011), a partir de comunicação pessoal de Gallego, originaria a uma tendência de concentração de pontos próximo das margens dos quadrados.

No referido trabalho de Brus *et al.* (2011), são apresentados dois exemplos das zonas que maximizam a distância relativamente ao ponto anterior caso se encontre localizado no canto inferior esquerdo ou no meio do lado esquerdo do quadrado (Figura 3.3).

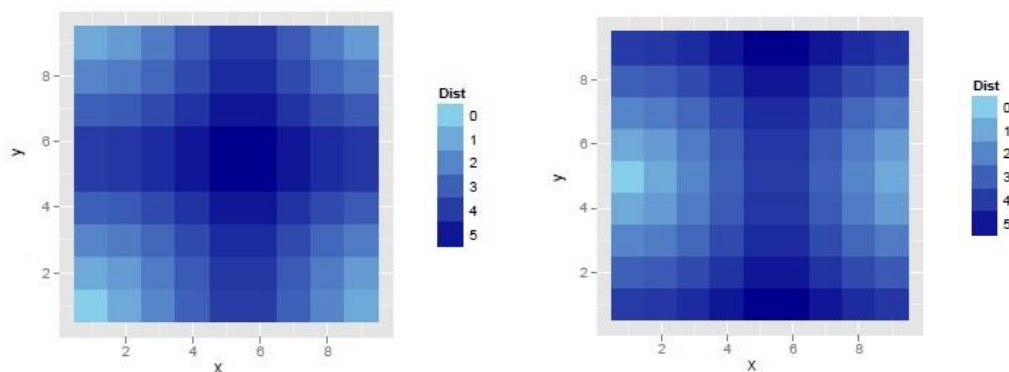


Figura 3.3: Indicação das localizações de acordo com as distâncias a um ponto localizado na extremidade à esquerda e no meio do lado esquerdo (adaptado de Brus *et al.*, 2011)

Utilizando este tipo de distância a probabilidade de inclusão de cada ponto na amostra é a mesma para cada uma das réplicas. A partir da Figura 3.4, obtida por simulação, observa-se que a probabilidade de inclusão das estimativas flutua em torno do valor 0.024691, podendo, portanto, considerar-se aleatório o padrão espacial.

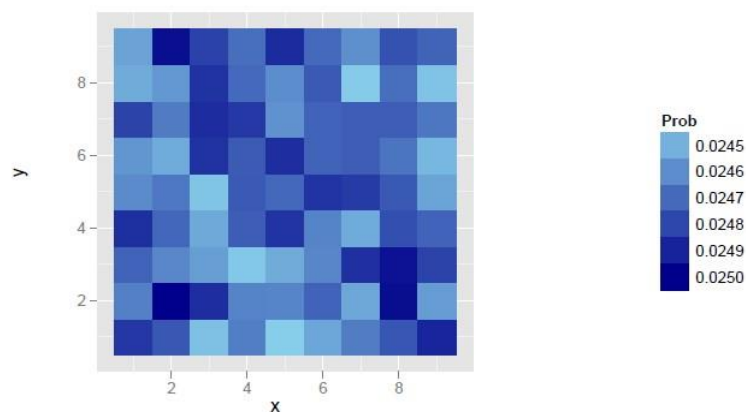


Figura 3.4: Probabilidades de inclusão estimadas usando o algoritmo implementado no projeto LUCAS (retirado de Brus *et al.*, 2011)

O algoritmo é implementado uma única vez e a posição dos pontos é utilizada com a mesma numeração em todos os segmentos. Aos números sombreados nas grelhas da Figura 3.2 significa que pertencem ao mesmo domínio.

As réplicas são, por fim, selecionadas sequencialmente a partir do número 1 até à réplica com o número da dimensão da amostra para cada domínio, dependendo das seguintes considerações:

- Para cada país, a dimensão da amostra da segunda fase é obtida através do valor do *Shannon Evenness Index* (SEI) por NUTS 2. Para países com um índice superior ao valor médio dos países europeus (SEI=0,64), houve um aumento da dimensão da amostra relativamente ao inquérito LUCAS anterior. Como para Portugal o valor SEI foi de 0.75, a dimensão da amostra da segunda fase correspondeu a 8 % do número de pontos da primeira fase (Palmieri, 2012);
- O número de pontos mínimos para cada domínio, ao nível das NUTS 2, foi fixado em 4 de forma a serem produzidas estimativas aceitáveis em termos de precisão (Martino e Palmieri, 2009).

3.4. Metodologia de estimação utilizada

Para este trabalho foi utilizada para a obtenção de estimativas das áreas das classes de ocupação e uso do solo uma metodologia de estimação baseada num estimador em duas fases proposto por Martino e Palmieri (2009) que tem sido implementado pelo Eurostat, desde 2006, para a produção de estimativas oficiais relativas aos dados provenientes do projeto LUCAS.

Para a metodologia de amostragem, o universo das localizações geográficas é composto pelo conjunto de todos os pontos obtidos através de uma grelha de pontos, obtidos a partir de uma grelha com pontos espaçados $1 \times 1 \text{ km}^2$ e que pertencem a uma determinada NUTS 1, NUTS 2 ou NUTS 3.

Portanto, a dimensão do espaço amostral pode variar conforme o nível a que se refere a NUTS. No caso deste estudo varia do nível 1, mais abrangente, correspondente a Portugal continental, até ao nível 3, nível mais pormenorizado, cuja descrição é apresentada no Anexo 1.

Em termos de terminologia, no que se refere à metodologia de estimação, há a necessidade de distinguir as seguintes notações:

- N' : número total de pontos no universo (composto pelos pontos sobrepostos numa grelha quadrada com pontos espaçados por $1 \times 1 \text{ km}$ pelos países participantes no projeto LUCAS);
- N : número total de pontos na NUTS correspondente à primeira fase do processo de amostragem (cujos pontos estão sobrepostos numa grelha quadrangular com pontos espaçados $2 \times 2 \text{ km}$);
- n : número total de pontos da amostrada segunda fase (na NUTS 1, NUTS 2 ou NUTS 3);
- N_h : número de pontos no universo da primeira fase pertencentes ao h -ésimo estrato (na NUTS 1, NUTS 2 ou NUTS 3);
- n_h : número de pontos na amostra da segunda fase pertencentes ao h -ésimo estrato (na NUTS 1, NUTS 2 ou NUTS 3);

3.4.1. Descrição dos tipos de pesos referentes à metodologia de estimação

Como foi referido no esquema de amostragem, para cada NUTS 1, NUTS 2 ou NUTS 3, os pontos da primeira fase são divididos em sete estratos. Na segunda fase, cada ponto é classificado segundo a nomenclatura do sistema de classificação de uso e ocupação do solo do projeto LUCAS.

Para a implementação do estimador para a obtenção das estimativas das áreas das classes de uso e ocupação do solo, há a necessidade de calcular os valores associados a dois tipos de pesos:

- Pesos dos estratos (*weights for stratification*), w_h , que compõem cada NUTS;
- Pesos da segunda fase do esquema de amostragem, $peso_f$, calculados segundo um conjunto de regras que envolvem a seleção de pontos.

Pesos dos estratos na primeira fase (w_h)

A partir da fotointerpretação dos pontos e da estratificação da população em sete estratos, para NUTS 1, NUTS 2 ou NUTS 3, é possível obter diferentes tipos de domínios (dados pela combinação estrato/NUTS).

A cada um destes domínios está associado um peso, designado neste trabalho por peso do estrato (*weights for stratification*) que é representado por w_h .

O valor w_h é obtido, segundo Martino e Palmieri (2009), por

$$w_h = \frac{N_h}{N} \quad (3.1)$$

em que N_h é o número de pontos no h -ésimo estrato na NUTS 1, 2 ou 3 e N é o número total de pontos que pertencem à NUTS 1, 2 ou 3.

Somando os valores dos pesos para cada estrato para uma NUTS, obviamente se observa que

$$\sum_{h=1}^H w_h = 1 \quad (3.2)$$

Peso do ponto na segunda fase ($peso_f2$)

Neste trabalho os pesos da amostra na segunda fase de amostragem, $peso_f2$, foram obtidos a partir das instruções fornecidas pelo Eurostat. Segundo Martino e Palmieri (2009), na primeira fase do processo de amostragem (para cada domínio) os pontos pertencem a dois subconjuntos $S_{h,1}$ e $S_{h,2}$:

- O subconjunto $S_{h,1}$ é composto por pontos localizados em áreas com altitudes acima de 1000 metros (em 2012, essa altitude foi alterada para 1500) e/ou pontos localizados em pequenas ilhas sem ligações dadas por pontes (número total de pontos de $S_{h,1}$ é representado por $N_{h,1}$);
- $S_{h,2}$ é formado pelos pontos que se localizam em áreas com altitude inferior a 1000 m (1500 m, no caso do LUCAS 2012) e que não

pertençam a pequenas ilhas sem ligação por pontes (número total de pontos de $S_{h,2}$ é representado por $N_{h,2}$).

Para além disso, $S_{h,2}$, para o esquema de amostragem do LUCAS para o ano de 2012, também se dividiu em outros dois subconjuntos:

- $S_{h,21}$: subconjunto de pontos já observados em 2009 (cujo número total de pontos é dado por $N_{h,21}$);
- $S_{h,22}$: subconjunto formado pelos pontos que não foram observados em 2009 (cujo número total de pontos é dado por $N_{h,22}$).

Consequentemente, o número total de pontos localizados em cada NUTS 1, NUTS 2 ou NUTS 3, para o h -ésimo estrato, é dado por

$$N_h = N_{h,1} + N_{h,2} = N_{h,1} + (N_{h,21} + N_{h,22}) \quad (3.3)$$

Segundo os mesmos autores, a segunda fase do esquema de amostragem foi conduzida de acordo com os seguintes critérios:

- Exclusão de pontos pertencentes ao subconjunto $S_{h,1}$ (*missing by design*);
- Inclusão de tantos pontos quanto possível que pertençam ao subconjunto $S_{h,21}$;
- Inclusão de pontos pertencentes ao subconjunto $S_{h,22}$ só depois da inclusão exaustiva de pontos no subconjunto $S_{h,21}$ (*partially missing by design*).

Para o cálculo do $peso_{f2}$, de cada ponto pertencente a um determinado estrato e com uma classe de ocupação do solo associada, é obtido pelo produto de peso de amostragem (*sampling weight*), $w_{h,21}$ e $w_{h,22}$, e *Weight for missing data adjustment* ($s_{h,2}$).

Por sua vez, os valores de $s_{h,2}$, podem ser obtidos de duas formas:

$$s_{h,2} = \begin{cases} \frac{N_h}{N_{h,2}}, & N_{h,21} = 0 \text{ ou } N_{h,21} < n_h \\ \frac{N_h}{N_{h,21}}, & N_{h,21} \geq n_h \end{cases} \quad (3.4)$$

Assim, sumariando, os pesos, para cada domínio, da segunda fase do esquema de amostragem ($peso_{f2}$) são calculados de acordo com a Tabela 3.4.

Tabela 3.4: Pesos da segunda fase, *peso_f2*, atribuído ao ponto, de acordo o subconjunto do esquema de amostragem LUCAS a que pertença (Martino e Palmieri, 2009).

Subconjunto		Dimensão(N)		Descrição	Pesos da segunda fase de amostragem
$S_{h,1}$		$N_{h,1}$		Pontos com altitudes acima de 1000 m (1500 m, no LUCAS) e/ou pertencentes a pequenas ilhas sem ligações através de pontes.	$p_{h,1} = 0$
$S_{h,2}$		$N_{h,2}$		Pontos com altitudes abaixo de 1000 m (1500 m, no LUCAS) e não pertencentes a pequenas ilhas sem ligações através de pontes.	
	$S_{h,21}$		$N_{h,21}$	Subconjunto de $S_{h,2}$ composto por pontos já observados no inquérito LUCAS realizado em 2009.	$p_{h,21} = w_{h,21} \times S_{h,2} = \begin{cases} 0, & N_{h,21} = 0 \\ 1 \times \frac{N_h}{N_{h,2}}, & N_{h,21} < n_h \\ \frac{N_{h,21}}{n_h} \times \frac{N_h}{N_{h,21}}, & N_{h,21} \geq n_h \end{cases}$
	$S_{h,22}$		$N_{h,22}$	Subconjunto de $S_{h,2}$ composto por pontos não observados no inquérito LUCAS realizado em 2012.	$p_{h,22} = w_{h,22} \times S_{h,2} = \begin{cases} \frac{N_{h,22}}{n_h} \times \frac{N_h}{N_{h,2}}, & N_{h,21} = 0 \\ \frac{N_{h,22}}{n_h - N_{h,21}} \times \frac{N_h}{N_{h,2}}, & N_{h,21} < n_h \\ 0, & N_{h,21} \geq n_h \end{cases}$

Distribuição de pontos pelos subconjuntos e cálculo dos pesos w_h e *peso_f2* por NUTS 2

LUCAS 2009

Na Tabela 3.5 é apresentada, a distribuição do número de pontos por cada domínio (estrato/NUTS 2), para os subconjuntos, $S_{h,1}$ e $S_{h,2}$, a partir da implementação das regras definidas na Tabela 3.4.

Tabela 3.5: Número de pontos para cada subconjunto $S_{h,1}$ e $S_{h,2}$ (para Portugal em 2009)

Estrato	Nh,2						Nh,1					
	NUTS2					TOTAL	NUTS2					TOTAL
	PT11	PT15	PT16	PT17	PT18		PT11	PT15	PT16	PT17	PT18	
1	667	110	777	152	2815	4521	4	0	5	0	0	9
2	562	191	628	53	532	1966	1	0	0	0	0	1
3	966	232	903	100	615	2816	115	0	38	1	1	155
4	2481	566	3940	258	3487	10732	110	0	76	0	0	186
5	85	59	250	18	162	574	15	2	16	0	1	34
6	265	53	348	104	199	969	4	0	2	0	0	6
7	44	21	69	22	73	229	1	2	1	30	25	59
TOTAL	5070	1232	6915	707	7883	21807	250	4	138	31	27	450

Uma vez que no inquérito LUCAS de 2009 foi implementado oficialmente pela primeira vez em Portugal, todos os 5428 pontos selecionados na segunda fase do esquema de amostragem pertencem ao subconjunto $S_{h,22}$ de $S_{h,2}$. Assim, no cálculo dos pesos da segunda fase, $peso_f2$, há apenas a considerar o tipo $p_{h,22}$.

De seguida é apresentado um exemplo de como se processa o cálculo dos pesos para as duas fases do esquema de amostragem, para o caso de um ponto que pertença à NUTS 2 Norte (PT11) e à classe das culturas Permanentes (estrato 1):

- O peso do estrato na primeira fase é obtido por

$$w_h = \frac{667 + 4}{5070 + 250} = \frac{671}{5320} \approx 0,126.$$

- Como no inquérito do projeto LUCAS realizado em 2009, não há a registar pontos que tenham sido observados em inquéritos anteriores tem-se que $N_{h,21} = 0$ e $p_{h,21} = 0$. Consequentemente, o peso da segunda fase, $peso_f2$, é dado por

$$p_{h,22} = \frac{667}{171} \times \frac{671}{667} \approx 3,923.$$

LUCAS 2012

Na Tabela 3.6 é apresentada, a distribuição do número de pontos por cada domínio (estrato/NUTS2), para os subconjuntos, $S_{h,1}$ e $S_{h,2}$, a partir da implementação das regras definidas na Tabela 3.4.

Tabela 3.6: Número de pontos para cada subconjunto $S_{h,1}$ e $S_{h,2}$ (para Portugal em 2012)

Estrato	Nh,2						Nh,1					
	NUTS2					TOTAL	NUTS2					TOTAL
	PT11	PT15	PT16	PT17	PT18		PT11	PT15	PT16	PT17	PT18	
1	585	91	673	103	1994	3446	86	19	109	49	821	1084
2	431	159	520	40	419	1569	132	32	108	13	113	398
3	634	111	662	76	410	1893	447	121	279	25	206	1078
4	1429	223	2579	155	2080	6466	1162	343	1437	103	1407	4452
5	50	24	158	14	108	354	50	37	108	4	55	254
6	250	43	318	99	173	883	19	10	32	5	26	92
7	20	9	31	13	34	107	25	14	39	39	64	181
TOTAL	3399	660	4941	500	5218	14718	1921	576	2112	238	2692	7539

Uma vez que no inquérito LUCAS de 2012, para cada domínios dado pela combinação estrato/NUTS 2, o conjunto $S_{h,2}$ passou a ser formado por pontos pertencentes aos subconjuntos $S_{h,21}$ (terem sido observados em 2009) e $S_{h,22}$ (não terem sido observados em 2009), conseqüente, para o cálculo dos pesos da segunda fase do esquema de amostragem, $peso_{f2}$, há a necessidade de considerar, respetivamente para $S_{h,21}$ e $S_{h,22}$, os pesos $p_{h,21}$ e $p_{h,22}$,

Para cada domínio, o número de pontos $N_{h,21}$ e $N_{h,22}$ de $N_{h,2}$, distribuem-se segundo os valores apresentados na Tabela 3.7.

Tabela 3.7: Número de pontos $N_{h,21}$ e $N_{h,22}$ para cada domínio (para Portugal, em 2012)

Estrato	NUTS 2															TOTAL
	PT11			PT15			PT16			PT17			PT18			
	Nh,21	Nh,22	Total	Nh,21	Nh,22	Total	Nh,21	Nh,22	Total	Nh,21	Nh,22	Total	Nh,21	Nh,22	Total	
1	170	415	585	27	64	91	188	485	673	33	70	103	643	1351	1994	3446
2	140	291	431	48	111	159	152	368	520	13	27	40	131	288	419	1569
3	191	443	634	40	71	111	192	470	662	23	53	76	135	275	410	1893
4	491	938	1429	108	115	223	783	1796	2579	53	102	155	690	1390	2080	6466
5	15	35	50	8	16	24	43	115	158	3	11	14	31	77	108	354
6	66	184	250	13	30	43	85	233	318	24	75	99	47	126	173	883
7	4	16	20	5	4	9	5	26	31	3	10	13	9	25	34	107
TOTAL	1077	2322	3399	249	411	660	1448	3493	4941	152	348	500	1686	3532	5218	14718

De seguida é apresentado um exemplo de como se processa o cálculos dos pesos para as duas fases do esquema de amostragem, para o caso de um ponto que pertença à NUTS 2 centro (PT16) e pastagens (estrato 3):

- O peso do estrato na primeira fase é obtido por

$$w_h = \frac{662 + 279}{4941 + 2112} = \frac{941}{7053} \approx 0,133.$$

- Caso o ponto selecionado na segunda fase do esquema de amostragem tenha sido observado também em 2009, como $N_{h,21}$ (192 pontos) < n_h (303 pontos), em que n_h é o número de pontos da segunda fase pertencentes ao domínio (PT 16/estrato 3), $peso_f2$, é obtido calculando

$$p_{h,21} = \frac{662 + 279}{662} = \frac{941}{662} \approx 1,421.$$

- Caso o ponto selecionado na segunda fase do esquema de amostragem tenha sido observado pela primeira vez em 2012, como $N_{h,21}$ (192 pontos) < n_h (303 pontos), $peso_f2$, é obtido calculando

$$p_{h,22} = \frac{470}{303 - 192} \times \frac{941}{662} \approx 6,019$$

3.4.2. Cálculo dos pesos w_h e $peso_f2$ por NUTS 1 e NUTS 3

Neste trabalho foram igualmente calculados os pesos para os estratos na primeira fase do esquema de amostragem (w_h) e os pesos da segunda fase, $peso_f2$, ao nível das NUTS 1 e NUTS 3 para a obtenção das estimativas das áreas das classes de uso e ocupação definidas na nomenclatura do sistema de classificação do projeto LUCAS.

Para tal, foram utilizados os procedimentos descritos no subcapítulo anterior, pela extensão do conceito de domínio, passando este a ser definido pelo conjunto de pontos que pertencem simultaneamente ao mesmo estrato e NUTS 1 ou NUTS 3.

Assim, recorrendo à linguagem de programação R, procedeu-se à obtenção dos pesos, cujos primeiros dez valores da tabela de dados LUCAS, para 2009 e 2012, se encontram nas Tabelas 3.8 e 3.9.

Tabela 3.8: Pesos w_h e peso_f2 ao nível das NUTS 3 (2009)

POINT_ID	STRATA (Estratos)	NUTS 1	NUTS 2	NUTS 3	wh (NUTS 1)	peso_f2 (NUTS 1)	wh (NUTS 3)	peso_f2 (NUTS 3)
26381958	4	PT1	PT17	PT171	0,491	4,089	0,279	4,318
26401764	4	PT1	PT15	PT150	0,491	4,089	0,458	4,072
26401954	6	PT1	PT17	PT171	0,044	4,046	0,179	4,692
26421762	3	PT1	PT15	PT150	0,133	4,226	0,188	4,070
26421766	3	PT1	PT15	PT150	0,133	4,226	0,188	4,070
26421960	4	PT1	PT17	PT171	0,491	4,089	0,279	4,318
26421964	3	PT1	PT17	PT171	0,133	4,226	0,185	3,316
26441952	3	PT1	PT17	PT171	0,133	4,226	0,185	3,316
26441956	3	PT1	PT17	PT171	0,133	4,226	0,185	3,316

Tabela 3.9: Pesos w_h e peso_f2 ao nível das NUTS 3 (2012)

POINT_ID	STRATA (Estratos)	NUTS 1	NUTS 2	NUTS 3	wh (NUTS 1)	peso_f2 (NUTS 1)	wh (NUTS 3)	peso_f2 (NUTS 3)
26541954	1	PT1	PT17	PT171	0,204	1,315	0,232	1,549
26541962	3	PT1	PT17	PT171	0,133	5,626	0,185	7,875
26541972	1	PT1	PT17	PT171	0,204	1,315	0,232	1,549
26541990	6	PT1	PT16	PT16B	0,044	1,104	0,070	1,026
26561766	1	PT1	PT15	PT150	0,204	1,315	0,089	1,209
26561770	4	PT1	PT15	PT150	0,491	1,689	0,458	2,538
26561778	4	PT1	PT15	PT150	0,491	1,689	0,458	2,538
26561782	4	PT1	PT15	PT150	0,491	4,913	0,458	3,649
26561784	4	PT1	PT15	PT150	0,491	1,689	0,458	2,538

3.4.3. Estimadores para a proporção e área total

A estimativa da proporção de uma l -ésima classe de uso ou ocupação do solo (p^l), para uma dada NUTS decorre da aplicação da seguinte fórmula (Martino e Palmieri, 2009)

$$p^l = \sum_{h=1}^7 w_h \left(\frac{\sum_{i=1}^{n_h} I_i^l \times p_{i,hj}}{\sum_{i=1}^{n_h} p_{i,hj}} \right) \quad (3.5)$$

em que:

- p^l : proporção de território para cada NUTS 1, NUTS 2 ou NUTS 3 coberto pela l -ésima classe de ocupação ou uso do solo;
- w_h : peso do h -ésimo estrato;

- $p_{i,hj}$: peso da i -ésima observação pertencente ao h -ésimo estrato e ao j -ésimo subconjunto;
- I_i^l : variável que tem valor 1 (se a ocupação do solo observada no i -ésimo ponto for da l -ésima classe) ou valor 0 (caso contrário).

A estimativa da área para l -ésima classe de ocupação ou uso do solo (a^l), para uma dada NUTS 1, NUTS 2 ou NUTS 3, é obtida multiplicando a proporção estimada para l -ésima classe pela área da NUTS correspondente (Gallego, 2013)

$$a^l = \text{Area}_{NUTS} \times p^l \quad (3.6)$$

onde Area_{NUTS} é área correspondente à NUTS.

3.4.4. Estimador para a variância da proporção e área total

De acordo Cochran (1977), a variância associada à estimativa para uma l -ésima classe de uso ou ocupação do solo, $\text{var}(p^l)$, para uma dada NUTS 1, NUTS 2 ou NUTS 3 é obtida a partir do estimador para duas fases com estratificação dada por

$$\text{var}(p^l) = \sum_{h=1}^7 \frac{w_h^2 \times s_h^2}{n_h} - \sum_{h=1}^7 \frac{w_h^2 \times s_h^2}{N'} + \frac{g'}{N} \sum_{h=1}^7 w_h (p_h^l - p^l), \quad (3.7)$$

em que

$$g' = \frac{N' - N}{N' - 1}. \quad (3.8)$$

$$s_h^2 = \frac{n_h p_h (1 - p_h)}{n_h - 1}, \quad (3.9)$$

e

- N' : corresponde ao número total de pontos no universo (sobrepastos na grelha de pontos espaçados 1x1 km² por todos os países onde o projeto LUCAS foi implementado);
- N : corresponde ao número total de pontos na NUTS 1, NUTS 2 ou NUTS 3 na primeira fase do processo de amostragem (cujos pontos estão sobrepostos numa grelha quadrangular com pontos espaçados 2x2 km);
- n_h : número de pontos da segunda fase pertencentes ao h -ésimo estrato na NUTS 1, NUTS 2 ou NUTS 3;
- p_h^l : proporção da l -ésima classe de uso ou ocupação do solo no h -ésimo estrato na NUTS 1, NUTS 2 ou NUTS 3.

Segundo Gallego (2013), a variância das estimativas da área para l -ésima classe de ocupação ou uso do solo, $\text{var}(a^l)$, para uma dada NUTS 1, NUTS 2 ou NUTS 3, de acordo com um estimador estratificado é dado por

$$\text{var}(a^l) = \text{Area}_{NUTS}^2 \times \text{var}(p^l). \quad (3.10)$$

No capítulo seguinte são apresentados os resultados da aplicação dos estimadores referidos em 3.6 e 3.10 para a obtenção das estimativas para as NUTS 1, NUTS 2 e NUTS 3 a partir dos dados recolhidos nos inquéritos LUCAS de 2009 e 2012.

Capítulo 4

Resultados obtidos

4.1. Análise exploratória da base de dados LUCAS *survey*

Os dados recolhidos em 2009 e 2012 no projeto LUCAS podem ser consultados a partir das bases de dados disponibilizadas na página *Web*² do Eurostat relacionado com o projeto (Eurostat, 2012a).

O número de registos das bases de dados corresponde ao número de pontos observados nos dois anos em que foram efetuadas recolha de dados para a obtenção de estatísticas oficiais. A descrição das variáveis e a forma como se procedeu ao preenchimento dos dados que compõem a base de dados é apresentada no documento de referência técnica do LUCAS (Eurostat, 2013b).

Aquando da observação *in situ* dos pontos georeferenciados, os dados recolhidos foram registados pelos observadores de acordo com o formulário de campo associado ao projeto LUCAS (Eurostat, 2012b). Portanto, a cada ponto (ou localização geográfica) é-lhe atribuído, para cada uma das duas fases do esquema de amostragem LUCAS, um valor associado às variáveis.

De entre as variáveis que compõem a base de dados LUCAS (designada neste estudo por LUCAS *survey*), destacam-se, pela sua importância na obtenção das estimativas das áreas das classes de ocupação e uso do solo de acordo com a nomenclatura de classificação definida pelo projeto LUCAS, as apresentadas na Tabela 4.1.

Pelos valores atribuídos às variáveis LC1 e LC2, observa-se que em ambos os questionários, os observadores atribuíram ao ponto dois tipos de classes de ocupação do solo: em 2009 cultura permanente de nozes (B74) e vegetação espontânea (E30) e em 2012 cultura permanente de nozes(B74) e prados com árvores esparsas (E10). O mesmo acontece com as variáveis LU1 e LU2 em que foram atribuídas duas classes de uso do solo, produção agrícola (U111) e infra-estruturas agrícolas (U112).

²http://epp.eurostat.ec.europa.eu/portal/page/portal/lucas/data/LUCAS_primary_data/2012

Tabela 4.1: Descrição das variáveis da base de dados LUCAS *survey* envolvidas no processo de estimação e exemplo dos valores para o ponto com ID=26601762

Variáveis	Descrição da variável	Tipo de variável	Categorias	2009	2012
<i>ID</i>	Número de identificação do ponto	Discreta		26601762	26601762
<i>STRATA</i>	Tipo de estrato	Categórica	1: Terras cultivadas; 2: Culturas permanentes; 3: Pastagens; 4: Areas de florestas e matos; 5: Solo nu, vegetação esparsa; 6: territórios artificializados; 7: corpos de água.	1	1
<i>NUTS 1</i> <i>NUTS 2</i> <i>NUTS 3</i>	Designação da NUTS	Categórica	Descrição das NUTS (ver Anexo 1)	PT1 PT15 PT150	PT1 PT15 PT150
<i>AREA_NUTS1</i> <i>AREA_NUTS2</i> <i>AREA_NUTS3</i>	Área de cada NUTS	Contínua	Descrição das áreas de cada NUTS (Anexo 1)	89089 4997 4997	89089 4997 4997
<i>LC1</i>	Primeira classe de ocupação do solo atribuída ao ponto	Categórica	Descrição das classes de ocupação do solo (Anexo 4)	B74	B74
<i>LC2</i>	Segunda classe de ocupação do solo atribuída ao ponto	Categórica	Descrição das classes de ocupação do solo (Anexo 4)	E30	E10
<i>LU1</i>	Primeira classe de uso do solo atribuída ao ponto	Categórica	Decrição das classes de uso do solo (Anexo 4)	U111	U111
<i>LU2</i>	Segunda classe de uso do solo atribuída ao ponto	Categórica	Descrição das classes de uso do solo (ver Anexo 5)	U112	U112
<i>OBS_RADIUS</i>	Dimensão do raio do círculo definido em torno do ponto	Categórica	1= 1.5m; 2=20m	2	2
<i>AREA_SIZE</i>	Dimensão da área em torno do ponto	Categórica	1: [0,0.5ha]; 2: [0.5ha, 1ha]; 3: [1ha, 10ha]; 4:[10ha, +∞]; 8: N.R.	2	3
<i>wh</i>	Peso do estrato na primeira fase (NUTS 2)	Contínua		0,0890	4,0732
<i>peso_f2</i>	Peso do ponto na segunda fase (NUTS 2)	Contínua		0,0890	1,2087

4.1.1. Tipo de estrato (variável STRATA)

4.1.1.1. Distribuição do tipo de estrato na amostra da primeira fase

Nos dois anos em que foi implementado em Portugal o projeto LUCAS, 2009 e 2012, os 22257 pontos que compõem a primeira fase do esquema de amostragem foram fotointerpretados e classificados de acordo com sete estratos, cuja distribuição pelos domínios (estrato/NUTS 2) é apresentada na Tabela 4.2.

Tabela 4.2: Distribuição de pontos por domínio (estrato/NUTS2) na primeira fase do esquema de amostragem LUCAS para Portugal

Estratos		PT11		PT15		PT16		PT17		PT18		Total	%
		Norte		Algarve		Centro		Lisboa		Alentejo			
1	<i>Terras cultivadas</i>	671	13%	110	9%	782	11%	152	21%	2815	36%	4530	20,4%
2	<i>Culturas permanentes</i>	563	11%	191	15%	628	9%	53	7%	532	7%	1967	8,8%
3	<i>Pastagens</i>	1081	20%	232	19%	941	13%	101	14%	616	8%	2971	13,3%
4	<i>Florestas e matos</i>	2591	49%	566	46%	4016	57%	258	35%	3487	44%	10918	49,1%
5	<i>Solo nu e vegetação esparsa</i>	100	2%	61	5%	266	4%	18	2%	163	2%	608	2,7%
6	<i>Territórios artificializados</i>	269	5%	53	4%	350	5%	104	14%	199	3%	975	4,4%
7	<i>Corpos de água</i>	45	1%	23	2%	70	1%	52	7%	98	1%	288	1,3%
Total		5320	100%	1236	100%	7053	100%	738	100%	7910	100%	22257	
Percentagem		23,9%		5,6%		31,7%		3,3%		35,5%			

Pela análise da Tabela 4.2, observa-se que aproximadamente 50% do território português é coberto por florestas e matos (estrato 4) e que existe igualmente uma presença relevante de terras cultivadas (estrato 1) com 20,4 %, e de pastagens (estrato 3) com 13,3 %.

Em termos de distribuição de pontos ao nível das NUTS 2, confirma-se a proeminência de florestas e matos (estrato 4) por todo país relativamente aos outros estratos. Observa-se um aumento da percentagem de ocupação de terras cultivadas (estrato 1) no Alentejo (PT18) e em Lisboa (PT17) comparativamente ao que ocorre no resto do país.

Pela análise da Tabela 4.3 onde é apresentada a distribuição de pontos pelos sete estratos ao nível das NUTS 3, observa-se que, excetuando o caso do Baixo Alentejo (PT184), em que terras cultivadas (estrato 1) tem uma maior predominância em relação aos outros estratos, para todas as outras NUTS 3,

florestas e matos (estrato 4) é o que apresenta maior distribuição (maioritariamente acima de 50%).

Terras cultivadas (estrato 1), culturas permanentes (estrato 2) e pastagens (estrato 3) têm distribuições que variam na maioria dos casos abaixo de 25%. Relativamente a terras cultivadas (estrato 1), a situação muda consideravelmente para as NUTS 3 pertencentes a Lisboa (PT17) e Alentejo (PT18). Nestes estes casos, observam-se os valores mais altos para este tipo de estrato.

Tabela 4.3: Distribuição de pontos por domínio (estrato/NUTS 3) na primeira fase do esquema do esquema de amostragem LUCAS para Portugal continental (em 2009 e 2012)

NUTS3	Estratos														Total	%
	1		2		3		4		5		6		7			
PT111	45	8%	16	3%	161	29%	277	50%	24	4%	26	5%	7	1%	556	2,5%
PT112	44	14%	7	2%	43	14%	176	57%	10	3%	28	9%	2	1%	310	1,4%
PT113	42	14%	8	3%	53	17%	167	54%	6	2%	31	10%	4	1%	311	1,4%
PT114	42	21%	7	3%	12	6%	80	40%	4	2%	54	27%	3	1%	202	0,9%
PT115	55	8%	45	7%	151	23%	351	53%	12	2%	39	6%	5	1%	658	3,0%
PT116	30	14%	2	1%	16	7%	140	65%	4	2%	23	11%	0	0%	215	1,0%
PT117	78	8%	226	22%	231	22%	453	44%	12	1%	22	2%	10	1%	1032	4,6%
PT118	335	16%	252	12%	414	20%	949	47%	28	1%	46	2%	14	1%	2038	9,2%
PT150	110	9%	191	15%	232	19%	566	46%	61	5%	53	4%	23	2%	1236	5,6%
PT161	59	13%	10	2%	45	10%	239	53%	30	7%	45	10%	24	5%	452	2,0%
PT162	89	17%	42	8%	47	9%	278	54%	22	4%	30	6%	6	1%	514	2,3%
PT163	41	9%	37	8%	56	13%	255	58%	16	4%	31	7%	1	0%	437	2,0%
PT164	26	4%	34	5%	54	8%	486	74%	24	4%	25	4%	4	1%	653	2,9%
PT165	116	13%	21	2%	131	15%	533	61%	35	4%	32	4%	2	0%	870	3,9%
PT166	10	2%	23	5%	26	5%	366	77%	21	4%	24	5%	4	1%	474	2,1%
PT167	25	11%	7	3%	42	19%	128	58%	7	3%	10	5%	1	0%	220	1,0%
PT168	99	10%	92	9%	271	27%	494	49%	22	2%	36	4%	2	0%	1016	4,6%
PT169	163	17%	95	10%	105	11%	506	54%	33	4%	18	2%	12	1%	932	4,2%
PT16A	54	16%	24	7%	39	11%	195	56%	11	3%	19	5%	4	1%	346	1,6%
PT16B	60	11%	142	26%	78	14%	212	38%	20	4%	39	7%	5	1%	556	2,5%
PT16C	40	7%	101	17%	47	8%	322	55%	25	4%	41	7%	5	1%	581	2,6%
PT171	79	23%	19	6%	63	19%	95	28%	8	2%	61	18%	15	4%	340	1,5%
PT172	73	18%	34	8%	39	10%	163	40%	11	3%	43	11%	43	11%	406	1,8%
PT181	329	25%	20	2%	121	9%	742	56%	37	3%	42	3%	31	2%	1322	5,9%
PT182	435	28%	120	8%	154	10%	748	48%	55	4%	35	2%	16	1%	1563	7,0%
PT183	704	39%	108	6%	132	7%	791	44%	20	1%	40	2%	14	1%	1809	8,1%
PT184	1099	51%	167	8%	145	7%	641	30%	38	2%	33	2%	23	1%	2146	9,6%
PT185	248	23%	117	11%	63	6%	565	53%	12	1%	49	5%	8	1%	1062	4,8%
TOTAL	4530	20%	1967	9%	2971	13%	10918	49%	608	3%	975	4%	288	1%	22257	100,0%

4.1.1.2. Distribuição do tipo de estrato na amostra da segunda fase

A distribuição do número de pontos por domínio na segunda fase do processo de amostragem foi realizada de forma a obter amostras espacialmente homogéneas e com uma menor autocorrelação entre os pontos (Brus *et al.* 2011).

Para Portugal, em 2009, foram observados na segunda fase 5428 pontos, cuja distribuição é apresentada na Tabela 4.4. Observa-se que os resultados obtidos não diferem consideravelmente, em percentagem, dos obtidos na primeira fase.

Tabela 4.4: Distribuição de pontos por domínio (estrato/NUTS 2) na segunda fase do esquema do esquema de amostragem LUCAS para Portugal continental (em 2009)

Estratos		PT11		PT15		PT16		PT17		PT18		Total	%
		Norte		Algarve		Centro		Lisboa		Alentejo			
1	<i>Terrenos cultiváveis</i>	171	13%	27	9%	194	11%	38	22%	686	36%	1116	20,6%
2	<i>Culturas permanentes</i>	144	11%	49	16%	156	9%	13	7%	135	7%	497	9,2%
3	<i>Pastagens</i>	247	19%	57	19%	225	13%	25	14%	149	8%	703	13,0%
4	<i>Florestas e matos</i>	636	49%	139	45%	983	57%	63	36%	849	44%	2670	49,2%
5	<i>Solo nu e vegetação esparsa</i>	21	2%	15	5%	59	3%	4	2%	41	2%	140	2,6%
6	<i>Territórios artificializados</i>	68	5%	13	4%	87	5%	25	14%	48	2%	241	4,4%
7	<i>Corpos de água</i>	13	1%	6	2%	18	1%	6	3%	18	1%	61	1,1%
Total		1300	100%	306	100%	1722	100%	174	100%	1926	100%	5428	
%		23,9%		5,6%		31,7%		3,2%		35,5%			

De acordo com Palmieri *et al.* (2011), Portugal é um país com uma distribuição espacial heterogénea em termos paisagísticos. Este facto é refletido no valor do Índice Shannon para Portugal (0.75) que é superior à média Europeia. Tendo em consideração este facto, para os dados recolhidos no LUCAS de 2012 houve um aumento do número de pontos observados na segunda fase de amostragem (7336 pontos). A distribuição dos pontos por tipo de estrato ao nível das NUTS 2 é apresentada na Tabela 4.5.

Tabela 4.5: Distribuição de pontos por domínio (estrato/NUTS 2) na segunda fase do esquema do esquema de amostragem LUCAS para Portugal continental (em 2012)

Estratos		PT11		PT15		PT16		PT17		PT18		Total	%
		Norte		Algarve		Centro		Lisboa		Alentejo			
1	<i>Terras cultivadas</i>	231	13%	36	9%	262	11%	51	22%	929	36%	1509	20,6%
2	<i>Culturas permanentes</i>	195	11%	66	16%	211	9%	17	7%	182	7%	671	9,1%
3	<i>Pastagens</i>	333	19%	77	19%	303	13%	33	14%	201	8%	947	12,9%
4	<i>Florestas e matos</i>	861	49%	188	46%	1333	57%	85	37%	1150	44%	3617	49,3%
5	<i>Solo nu e vegetação esparsa</i>	28	2%	20	5%	79	3%	5	2%	55	2%	187	2,5%
6	<i>Territórios artificializados</i>	92	5%	17	4%	117	5%	33	14%	65	2%	324	4,4%
7	<i>Corpos de água</i>	17	1%	8	2%	24	1%	8	3%	24	0,9%	81	1,1%
Total		1757	100%	412	100%	2329	100%	232	100%	2606	100%	7336	
%		24,0%		5,6%		31,7%		3,2%		35,5%			

Os resultados da Tabela 4.6 confirmam que com a desagregação das NUTS 2 por NUTS 3, o número de pontos para alguns estratos diminuiu consideravelmente, observando-se a existência de vários domínios (estrato/NUTS 3) em que não foram extraídos pontos na segunda fase do processo de amostragem. Esta situação deve-se ao facto de amostra ter sido definida para a produção de estimativas ao nível das NUTS2 (Gallego, 2013). Tal facto, como se verá neste capítulo, condicionará a precisão das estimativas obtidas ao nível das NUTS 3 neste trabalho.

Tabela 4.6: Distribuição de pontos por domínio (estrato/NUTS 3) na segunda fase do esquema do esquema de amostragem LUCAS para Portugal continental

2009								Total	2012								Total
NUTS3	Estratos								NUTS3	Estratos							
	1	2	3	4	5	6	7		1	2	3	4	5	6	7		
PT111	11	5	35	67	7	7	2	134	PT111	17	5	35	86	10	11	4	168
PT112	10	0	8	43	1	6	1	69	PT112	16	2	13	61	0	9	0	101
PT113	10	4	12	39	2	11	2	80	PT113	15	4	23	61	4	14	1	122
PT114	14	4	4	25	2	8	1	58	PT114	17	5	6	37	3	12	2	82
PT115	12	15	35	84	2	11	0	159	PT115	17	20	45	105	2	14	1	204
PT116	10	1	3	30	0	7	0	51	PT116	13	2	8	48	0	8	0	79
PT117	16	54	48	127	4	8	3	260	PT117	17	68	73	147	1	9	4	319
PT118	88	61	102	221	3	10	4	489	PT118	119	89	130	317	8	15	5	683
PT150	27	49	57	139	15	13	6	306	PT150	36	66	77	188	20	17	8	412
PT161	15	1	15	62	7	12	6	118	PT161	18	3	24	89	8	13	7	162
PT162	19	14	12	65	7	8	2	127	PT162	24	20	16	112	10	13	4	199
PT163	8	14	16	59	4	9	0	110	PT163	15	18	18	98	7	16	1	173
PT164	5	6	14	129	3	4	1	162	PT164	8	8	22	150	7	6	1	202
PT165	31	10	20	127	10	10	0	208	PT165	44	12	32	178	13	10	1	290
PT166	2	8	5	90	5	4	2	116	PT166	2	9	6	121	8	5	0	151
PT167	6	0	7	23	0	1	0	37	PT167	12	1	13	33	0	1	1	61
PT168	29	22	66	116	3	7	0	243	PT168	39	27	82	150	5	9	0	312
PT169	39	27	30	127	7	2	4	236	PT169	44	35	33	132	5	5	4	258
PT16A	13	4	10	45	2	8	1	83	PT16A	19	6	13	62	1	8	2	111
PT16B	14	30	23	58	3	12	0	140	PT16B	21	43	30	82	6	15	2	199
PT16C	13	20	7	82	8	10	2	142	PT16C	16	29	14	125	9	16	1	210
PT171	17	3	19	22	2	13	4	80	PT171	25	5	23	35	2	15	2	107
PT172	21	10	6	41	2	12	2	94	PT172	26	12	10	50	3	18	6	125
PT181	81	3	25	190	7	9	1	316	PT181	122	7	35	264	12	11	2	453
PT182	118	32	29	180	15	8	7	389	PT182	161	43	47	257	24	10	7	549
PT183	169	24	34	199	3	10	0	439	PT183	221	33	45	257	6	14	6	582
PT184	249	47	44	156	10	9	8	523	PT184	346	60	50	225	9	13	9	712
PT185	69	29	17	124	6	12	2	259	PT185	79	39	24	147	4	17	0	310
TOTAL	1116	497	703	2670	140	241	61	5428	TOTAL	1509	671	947	3617	187	324	81	7336

4.1.2. Classes de ocupação do solo (variáveis LC1 e LC2)

Como foi referido aquando da descrição da Tabela 4.1, as variáveis LC1 e LC2 dizem respeito aos tipos de classes de ocupação do solo que podem ser atribuídos a cada ponto observado *in situ* na segunda fase do esquema de amostragem por parte dos observadores de acordo com o documento de referência técnica (Eurostat, 2012a). A variável LC2 apenas tem uma classe atribuída nos casos em que existe heterogeneidade do solo.

No caso do exemplo do ponto apresentado na Tabela 4.1, com ID 26601762, é possível observar que este se localiza numa zona heterogénea em termos de classes de ocupação do solo que o envolvem. Por isso, o observador, atribuiu nas variáveis LC1 e LC2, para o nível 3 do sistema de classificação do LUCAS, as classes de ocupação do solo B74 e E30 (em 2009) e B74 e E10 (em 2012).

De seguida é efetuada uma análise exploratória sobre distribuição de pontos da variável LC1 pelos os três níveis que compõem as classes de ocupação do solo do sistema de classificação usado no projeto LUCAS. A escolha desta variável, em detrimento da LC2, prende-se com o facto de serem os valores destaque são usados atualmente pelo Eurostat para a obtenção de estimativas para as classes de ocupação do solo através dos estimadores descritos em 3.6 e 3.10.

4.1.2.1. Distribuição do tipo de classe de ocupação do solo por NUTS 1, NUTS 2 e NUTS 3

Nível 1 das classes de ocupação do solo do LUCAS

Pela análise da Tabela 4.7, observa-se, para o nível 1 da nomenclatura de classificação LUCAS, que a maior parte dos pontos da segunda fase do processo de amostragem foram classificados como pertencendo a florestas (C), com aproximadamente 45 % do número de pontos para a Portugal continental (NUTS 1). Os restantes pontos, para as NUTS 1 e todas as NUTS 2, pertencem fundamentalmente a três classes: terrenos agrícolas (B), matos (D) e prados (E) com proporções muito semelhantes.

As restantes classes, solo nu e líquenes (F), corpos de água (G) e zonas húmidas (H) apresentam um número de pontos para todas as NUTS 2 abaixo de 5% do total de pontos.

Tabela 4.7: Distribuição de pontos, por tipo de classe de ocupação do solo ao nível 1 da nomenclatura de classificação LUCAS por NUTS 1 e NUTS 2

		Classes de ocupação do solo (Nível 1)								Total	
		Ano	A	B	C	D	E	F	G		H
PT11	Norte	2009	89	279	494	197	154	71	14	1	1299
			7%	21%	38%	15%	12%	5%	1%	0%	100%
PT15	Algarve	2009	150	357	602	315	253	63	17	0	1757
			9%	20%	34%	18%	14%	4%	1%	0%	100%
PT16	Centro	2009	18	54	78	98	31	13	5	9	306
			6%	18%	25%	32%	10%	4%	2%	3%	100%
PT17	Lisboa	2009	25	93	133	79	57	12	6	7	412
			6%	23%	32%	19%	14%	3%	1%	2%	100%
PT18	Alentejo	2009	103	277	899	178	175	68	17	6	1723
			6%	16%	52%	10%	10%	4%	1%	0%	100%
PT1	Portugal continental	2009	150	369	1102	318	261	98	27	4	2329
			6%	16%	47%	14%	11%	4%	1%	0%	100%
PT17	Lisboa	2009	28	39	43	13	28	15	4	4	174
			16%	22%	25%	7%	16%	9%	2%	2%	100%
PT18	Alentejo	2009	39	44	56	29	52	3	6	3	232
			17%	19%	24%	13%	22%	1%	3%	1%	100%
PT18	Alentejo	2009	49	374	980	82	359	47	33	2	1926
			3%	19%	51%	4%	19%	2%	2%	0%	100%
PT1	Portugal continental	2009	70	498	1342	101	495	54	45	1	2606
			3%	19%	51%	4%	19%	2%	2%	0%	100%
PT1	Portugal continental	2009	287	1023	2494	568	747	214	73	22	5428
			5%	19%	46%	10%	14%	4%	1%	0%	100%
PT1	Portugal continental	2012	434	1361	3235	842	1118	230	101	15	7336
			6%	19%	44%	11%	15%	3%	1%	0%	100%

Nível 2 e nível 3 das classes de ocupação do solo do LUCAS

Pela análise da Tabela 4.8, em que é apresentada a distribuição do número de pontos, por NUTS 2, observa-se que aproximadamente 25% dos pontos extraídos na segunda fase da amostra, para 2009 e 2012, pertencem à classe floresta de folhas caducifólias (C11).

Outras classes com alguma representatividade, com aproximadamente 10% dos pontos, há a referir: prados sem árvores (E20), em 2012; vegetação espontânea (E30), em 2009; vinha, olival, viveiros e culturas permanentes industriais (B80), outras superfícies arborizadas (C20) e matos sem árvores (D20). Todas as outras classes têm uma percentagem de pontos abaixo de 5% .

Analisando a distribuição do número de pontos por NUTS 3, para as classes LUCAS de ocupação do Solo (para o nível 2 e 3), observam-se 208 casos em que não existem pontos. Este número corresponde a 28% das situações (Tabela 4.9).

Tabela 4.8: Distribuição dos pontos da amostra da segunda fase para o nível 3 das classes de ocupação do solo por NUTS 1 e NUTS 2

Classes de ocupação do solo	2009							Portugal continental	PT1	2012							Portugal continental	PT1
	PT11	PT15	PT16	PT17	PT18	PT1				PT11	PT15	PT16	PT17	PT18	PT1			
	Norte	Algarve	Centro	Lisboa	Alentejo					Norte	Algarve	Centro	Lisboa	Alentejo				
A10	27	4	43	5	15	94	1,7%	A10	59	11	47	13	23	153	2%			
A20	62	14	60	23	34	193	3,6%	A20	91	14	103	26	47	281	3,8%			
B10	62	3	85	8	146	304	5,6%	B10	64	1	100	11	202	378	5,2%			
B20	9	0	8	1	4	22	0,4%	B20	17	1	12	4	1	35	0,5%			
B30	0	0	1	0	16	17	0,3%	B30	0	0	0	0	8	8	0,1%			
B40	13	0	12	2	16	43	0,8%	B40	13	0	12	5	15	45	0,6%			
B50	16	0	13	14	33	76	1,4%	B50	21	1	21	2	48	93	1,3%			
B70	43	29	20	3	4	99	1,8%	B70	70	46	40	4	12	172	2,3%			
B80	136	22	138	11	155	462	8,5%	B80	172	44	184	18	212	630	8,6%			
C11	188	29	375	14	668	1274	23,5%	C11	181	39	470	27	1062	1779	24,3%			
C12	98	16	230	13	78	435	8,0%	C12	106	31	345	16	111	609	8,3%			
C13	63	1	87	4	64	219	4,0%	C13	22	4	46	4	34	110	1,5%			
C20	114	24	160	11	155	464	8,5%	C20	167	46	164	5	93	475	6,5%			
C30	31	8	47	1	15	102	1,9%	C30	126	13	77	4	42	262	3,6%			
D10	31	42	57	7	20	157	2,9%	D10	47	28	135	11	35	256	3,5%			
D20	166	56	121	6	62	411	7,6%	D20	268	51	183	18	66	586	8,0%			
E10	4	9	7	0	16	36	0,7%	E10	13	19	34	5	47	118	1,6%			
E20	23	2	9	3	157	194	3,6%	E20	188	28	159	21	384	780	10,6%			
E30	127	20	159	25	186	517	9,5%	E30	52	10	68	26	64	220	3,0%			
F00	71	13	68	15	47	214	3,9%	F10	25	1	37	1	11	75	1,0%			
								F20	1	0	4	0	1	6	0,1%			
								F40	37	11	57	2	42	149	2,0%			
G10	9	2	7	1	21	40	0,7%	G10	5	2	14	3	27	51	0,7%			
G20	5	2	5	1	11	24	0,4%	G20	11	3	8	0	18	40	0,5%			
G30	0	1	5	2	1	9	0,2%	G30	1	1	5	3	0	10	0,1%			
H10	0	0	3	0	0	3	0,1%	H10	0	0	1	0	0	1	0,0%			
H20	1	9	3	4	2	19	0,4%	H20	0	7	3	3	1	14	0,2%			

Tabela 4.9: Distribuição dos pontos da amostra da segunda fase para o nível 3 das classes de ocupação do solo por NUTS 3 (em 2012)

Classes de ocupação do solo (Nível 2 e 3)	NUTS 3																												
	PT111	PT 112	PT 113	PT 114	PT 115	PT 116	PT 117	PT 118	PT 150	PT 161	PT 162	PT 163	PT 164	PT 165	PT 166	PT 167	PT 168	PT 169	PT 16A	PT 16B	PT 16C	PT 171	PT 172	PT 181	PT 182	PT 183	PT 184	PT 185	
A10	8	7	4	9	12	2	6	11	11	7	7	6	3	5	0	1	2	2	2	6	6	5	8	4	3	4	4	8	
A20	11	4	17	8	16	10	12	13	14	9	13	14	3	16	3	1	6	3	5	18	12	12	14	9	5	9	8	16	
B10	4	4	11	11	10	3	1	20	1	8	15	8	4	11	1	1	14	10	10	8	10	7	4	27	40	33	73	29	
B20	1	1	3	0	4	1	1	6	1	2	1	1	0	2	1	0	1	0	0	4	0	2	2	0	0	0	0	1	
B30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	1
B40	0	3	2	1	2	1	2	2	0	0	0	2	0	4	0	1	0	0	1	2	2	5	0	3	1	1	0	10	
B50	2	6	1	2	3	5	2	0	1	5	2	0	2	7	0	1	0	1	0	3	0	1	1	8	8	10	16	6	
B70	2	3	2	3	5	0	20	35	46	2	1	2	1	1	2	2	7	1	2	15	4	3	1	1	2	1	4	4	
B80	9	3	7	2	19	1	61	70	44	6	11	13	7	20	13	2	23	38	13	17	21	7	11	8	45	45	94	20	
C11	20	5	5	10	18	26	19	79	39	56	47	22	49	49	26	1	20	80	12	30	77	11	16	198	261	285	209	109	
C12	6	0	1	1	7	3	35	53	31	21	41	42	32	54	53	9	14	27	18	17	17	0	16	45	15	11	23	17	
C13	4	1	2	1	5	3	1	5	4	3	4	8	6	9	5	0	2	0	2	2	5	1	3	10	6	7	5	6	
C20	0	0	5	0	6	3	34	119	46	4	11	18	20	15	6	4	24	29	11	12	10	4	1	10	27	11	36	9	
C30	14	23	22	14	34	1	4	14	13	3	1	3	16	8	8	3	11	6	4	2	12	1	3	15	6	8	5	8	
D10	13	3	5	1	7	0	13	5	28	0	6	1	21	9	17	6	49	9	4	6	7	7	4	11	6	1	14	3	
D20	36	14	12	4	32	6	67	97	51	4	9	7	14	31	9	15	52	10	12	13	7	11	7	8	18	6	27	7	
E10	1	5	0	1	1	1	2	2	19	2	3	0	2	0	1	2	11	5	1	2	5	2	3	13	5	9	16	4	
E20	17	10	16	6	11	3	20	105	28	6	11	18	6	12	0	6	38	30	5	22	5	19	2	54	73	102	129	26	
E30	4	5	2	3	5	10	5	18	10	8	5	2	4	13	0	3	10	1	5	13	4	4	22	16	8	17	9	14	
F10	6	2	1	0	3	0	6	7	1	0	0	1	5	10	1	1	18	0	1	0	0	1	0	0	8	1	2	0	
F20	0	0	0	1	0	0	0	0	0	2	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	
F40	4	2	3	2	3	0	5	18	11	5	4	4	6	12	4	1	9	2	2	3	5	0	2	8	3	8	15	8	
G10	1	0	0	1	0	0	1	2	2	1	1	1	1	2	0	1	0	3	0	3	1	2	1	1	6	10	9	1	
G20	4	0	1	1	1	0	2	2	3	1	4	0	0	0	1	0	0	1	1	0	0	0	0	3	3	3	7	2	
G30	1	0	0	0	0	0	0	0	1	3	2	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	
H10	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H20	0	0	0	0	0	0	0	0	7	3	0	0	0	0	0	0	0	0	0	0	0	1	2	1	0	0	0	0	

4.1.2.2. Distribuição de pontos com duplos registos para as variáveis LC1 e LC2

Como foi referido anteriormente, em situações em que um ponto se localiza em zonas heterogéneas em termos de composição do solo, pode ocorrer a necessidade, por parte do observador na segunda fase do esquema de amostragem, em atribuir duas classes ao ponto registadas nas variáveis LC1 e LC2.

Nas Tabelas 4.10 e 4.11 observa-se que, para o nível 1 das classes de ocupação do solo LUCAS, em 2009 houve 1516 pontos com dois tipos de classes e em 2012 o número foi de 1565.

Observa-se, igualmente, que maioritariamente os casos de pontos com duplos registos ocorreram em terrenos agrícolas (B), florestas (C), prados (E) e solos nu e líqüenes (F).

As células coloridas das Tabelas 4.10 e 4.11, representam as combinações dadas pelas classes C/D, B/E, C/E e B/F com maior representação de registos duplos e conseqüentemente, as classes que apresentam maior heterogeneidade em termos de ocupação do solo. No total, estas combinações representam cerca 90 % dos casos.

Tabela 4.10: Distribuição de pontos em 2009 com duplos registos na segunda fase do esquema de amostragem para as classes de ocupação do solo (ao nível da nomenclatura LUCAS)

LC1	LC 2				Total	%
	B	D	E	F		
A	3				3	0,2%
B	36	17	312	168	533	35,2%
C	26	306	565	83	980	64,6%
Total	65	323	877	251	1516	
%	4,3%	21,3%	57,8%	16,6%		

Tabela 4.11: Distribuição de pontos em 2012 com duplos registos na segunda fase do esquema de amostragem para as classes de ocupação do solo (ao nível da nomenclatura LUCAS)

LC 1	LC 2				Total	%
	B	D	E	F		
A	1				1	0,1%
B	38	33	504	137	712	45,5%
C	35	203	593	20	851	54,4%
F		1			1	0,1%
Total	74	237	1097	157	1565	
%	4,7%	15,1%	70,1%	10,0%		

Uma vez que para a obtenção de estimativas das áreas das classes de ocupação do solo foram utilizadas exclusivamente as categorias indicados na variável LC1, tal facto poderá originar um fenómeno de subjetividade na atribuição das classes entre as fontes de classificação nacionais (NDS) para pontos situados em regiões com diversidade paisagística.

4.1.2.3. Alteração do tipo de classe de ocupação do solo entre 2009 e 2012

Outro aspeto que pode contribuir prejudicialmente na comparação entre as estimativas obtidas no projeto LUCAS e resultados provenientes de outras fontes nacionais é a questão temporal.

Por exemplo, analisando os resultados da Tabela 4.12, para os dados recolhidos nos questionários do LUCAS, entre 2009 e 2012, observa-se que 1032 pontos (20% do total de pontos observados em 2009) foram classificados com um tipo de nomenclatura e no inquérito realizado em 2012 foi-lhes atribuído uma classe diferente (ao nível 1 do sistema de classificação LUCAS).

Perante esta situação, o processo de comparação, necessariamente, poderá ficar condicionado não apenas pela forma como se recolhem os dados (bastante diferentes entre as diversas fontes de sistemas de classificação), mas em particular pelo momento em que foi efetuada essa recolha.

Uma vez que existem apenas dados para dois momentos para os inquéritos LUCAS (2009 e 2012) é difícil modelar a relação temporal utilizando as estimativas obtidas até este momento.

Tabela 4.12: Descrição do número de pontos cuja classe de ocupação do solo em 2012 foi diferente da atribuída em 2009 (ao nível 1 da nomenclatura do sistema de classificação LUCAS)

LC 1 (em 2009)	LC 1 (em 2012)							Total
	A	B	C	D	E	F	G	
A		9	3	4	11	19		46
B	5		11	16	167	15	1	215
C	8	20		111	32	26	1	198
D	6	27	96		41	16	5	191
E	12	123	53	73		13	1	275
F	10	17	15	20	32		2	96
G			2	3	1			6
H		1			1		3	5
Total	41	197	180	227	285	89	13	1032

4.1.3. Classes de uso do solo (variáveis LU1 e LU2)

As variáveis LU1 e LU2 dizem respeito aos tipos de classes de uso do solo que podem ser atribuídas a cada ponto observado *in situ* na segunda fase do esquema de amostragem do projeto LUCAS.

No caso do exemplo do ponto apresentado na Tabela 4.1, com ID 26601762, é possível observar que este se localiza numa zona heterogénea em termos de classes de uso do solo. Por isso, foi-lhe atribuído nas variáveis LU1 e LU2, ao nível 3 do sistema de classificação, duas classes, produção agrícola comercial (U111) e infra-estruturas agrícolas (U112).

De seguida é apresentada a distribuição do número de pontos pertencentes às classes de agricultura (U110), floresta (U120), indústria extrativa (U130), serviços culturais, entretenimento e recreativos (U340) e áreas naturais ou com outros usos socioeconómicos (U612) que foram usadas neste estudo na obtenção de estimativas para classes de uso do solo a partir dos valores registados na variável LU1.

A escolha da variável LU1 prende-se com o facto de ser os valores desta variável que são usados atualmente pelo Eurostat para a obtenção de estimativas para as classes de ocupação do solo através dos estimadores apresentados em 3.6 e 3.10.

4.1.3.1. Distribuição do tipo de classe de uso do solo por NUTS 1, NUTS 2 e NUTS 3

Pela análise das Tabelas 4.13 e 4.14, observa-se a maior presença de pontos pelas classes U110 e U120 por todo o país. Observa-se que os dados da base de dados LUCAS *survey* para o ano 2009 não houve registo de pontos pertencentes à classe U612.

Ao nível das NUTS 3 observa-se que as classes U130 e U340 apresentam um número de pontos baixo o que originará estimativas com pouca precisão (Tabela 4.14).

Tabela 4.13: Distribuição dos pontos da amostra da segunda fase para as classes U110, U120, U130, U340 e U612 por NUTS 1 e 2 (em 2009 e 2012)

LU	PT11		PT15		PT16		PT17		PT18		PT1	
	2009	2012	2009	2012	2009	2012	2009	2012	2009	2012	2009	2012
U110	475	652	88	167	513	708	60	99	880	1572	2016	3198
U120	357	488	54	116	767	965	27	43	830	812	2035	2424
U130	2	1	3	3	6	9	2	1	2	2	15	16
U340	3	3	2	3	7	6	5	5	5	3	22	20
U612	0	213	0	25	0	205	0	19	0	56	0	518
TOTAL	837	1357	147	314	1293	1893	94	167	1717	2445	4088	6176

Tabela 4.14: Distribuição dos pontos da amostra da segunda fase para as classes U110, U120, U130, U340 e U612 por NUTS 3 (em 2009 e 2012)

NUTS 3	LUCAS 2009				LUCAS 2012				
	U110	U120	U130	U340	U110	U120	U130	U340	U612
PT111	31	44	0	0	41	53	0	0	39
PT112	24	23	0	0	40	36	0	0	10
PT113	25	28	1	1	38	36	0	0	15
PT114	20	16	0	1	25	24	0	1	1
PT115	40	58	1	1	58	72	1	0	33
PT116	11	23	0	0	23	32	0	0	7
PT117	108	57	0	0	117	62	0	0	51
PT118	216	108	0	0	310	173	0	2	57
PT150	88	54	3	2	167	116	3	3	25
PT161	31	63	0	2	40	83	2	3	15
PT162	41	62	1	1	48	98	1	0	17
PT163	32	53	0	0	48	83	0	1	3
PT164	21	106	0	0	26	106	0	1	34
PT165	48	88	0	1	73	111	0	0	56
PT166	19	75	0	0	19	107	0	0	7
PT167	11	14	0	0	16	14	0	0	10
PT168	96	57	1	1	112	55	2	1	46
PT169	83	112	1	1	120	123	1	0	2
PT16A	30	36	0	0	45	39	0	0	7
PT16B	60	33	3	0	89	51	3	0	7
PT16C	41	68	0	1	72	95	0	0	1
PT171	24	4	0	3	47	12	1	3	10
PT172	36	23	2	2	52	31	0	2	9
PT181	97	166	0	3	268	151	0	1	16
PT182	162	183	0	0	336	174	0	0	19
PT183	209	202	1	0	348	201	2	0	8
PT184	298	166	0	0	453	182	0	0	10
PT185	114	113	1	0	167	104	0	2	3

4.1.4. Raio atribuído ao ponto (variável *OBS_RADIUS*)

No projeto LUCAS, considera-se que um ponto é uma representação geométrica definida por um círculo com 1,5 metros de raio o que corresponde a uma área aproximada de 7 m². Como foi descrito na Tabela 4.1, a variável *OBS_RADIUS* pode ter dois valores: 1 (para os casos em que o raio do círculo que representa o ponto é menor que 1,5 m) ou 2 (é atribuído de um modo genérico o valor 20 metros).

Na maior parte dos casos, os pontos observados na segunda fase estão localizados em superfícies homogéneas, cuja classificação do ponto de acordo com as classes de ocupação e uso do solo é fácil de ser realizada, como ocorre na Figura 4.1A. Para estes casos, o observador regista na variável *OBS_RADIUS* o valor 1.

Para casos de solos heterogéneos, o observador necessita de estender a área em torno do ponto (*extend windows*) para um círculo de raio 20 metros (o que corresponde uma área aproximadamente de 0.13 ha), como ocorre no caso apresentado na Figura 4.1B, em que se tem uma mistura de duas classes: florestas (C) e matos (D). Para estes casos, a variável *OBS_RADIUS* toma a categoria 2.



Figura 4.1A: Exemplo de uma superfície homogénea



Figura 4.1B: Exemplo de uma superfície heterogénea

(retirado a partir de Eurostat 2013b)

4.1.4.1. Distribuição dos pontos por tipo de raio observado (NUTS 2 e NUTS 3)

Pela análise do gráfico da Figura 4.2 relativos à distribuição em percentagem de pontos de acordo com a categoria de raio observado (*OBS_RADIUS*), 1 ou 2, pode concluir-se que para Portugal Continental existe uma grande diversidade paisagística, com quase 80 % dos pontos (ao nível das NUTS 2) a estarem localizados em superfícies heterogéneas.

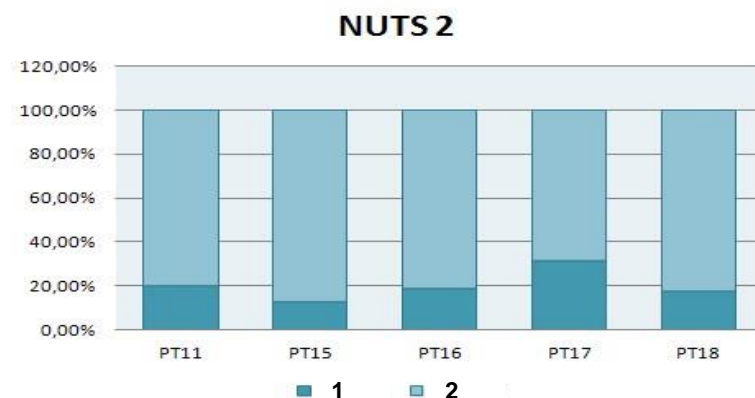


Figura 4.2: Distribuição da variável *OBS_RADIUS* por NUTS 2 (em 2012)

Para cada NUTS 3, observa-se igualmente que a maioria dos pontos localiza-se em superfícies com solos heterogéneos. Realce-se que para Grande Porto (PT114) e Grande Lisboa (PT171) há uma maior aproximação entre a percentagem de pontos com a categoria 1 relativamente aos com a categoria 2 (Figura 4.3).

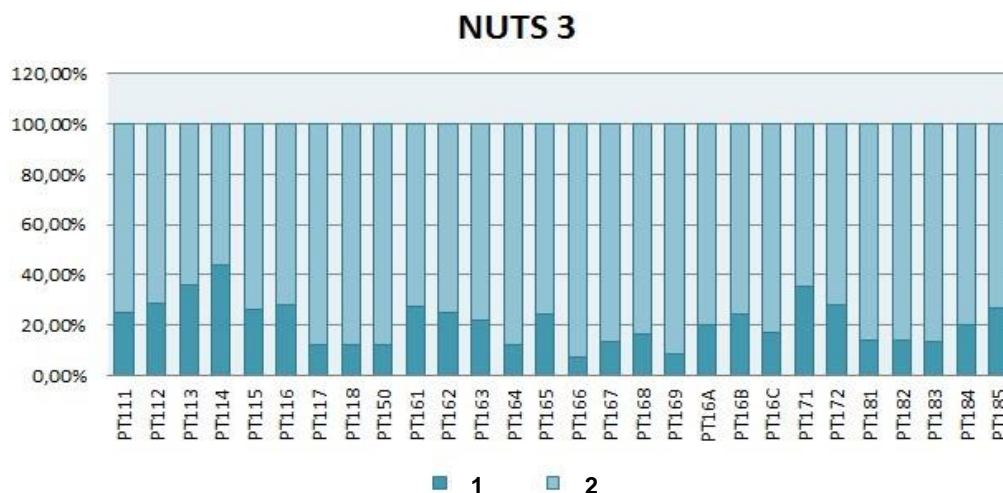


Figura 4 3: Distribuição da variável *OBS_RADIUS* por NUTS 3 (em 2012)

4.1.4.2. Distribuição dos pontos por tipo de raio (por classe de ocupação do solo ao nível 1)

Pela Figura 4.4, para as classes de ocupação do solo para o nível 1, observa-se que os pontos pertencentes a terrenos artificiais (A), solos nus e líquenes (F) e corpos de água (G), se localizam em superfícies homogéneas (com categoria *OBS_RADIUS* igual a 1).

Pelo contrário, os pontos classificados como pertencendo a florestas (classe C), matos (D) e prados (E), têm categoria para *OBS_RADIUS* igual a 2, o que significa que pertencem a zonas com heterogeneidade do solo.

Relativamente aos pontos classificados como pertencendo a terrenos de cultivo (B), ao nível 2 da nomenclatura definida no projeto LUCAS, observa-se uma distinção clara entre as classes cereais (B10), culturas de raízes comestíveis (B20), culturas temporárias industriais (B30), leguminosas secas, hortícolas e flores (B40) e culturas forrageiras (B50), localizadas em regiões homogéneas e os pontos classificados como sendo frutos secos, frutos de casca rija, frutos pequenos de baga, citrinos e frutos sub-tropicais (B70) e vinha, olival, viveiros e culturas permanentes industriais (B80), situados em regiões heterogéneas.

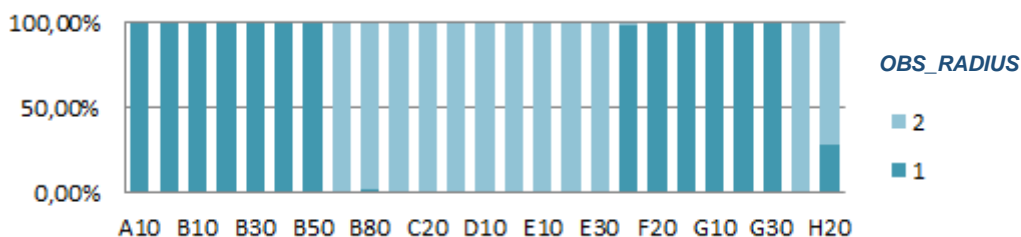


Figura 4.4: Distribuição da variável *OBS_RADIUS* pelas classes de ocupação do solo (nível 2)

4.1.5. Variável *AREA_SIZE* (dimensão da área em torno do ponto)

A variável *AREA_SIZE* corresponde à estimativa que é atribuída pelo observador relativamente à área que envolve o ponto. Neste processo de estimação é tido em consideração que um círculo com 40 m de raio representa uma área de 0.5027 *ha* e que um círculo com 56 m de raio equivale a 1 *ha*. Para mais informação, consultar Eurostat (2013b). As categorias da variável *AREA_SIZE* são ordenadas de acordo com a Tabela 4.15.

Tabela 4.15: Descrição das categorias da variável *AREA_SIZE*

Categoria <i>AREA_SIZE</i>	Descrição
1	Menos do que 0.5 <i>ha</i>
2	Maior ou igual a 0.5 <i>ha</i> e menor do que 1 <i>ha</i>
3	Maior ou igual a 1 <i>ha</i> e menor do que 10 <i>ha</i>
4	Superior a 10 <i>ha</i>

4.1.5.1. Distribuição dos pontos por tipo de *area size* (ao nível das NUTS 2)

Pelos gráficos das Figuras 4.5 e 4.6, observa-se que é no centro de Portugal (PT16) e Alentejo (PT18), que se observam pontos com dimensão de áreas superiores (categorias 3 e 4).

A distribuição do número de pontos é semelhante para os dados LUCAS recolhidos em 2009 e 2012 para o Norte (PT11) e Alentejo (PT18). Para Lisboa (PT17) e Algarve (PT15), observa-se um decréscimo da dimensão da área em torno do ponto nas categorias 1 e 2.

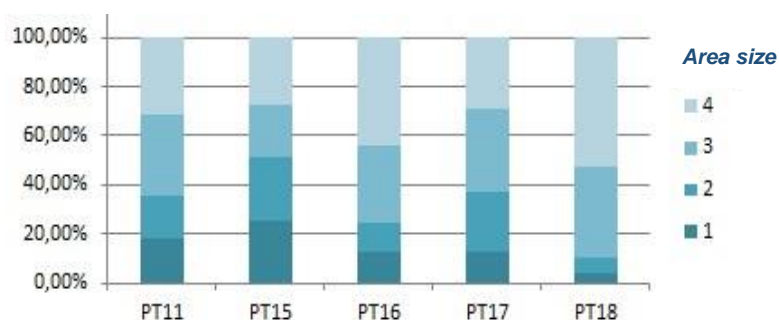


Figura 4 5:Distribuição dos pontos por AREA_SIZE ao nível das NUTS 2 (em 2009)

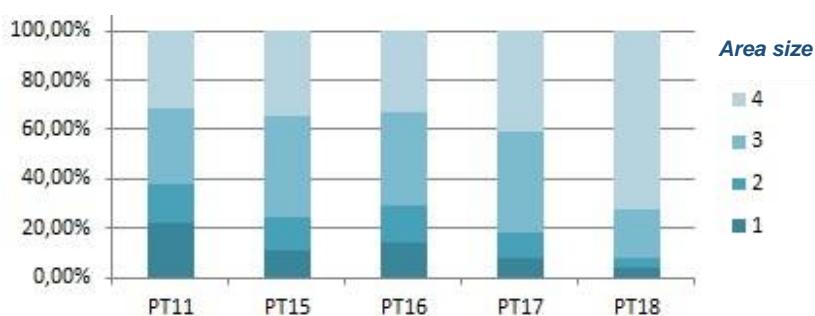


Figura 4 6: Distribuição dos pontos por AREA_SIZE ao nível das NUTS 2 (em 2012)

4.1.5.2. Distribuição dos pontos por tipo de area size (ao nível 1 das classes de ocupação)

Pelo gráfico da Figura 4.7, observa-se que para o nível 1 da nomenclatura de classificação do LUCAS, as classes de ocupação do solo com maiores dimensões de área em torno dos pontos são florestas (C), matos (D) e prados (E), corpos de água (G) e Zonas húmidas (H). Os pontos pertencentes a territórios artificializados (A), terrenos agrícolas (B) e solo nu e líquenes (F), são os que apresentam dimensões de áreas com valores mais próximos.

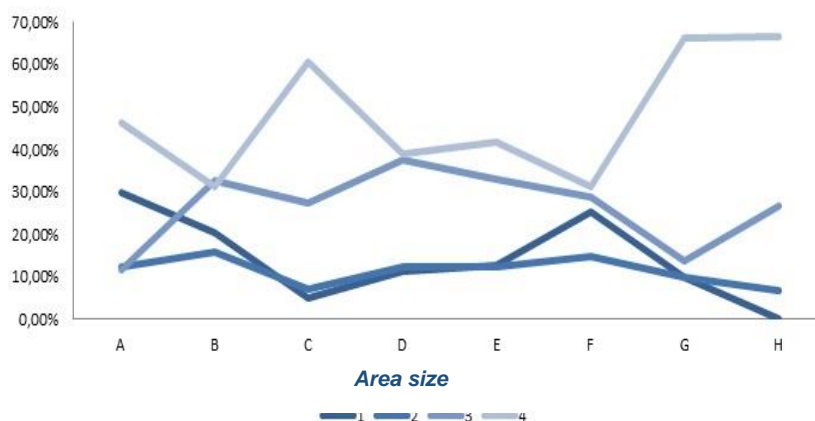


Figura 4.7: Distribuição do número de pontos por AREA_SIZE e por classe de ocupação do solo para o nível 1 da nomenclatura de classificação LUCAS

4.2. Estimativas para as classes de ocupação do solo

De seguida são apresentadas as estimativas obtidas para as áreas, em Km², das classes de ocupação do solo para os três níveis da nomenclatura do sistema de classificação LUCAS.

Para a obtenção das estimativas das classes foram utilizados os estimadores para a área e variância apresentados em 3.6 e 3.10 para cada NUTS 1, NUTS 2 ou NUTS 3, a partir dos dados recolhidos nos inquéritos LUCAS em 2009 e 2012 para Portugal, registados nas bases de dados LUCAS *survey*. Portanto, as estimativas para a NUTS 1 e NUTS 2 não foram obtidas por agregação de estimativas provenientes de níveis inferiores.

Para além das estimativas produzidas é igualmente efetuada uma avaliação da qualidade das mesmas pela análise dos coeficientes de variação (CV) obtidos tendo em consideração os limites superiores para os errosestipulados pelo Eurostat (Anexo 5).

4.2.1. Avaliação das estimativas para a NUTS 1

4.2.1.1. Nível 1 (da nomenclatura de classificação LUCAS)

Na Tabela 4.16 são apresentados para Portugal continental (PT1) as estimativas das áreas das classes de ocupação do solo e os coeficientes de variação utilizando os estimadores apresentados em 3.6 e 3.10.

Pelos resultados obtidos, observa-se que, excetuando o caso de zonas húmidas (H), todas as outras classes ao nível 1 da nomenclatura de classificação LUCAS, apresentam valores de coeficientes de variação que estão de acordo com os limites máximos de precisão apresentados no Anexo 5. Os CV superiores a a 20% para a classe zonas húmida prendem-se, sobretudo, pela menor dimensão da amostra da segunda fase que pertencem a esta classe.

Comparando os CV para os dados recolhidos no LUCAS de 2009 e 2012, observa-se uma melhoria na precisão das estimativas, excetuando o caso da classe H.

Tabela 4.16: Estimativas da área (km²) e respetivos CV, ao nível 1 das classes de ocupação do solo, para os dados LUCAS recolhidos para Portugal continental (PT1)

Classes de ocupação do solo (Nível 1)	2009		2012	
	ESTIM_AREA	CV	ESTIM_AREA	CV
A	4691	4,90	5545	4,03
B	16555,6	2,43	15714	2,23
C	40929,4	1,27	39317	1,16
D	9390,6	3,85	10363	3,12
E	12267	3,20	13465	2,61
F	3553	6,57	2999	6,18
G	1312,5	8,72	1514	7,32
H	389,9	20,70	173	27,26
Total	89089		89089	

4.2.1.2. Nível 2 (da nomenclatura de classificação LUCAS)

Na Tabela 4.17 são apresentadas as estimativas das áreas, para o nível 2 das classes de ocupação do solo (LC) e os respetivos CV para Portugal continental (PT1). Assinalados a cinzento são indicados os CV cujo valor é superior ao valor máximo aceitável para a precisão das estimativas (Anexo 5).

Para os dados recolhidos em 2009, excetuando as classes corpos de água costeiros (G30), zonas húmidas interiores (H10) e zonas húmidas costeiras (H20) todas as outras classes apresentam CV abaixo dos limites máximos referenciados no Anexo 5. Para o caso das estimativas obtidas a partir dos dados de 2012, para além das classes G30, H10 e H20 referidas anteriormente, também as classes culturas temporárias industriais (B30) e areias (F20) e outras áreas de solo nu (F40) apresentam CV acima dos limites máximos para essas classes. Tal facto, deve-se à redução do número de pontos de 2009 para 2012 na classe B30 e também à desagregação da classe F00 em F10, F20 e F40, o que origina uma diminuição do número de pontos amostrados na classe F20.

4.2.1.3. Nível 3 (da nomenclatura de classificação LUCAS)

Para o nível 3 da nomenclatura do sistema de classificação LUCAS observa-se uma desagregação da classe floresta (C10) em floresta de folhosas caducifólias (C11), floresta de coníferas (C12) e floresta mista (C13), mantendo-se as restantes classes com a mesma designação. Portanto, as conclusões retiradas relativamente à precisão das estimativas apresentadas na Tabela 4.17 são as mesmas.

Tabela 4.17: Estimativas da área (km²) e respetivos CV, ao nível 2 das classes de ocupação do solo, para os dados LUCAS recolhidos para Portugal Continental (PT1)

NUTS 1	Classes de ocupação do solo(Nível 2)	2009		Classes LC (Nível 2)	2012	
		ESTIM_AREA	CV		ESTIM_AREA	CV
PT1	A10	1534	9,62	A10	2123	6,95
PT1	A20	3157	6,48	A20	3422	5,56
PT1	B10	4964	5,25	B10	4450	4,84
PT1	B20	361	21,11	B20	398	17,24
PT1	B30	276	23,95	B30	111	32,45
PT1	B40	703	15,06	B40	594	14,07
PT1	B50	1240	11,19	B50	1154	10,00
PT1	B70	1596	9,74	B70	2025	7,37
PT1	B80	7416	3,83	B80	6982	3,57
PT1	C10	31617	1,62	C10	30474	1,47
PT1	C20	7635	4,36	C20	5663	4,4
PT1	C30	1677	9,71	C30	3180	5,94
PT1	D10	2582	7,75	D10	2923	6,24
PT1	D20	6809	4,61	D20	7440	3,77
PT1	E10	593	16,46	E10	1523	8,77
PT1	E20	3174	6,81	E20	9306	3,27
PT1	E30	8500	4,05	E30	2636	6,55
PT1	F00	3553	6,57	F10	968	11,07
PT1				F20	62	44,55
PT1				F40	1968	7,69
PT1	G10	717	13,81	G10	824	11,35
PT1	G20	429	19,11	G20	536	14,65
PT1	G30	167	32,20	G30	155	29,10
PT1	H10	52	57,70	H10	23	74,26
PT1	H20	338	22,34	H20	151	29,31

Pela análise da tabela 4.18, observa-se que as estimativas das áreas das classes C11, C12 e C13 apresentam CV abaixo dos limites máximos de referência. Para as classes C11 e C12 observa-se uma ligeira melhoria na qualidade da precisão das estimativas.

Na Tabela 4.18 são apresentados para o nível 3 do sistema de classificação LUCAS as estimativas das áreas e os coeficientes de variação obtidos para as classes de ocupação do solo.

Tabela 4.18: Estimativas da área (km²) e respetivos CV, ao nível 3 das classes de ocupação do solo, para os dados LUCAS recolhidos para Portugal continental (PT1)

NUTS 1	Classes de ocupação do solo(Nível 3)	2009		Classes LC (Nível 3)	2012	
		ESTIM_AREA	CV		ESTIM_AREA	CV
PT1	C11	20889.7	2,29	C11	21684.1	1,93
PT1	C12	7138.6	4,47	C12	7452.1	3,74
PT1	C13	3588.9	6,46	C13	1337.9	9,26

4.2.2. Avaliação das estimativas para a NUTS 2

4.2.2.1. Nível 1 (da nomenclatura de classificação LUCAS)

Na Tabela 4.19 são apresentadas as estimativas das áreas, ao nível 1 das classes de ocupação do solo, e os respetivos CV para NUTS 2. Pelos valores dos coeficientes de variação obtidos, pode concluir-se que para a maioria das classes houve uma melhoria da precisão das estimativas do ano de 2009 para 2012. Observam-se, no entanto, algumas exceções, como acontece no caso de matos (D), para o Algarve (PT15), solo nu e líquenes (F), para a região Norte (PT11) e Lisboa (PT17) e zonas húmidas (H), para as cinco NUTS 2, devido a decréscimo do número de pontos observados na amostra da segunda fase do esquema de amostragem.

Tabela 4.19: Estimativas da área (km²) e respetivos CV, ao nível 1 das classes de ocupação do solo (NUTS 2)

NUTS 2	Classes de ocupação do solo (Nível 1)	2009		2012	
		ESTIM_AREA	CV	ESTIM_AREA	CV
PT11	A	1449	8,82	1869	7,22
	B	4447	4,46	3976	4,39
	C	8135	3,14	7264	3,05
	D	3301	6,29	4035	4,79
	E	2541	7,35	3083	5,59
	F	1200	11,24	852	11,57
	G	200	17,36	207	17,93
	H	14	100,79		
PT15	A	296	20,59	343	14,65
	B	867	10,56	1088	8,26
	C	1280	9,40	1605	6,68
	D	1610	7,81	920	10,01
	E	507	16,59	778	11,07
	F	213	26,56	125	30,53
	G	80	43,23	65	42,09
	H	144	29,04	73	36,51
PT16	A	1685	8,50	1960	6,75
	B	4504	4,80	4346	4,29
	C	14715	1,94	13389	1,92
	D	2935	6,94	3844	5,08
	E	2869	6,77	3105	5,60
	F	1124	11,73	1175	9,82
	G	268	17,23	329	14,53
	H	99	40,53	51	48,26
PT17	A	471	11,78	539	9,31
	B	643	12,49	485	13,36
	C	715	10,81	649	11,17
	D	215	25,30	380	16,11
	E	461	16,94	742	10,09
	F	253	22,97	17	89,43
	G	141	33,48	136	31,16
	H	103	48,92	54	64,63
PT18	A	808	11,73	842	10,39
	B	6063	3,98	5823	3,64
	C	16076	1,87	16452	1,58
	D	1357	10,60	1140	9,96
	E	5886	4,39	5739	3,85
	F	767	14,09	832	11,73
	G	616	13,74	771	10,34
	H	33	70,17	6	142,59

Pela análise das medidas de localização associadas aos CV das estimativas das áreas das classes de ocupação do solo da Tabela 4.20, confirma-se a melhoria da precisão das estimativas proveniente dos dados recolhidos em 2012 relativamente aos de 2009.

Tabela 4.20: Comparação das medidas de localização para os valores CV ao nível 1 das classes de ocupação do solo, por NUTS 2, para os dados recolhidos nos LUCAS de 2009 e 2012

LUCAS	Min.	Q1	Mediana	Média	Q3	Máx.
2009	1,9	7,0	11,7	18,4	22,4	100,8
2012	1,6	5,6	10,1	19,2	16,1	142,6

4.2.2.2. Nível 2 e 3 (da nomenclatura de classificação LUCAS)

Nos Anexos 6 e 7 são apresentadas, para o nível 2 e 3 do sistema de classificação LUCAS, as estimativas das áreas das classes de ocupação do solo.

Os CV cujos valores estão acima dos limites máximos estipulados pelo Eurostat (Anexo 5) são indicados na tabela a cinzento.

Analisando as estimativas obtidas provenientes dos dados de 2009, observa-se que floresta (C10) e vinha, olival, viveiros e culturas permanentes industriais (B80) são as únicas cujos CV não estão acima dos limites de referência para a precisão das estimativas das áreas (Anexo 5).

Para o caso das estimativas obtidas a partir dos dados recolhidos em 2012, as classes floresta de folhosas caducifólias (C11) e prados sem árvores (E20), apresentam estimativas com precisão abaixo dos limites para todas as NUTS 2.

Observa-se, para ambos os anos, um número elevado de estimativas com baixa precisão, o que, considerando a que amostra escolhida na segunda fase do esquema de amostragem é representativa ao nível das NUTS 2, seria de esperar uma maior precisão das estimativas produzidas.

4.2.3. Avaliação das estimativas para a NUTS 3

4.2.3.1. Nível 1 (da nomenclatura de classificação LUCAS)

Tendo em consideração que a amostra da segunda fase do esquema de amostragem do projeto LUCAS foi recolhida para a obtenção de estimativas das áreas das classes de ocupação do solo ao nível das NUTS 2, confirma-se, pelas medidas de localização dos CV das estimativas para as NUTS 3 (Tabela 4.21) uma diminuição da precisão das estimativas comparando com os valores obtidos para NUTS 2 (Tabela 4.20).

Ainda pela Tabela 4.23, observa-se que as estimativas obtidas com base nos dados recolhidos em 2012 apresentam uma melhor precisão relativamente às provenientes dos dados de 2009.

Tabela 4.21: Comparação das medidas de localização para os valores CV ao nível 1 das classes de ocupação do solo, por NUTS 3, para os dados recolhidos nos LUCAS de 2009 e 2012

LUCAS	Min.	Q1	Mediana	Média	Q3	Máx.
2009	3,6	14,3	24,5	29,6	36,0	104,3
2012	3,0	12,5	20,1	27,1	32,3	173,7

Nos Anexos 8 e 9 são apresentadas as estimativas para o nível 1 da nomenclatura de classificação LUCAS, para os anos 2009 e 2012, por NUTS 3. Observa-se que nenhuma das estimativas obtidas para corpos de água (G) e zonas húmidas (H), apresenta precisões para 2009 e 2012 abaixo dos limites máximos de referência (Anexo 5).

Para ambos os anos, como seria de esperar pela dimensão do número de pontos das classes, terrenos agrícolas (B) e florestas (C), são as classes de ocupação que apresentam melhores precisões para as estimativas das áreas.

4.2.3.2. Para o nível 2 e 3 (da nomenclatura de classificação LUCAS)

À medida que se procede a uma análise mais desagregada dos níveis da nomenclatura LUCAS, o número de pontos observados na segunda fase do esquema de amostragem é menor e, conseqüentemente, a precisão das estimativas das áreas obtidas diminui. Nas Tabelas 4.22 e 4.23 são apresentadas as NUTS 3 para as quais as estimativas das áreas das classes de ocupação do solo, ao nível 2 e 3 do sistema de classificação LUCAS apresentam CV dentro dos limites máximos de erro definidos pelo Eurostat para o estudo piloto (Anexo 5).

Tabela 4.22: Estimativas da área (km²) e respetivos CV, para os níveis 2 e 3 das classes de ocupação do solo, por NUTS 3, cujos CV estão dentro dos limites máximos de precisão (em 2009)

Classes LC (Nível 2 e 3)	NUTS3	AREA_ESTIM	CV	Classes LC (Nível 2 e 3)	NUTS3	AREA_ESTIM	CV	Classes LC (Nível 2 e 3)	NUTS3	AREA_ESTIM	CV
B10	PT184	953	12,07	C10	PT171	83	42	C11	PT183	2945	4,81
B70	PT117	286	23,06	C10	PT172	411	13,09	C11	PT184	2184	7,08
B70	PT118	329	21,18	C10	PT181	2640	4,54	C11	PT185	1435	7,69
B70	PT150	463	15,86	C10	PT182	2860	4,59	C12	PT117	434	17,12
C10	PT111	648	12,45	C10	PT183	3218	4,25	C12	PT118	623	15,52
C10	PT112	359	16,65	C10	PT184	2649	6,12	C12	PT150	262	24,15
C10	PT113	439	12,64	C10	PT185	1959	5,07	C12	PT162	359	18,92
C10	PT114	200	18,33	C11	PT113	300	18,12	C12	PT163	550	12,46
C10	PT115	942	9,53	C11	PT115	488	15,69	C12	PT164	501	15,85
C10	PT116	386	14,49	C11	PT117	317	19,95	C12	PT165	489	16,85
C10	PT117	823	11,2	C11	PT118	1104	11,12	C12	PT166	478	16,26
C10	PT118	1934	7,66	C11	PT150	476	17,4	C12	PT169	354	19,77
C10	PT150	754	13,31	C11	PT161	599	11,55	C12	PT181	576	15,75
C10	PT161	923	6,44	C11	PT162	407	16,55	C20	PT117	378	19,14
C10	PT162	990	6,99	C11	PT164	679	11,82	C20	PT118	1132	11,05
C10	PT163	794	8,67	C11	PT165	603	15,02	C20	PT150	395	19,01
C10	PT164	1437	6,07	C11	PT166	387	17,42	C20	PT168	694	13,76
C10	PT165	1429	7,47	C11	PT168	408	18,92	C20	PT181	499	17,2
C10	PT166	980	8,79	C11	PT169	1332	8,64	C20	PT182	638	14,96
C10	PT168	691	13,6	C11	PT16B	383	14,96	C20	PT183	554	16,18
C10	PT169	1755	6,19	C11	PT16C	778	10,82	C20	PT184	716	14,15
C10	PT16A	509	12,04	C11	PT172	175	26,06	F00	PT118	443	19,89
C10	PT16B	518	12,44	C11	PT181	1872	6,56				
C10	PT16C	1013	8,27	C11	PT182	2534	5,39				

Tabela 4.23: Estimativas da área (km²) e respetivos CV, para os níveis 2 e 3 das classes de ocupação do solo, por NUTS 3, cujos CV estão dentro dos limites máximos de precisão (em 2012)

Classes LC (Nível 2 e 3)	NUTS3	AREA_ESTIM	CV	Classes LC (Nível 2 e 3)	NUTS3	AREA_ESTIM	CV	Classes LC (Nível 2 e 3)	NUTS3	AREA_ESTIM	CV
A10	PT163	56	7,52	C10	PT169	1681	8,57	C12	PT166	642	10,89
A10	PT164	50	3,4	C10	PT16C	1083	8,83	C12	PT169	498	16,31
A20	PT162	198	11,63	C10	PT171	131	7,42	C12	PT181	481	14,25
A20	PT163	142	10,22	C11	PT111	286	19,53	C20	PT115	69	11,11
A20	PT164	26	3,37	C11	PT115	280	19,32	C20	PT162	132	11,66
B10	PT163	80	6,01	C11	PT116	290	14,35	C20	PT164	244	15,72
B10	PT164	101	4,89	C11	PT118	891	10,61	C20	PT169	406	13,19
B20	PT111	5	21,03	C11	PT150	468	14,99	C20	PT16C	89	18,02
B20	PT162	4	21,63	C11	PT161	612	9,25	C20	PT171	36	19,68
B20	PT163	19	6,61	C11	PT162	459	11,71	C30	PT115	457	16,73
B20	PT165	46	21,86	C11	PT164	671	11,38	C30	PT162	5	13,21
B20	PT166	6	26,8	C11	PT165	579	12,67	C30	PT163	24	14,83
B20	PT171	25	16,34	C11	PT166	398	15,25	C30	PT164	169	8,07
B40	PT113	21	25	C11	PT169	1183	8,73	C30	PT165	76	16,17
B40	PT163	11	8,09	C11	PT16B	317	15,12	E30	PT163	57	6,45
B40	PT167	8	24,09	C11	PT16C	815	8,6	F10	PT165	140	19,27
B50	PT150	35	21,31	C11	PT172	229	19,5	F10	PT167	27	17,94
B50	PT162	61	15,07	C11	PT181	2351	4,51	F40	PT163	44	12,75
B50	PT164	27	4,6	C11	PT182	3053	3,8	F40	PT165	141	14,86
B50	PT169	41	15,01	C11	PT183	3539	3,7	F40	PT168	70	12,81
B50	PT171	6	12,05	C11	PT184	2470	5,37	F40	PT16C	61	12,55
B50	PT184	197	21,35	C11	PT185	1742	5,82	G10	PT150	21	19,89
B70	PT115	57	18,18	C12	PT117	411	15,61	G10	PT162	6	14,98
B70	PT150	492	14,88	C12	PT118	631	13	G10	PT163	4	3,42
B70	PT163	36	7,3	C12	PT150	387	16,63	G10	PT169	39	12,16
B70	PT168	60	12,65	C12	PT161	226	18,8	G10	PT16C	20	15,92
B70	PT171	39	10,66	C12	PT162	421	13,1	G20	PT162	21	12,12
B70	PT172	45	28,8	C12	PT163	449	12,03	G30	PT162	39	11,8
C10	PT115	428	14,91	C12	PT164	408	15,85				
C10	PT163	730	6,46	C12	PT165	595	12,53				

4.3. Comparação de estimativas para as classes de ocupação do solo (Entre o projeto LUCAS e outros sistemas de classificação nacionais)

De seguida são comparadas as estimativas das áreas das classes de ocupação do solo, em Km², obtidas a partir dos dados recolhidos nos inquéritos LUCAS realizados em 2009 e 2012 (registados nas bases de dados LUCAS *survey*) e as áreas obtidas a partir de vários sistemas de classificação nacionais (NDS). A finalidade deste processo de comparação de estimativas prende-se com os objetivos propostos para o estudo piloto desenvolvido pelo INE apresentados na introdução deste trabalho de investigação.

As principais fontes de dados nacionais a partir do qual se obtiveram as estimativas das áreas foram a COS 2007, recenseamento agrícola (RA), inventário de florestas nacionais (IFN), Parcelário agrícola, Informação de fontes hídricas do sistema nacional e resultados obtidos através de deteção remota (a partir da classificação automática de imagens (OBIA-*Object based image analysis*)).

A comparação dos resultados nesta secção é realizada para os três níveis da nomenclatura LUCAS para as NUTS 1, NUTS 2 e NUTS 3.

4.3.1. Nível 1 da nomenclatura de classificação LUCAS

4.3.1.1. Comparação de resultados por NUTS 1 (Portugal Continental)

As diferenças entre as estimativas obtidas a partir dos dados recolhidos nos inquéritos do projeto LUCAS em 2009 e 2012 para Portugal continental (PT1) relativamente aos resultados provenientes da COS 2007 e de outros sistemas de classificação nacionais (NDS) são apresentadas na Tabela 4.24.

Tabela 4.24: Diferenças entre os resultados estimados com base nos dados LUCAS de 2009 e 2012 com os resultados obtidos de outras fontes de classificação (NDS) e resultados exclusivos da COS 2007 (por NUTS 1 ao nível 1 da nomenclatura de classificação LUCAS)

NUTS 1	LC_L1	ESTIM_AREAS				Diferenças (%)				Diferenças (%)			
		LUCAS 2009	LUCAS 2012	NDS	COS 2007	LUCAS 2009		LUCAS 2012		LUCAS 2009		LUCAS 2012	
						NDS	COS2007	NDS	cos 2007	NDS	COS2007	NDS	COS 2007
PT1	A	4691	5545	4266	4266	425	425	1278	1278	10,0%	10,0%	30,0%	30,0%
PT1	B	16556	15714	15553	30493	1002	-13938	160	-14779	6,4%	-45,7%	1,0%	-48,5%
PT1	C	40929	39317	28718	21214	12211	19715	10599	18103	42,5%	92,9%	36,9%	85,3%
PT1	D	9391	10363	26934	25591	-17544	-16201	-16572	-15229	-65,1%	-63,3%	-61,5%	-59,5%
PT1	E	12267	13465	10579	4414	1688	7853	2886	9050	16,0%	177,9%	27,3%	205,0%
PT1	F	3553	2999	1405	1431	2148	2122	1593	1567	152,8%	148,2%	113,4%	109,5%
PT1	G	1313	1514	1388	1388	-76	-76	126	126	-5,4%	-5,4%	9,1%	9,1%
PT1	H	390	173	244	290	146	100	-71	-116	59,8%	34,6%	-29,0%	-40,2%

Pela análise dos resultados, para a maior parte das classes de ocupação do solo, ao nível 1, observam-se diferenças acentuadas entre as estimativas obtidas pelos dados do LUCAS e os obtidos a partir da COS 2007 e NDS.

Para o caso dos resultados provenientes da COS 2007, apesar da distância temporal relativamente aos dados recolhidos no inquérito LUCAS de 2012, para Portugal continental (PT1) não se pode considerar que existe uma tendência de aumento dos desvios em relação às estimativas obtidas com base nos dados do LUCAS recolhidos em 2009.

Para os caso dos resultados provenientes de outras fontes de classificação, na Tabela 4.24 com designação NDS, exetuoando para terrenos artificializados (A), prados (E) e corpos de água (G), observa-se uma diminuição das diferenças relativamente às estimativas obtidas no projeto LUCAS de 2009 para 2012.

Comparando com as estimativas obtidas a partir dos dados LUCAS, observa-se para as restantes classes, que os resultados dados por NDS são ligeiramente mais próximos, para a maior parte dos casos, dos obtidos através dos dados do LUCAS.

4.3.1.2. Comparação de resultados por NUTS 2

No Anexo 10 são apresentadas para as NUTS 2 as estimativas das áreas das classes de ocupação do solo usando a metodologia de estimação descrita no subcapítulo 3.4, assim como os resultados das áreas provenientes de outras fontes de classificação nacionais.

Analisando as diferenças das estimativas apresentadas no Anexo 11, pode concluir-se que ao nível 1 do sistema de classificação LUCAS, praticamente para todas as classes de ocupação do solo, as estimativas entre os diferentes sistemas de classificação são muito diferentes. Efetivamente, excetuando os casos de territórios artificializados (A), terrenos agrícolas (B), prados (classe E) e zonas húmidas (classe H), para as NUTS 2 assinaladas a cinzento na Tabela 4.30, todas as outras diferenças apresentam valores superiores a 10%.

4.3.1.3. Comparação de resultados por NUTS 3

As diferenças para as NUTS 3 entre as estimativas das áreas obtidas a partir dos dados recolhidos nos inquéritos de 2009 e 2012 do projeto LUCAS e as áreas

provenientes da COS 2007 e outras fontes de classificação nacionais são apresentadas nos Anexos 12 e 13. Pelos resultados obtidos, observa-se que para a maioria das classes de ocupação do solo, ao nível da nomenclatura de classificação LUCAS, as áreas obtidas são bastante diferentes.

Nas Tabela 4.25 e 4.26 são apresentadas as classes de ocupação do solo cujas estimativas das áreas usando a metodologia de estimação apresentada em 3.6 e 3.10 do projeto LUCAS relativamente às áreas de outras fontes de classificação NDS apresentam diferenças inferiores a 10% e que têm limites de precisão abaixo dos limites máximos estabelecidos pelo Eurostat (Anexo 5).

Tabela 4.25: Diferenças entre os resultados estimados com base nos dados LUCAS de 2009 com os resultados obtidos de outras fontes de classificação (NDS), por NUTS 3 ao nível 1 da nomenclatura de classificação LUCAS (com diferenças inferiores a 10 % e dentro do limites de qualidade de precisão)

NUTS3	Classes de ocupação do solo	LUCAS/Est_Area (2009)	Área/COS2007	Dif.(%)	Área/NDS	Dif.(%)
PT16B	B	546			541	1,0
PT184	B	2268			2513	-9,7
PT112	C	426	424	0,5		
PT114	C	200	220	-9,3	210	0,0
PT116	C	444	414	7,2		
PT168	C	1551			1434	8,2
PT16B	C	625	571	9,5		
PT118	E	1226			1223	0,2
PT150	E	507			496	2,1
PT181	E	913			1003	-9,0
PT182	E	1193			1317	-9,5
PT184	E	1826			1691	8,0

Tabela 4.26: Diferenças entre os resultados estimados com base nos dados LUCAS de 2012 com os resultados obtidos de outras fontes de classificação (NDS), por NUTS 3 ao nível 1 da nomenclatura de classificação LUCAS (com diferenças inferiores a 10 % e dentro do limites de qualidade de precisão)

NUTS3	Classes de ocupação do solo	LUCAS/Est_Area (2012)	Area/COS2007	Dif.(%)	Area/NDS	Dif.(%)
PT117	B	982			1025	-4,2
PT118	B	1448			1558	-7,0
PT150	B	1088	1037	4,8		
PT162	B	359			343	4,7
PT166	B	187	189	-1,0		
PT169	B	611			583	4,8
PT16A	B	335			308	8,6
PT182	B	1121			1106	1,4
PT184	B	2296			2513	-8,6
PT185	B	827			906	-8,7
PT112	C	415	424	-2,2		
PT115	C	954	884	7,9	973	-2,0
PT116	C	404	414	-2,5	376	7,4
PT150	C	1605			1755	-8,5
PT165	C	1486	1503	-1	1506	-1,3
PT184	C	3263			3013	8,3
PT181	G	150	153	-2	153	-1,8

4.3.2. Nível 2 e 3 da nomenclatura de classificação LUCAS

4.3.2.1. Comparação de resultados por NUTS 1 (Portugal Continental)

Na Tabela 4.27 são apresentadas as estimativas das áreas para as classes de ocupação do solo, para os níveis 2 e 3 da nomenclatura de classificação LUCAS, obtidas através dos dados recolhidos no LUCAS em 2009 e 2012 e as áreas provenientes de outras fontes de classificação.

Na Tabela 4.27 são, igualmente, assinaladas a cinzento as classes para as quais os resultados das áreas são mais próximos (com diferenças inferiores a 20%). Para todos estas classes os CV associados às estimativas das áreas baseadas nos dados LUCAS não excedem os limites máximos apresentados no Anexo 5.

Pelo valores das medidas de localização apresentados na Tabela 4.28, confirma-se, mesmo para Portugal Continental, que as áreas obtidas a partir dos dados LUCAS e as provenientes de NDS diferem acentuadamente para as classes de ocupação do solo para os níveis 2 e 3.

Tabela 4.27: Diferenças entre os resultados estimados com base nos dados LUCAS de 2012 com os resultados obtidos de outras fontes de classificação (NDS), por NUTS 3 ao nível 1 da nomenclatura de classificação LUCAS

NUTS 1	Classes de ocupação do solo	Area / NDS	LUCAS 2009		LUCAS 2012	
			ESTIM_AREA	Diferença LUCAS / NDS	ESTIM_AREA	Diferença LUCAS / NDS
PT1	A10	2918,2	1534,2	-47,4%	2123	-27,3%
PT1	A20	1348,2	3156,8	134,2%	3422	153,8%
PT1	B10	3455,6	4963,6	43,6%	4450	28,8%
PT1	B20	173,3	361,2	108,4%	398	129,4%
PT1	B30	247,6	276,2	11,5%	111	-55,0%
PT1	B40	610,5	703,0	15,2%	594	-2,7%
PT1	B50	4739,8	1240,0	-73,8%	1154	-75,7%
PT1	B70	1193,4	1595,8	33,7%	2025	69,7%
PT1	B80	5133,4	7415,7	44,5%	6982	36,0%
PT1	C10	26273,3	31617,2	20,3%	30474	16,0%
PT1	C11	18197,7	20889,7	14,8%	21684	19,2%
PT1	C12	8068,2	7138,6	-11,5%	7452	-7,6%
PT1	C13	7,4	3588,9	48669,3%	1338	18080,6%
PT1	C20	2178	7635,4	250,6%	5663	160,0%
PT1	C30	267	1676,7	527,8%	3180	1090,7%
PT1	D10	7475	2582,0	-65,5%	2923	-60,9%
PT1	D20	19459	6808,7	-65,0%	7440	-61,8%
PT1	E10	2341	593,3	-74,7%	1523	-35,0%
PT1	E20	3571	3173,8	-11,1%	9306	160,6%
PT1	E30	4667	8499,9	82,1%	2636	-43,5%
PT1	F00	1405,2	3553,0	152,8%	—	—
PT1	F10	495	—	—	968	95,5%
PT1	F20	84	—	—	62	-25,5%
PT1	F30	826	—	—	1968	138,2%
PT1	G10	772	716,8	-7,2%	824	6,7%
PT1	G20	305	428,6	40,6%	536	75,7%
PT1	G30	311	167,1	-46,3%	155	-50,3%
PT1	H10	41	52,1	27,2%	23	-45,1%
PT1	H20	203	337,9	66,4%	151	-25,8%

Tabela 4.28: Medidas de localização das diferenças (em %) entre dados LUCAS, por nível 2 e 3 da nomenclatura do projeto LUCAS, com os resultados obtidos através de outras fontes de classificação nacionais para Portugal Continental (PT1)

LUCAS	Min.	Q1	Mediana	Média	Q3	Máx.
2009	7,2%	19,0%	46,9%	1947,9%	88,7%	48669,3%
2012	2,7%	28,4%	58,0%	797,5%	131,6%	18080,6%

4.3.2.2. Comparação de resultados por NUTS 2

Na Tabela 4.29 são apresentadas as estimativas das áreas para as classes de ocupação do solo ao nível 2 e 3, para as quais as diferenças entre as estimativas do LUCAS e as áreas dadas por NDS são mais próximas (valores inferiores a 20%) e cujos CV estão dentro dos limites de erro de precisão para as estimativas usando a metodologia de estimação do LUCAS (Anexo 5).

Tabela 4.29: Diferenças entre os resultados estimados com base nos dados LUCAS de 2012 com os resultados obtidos de outras fontes de classificação (NDS), por NUTS 2 para os níveis 2 e 3 da nomenclatura de classificação LUCAS (com diferenças inferiores a 20 % e dentro do limites de qualidade de precisão)

NUTS2	Classes de ocupação do solo	LUCAS 2009 ESTIM_AREA	Area/NDS	Dif.(%)	NUTS2	Classes de ocupação do solo	LUCAS 2012 ESTIM_AREA	Area/NDS	Dif.(%)
PT11	B70	686	592	16,0	PT11	B80	1742	1591	9,5
PT11	C10	5739	5504	4,3	PT11	F40	529	444	19,1
PT11	C11	3101	3750	-17,3	PT15	D10	343	360	-4,8
PT11	C12	1607	1752	-8,3	PT16	C11	5723	6075	-5,8
PT16	B70	327	289	13,0	PT16	E10	356	322	10,4
PT16	C11	6125	6075	0,8	PT18	B10	2389	2114	13,0
PT18	B10	2393	2114	13,2	PT18	B80	2420	2064	17,2
PT18	B30	262	239	9,8	PT18	G10	504	503	0,2
PT18	E20	2576	2799	-8,0					
PT18	E30	3048	2575	18,4					

Como se observa pelos valores das medidas de localização da Tabela 4.30, para as NUTS 2, as diferenças em percentagem entre as estimativas com base nos dados LUCAS para 2009 e 2012 e as provenientes de outras fontes de classificação nacionais (NDS) são acentuadas. Pode concluir-se que no ano 2012 as estimativas com base nos dados LUCAS se aproximam um pouco mais dos resultados de NDS. Os valores elevados em termos médios e máximos devem-se à existência de classes com um número bastante reduzido de pontos para as NUTS 2.

Tabela 4.30: Medidas de localização das diferenças (em %) entre dados LUCAS, por nível 2 e 3 das classes de ocupação do solo, para NUTS 2, com os resultados obtidos de outras fontes de classificação nacionais

LUCAS	Min.	Q1	Mediana	Média	Q3	Máx.
2009	0,8%	43,1%	72,7%	3676,1%	133,5%	274109,3%
2012	0,2%	40,5%	66,3%	1716,2%	188,7%	117301,3%

4.3.2.3. Comparação de resultados por NUTS 3

Na Tabela 4.31 são apresentadas as estimativas das áreas para as classes de ocupação do solo ao nível 2 e 3, para as quais as diferenças entre as estimativas do LUCAS e as áreas dadas por NDS são inferiores a 20%.

Uma vez que todas as estimativas dadas pela metodologia LUCAS, para estas classes, apresentam CV abaixo dos limites máximos especificados pelo Eurostat (Anexo 5), pode assumir-se para estes casos, que as aproximações dos resultados entre os sistemas de classificação são aceitáveis.

Tabela 4.31: Diferenças entre os resultados estimados com base nos dados LUCAS de 2012 com os resultados obtidos de outras fontes de classificação (NDS), por NUTS 3 para os níveis 2 e 3 da nomenclatura de classificação LUCAS (com diferenças inferiores a 20 % e dentro do limites de qualidade de precisão)

NUTS3	Classes de ocupação do solo	LUCAS 2009 ESTIM_AREA	Area/NDS	Dif.(%)	NUTS2	Classes de ocupação do solo	LUCAS 2012 ESTIM_AREA	Area/NDS	Dif.(%)
PT184	B10	953	940	1,4	PT111	B2	5	5	-12,3
PT118	B70	329	357	-7,8	PT164	C11	671	778	-13,8
PT112	C10	359	340	5,6	PT184	C11	2470	2250	9,8
PT114	C10	200	206	-3,1	PT164	C12	408	343	19,0
PT115	C10	942	893	5,5	PT181	C12	481	411	16,9
PT116	C10	386	367	5,2	PT115	C20	69	75	-9,0
PT165	C10	1429	1490	-4,1	PT162	G1	6	7	-14,9
PT182	C10	2860	2424	18,0	PT169	G1	39	32	19,7
PT184	C10	2649	2991	-11,4					
PT164	C11	679	778	-12,7					
PT184	C11	2184	2250	-2,9					
PT165	C12	488	433	12,7					

Como se observa pelos valores das medidas de localização da Tabela 4.32, as diferenças em percentagem entre as estimativas com base nos dados LUCAS, para 2009 e 2012, e as provenientes de outras fontes de classificação nacionais (NDS) para as NUTS 3 são ainda mais acentuadas do que as obtidas para NUTS 2 (Tabela 4.30).

Tabela 4.32: Medidas de Localização das diferenças (em %) entre dados LUCAS, por nível 2 e 3 das classes de ocupação do solo, para NUTS 3, com os resultados obtidos de outras fontes de classificação nacionais

LUCAS	Min.	Q1	Mediana	Média	Q3	Máx.
2009	0,8%	45,0%	86,0%	9994,8%	339,5%	1563096,3%
2012	0,4%	43,0%	83,3%	4943,1%	426,2%	730340,2%

4.4. Estimativas para as classes de uso do solo

De seguida são apresentadas as estimativas obtidas para as áreas, em Km², das classes de uso do solo, ao nível 2 da nomenclatura do sistema de classificação LUCAS, agricultura (U110), floresta (U120), indústrias extrativas (U130), serviços culturais, entretenimento e recreativos (U340) e áreas naturais ou com outros usos socioeconómicos (U612), utilizadas no âmbito do estudo piloto “sinergias entre o projeto LUCAS e sistemas de classificação nacionais” desenvolvido pelo INE.

Para a obtenção das estimativas das classes foram utilizados os estimadores para a área e variância descritos em 3.6 e 3.10 a partir de cada NUTS 1, NUTS 2 e NUTS 3 e não através de um processo de agregação de NUTS de níveis inferiores para níveis superiores.

Para além das estimativas das áreas obtidas a partir dados de 2009 e 2012 registados nas bases de dados LUCAS *survey para Portugal*, é igualmente efetuada uma avaliação da qualidade das mesmas pela análise dos coeficientes de variação (CV) obtidos tendo em consideração os limites superiores para os erros estipulados pelo Eurostat (Anexo 5).

4.4.1. Avaliação das estimativas para a NUTS 1

Na Tabela 4.33 são apresentados para Portugal continental (PT1) as estimativas das áreas e os coeficientes de variação obtidos para as cinco classes de uso do solo referidas anteriormente.

Pelos resultados obtidos, observa-se que apenas as estimativas para as áreas das classes agricultura (U110), floresta (U120) e áreas naturais ou com outros usos socioeconómicos (U612) apresentam coeficientes de variação baixos para as estimativas das áreas.

Comparando as estimativas obtidas para os dados recolhidos em 2009 e 2012 nos inquéritos LUCAS, pode concluir-se que o aumento do número de pontos da amostra LUCAS 2012 não teve implicações para uma melhoria da precisão das estimativas das cinco classes de uso do solo em estudo.

Tabela 4.33: Estimativas da área (km²) e respetivos CV, para as classes de uso do solo para os dados LUCAS recolhidos em 2009 e 2012 para Portugal continental (PT1)

Classes de uso do solo	2009		2012	
	ESTIM_AREA	CV	ESTIM_AREA	CV
U110	32875,9	1,46	38657	1,18
U120	33389,9	1,55	29401	1,53
U130	254,2	25,63	204	24,79
U340	362,5	21,13	246	22,16
U612	--	--	6417	4,08

4.4.2. Avaliação das Estimativas ao nível da NUTS 2

Na Tabela 4.34 são apresentados para as NUTS 2 as estimativas das áreas e os coeficientes de variação obtidos para as cinco classes de uso do solo referidas anteriormente utilizando os estimadores apresentados em 3.6. e 3.10.

Tabela 4.34: Estimativas da área (km²) e respetivos CV, para as classes de uso do solo para os dados LUCAS recolhidos ao nível das NUTS 2 em 2009 e 2012

Classes de uso do solo	NUTS 2	2009		2012	
		ESTIM_AREA	CV	ESTIM_AREA	CV
U110	PT11	7683	3,15	7597	2,81
U110	PT15	1425	8,13	1994	5,46
U110	PT16	8371	3,07	8382	2,70
U110	PT17	979	8,68	1221	6,90
U110	PT18	14392	1,99	19043	1,48
U120	PT11	5855	4,14	5906	3,55
U120	PT15	886	12,1	1406	7,41
U120	PT16	12568	2,35	11684	2,18
U120	PT17	444	15,9	539	13,33
U120	PT18	13612	2,19	9849	2,74
U130	PT11	33	69,6	7	126,22
U130	PT15	48	55,2	29	61,51
U130	PT16	100	40,6	106	33,89
U130	PT17	36	71,3	5	181,21
U130	PT18	32	70,3	38	56,05
U340	PT11	53	57,6	23	71,57
U340	PT15	33	70,5	24	71,3
U340	PT16	114	37,4	64	43,78
U340	PT17	82	43,6	58	45,17
U340	PT18	82	44,2	35	58,58
U612	PT11	—	—	2756	6
U612	PT15	—	—	282	19,3
U612	PT16	—	—	2591	6,34
U612	PT17	—	—	284	20,43
U612	PT18	—	—	754	12,34

Pela análise das medidas de localização associadas aos CV das estimativas das áreas das Tabelas 4.35 e 4.36, observa-se que, como ocorre para Portugal continental (PT 1), para as NUTS 2, as únicas classes que apresentam estimativas com precisões mais altas são agricultura (U110) e floresta (U120) e áreas naturais ou com outros usos socioeconómicos (U612). Os resultados obtidos para as classes indústrias extrativas (U130), serviços culturais, entretenimento e recreativos (U340) apresentam CV altos e conseqüentemente as precisões são baixas.

Tabela 4.35: Medidas de localização para os valores CV das classes de uso do solo, por NUTS 2, para os dados recolhidos nos LUCAS de 2009

Classes de uso do solo	Min.	Q1	Mediana	Média	Q3	Máx.
U110	2,0	3,1	3,2	5,0	8,1	8,7
U120	2,2	2,4	4,1	7,3	12,1	15,9
U130	40,6	55,2	69,6	61,4	70,3	71,3
U340	37,4	43,6	44,2	50,7	57,6	70,5

Tabela 4.36: Medidas de localização para os valores CV das classes de uso do solo, por NUTS 2, para os dados recolhidos nos LUCAS de 2012

Classes de uso do solo	Min.	Q1	Mediana	Média	Q3	Máx.
U110	1,5	2,7	2,8	3,9	5,5	6,9
U120	2,2	2,7	3,6	5,8	7,4	13,3
U130	33,9	56,1	61,5	91,8	126,2	181,2
U340	43,8	45,2	58,6	58,1	71,3	71,6
U612	6,0	6,3	12,3	12,9	19,3	20,4

4.4.3. Avaliação das Estimativas ao nível da NUTS 3

Nos Anexos 14 e 15 são apresentadas ao nível das NUTS 3 as estimativas das áreas e os coeficientes de variação obtidos utilizando os estimadores descritos na secção 3.4 para as cinco classes de uso do solo analisadas neste trabalho agricultura (U110), floresta (U120), indústrias extrativas (U130), serviços culturais, entretenimento e recreativos (U340) e áreas naturais ou com outros usos socioeconómicos (U612).

Pela análise das medidas de localização associadas aos CV das estimativas das áreas das classes de uso do solo, apresentadas na Tabela 4.37 e Tabela 4.38, ao nível das NUTS 3, observa-se que continuam a ser agricultura (U110), floresta (U120) e áreas naturais ou com outros usos socioeconómicos (U612) as classes que apresentam as melhores precisões.

Observa-se, no entanto, um aumento generalizado dos CV para todas as medidas de localização comparando com os resultados obtidos para as NUTS 2 (Tabelas 4.35 e 4.36). Para os casos das classes U110 e U120, os valores máximos passam a ser, respetivamente, 25.3 e 48.1 (para os dados do ano 2009) e 19.1 e 27.5 (para 2012), enquanto para NUTS 2 o valor mais elevado é 15.9 (para a classe U120).

Para o caso da classe U612 é possível observar que o valor mediano (com CV=26.93) corresponde a uma estimativa com menor precisão quando comparado com o valor máximo alcançado para NUTS 2 (CV=20.4) apresentado na Tabela 4.42.

Tabela 4.37: Medidas de localização para os valores dos CVs das estimativas LUCAS 2009 para uso do solo, ao nível das NUTS 3

Classes de uso do solo	Min.	Q1	Mediana	Média	Q3	Máx.
U100	3,3	6,3	11,5	10,3	13,3	25,3
U120	4,2	6,2	8,7	10,6	12,0	48,1
U130	55,2	79,1	98,4	88,3	100,3	101,8
U340	54,8	70,4	98,4	85,4	99,9	102,9

Tabela 4.38: Medidas de localização para os valores dos CVs das estimativas LUCAS 2012 para uso do Solo, ao nível das NUTS 3

Classes de uso do solo	Min.	Q1	Mediana	Média	Q3	Máx.
U100	2,4	5,6	8,9	8,4	10,9	19,1
U120	4,6	6,3	7,3	9,2	11,0	27,5
U130	57,8	61,5	75,8	92,2	130,2	146,0
U340	60,8	69,0	73,5	99,7	140,0	179,2
U612	11,7	15,8	26,9	32,8	33,9	129,4

4.5. Comparação de estimativas para as áreas de uso do solo (Entre o projeto LUCAS e outros sistemas de classificação nacionais)

De seguida são comparadas as estimativas das áreas, em Km², obtidas a partir dos dados recolhidos nos inquéritos LUCAS realizados em 2009 e 2012 (presentes nas bases de dados LUCAS *survey*) e as áreas obtidas a partir de vários sistemas de classificação nacionais (NDS). A finalidade deste processo de comparação de estimativas prende-se com os objetivos propostos para o estudo piloto desenvolvido pelo INE apresentados na introdução deste trabalho de investigação.

Como foi referido anteriormente as principais fontes de dados nacionais (NDS) foram a COS 2007, recenseamento agrícola (RA), inventário de florestas nacionais (IFN), parcelário agrícola, informação de fontes hídricas do sistema nacional e resultados obtidos através de deteção remota (a partir da classificação automática de imagens (OBIA- *Object based image analysis*)).

A comparação das áreas incidu exclusivamente sobre as classes de uso do solo, ao nível 2 da nomenclatura do sistema de classificação LUCAS, agricultura (U110), floresta (U120), indústrias extrativas (U130), serviços culturais, entretenimento e recreativos (U340) e áreas naturais ou com outros usos socioeconómicos (U612) para as NUTS 1, NUTS 2 e NUTS 3.

4.5.1. Comparação de resultados ao nível das NUTS 1

Na Tabela 4.39 são apresentadas para Portugal continental as estimativas das áreas, os CV e as diferenças entre os resultados obtidos no projeto LUCAS e as estimativas provenientes por outras fontes nacionais, para as cinco classes de uso do solo.

Pelos resultados obtidos, constata-se que as estimativas das áreas das classes de uso do solo obtidas com base nos dados LUCAS *survey* dos anos 2009 e 2012 apresentam diferenças significativas relativamente aos resultados de NDS.

Para os dados recolhidos no LUCAS de 2009, observa-se que a classe agricultura (U110) é a única que apresenta estimativas obtidas através da metodologia de estimação usada no LUCAS que são semelhantes com as provenientes de outras fontes nacionais de dados.

Para os caso dos dados obtidos em 2012, as diferenças entre as estimativas para as classes agricultura (U110) e indústrias extrativas (U130) aumentaram. As diferenças para as classes), floresta (U120) e serviços culturais, entretenimento e recreativos (U340) diminuíram. Observa-se que áreas naturais ou com outros usos socioeconómicos (U612) é a única cujas estimativas LUCAS se aproximam ligeiramente das provenientes de outros sistemas de classificação nacionais (diferença de 8,3%).

Tabela 4.39: Diferenças entre os resultados estimados através dos dados LUCAS 2009 e 2012, para as classes LU, para Portugal continental, com os resultados obtidos de outras fontes de dados nacionais (NDS)

NUTS 1	Classes de uso do solo	Dados LUCAS				NDS	Diferenças (%)	
		2009		2012			LUCAS 2009/NDS	LUCAS 2012/NDS
		ESTIM_AREA	CV	ESTIM_AREA	CV			
PT1	U110	32875,9	1,46	38657,2	1,18	32184,5	2,1%	20,1%
PT1	U120	33389,9	1,55	29400,6	1,53	8374,3	298,7%	251,1%
PT1	U130	254,2	25,63	203,8	24,79	468,1	-45,7%	-56,5%
PT1	U340	362,5	21,13	246,4	22,16	130,9	176,9%	88,2%
PT1	U612	–	–	6416,5	4,08	6998,4	–	-8,3%

4.5.2. Comparação de resultados ao nível das NUTS 2

Na Tabela 4.40 são apresentadas as estimativas das áreas e as diferenças entre os resultados obtidos no projeto LUCAS e dados provenientes de NDS ao nível das NUTS 2.

Pelos resultados obtidos, para todas as NUTS 2 observa-se que as estimativas das áreas das classes de uso do solo baseadas nos dados recolhidos no projeto LUCAS de 2009 e 2012, apresentam diferenças consideráveis relativamente aos valores das áreas recolhidos através de outras fontes de dados nacionais.

Em termos médios, a classe agricultura (U110) é a que apresenta menores diferenças (44%). Todas as restantes classes têm diferenças superiores a 100%.

Tabela 4.40: Comparação de resultados ao nível das NUTS 2 entre os dados LUCAS 2009 e os dados provenientes de outras fontes de dados nacionais (NDS)

Classes de uso do solo	ESTIM_AREA (LUCAS 2009)					NDS					Diferenças entre áreas LUCAS e NDS (%)				
	PT11	PT15	PT16	PT17	PT18	PT11	PT15	PT16	PT17	PT18	PT11	PT15	PT16	PT17	PT18
U110	7683	1425	8371	979	14392	6031	743	5395	824	19191	27%	92%	55%	19%	-25%
U120	5855	886	12568	444	13612	2323	466	2608	189	2789	152%	90%	382%	135%	388%
U130	33	48	100	36	32	234	12	130	0	91	-86%	283%	-23%	—	-65%
U340	53	33	114	82	82	26	23	33	26	23	104%	44%	246%	212%	261%
U612	—	—	—	—	—	2272	471	1889	448	1918	—	—	—	—	—

Analisando os resultados da Tabela 4.41, observa-se que os desvios, de um modo geral, também são elevados e que as únicas classes que tiveram estimativas próximas das dadas pela NDS foram, para o Alentejo (PT18), a classe agricultura (U110) e no centro de Portugal (PT15), a classe serviços culturais, entretenimento e recreativos (U340), com 1% e 3% de desvio, respetivamente.

Em termos médios, considerando todas as NUTS 2, a classe U612 é a que apresenta resultados mais próximos dos obtidos por outras fontes (25% de desvios).

Tabela 4.41: Comparação de resultados ao nível das NUTS 2 entre os dados LUCAS 2012 e os dados provenientes de outras fontes de dados nacionais (NDS)

Classes de uso do solo	ESTIM_AREA (LUCAS 2009)					NDS					Diferenças entre áreas LUCAS e NDS (%)				
	PT11	PT15	PT16	PT17	PT18	PT11	PT15	PT16	PT17	PT18	PT11	PT15	PT16	PT17	PT18
U110	7597	1994	8382	1221	19043	6031	743	5395	824	19191	26%	168%	55%	48%	-1%
U120	5906	1406	11684	539	9849	2323	466	2608	189	2789	154%	202%	348%	186%	253%
U130	7	29	106	5	38	234	12	130	0	91	-97%	135%	-18%	—	-58%
U340	23	24	64	58	35	26	23	33	26	23	-13%	3%	93%	120%	55%
U612	2756	282	2591	284	754	2272	471	1889	448	1918	21%	-40%	37%	-37%	-61%

4.5.3. Comparação de resultados ao nível das NUTS 3

Nos Anexos 14 e 15 são apresentadas as estimativas das áreas, em Km², os valores dos CV e as diferenças entre os resultados das áreas entre os dados recolhidos em 2009 e 2012 no projeto LUCAS e os dados provenientes de outros sistemas de classificação nacionais (NDS) para as NUTS 3.

Como se observa a partir dos valores das medidas de localização apresentadas na Tabela 4.42 e Tabela 4.43, as diferenças ao nível das NUTS 3 são elevadas para todas as classes de uso do solo.

Para ambos os anos, a classe agricultura (U110) é a que apresenta menor dispersão nas diferenças. De realçar que os valores mais baixos de CV para as classes serviços culturais, entretenimento e recreativos (U340) e áreas naturais ou com outros usos socioeconómicos (U612) são inferiores ao da classe U110 para o ano de 2012.

Tabela 4.42: Medidas de localização das diferenças (em %) entre dados LUCAS 2009, por Uso do Solo, ao nível das NUTS 2, com os resultados obtidos de outras fontes de classificação nacionais

Classes de uso do solo	Min.	Q1	Mediana	Média	Q3	Máx.
U100	4,6	29,8	42,7	63,1	78,6	249,0
U120	53,4	184,7	307,9	352,4	467,8	1096,8
U130	34,3	54,9	266,8	583,1	1070,6	1715,0
U340	38,4	202,2	344,9	417,2	599,3	1199,8

Tabela 4.43: Medidas de localização das diferenças (em %) entre dados LUCAS 2012, por Uso do Solo, ao nível das NUTS 2, com os resultados obtidos de outras fontes de Classificação nacionais

Classes de uso do solo	Min.	Q1	Mediana	Média	Q3	Máx.
U100	5,9	18,5	35,5	67,3	100,0	267,7
U120	63,8	201,3	259,0	313,2	393,8	1037,3
U130	28,8	41,2	135,3	217,3	289,3	695,8
U340	3,3	60,6	182,8	312,8	371,3	1563,9
U612	5,6	38,7	61,6	1222,7	102,8	12087,0

Capítulo 5

Conclusões

O LUCAS é um projeto de amostragem espacial que tem sido implementado pelo Eurostat com o intuito da obtenção regular de estatísticas ao nível do espaço europeu, tendo por base uma distinção clara entre os conceitos de uso e ocupação do solo. Para tal, tem sido utilizado um esquema de amostragem areolar baseado em pontos, recolhidos num esquema em duas fases (por estratificação). Os pontos são classificados de acordo com a nomenclatura de classificação do LUCAS definida com o propósito de uniformizar a multiplicidade de sistemas de classificação existentes nos países da União Europeia.

Em Portugal continental o projeto foi implementado em 2009 e 2012. Até ao momento foram publicados por parte do Eurostat para cada um dos anos, exclusivamente, estimativas das áreas para as classes de ocupação do solo para Portugal continental (PT1), ao nível 1 da nomenclatura do projeto LUCAS.

Com o propósito de examinar e comparar os resultados baseados nos dados LUCAS com os provenientes de outros sistemas de classificação nacionais, assim como determinar a concordância entre as diferentes nomenclaturas de classificação, o Eurostat empreendeu o desenvolvimento de um estudo piloto pelos estados membros participantes designado por “sinergias entre o projeto LUCAS e outros sistemas de classificação nacionais”.

Tendo em consideração que neste trabalho, assim como no relatório interno do INE enviado em Julho de 2014 para o Eurostat, no âmbito do referido estudo LUCAS piloto, são apresentadas as estimativas das áreas para todas as regiões de Portugal continental (ao nível das NUTS 3), pode concluir-se que os resultados obtidos revestem-se da maior importância para o conhecimento da

composição atual do tipo de solo em Portugal Continental e as funções socioeconómicas que lhe estão associadas.

De seguida são apresentadas as principais conclusões retiradas dos resultados obtidos neste trabalho, tendo em consideração os objetivos propostos inicialmente.

5.1. Avaliação das Estimativas obtidas através da metodologia de estimação usada

Para Portugal continental (PT1), as estimativas obtidas para as áreas das classes de ocupação do solo baseadas nos dados recolhidos em 2009 e 2012, evidenciaram para a maioria da classes coeficientes de variação dentro dos limites máximos estabelecidos pelo Eurostat (Anexo 5).

Efetivamente, para os dados recolhidos em 2009 e 2012, os únicos casos de classes com valores de precisão inferiores aos desejados são, ao nível 1 da nomenclatura LUCAS, zonas húmidas (H) com CV acima de 20%.

Ao nível 2 da nomenclatura, para os dados recolhidos em 2009, corpos de água costeiros (G30), zonas húmidas interiores (H10) e zonas húmidas costeiras (classe H20) apresentam CV superiores aos limites máximos. A estas classes, para os dados recolhidos em 2012, acrescentam-se culturas temporárias industriais (B30) e areias (F20) e outras áreas de solo nu (F40) também com CV superiores aos limites máximos.

Para as NUTS 2, ao nível 1 da nomenclatura proposta, observa-se uma diminuição generalizada da precisão das estimativas. Destacam-se o caso da classe zonas húmidas (H) em que para todas as NUTS 2 os CV estão acima dos limites máximos definidos pelo Eurostat.

As classes corpos de água (G) e solo nu e líquenes (F), para as regiões do Algarve (PT15) e Lisboa (PT17) também apresentam estimativas com CV elevados. Para todos os outros casos o estimador utilizado evidenciou precisões dentro dos limites de referência.

Ainda no que se refere às NUTS 2, para os níveis 2 e 3 da nomenclatura de classificação proposta pelo LUCAS, observa-se uma natural diminuição da

precisão das estimativas devido ao menor número de pontos pertencentes a cada classe. Algarve e Lisboa apresentam-se como as regiões com valores mais elevados de CV, enquanto que a região Norte (PT11) e Alentejo (PT18) são as que apresentam melhores precisões.

Excetuando as classes floresta (C10) e vinha, olival, viveiros e culturas permanentes industriais (B80), para os dados recolhidos em 2009 e 2012, e floresta de folhosas caducifólias (C11) e prados sem árvores (E20), para os dados de 2012, que apresentam estimativas com precisões abaixo dos limites de referência para todas as NUTS 2. Todas as outras classes têm pelo menos algumas NUTS 2 com CV elevados.

No que se refere a NUTS 3, observa-se que o esquema de amostragem LUCAS fornece estimativas com baixas precisões devido ao menor número de pontos selecionados. Efetivamente, a maior parte das estimativas com CV abaixo dos limites máximos, para os três níveis da nomenclatura de classificação, pertencem às classes terrenos terreno agrícolas (B) e florestas (C), enquanto que para as outras classes os CV na generalidade dos casos são elevados.

Para os dados recolhidos em 2012, observa-se para algumas classes, tais como outras áreas de solo nu (F40) e planos de água interiores (G10), CV dentro dos limites máximos para NUTS 3.

Ao nível das classes de uso do solo, neste estudo foram efetuadas estimativas para as classes agricultura (U110), floresta (U120), indústrias extrativas (U130), serviços culturais, entretenimento e recreativos (U340) e áreas naturais ou com outros usos socioeconómicos (U612), para os três níveis NUTS.

Para as classes U110 e U120, para os dados recolhidos em 2009 e 2012, para as NUTS 1 e 2, as estimativas das áreas obtidas evidenciam CV com valores baixos. Para as NUTS 3, apesar da diminuição das precisões, os valores médios dos CV são próximos de 10%, evidenciando precisões aceitáveis tendo em consideração os resultados obtidos para as classes de ocupação do solo.

Para os dados recolhidos em 2012, áreas naturais ou com outros usos socioeconómicos (U612) também apresenta CV baixos. No entanto, para as NUTS 3, os CV são aproximadamente de 30%.

As restantes classes, para ambos os anos, devido ao baixo número de pontos recolhidos na segunda fase do esquema de amostragem, apresentam CV bastante elevados.

Assim, pode concluir-se, pela avaliação efetuada das estimativas para as áreas das classes de ocupação e uso do solo, que a metodologia de estimação utilizada no projeto LUCAS, apenas é recomendável a sua implementação para casos em que a dimensão da amostra é suficientemente grande. Segundo Martino e Palmieri (2009) o esquema de amostragem do projeto LUCAS foi implementado tendo por base a necessidade de em cada domínio, ao nível das NUTS 2, o número mínimo de pontos ser 4. O que não acontece para a generalidade dos casos para NUTS 3.

5.2. Comparação de resultados com outras fontes nacionais de classificação de dados (NDS)

Para além da obtenção e avaliação das estimativas das áreas através dos dados recolhidos no projeto LUCAS em 2009 e 2012, procedeu-se igualmente à comparação das mesmas com os resultados das áreas provenientes de outras metodologias de estimação obtidas através de sistemas de classificação nacionais.

Pelas diferenças obtidas apresentadas nos subcapítulos 4.3 e 4.5 pode concluir-se que os resultados para a grande maioria dos casos não são comparáveis entre si.

A justificação para as diferenças observadas podem dever-se a alguns dos seguintes factores:

- Problemas relacionados com a uniformização das nomenclaturas usadas nos sistemas de classificação nacionais comparativamente com a proposta no projeto LUCAS. Efetivamente, na maior parte dos casos, existem diferenças, relativamente ao número de níveis associados,

assim como à não separação entre os conceitos de uso e ocupação do solo;

- Os sistemas de classificação baseiam-se igualmente em abordagens de amostragem espacial muito diferente. Por exemplo, para a obtenção de resultados foram recolhidos dados provenientes de recenseamentos agrícolas, métodos estatísticos (como o IFN) ou obtidos através de deteção remota (como a COS 2007);
- A questão temporal da recolha de dados utilizados ser muito diferente comparativamente com o momento em que foram os dados recolhidos nos inquéritos LUCAS. Como foi realçado aquando da análise exploratória da distribuição dos pontos nos questionários LUCAS, para além da existência de uma percentagem elevada de pontos com mais do que uma classificação de tipo de classe de uso e ocupação do solo (devido à heterogeneidade do solo), observaram-se alterações significativas, mesmo para o nível 1 da nomenclatura proposta pelo projeto LUCAS, no tipo de classificação de classe de ocupação do solo entre 2009 e 2012 (aproximadamente 20% dos pontos).

5.3. Sugestões para trabalho futuro

De entre as abordagens de estimação espacial, neste estudo foi utilizada uma metodologia *design based* tendo em consideração o esquema de amostragem implementado pelo Eurostat através da utilização de um estimador em duas fases (por estratificação).

Como foi referido na avaliação das estimativas obtidas, observou-se que excetuando algumas classes com um menor número de pontos, tais como as classificadas como corpos de água (G) e zonas húmidas (H), as estimativas para as NUTS 1 e para a generalidade das NUTS 2 evidenciaram CV abaixo dos máximos de referência apresentados pelo Eurostat para o estudo piloto (Anexo 5).

Contudo, ao nível das NUTS 3, excetuando algumas classes a terrenos terreno agrícolas (B) e florestas (C), para os três níveis, uma vez que o número de pontos por classe é muito pequeno, observam-se CV muito altos e fora dos limites recomendados pelo referido.

Tendo em consideração este aspeto, considera-se relevante, em futuros trabalhos de investigação, a implementação de uma metodologia *model based* como por exemplo a técnica em pequenos domínios, pela utilização de informação auxiliar dada, por exemplo, a partir de outras fontes de classificação nacionais (Rao, 2003).

Capítulo 6

Bibliografia

1. AFN, (2012). 6º Inventário Florestal Nacional (Plano de Projeto). Autoridade Florestal Nacional. Versão 2.0, Junho de 2012, Lisboa.
2. Barnett, V. (2004). *Environmental Statistics, Methods and Applications*, John Willey & Sons, Ltd.
3. Bettio M., Delincé J., Bruyas P., Croi W., Eiden G. (2002). *Area frame surveys: Aim, Principals and Operational Surveys, Building Agro Environmental Indicators, focussing on the European area frame survey LUCAS*, European Commission.
4. Brus, D. J. (2010). *Design-based and model-based sampling strategies for soil monitoring*, 19 th World Congress of Soil Science, Soil Solutions for a Changing World.
5. Brus, D. J. (2011). *Dealing with Contaminated Sites: From Theory towards Practical Application* (chap. 4), National Institute of Public Health and the Environment (RIVM), Bilthoven.
6. Brus, D. J., Knotters, M. (2008). *Sampling design for compliance monitoring of surface water quality: A case study in a Polder area*, *Water resources*, 44, W11410.
7. Brus, D.J., Knotters M., Metzger M. J., Walvoort, D.J.J. (2011). *Towards a European-wide sampling design for statistical monitoring of common habitats*, Alterra report 2213, Wageningen.
8. Cadima, E. L., Caramelo, A.M., Afonso-Dias, M., Barros, P. C., Tandstad, M.O., de Leiva-Moreno, J.I. (2005). *Sampling methods applied to fisheries science: a manual*. FAO Fisheries department Rome.
9. Chun, Y., Griffith, D.A. (2013). *Spatial Statistics & Geostatistics: Theory and Applications for Geographic Information Science & Technology (GIST)*, SAGE Publications Ltd.
10. Cochran W., (1977). *Sampling Techniques*. New York: John Wiley and Sons.

11. DGT (2013). Relatório Anual 2012-2013. LANDYN - Alterações de uso e ocupação do solo em Portugal Continental: caracterização, forças motrizes e cenários futuros, Relatório NUT II, Direção Geral do Território
12. Eurostat (2012a). <http://epp.eurostat.ec.europa.eu/portal/page/portal/lucas>, página com as bases de dados LUCAS 2009 e 2012 para os 27 União Europeia nos quais o projeto LUCAS foi implementado.
13. Eurostat (2012b). LUCAS 2012-Technical Reference Document: C-2: Field form”, European Commission.
14. Eurostat (2013a). Land cover and land use, landscape (LUCAS). Reference Metadata in Euro SDMX Metadata Structure (ESMS). Compiling agency: Eurostat, the statistical office of the European Union. Last update: 25 October 2013.
15. Eurostat (2013b). LUCAS 2012-Technical Reference Document:C-1 Instructions for Surveyors (Revised), European Commission.
16. Eurostat (2013c). LUCAS 2012-Technical Reference Document: C-3: Land use and Land Cover Classification (revised), European Commission.
17. FAO (1996). Multiple frame agricultural surveys, volume 1: Current surveys based on area and list sampling methods, FAO statistical development series.
18. Ferreira, M. C., Campos, P. (2013). Inventário sobre as Principais Metodologias no Domínio da Amostragem sobre o Uso/Ocupação do Solo, DMSI/ME – Instituto Nacional de Estatística.
19. Fuller, W. A. (1999). Estimation Procedures for the United States National Resources Inventory, SSC Annual Meeting, June 1999, Proceeding of the Survey Methods Section.
20. Gallego, J. (1995). Sampling frames of square segments an agricultural information system for the European Union, Joint Research Centre, European Commission.
21. Gallego, J. (2007). Sampling efficiency of the EU point survey LUCAS 2006, Paper presented at the 56th ISI session, Lisbon, 22-29 September 2007.
22. Gallego, J. (2013). The use of a Point Sample as a Master Frame for Agricultural Statistics, ICAS VI, Sixth International conference on Agricultural Statistics.

23. Gallego, J., Bamps, C., (2008). Using CORINE land cover and the point survey LUCAS for area estimation. *International Journal of Applied Earth Observation and Geoinformation*, 10 (2008): 467-475.
24. Gruijter J., Brus, D.J., Bierkens, M., Knotters, M. (2006). *Sampling for Natural Resource Monitoring*, Springer.
25. Gruijter, J., ter Braak, C. J. F. (1990). Model-Free Estimation from Spatial: A reappraisal of classical sampling theory. *Mathematical Geology* 22 (4): 407 – 415.
26. Jacques, P., Gallego, J. (2005). The LUCAS project – The new methodology in the 2005/2006 surveys. *Agri-environment workshop: Belgirate, September de 2005*.
27. Kim, S. D., Lee, S. E. (2013). A note on double sampling estimate and post estimation with BLS outlier adjustment, *Proceedings 59th ISI World Statistics Congress, 25-30 August 2013, Hong Kong*.
28. Kincaid T., Olsen T., Stevens D., Platt C., White C., Remington R. (2014). *spsurvey package: Spatial Survey Design and Analysis, CRAN r-project*.
29. Landyn (2013). <http://landyn.isegi.unl.pt/>, página do projeto *Landyn*, acedida em 16/07/2014.
30. Manly B. F.J. (2001). *Statistics for Environmental Science and Management*, Chapman & Hall / CRC.
31. Martino, L., Palmieri, A., Gallego, J. (2009). Use of auxiliary information in the sampling strategy of a European area frame agro-environmental survey, *First Italian Conference on Survey Methodology. Siena: 10-12 June 2009*.
32. Martino, L, Palmieri, A. (2009). LUCAS 2009 Sampling weights, Eurostat
33. Martino, L., Fritz, M, (2008). New Insight into Land Cover and Land Use in Europe, Eurostat – *Statistics in focus*, 33/2008.
34. Opsomer, J. D., Fuller, W. A., Li, X. (2003). Replication Variance Estimation for the National Resources Inventory, 2003 Joint Statistical Meetings - Section on Survey Research Methods.
35. Palmieri, L. Martino, P. Dominici and M. Kasanko (2011). Land Cover and Land Use Diversity Indicators in LUCAS 2009 data. *Conference on Land quality and land use information in the European Union - Keszthely (HU) 2011*.

36. Palmieri, A. (2012). LUCAS 2012 Sampling design, Internal Note. Eurostat
37. Rao, J. N. K. (2003). Small Area Estimation, Wiley.
38. Ruiz-Medina, M.D. (2012). New challenges in spatial and spatiotemporal functional statistics for high-dimensional data, *Spatial Statistics*, 1: 82-91.
39. Sahoo, P. M., Singh, R., Rai, A. (2006). Spatial Sampling Procedures for Agricultural Surveys using Geographical Information System, *J. Ind. Soc. Agril. Statist*, 60 (2): 134-143.
40. Sanders, D. H. (1995). *Statistics, A First Course – Fifth Edition*, McGraw-Hill, Inc.
41. Scheaffer, R. L., Mendelhall, W., Ott, R. L. (1996). *Elementary Survey Sampling – Fifth Edition*, Wadsworth Publishing Company.
42. Schoenmakers, B. J. (2005). Desenvolvimento de uma Metodologia de Amostragem para Caracterização da Ocupação do Solo de Portugal Continental, Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Estatística e Gestão de Informação pelo Instituto Superior de Estatística e Gestão de Informação da Universidade Nova de Lisboa.
43. Scott, W. (2008). Cs technical report no.4/07, statistical report. Technical report, Centre for Ecology and Hydrology.
44. Stevens, Jr., Olsen, D.L., (2004). Spatially-balanced sampling of natural resources. *Journal of the American Statistical Association*, 99:262-277.
45. Storch, H.V., Zwiers, F. W. (1999). *Statistical Analysis in Climate Research*, Cambridge University Press.
46. Theobald, D. M., Stevens, Jr., White, D., Urquhart, N. S., Olsen, A. R., Norman, J. B. (2007). Using GIS to Generate Spatially Balanced Random Survey Designs for Natural Resource Applications. *Environmental Mangement*, 40: 134 – 146.
47. Wang, J.-F., Stein, A., Gao, B.-B., Ge Y. (2012). A review of spatial sampling, *Spatial Statistics*, 2: 1-14.

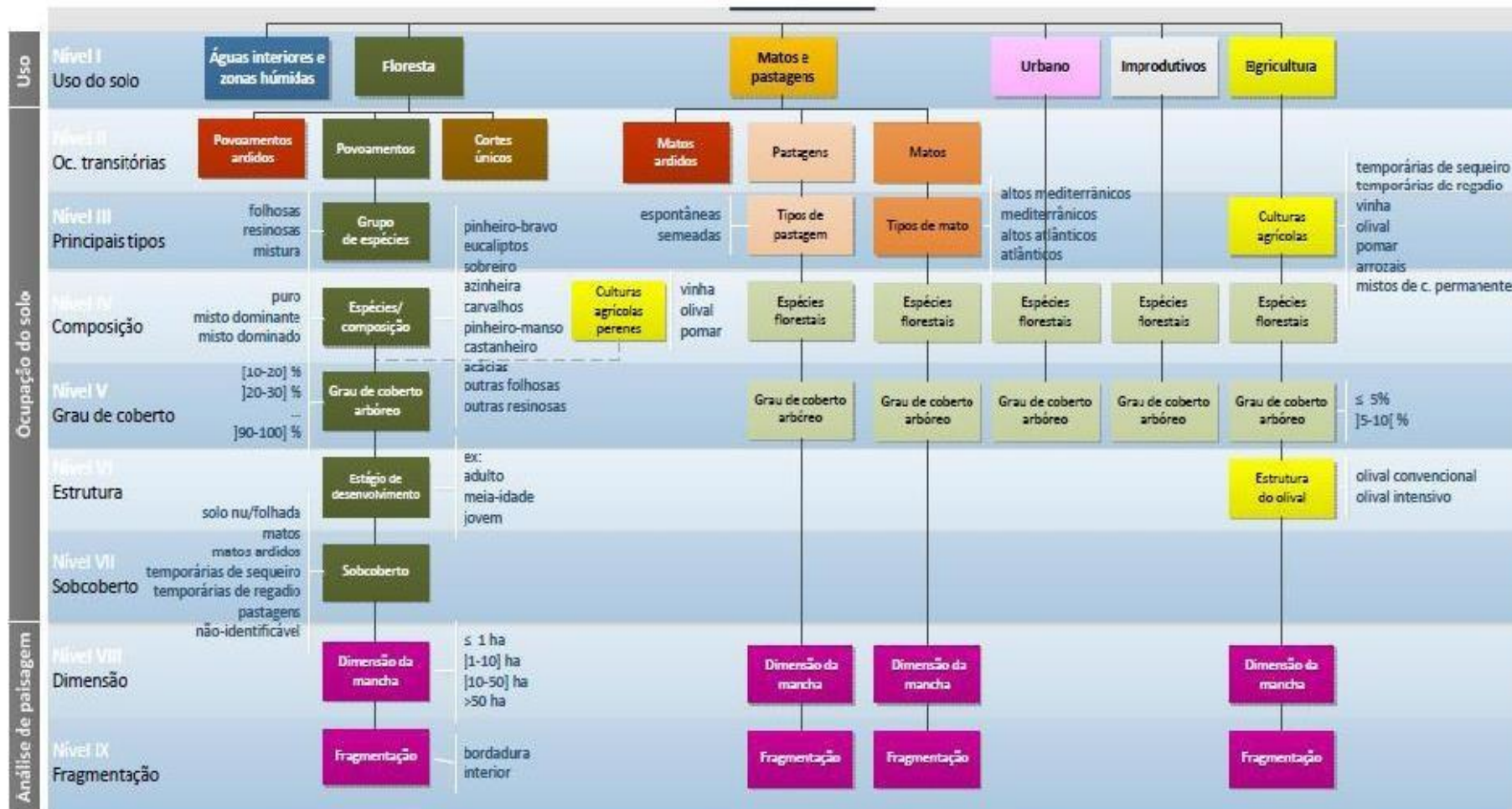
Anexos

Anexo 1: Descrição das NUTS 1, NUTS 2 e NUTS 3 ao nível do território de Portugal (utilizadas no projeto LUCAS)

NUTS 1	Designação	NUTS 2	Designação	NUTS 3	Designacao
PT1	Portugal Continental	PT11	Norte	PT111	Minho-Lima
				PT112	Cávado
				PT113	Ave
				PT114	Grande Porto
				PT115	Tâmega
				PT116	Entre Douro e Vouga
				PT117	Douro
				PT118	Alto Trás-os-Montes
		PT15	Algarve	PT150	Algarve
		PT16	Centro	PT161	Baixo Vouga
				PT162	Baixo Mondego
				PT163	Pinhal Litoral
				PT164	Pinhal Interior Norte
				PT165	Dão-Lafões
				PT166	Pinhal Interior Sul
				PT167	Serra da Estrela
				PT168	Beira Interior Norte
				PT169	Beira Interior Sul
				PT16A	Cova da Beira
		PT16B	Oeste		
		PT16C	Médio Tejo		
		PT17	Lisboa	PT171	Grande Lisboa
				PT172	Península de Setúbal
		PT18	Alentejo	PT181	Alentejo Litoral
				PT182	Alto Alentejo
PT183	Alentejo Central				
PT184	Baixo Alentejo				
PT185	Lezíria do Tejo				

Anexo 2: Nomenclatura do IFN6 (Fonte: AFN, 2012)

100



Anexo 3: Nomenclatura utilizada no projeto *Landyn* (Fonte: DGT, 2013)

Código	Classificação	Descrição	Simplificada
L1	TUC	Tecido urbano contínuo	Artificializados
L2	TUD	Tecido urbano descontínuo	Artificializados
L3	ICE	Indústria, comércio e equipamentos gerais	Artificializados
L4	RVF	Redes viárias e ferroviárias e espaços associados	Artificializados
L5	APO	Áreas portuárias	Artificializados
L6	AER	Aeroportos e aeródromos	Artificializados
L7	AEI	Áreas de extração de inertes	Artificializados
L8	ADR	Áreas de deposição de resíduos	Artificializados
L9	ACO	Áreas em construção	Artificializados
L10	CTS	Culturas temporárias de sequeiro	Agrícolas
L11	CTR	Culturas temporárias de regadio	Agrícolas
L12	ARR	Arrozais	Agrícolas
L13	VIN	Vinhas	Agrícolas
L14	POM	Pomares	Agrícolas
L15	OLI	Olivais	Agrícolas
L16	PAP	Pastagens permanentes	Agrícolas
L17	AAH	Áreas agrícolas heterogéneas	Agrícolas
L18	SAF	Sistemas agro-florestais	Agro-florestais
L19	FFO	Florestas de folhosas (excluindo o eucalipto e espécies invasoras)	Florestas
L20	FRE	Florestas de resinosas	Florestas
L21	FEE	Florestas de eucalipto e espécies invasoras	Florestas
L22	VHN	Vegetação herbácea natural	Incultos
L23	MAT	Matos	Incultos
L24	OUT	Outras formações lenhosas; Cortes e novas plantações; Viveiros florestais; Aceiros e/ou corta-fogos	Florestas
L25	ZDE	Zonas descobertas e com pouca vegetação	Incultos
L26	AAR	Áreas áridas	Incultos
L27	FAF	Florestas abertas de folhosas (excluindo o eucalipto e espécies invasoras)	Florestas
L28	FAR	Florestas abertas de resinosas	Florestas
L29	FAE	Florestas abertas de eucalipto e espécies invasoras	Florestas
L30	ZHU	Zonas húmidas	Zonas húmidas
L31	H20	Corpos de água	Corpos de água
L32	GLF	Campos de Golfe	Artificializados

Anexo 4: Nomenclatura do sistema de classificação do projeto LUCAS 2012

Tipo	Nomenclatura original			Nomenclatura proposta		
	Classes			Classes		
	Nível 1	Nível 2	Nível 3	Nível 1	Nível 2	Nível 3
Land Cover (3 níveis obrigatórios)	A ARTIFICIAL LAND	A10 Built-up areas		A TERRITÓRIOS ARTIFICIALIZADOS	A10 Áreas edificadas	
		A20 Artificial non-built up areas			A20 Áreas não construídas artificiais	
	B CROPLAND	B10 Cereals		B TERRENOS AGRÍCOLAS	B10 Cereais	
		B20 Root crops			B20 Culturas de raízes comestíveis	
		B30 Non-permanent industrial crops			B30 Culturas temporárias industriais	
		B40 Dry pulses, vegetables and flowers			B40 Leguminosas secas, hortícolas e flores	
		B50 Fodder crops			B50 Culturas forrageiras	
		B70 Fruit, nuts trees and berries			B70 Frutos secos, frutos de casca rija, frutos pequenos de baga, citrinos e frutos subtropicais	
		B80 Other permanent crops			B80 Vinha, olival, viveiros e culturas permanentes industriais	
	C WOODLAND	C10 Forest FAO	C11 broadleaved (75%)	C FLORESTA	C10 Floresta (de acordo com a FAO)	C11 Floresta de folhosas caducifólias (75%)
			C12 coniferous (75%)			C12 Floresta de coníferas (75%)
			C13 mixed			C13 Floresta mista
		C20 Other wooded land (FAO)		C20 Outras superfícies arborizadas (FAO)		
		C30 Other		C30 Outras superfícies		

Tipo	Nomenclatura original			Nomenclatura proposta			
	Classes			Classes			
	Nível 1	Nível 2	Nível 3	Nível 1	Nível 2	Nível 3	
		wooded land (non- FAO)			arborizadas (não-FAO)		
	D SHRUBLAND	D10 Shrubland with sparse trees		D MATOS	D10 Matos com árvores esparsas		
		D20 Shrubland without trees			D20 Matos sem árvores		
	E GRASSLAND	E10 Grassland with sparse trees		E PRADOS	E10 Prados com árvores esparsas		
		E20 Grassland without trees			E20 Prados sem árvores		
		E30 Spontaneous vegetation			E30 Vegetação espontânea		
	F BARE LAND AND LICHENS	F10 Rocks and stones		F SOLO NU E LÍQUENES	F10 Afloramentos rochosos e zonas pedregosas		
		F20 Sand			F20 Areias		
		F30 Lichens			F30 Líquenes		
		F40 Other bare soil			F40 Outras áreas de solo nu		
	G WATER	G10 Inland water bodies		G CORPOS DE ÁGUA	G10 Planos de água interiores		
		G20 Inland running water			G20 Águas interiores correntes		
		G30 Coastal water bodies			G30 Corpos de água costeiros		
		G50 Glaciers, permanent snow			G50 Glaciares e neves permanentes		
	H WETLAND	H10 Inland wetlands		H ZONAS HÚMIDAS	H10 Zonas húmidas interiores		
		H20 Coastal wetlands			H20 Zonas húmidas costeiras		
	Land Use (nível 3 é opcional)	U100 PRIMARY PRODUCTION	U110 Agriculture	U111 Commercial agricultural production	U100 PRODUÇÃO PRIMÁRIA	U110 Agricultura	U111 Produção agrícola comercial
				U112 Farming infrastructure			

Tipo	Nomenclatura original Classes			Nomenclatura proposta Classes			
	Nível 1	Nível 2	Nível 3	Nível 1	Nível 2	Nível 3	
				U113 Agricultural production for own consumption			U113 Produção agrícola para consumo próprio
		U120 Forestry	U121 Forestry based on short rotation	U120 Floresta	U120 Floresta	U121 Florestas de rotação curta	
			U122 Forestry based on intermediate or long rotation			U122 Florestas de rotação intermédia ou longa	
			U123 Forestry based on continuous cover			U123 Floresta de coberto contínuo	
		U130 Mining and quarrying	U131 Mining of energy producing materials			U130 Indústrias extractivas	U131 Extração de produtos energéticos
			U132 Mining of metal ores	U132 Extração de minérios metálicos			
		U140 Aquaculture and fishing	U141 Aquaculture	U140 Aquicultura e Pesca	U140 Aquicultura e Pesca	U141 Aquicultura	
			U142 Professional fishing			U142 Pesca profissional	
		U150 Other primary production	U151 Hunting	U150 Outra produção primária	U150 Outra produção primária	U151 Caça	
			U152 Management of migratory animals			U152 Gestão de animais migradores (repovoamento cinegético)	
			U153 Picking of natural products			U153 Colheita de produtos naturais	
	U200 SECONDARY PRODUCTION	U210 Raw industry	U211 Manufacturing of textile products			U200 PRODUÇÃO SECUNDÁRIA	U210 Indústria de Matérias Primas
				U212 Manufacturing of wood and wood based products	U212 Fabricação de madeira e dos produtos à base de madeira		
				U213 Manufacturing of pulp, Paper and paper products	U213 Fabricação de pasta, de papel, cartão e seus artigos		
				U214 Manufacturing of coke, Refined petroleum products and nuclear fuel	U214 Fabricação de coque, de produtos petrolíferos refinados e aglomerados de combustíveis		
				U215 Manufacturing of chemicals chemical products man- made fibbers	U215 Fabricação de produtos químicos e de fibras sintéticas ou artificiais		
				U216 Manufacturing of basic	U216 Indústrias metalúrgicas		

Tipo	Nomenclatura original Classes			Nomenclatura proposta Classes		
	Nível 1	Nível 2	Nível 3	Nível 1	Nível 2	Nível 3
				metals and fabricated metals		
			U217 Manufacturing of non-metallic mineral products			U217 Fabricação de outros produtos minerais não metálicos
			U218 Manufacturing of rubber plastic products			U218 Fabricação de artigos de borracha e de matérias plásticas
		U220 Heavy end product industry	U221 Manufacturing of machinery		U220 Indústria pesada	U221 Fabrico de máquinas
			U222 Manufacturing of vehicles and transport equipment			U222 Fabrico de veículos automóveis e de outros equipamentos de transporte
		U230 Light end product industry	U231 Manufacturing of food, beverages and tobacco products		U230 Indústria ligeira	U231 Indústrias alimentares, de bebidas e de tabaco
			U232 Manufacturing of clothes and leather			U232 Indústrias do vestuário e do couro
			U233 Publishing and printing			U233 Edição e impressão
			U234 Manufacturing of electrical and optical equipment			U234 Fabricação de equipamento eléctrico e óptico
		U240 Energy production	U241 Nuclear based energy production		U240 Produção de Energia	U241 Produção de energia de origem nuclear
			U242 Fossil fuel based energy production			U242 Produção de energia a partir de combustíveis fósseis
			U243 Biomass based energy production			U243 Produção de energia a partir de biomassa
			U244 Renewable energy production			U244 Produção de energia a partir de fontes renováveis
	U300 TERTIARY PRODUCTION	U310 Commercial services	U311 Wholesale and retail trade and repair of vehicles and personal and household goods	U300 PRODUÇÃO TERCIÁRIA	U310 Serviços Comerciais	U311 Comércio por grosso e a retalho, reparação de veículos e de bens de uso pessoal e doméstico
			U312 Real estate services			U312 Serviços imobiliários

Tipo	Nomenclatura original			Nomenclatura proposta		
	Classes			Classes		
	Nível 1	Nível 2	Nível 3	Nível 1	Nível 2	Nível 3
		U320 Financial, professional and information services	U313 Accommodation and food services	U320 Serviços financeiros, profissionais e de informação	U313 Serviços de alojamento e restauração	
			U321 Financial and insurance services		U321 Serviços financeiros e de seguros	
			U322 Professional technical and scientific services		U322 Serviços técnico-profissionais e científicos	
			U323 Information and communication services		U323 Serviços de informação e de comunicação	
		U324 Administrative and support services	U324 Actividades administrativas e dos serviços de apoio			
		U330 Community services	U331 Public administration defence and social security services		U331 Serviços de administração pública, defesa e segurança social	
			U332 Educational services		U332 Serviços educativos	
			U333 Health and social services		U333 Serviços de saúde e sociais	
			U334 Religious services		U334 Serviços religiosos	
		U340 Cultural entertainment and recreational services	U341 Cultural services		U341 Serviços culturais	
			U342 Entertainment services		U342 Serviços de entretenimento	
			U343 Sports infrastructure		U343 Infra estruturas desportivas	
	U344 Open air recreational areas		U344 Espaços de divertimento ao ar livre			
	U400 TRANSPORT NETWORKS LOGISTICS AND UTILITIES	U410 Transport networks	U411 Road transport	U400 REDES DE TRANSPORTE E SERVIÇOS DE UTILIDADE PÚBLICA	U410 Redes de transportes	U411 Transporte rodoviário
			U412 Railway transport			U412 Transportes por caminhos de ferro
			U413 Air transport			U413 Transportes Aéreos
U414 Water transport			U414 Transportes por água			
U420 Logistical and storage services			U420 Serviços de logística e armazenagem			

Tipo	Nomenclatura original Classes			Nomenclatura proposta Classes		
	Nível 1	Nível 2	Nível 3	Nível 1	Nível 2	Nível 3
			U430 Utilities	U431 Electricity, gas and thermal power distribution services U432 Water and sewage infrastructure U433 Waste treatment		U430 Serviço de utilidade pública
	U500 RESIDENTIAL USE	U510 Permanent residential use U520 Temporary residential use		U500 UTILIZAÇÃO RESIDENCIAL	U510 Utilização residencial permanente U520 Utilização Residencial temporária	
	U600 OTHER USES OR NO SOCIO-ECONOMIC USE	U610 Transitional use U611 Abandoned areas U612 Natural areas not in other economic use		U600 OUTRAS UTILIZAÇÕES OU UTILIZAÇÕES NÃO SÓCIO-ECONÓMICAS	U610 Utilização temporária U611 Áreas abandonadas U612 Áreas naturais ou com outros usos sócio-económicos	

Anexo 5: Precisão (Limites máximos de erro esperados por classe de ocupação do solo por NUTS definidos pelo Eurostat)

Land Cover classes	Land Cover code	Upper-bound of the expected error
ARTIFICIAL LAND	A	
Built-up areas	A10	15%
Artificial non-built up areas	A20	15%
CROPLAND	B	
Cereals	B10	15%
Root crops	B20	25%
Non-permanent industrial crops	B30	25%
Dry pulses, vegetables and flowers	B40	25%
Fodder crops	B50	25%
Fruit, nuts trees and berries	B70	25%
Other permanent crops	B80	
WOODLAND	C	20%
Forest FAO	C10	20%
Broadleaved forest (FAO)	C11	20%
Coniferous forest (FAO)	C12	20%
Mixed forest (FAO)	C13	20%
Other coniferous wooded land (FAO)	C20	20%
Other coniferous wooded land (non-FAO)	C30	20%
SHRUBLAND	D	
Shrubland with sparse trees	D10	
Shrubland without trees	D20	
GRASSLAND	E	20%
Grassland with sparse trees	E10	
Grassland without trees	E20	7.5%
Spontaneous vegetation	E30	7.5%
BARE LAND AND LICHENS	F	
Rocks and stones	F10	20%
Sand	F20	20%
Lichens	F30	20%
Other bare soil	F40	20%
WATER	G	
Inland water bodies	G10	20%
Inland running water	G20	20%
Coastal water bodies	G30	20%
Glaciers, permanent snow	G50	20%
WETLAND	H	
Inland wetlands	H10	20%
Coastal wetlands	H20	20%

Anexo 6: Estimativas da área (km²) e respetivos CV, ao nível 2 e 3 das classes de ocupação do solo, por NUTS 2, para os dados recolhidos em 2009

Classes de ocupação do solo	PT 11	CV	PT 15	CV	PT 16	CV	PT 17	CV	PT18	CV
(Nível 2 e 3)	ESTIM_AREA		ESTIM_AREA		ESTIM_AREA		ESTIM_AREA		ESTIM_AREA	
A10	439	18,43	66	47,46	698	13,81	85	41,29	247	24,61
A20	1010	11,18	230	25,09	987	12,19	386	15,52	561	15,06
B10	992	11,67	49	55,83	1386	9,90	131	32,88	2393	7,50
B20	145	32,83	—	—	132	35,05	16	99,48	66	49,36
B30	—	—	—	—	16	99,69	—	—	262	24,48
B40	213	27,45	—	—	195	28,51	35	71,44	261	24,60
B50	251	24,07	—	—	213	27,40	229	24,73	540	16,90
B70	686	14,61	463	15,86	327	22,02	50	57,50	64	49,59
B80	2160	7,19	354	19,94	2236	6,97	182	22,10	2477	6,17
C10	5739	4,19	754	13,31	11319	2,58	516	13,90	13276	2,22
C11	3101	6,40	476	17,40	6125	4,31	233	24,15	10951	2,75
C12	1607	9,49	262	24,15	3772	5,91	216	25,29	1276	10,89
C13	1031	12,02	17	98,62	1422	10,17	67	48,36	1049	12,01
C20	1889	8,76	395	19,01	2622	7,36	183	28,35	2548	7,57
C30	507	17,48	131	34,89	775	14,23	16	99,48	252	25,59
D10	514	17,55	688	13,91	937	12,85	116	36,38	328	21,96
D20	2788	6,97	922	11,58	1999	8,59	99	38,98	1030	12,28
E10	68	49,46	148	32,63	115	37,36	—	—	262	24,67
E20	379	20,37	33	70,52	147	32,96	49	56,82	2576	7,28
E30	2094	8,26	326	21,26	2607	7,14	412	18,22	3048	6,75
F00	1200	11,24	213	26,56	1124	11,73	253	22,97	767	14,09
G10	125	23,11	32	70,70	109	33,86	35	100,51	387	19,73
G20	75	43,37	32	70,70	79	42,12	35	100,51	207	28,41
G30	—	—	16	100,41	80	42,48	70	64,15	22	100,34
H10	—	—	—	—	48	57,70	—	—	—	—
H20	14	100,79	144	29,04	50	57,85	103	48,92	33	70,17

Anexo 7: Estimativas da área (km²) e respetivos CV, ao nível 2 e 3 das classes de ocupação do solo, por NUTS 2, para os dados recolhidos em 2012

Classes de ocupação do solo (Nível 2 e 3)	PT 11 ESTIM_AREA	CV	PT 15 ESTIM_AREA	CV	PT 16 ESTIM_AREA	CV	PT 17 ESTIM_AREA	CV	PT18 ESTIM_AREA	CV
A10	837	11,21	187	20,94	617	12,56	200	21,07	279	20,08
A20	1032	10,35	156	26,87	1342	8,91	339	15,79	563	13,29
B10	779	11,60	9	119,46	1136	9,67	136	27,59	2389	6,50
B20	181	25,30	16	85,94	155	27,40	40	54,71	6	144,37
B30	—	—	—	—	—	—	—	—	110	32,52
B40	183	25,39	—	—	225	22,74	46	51,70	145	28,37
B50	243	22,13	35	58,33	267	20,69	12	99,75	604	13,70
B70	849	11,17	493	13,35	480	15,35	85	34,50	149	27,90
B80	1742	7,22	535	13,82	2082	6,52	167	25,80	2420	5,71
C10	3660	4,97	899	10,22	10509	2,41	577	12,08	14936	1,75
C11	2187	6,76	468	14,99	5723	3,90	359	15,94	13085	2,03
C12	1209	9,50	387	16,63	4264	4,67	176	27,83	1402	8,89
C13	264	21,00	45	50,97	522	14,73	42	52,82	449	16,02
C20	2016	7,25	548	13,40	2038	7,29	46	51,04	997	10,67
C30	1588	8,13	158	26,80	843	11,60	26	67,33	519	14,92
D10	511	14,96	343	17,94	1478	8,64	128	29,51	456	16,07
D20	3524	5,21	576	13,19	2366	6,71	252	20,64	684	12,96
E10	189	24,61	290	19,44	356	18,08	82	38,07	591	13,98
E20	2261	6,77	339	18,02	1989	7,27	271	18,88	4441	4,53
E30	633	13,27	149	27,13	761	12,18	389	15,66	707	12,69
F10	315	19,47	14	94,99	477	15,67	5	181,21	161	27,01
F20	8	134,89	—	—	51	49,41	—	—	6	140,80
F40	529	14,79	112	32,40	647	13,40	12	101,46	664	13,17
G10	61	40,56	21	76,06	185	22,07	81	49,01	504	14,38
G20	136	25,25	31	61,88	70	39,98	—	—	268	21,20
G30	11	98,43	14	94,99	74	38,85	54	64,63	—	—
H10	—	—	—	—	24	71,38	—	—	—	—
H20	—	—	73	36,51	27	65,56	54	64,63	6	142,59

Anexo 8: Estimativas da área (km²) e respetivos CV, ao nível 1 das classes de ocupação do solo, por NUTS 3, para os dados recolhidos no LUCAS de 2009

NUTS3	Classes de ocupação do solo (nível 1)															
	A		B		C		D		E		F		G		H	
	AREA	CV	AREA	CV	AREA	CV	AREA	CV	AREA	CV	AREA	CV	AREA	CV	AREA	CV
PT111	143	30,34	295	17,79	766	10,74	571	14,32	180	29,43	236	24,56	28	32,31	—	—
PT112	175	25,58	296	20,15	426	14,06	152	29,6	148	33,44	41	—	8	—	—	—
PT113	132	26,58	217	22,63	508	10,34	138	32,47	183	24,54	52	55,79	16	42,82	—	—
PT114	274	15,67	184	16,13	200	18,33	12	102,5	111	35,83	21	73,49	12	—	—	—
PT115	261	20,32	349	17,57	1173	7,57	386	17,96	299	21,87	152	31,85	—	—	—	—
PT116	173	23,54	94	—	444	11,98	38	68,29	56	48,05	57	54,78	—	—	—	—
PT117	106	34,4	1175	8,01	1314	8,09	934	11,03	294	22,89	231	25,62	54	32,66	—	—
PT118	232	26,91	1836	6,85	3305	4,93	1058	11,69	1226	10,19	443	19,89	58	40,23	14	102,59
PT150	296	20,59	866	10,56	1280	9,4	1610	7,81	507	16,59	213	26,56	80	43,23	144	29,04
PT161	164	23,63	279	—	995	5,84	—	—	181	25,08	31	70,1	81	35,2	73	43,22
PT162	180	22,02	462	11,88	1108	5,66	120	34,6	120	36,47	49	58,19	24	34,97	—	—
PT163	128	27,54	295	18,49	920	7,51	141	32,55	194	26,44	45	57,68	—	—	21	100,52
PT164	132	35,77	182	31,68	1794	4,9	320	19,07	110	39,28	47	75,65	31	—	—	—
PT165	358	20,82	391	18,38	1800	5,93	378	20,43	242	22,46	305	23,95	15	99,8	—	—
PT166	85	46,25	184	24,4	1306	6,09	255	23,52	42	63,23	16	97,62	16	42,99	—	—
PT167	42	—	65	56,61	438	14,52	170	33,09	153	32,61	—	—	—	—	—	—
PT168	74	47,94	735	11,46	1551	7,42	803	12,55	602	14,64	297	23,91	—	—	—	—
PT169	30	70,13	663	11,67	2161	4,9	314	20,99	446	16,28	99	40,52	36	41,47	—	—
PT16A	47	34,7	305	—	646	9,14	142	33,81	128	37,45	107	40,6	—	—	—	—
PT16B	288	21,13	546	14,08	625	10,54	200	26,63	451	16,83	110	38,82	—	—	—	—
PT16C	133	32,85	342	17,35	1451	5,74	103	36,73	167	31,59	91	43,92	20	38,37	—	—
PT171	290	14,1	369	15,96	207	23,5	120	32,75	288	21,04	16	104,29	46	39,54	42	57,23
PT172	171	18,47	290	19,38	487	11,44	100	42,7	170	29,45	235	22,64	86	100,61	86	100,61
PT181	177	27,57	585	14,31	3154	3,58	193	27,96	913	11,11	144	34,57	125	—	19	100,02
PT182	122	20,21	837	10,29	3587	3,84	318	21,85	1193	9,63	89	44,77	103	30,23	—	—
PT183	131	32,69	1279	9,01	3821	3,72	95	40,3	1523	8,32	199	27,91	180	29,44	—	—
PT184	151	26,82	2268	6,47	3394	4,95	549	15,74	1826	8,12	246	24,6	109	32,43	—	—
PT185	222	22,2	1092	8,39	2146	4,59	160	30,02	476	16,5	95	38,73	65	38,53	18	98,55

Anexo 9: Estimativas da área (km²) e respetivos CV, ao nível 1 das classes de ocupação do solo, por NUTS 3, para os dados recolhidos no LUCAS de 2012

NUTS3	Classes de ocupação do solo (nível 1)															
	A		B		C		D		E		F		G		H	
	AREA	CV	AREA	CV	AREA	CV	AREA	CV	AREA	CV	AREA	CV	AREA	CV	AREA	CV
PT111	214	22,09	234	20,18	549	12,49	759	10,5	281	19,85	122	31,96	58	35,34	—	—
PT112	95	33,15	240	18,54	415	12,39	170	23,4	269	18,71	57	45,72	—	—	—	—
PT113	212	19,07	253	16,47	371	12,88	184	—	188	20,55	22	62,12	16	81,59	—	—
PT114	242	10,56	131	21,09	260	12,9	56	38,21	89	29,62	24	58,3	12	49,35	—	—
PT115	342	16,1	469	13,76	954	8,19	502	13,96	237	22,17	96	35,87	20	—	—	—
PT116	80	35,3	138	20,88	404	10,04	76	36,56	164	21,45	—	—	—	—	—	—
PT117	302	—	982	8,58	1124	8,03	1152	8,69	274	20,62	244	22,62	30	42,95	—	—
PT118	367	16,81	1448	7,15	3187	4,52	1344	8,41	1438	7,87	343	18,18	44	34,44	—	—
PT150	343	14,65	1088	8,26	1605	6,68	920	10,01	778	11,07	125	30,53	65	42,09	73	36,51
PT161	150	27,47	333	13,66	923	5,23	26	57,51	161	23,28	98	31,69	68	34,12	43	48,80
PT162	241	17,88	359	14,58	1036	5,81	142	24,86	187	23,17	31	54,54	66	33,77	—	—
PT163	198	17,9	226	18,66	981	5,69	52	42,34	233	18,33	50	43,65	4	—	—	—
PT164	76	38,12	193	22,67	1565	5,11	465	14,54	162	26,78	140	29,04	16	—	—	—
PT165	223	22,29	514	13,03	1486	6,28	591	13,11	382	15,72	281	19,66	12	—	—	—
PT166	34	70,57	187	19,59	1250	5,34	306	18,12	6	136,29	116	28,13	7	173,68	—	—
PT167	53	—	95	—	273	18,17	278	18,45	116	30,63	48	53,63	4	—	—	—
PT168	81	39,87	374	16,64	988	9,81	1385	7,73	920	9,99	314	19,78	—	—	—	—
PT169	98	43,03	611	11,02	2172	4,8	295	21,47	511	14,27	15	95,67	47	39,25	—	—
PT16A	107	29,65	335	14,75	634	—	169	24,56	67	41,83	51	46,9	11	82,58	—	—
PT16B	233	16,17	604	10,3	773	8,12	143	26,57	380	15,07	40	49,89	46	38,59	—	—
PT16C	244	18,06	446	12,27	1317	5,05	122	27,99	97	30,56	61	39,71	20	—	—	—
PT171	277	11,49	300	17,25	173	23,1	199	21,18	340	14,33	16	104,29	66	22,79	5	145,18
PT172	257	15,36	176	21,23	481	12,18	160	26,31	402	13,62	12	102,61	86	46,13	51	69,90
PT181	140	29,15	474	13,52	3255	3,01	233	21,71	923	9	127	30,24	150	16,71	7	147,21
PT182	86	32,35	1121	8,03	3638	3,05	208	22,78	918	9,52	151	27,33	127	24,64	—	—
PT183	162	22,13	1136	8,8	4039	3,17	99	33,92	1402	7,81	150	27,94	241	20,65	—	—
PT184	108	27,35	2296	5,41	3263	4,27	438	15,54	1927	6,67	284	19,92	226	18,88	—	—
PT185	346	16,43	827	9,6	2318	4,05	144	28,72	496	14,6	123	32,28	21	82,26	—	—

Anexo 10: Estimativas da área (km²) para o nível 1 das classes de ocupação do solo, por NUTS 2, para os dados recolhidos no LUCAS de 2009 e 2012 e resultados obtidos de “outras fontes” e resultados exclusivos da COS 2007

Classe ocupação solo	ESTIM_AREA (LUCAS 2009)					ESTIM_AREA (LUCAS 2012)					Outras fontes dados (NDS)					COS 2007				
	PT11	PT15	PT16	PT17	PT18	PT11	PT15	PT16	PT17	PT18	PT11	PT15	PT16	PT17	PT18	PT11	PT15	PT16	PT17	PT18
A	1449	296	1685	471	808	1869	343	1960	539	842	1440	252	1455	609	511	1440	252	1455	609	511
B	4447	867	4504	643	6063	3976	1088	4346	485	5823	4242	395	3689	473	6754	6366	1037	6974	976	15140
C	8135	1280	14715	715	16076	7264	1605	13389	649	16452	5723	1755	10322	1248	9670	5267	664	9087	543	5653
D	3301	1610	2935	215	1357	4035	920	3844	380	1140	7558	1877	10671	346	6482	7242	2717	9296	560	5776
E	2541	507	2869	461	5886	3083	778	3105	742	5739	1522	496	1293	121	7146	17	102	565	56	3675
F	1200	213	1124	253	767	852	125	1175	17	832	617	38	425	15	311	771	40	473	30	118
G	200	80	268	141	616	207	65	329	136	771	178	100	237	176	697	178	100	237	176	697
H	14	144	99	103	33	—	73	51	54	6	6	83	107	14	34	6	83	114	53	34

113

Anexo 11: Diferenças entre os resultados estimados com base nos dados LUCAS de 2009 e 2012 com os resultados obtidos de outras fontes de classificação (NDS) e resultados exclusivos da COS 2007 (por NUTS 2 ao nível 1 da nomenclatura de classificação LUCAS)

Classe ocupação solo	Diferenças entre LUCAS 2009 e outras fontes de dados (%)					Diferenças entre LUCAS 2009 e COS (%)					Diferenças entre LUCAS 2012 e outras fontes de dados (%)					Diferenças entre LUCAS 2012 e COS (%)				
	PT11	PT15	PT16	PT17	PT18	PT11	PT15	PT16	PT17	PT18	PT11	PT15	PT16	PT17	PT18	PT11	PT15	PT16	PT17	PT18
A	1%	17%	16%	-23%	58%	1%	17%	16%	-23%	58%	30%	36%	35%	-12%	65%	30%	36%	35%	-12%	65%
B	5%	120%	22%	36%	-10%	-30%	-16%	-35%	-34%	-60%	-6%	176%	18%	3%	-14%	-38%	5%	-38%	-50%	-62%
C	42%	-27%	43%	-43%	66%	54%	93%	62%	32%	184%	27%	-9%	30%	-48%	70%	38%	142%	47%	19%	191%
D	-56%	-14%	-72%	-38%	-79%	-54%	-41%	-68%	-62%	-77%	-47%	-51%	-64%	10%	-82%	-44%	-66%	-59%	-32%	-80%
E	67%	2%	122%	280%	-18%	14978%	395%	408%	726%	60%	103%	57%	140%	511%	-20%	18197%	660%	450%	1229%	56%
F	94%	456%	165%	1622%	147%	56%	429%	138%	754%	549%	38%	226%	177%	16%	168%	11%	211%	148%	-41%	604%
G	13%	-20%	13%	-20%	-12%	13%	-20%	13%	-20%	-12%	17%	-35%	39%	-23%	11%	17%	-35%	39%	-23%	11%
H	122%	75%	-8%	648%	-4%	122%	75%	-13%	97%	-4%	—	-12%	-52%	290%	-83%	—	-12%	-55%	3%	-83%

Anexo 12: Diferenças em percentagem entre as estimativas LUCAS 2009 e os resultados provenientes de NDS e COS 2007

NUTS 3	Classes de ocupação do solo (Nível 1)															
	A		B		C		D		E		F		G		H	
	Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /	
	COS07	NDS	COS07	NDS	COS07	NDS	COS07	NDS	COS07	NDS	COS07	NDS	COS07	NDS	COS07	NDS
PT111	-14%	-14%	-32%	19%	19%	-16%	-14%	-27%	33950%	969%	-14%	273%	-3%	-3%	—	—
PT112	-7%	-7%	-13%	-19%	0%	20%	-18%	-48%	143943%	1185%	-56%	151%	-41%	-41%	—	—
PT113	-37%	-37%	-33%	-24%	23%	61%	-49%	-64%	20027%	1301%	163%	107%	129%	129%	—	—
PT114	-2%	-2%	-14%	-28%	-9%	-5%	-85%	-75%	39609%	2646%	207%	457%	-14%	-14%	—	—
PT115	15%	15%	-45%	-17%	33%	21%	-52%	-57%	36877%	641%	195%	258%	—	—	—	—
PT116	40%	40%	-29%	10%	7%	18%	-79%	-85%	—	1316%	419%	346%	—	—	—	—
PT117	5%	5%	-19%	15%	75%	-11%	-44%	-15%	5760%	40%	166%	53%	30%	30%	—	—
PT118	62%	62%	-35%	18%	118%	199%	-69%	-72%	13216%	0%	96%	47%	17%	17%	19143%	19183%
PT150	17%	17%	-16%	120%	93%	-27%	-41%	-14%	395%	2%	429%	456%	-20%	-20%	75%	75%
PT161	-21%	-21%	-40%	0%	22%	77%	—	—	13599%	900%	149%	70%	85%	85%	-19%	-19%
PT162	-3%	-3%	-19%	35%	19%	196%	-62%	-89%	15920%	233%	146%	192%	17%	17%	—	—
PT163	-34%	-34%	-20%	162%	30%	234%	-69%	-87%	16924%	514%	207%	-26%	—	—	1285%	1285%
PT164	39%	39%	-47%	45%	37%	55%	-60%	-73%	2413%	244%	52%	473%	34%	34%	—	—
PT165	110%	110%	-44%	18%	20%	20%	-64%	-72%	7390%	301%	711%	735%	-8%	-8%	—	—
PT166	131%	131%	-3%	95%	117%	38%	-75%	-68%	593%	326%	-21%	324%	-41%	-41%	—	—
PT167	90%	90%	-62%	-48%	141%	178%	-62%	-67%	9114%	387%	—	—	—	—	—	—
PT168	6%	6%	-40%	33%	185%	8%	-60%	-42%	1155%	33%	120%	82%	—	—	—	—
PT169	-31%	-31%	-32%	14%	159%	54%	-78%	-76%	15%	31%	232%	99%	-5%	-5%	—	—
PT16A	21%	21%	-26%	-1%	105%	-13%	-71%	-29%	139%	128%	65%	338%	—	—	—	—
PT16B	33%	33%	-42%	1%	10%	-36%	-52%	-46%	1313%	373%	278%	589%	—	—	—	—
PT16C	-23%	-23%	-46%	16%	89%	81%	-84%	-88%	579%	30%	141%	2309%	-41%	-41%	—	—
PT171	-23%	-23%	-23%	48%	32%	-62%	-57%	32%	3274%	647%	64%	208%	-15%	-15%	204%	204%
PT172	-27%	-27%	-42%	30%	26%	-30%	-64%	-61%	259%	105%	1094%	2397%	-29%	-29%	122%	—
PT181	151%	151%	-63%	-14%	95%	243%	-85%	-92%	54%	-9%	274%	172%	-18%	-18%	8%	8%
PT182	79%	79%	-68%	-24%	198%	37%	-73%	-69%	12%	-9%	187%	284%	11%	11%	—	—
PT183	21%	21%	-71%	-18%	386%	111%	-87%	-91%	64%	-35%	1325%	34%	-24%	-24%	—	—
PT184	54%	54%	-53%	-10%	426%	13%	-70%	-45%	98%	8%	1777%	292%	-34%	-34%	—	—
PT185	34%	34%	-33%	21%	53%	63%	-80%	-84%	189%	-40%	337%	318%	23%	23%	15%	15%

Anexo 13: Diferenças em percentagem entre as estimativas LUCAS 2012 e os resultados provenientes de NDS e COS 2007

NUTS 3	Classes de ocupação do solo (Nível 1)															
	A		B		C		D		E		F		G		H	
	Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /		Dif. LUCAS /	
	COS07	NDS	COS07	NDS	COS07	NDS	COS07	NDS	COS07	NDS	COS07	NDS	COS07	NDS	COS07	NDS
PT111	28%	28%	-46%	-6%	-14%	-39%	14%	-3%	53004%	1567%	-56%	93%	103%	103%	—	—
PT112	-50%	-50%	-30%	-35%	-2%	17%	-8%	-42%	262674%	2243%	-39%	247%	—	—	—	—
PT113	1%	1%	-22%	-12%	-10%	17%	-32%	-52%	20589%	1340%	13%	-11%	124%	124%	—	—
PT114	-14%	-14%	-39%	-48%	18%	24%	-29%	17%	31542%	2088%	254%	542%	-14%	-14%	—	—
PT115	50%	50%	-26%	12%	8%	-2%	-37%	-44%	29186%	487%	87%	126%	-8%	-8%	—	—
PT116	-35%	-35%	5%	61%	-2%	7%	-58%	-71%	—	4058%	—	—	—	—	—	—
PT117	199%	199%	-32%	-4%	50%	-24%	-31%	5%	5361%	31%	181%	61%	-28%	-28%	—	—
PT118	157%	157%	-49%	-7%	110%	188%	-60%	-65%	15523%	18%	51%	14%	-11%	-11%	—	—
PT150	36%	36%	5%	176%	142%	-9%	-66%	-51%	660%	57%	211%	227%	-35%	-35%	-12%	-12%
PT161	-28%	-28%	-28%	20%	13%	64%	-84%	-95%	12130%	793%	690%	437%	56%	56%	-52%	-52%
PT162	29%	29%	-37%	5%	11%	177%	-55%	-87%	24792%	418%	58%	87%	223%	223%	—	—
PT163	1%	1%	-39%	100%	39%	256%	-88%	-95%	20398%	639%	238%	-19%	242%	242%	—	—
PT164	-20%	-20%	-44%	53%	19%	35%	-42%	-61%	3595%	406%	349%	1600%	-31%	-31%	—	—
PT165	31%	31%	-27%	55%	-1%	-1%	-44%	-57%	11692%	531%	647%	670%	-23%	-23%	—	—
PT166	-8%	-8%	-1%	98%	107%	32%	-70%	-61%	4%	-36%	466%	2918%	-76%	-76%	—	—
PT167	140%	140%	-44%	-23%	50%	74%	-38%	-45%	6918%	271%	21%	112%	97%	97%	—	—
PT168	16%	16%	-69%	-32%	81%	-31%	-32%	1%	1817%	103%	133%	92%	—	—	—	—
PT169	123%	123%	-37%	5%	160%	55%	-80%	-77%	32%	50%	-51%	-71%	25%	25%	—	—
PT16A	175%	175%	-18%	9%	101%	-14%	-65%	-16%	26%	20%	-22%	106%	109%	109%	—	—
PT16B	8%	8%	-36%	12%	35%	-20%	-65%	-61%	1090%	299%	37%	150%	300%	300%	—	—
PT16C	41%	41%	-29%	51%	72%	64%	-81%	-86%	296%	-24%	60%	1503%	-41%	-41%	—	—
PT171	-26%	-26%	-37%	20%	11%	-68%	-29%	119%	3887%	783%	64%	208%	23%	23%	-62%	-62%
PT172	10%	10%	-65%	-21%	25%	-31%	-43%	-37%	750%	385%	-40%	25%	-29%	-29%	31%	—
PT181	99%	99%	-70%	-30%	101%	254%	-82%	-90%	56%	-8%	229%	140%	-2%	-2%	-64%	-64%
PT182	26%	26%	-57%	1%	203%	39%	-82%	-80%	-14%	-30%	388%	554%	38%	38%	—	—
PT183	49%	49%	-74%	-27%	414%	123%	-86%	-90%	51%	-40%	976%	1%	2%	2%	—	—
PT184	11%	11%	-53%	-9%	406%	8%	-76%	-56%	108%	14%	2065%	352%	38%	38%	—	—
PT185	109%	109%	-50%	-9%	65%	76%	-82%	-86%	201%	-37%	468%	443%	-60%	-60%	—	—

Anexo 14: Comparação das Estimativas para as classes de uso do solo U110, U120, U130, U340 e U612, dadas através dos dados LUCAS 2009, ao nível das NUTS 3, com os resultados obtidos de outras fontes de dados

NUTS3	Classes de uso do solo	EST_AREA (LUCAS)	CV	AREA_NDS	DIF. (%)	Classes de uso do solo	EST_AREA (LUCAS)	CV	AREA_NDS	DIF. (%)	Classes de uso do solo	EST_AREA (LUCAS)	CV	AREA_NDS	DIF. (%)	Classes de uso do solo	EST_AREA (LUCAS)	CV	AREA_NDS	DIF. (%)
111	U110	493	14	721	-32%	U120	734	11	253	190%	U130	--	--	36	--	--	--	--	3	--
112	U110	440	13	310	42%	U120	384	15	145	165%	U130	--	--	7	--	--	--	--	2	--
113	U110	383	13	248	54%	U120	467	12	137	240%	U130	17	98,48	1	1715%	U340	17	98,48	3	461%
114	U110	249	17	151	65%	U120	200	18	36	457%	U130	--	--	1	--	--	12	102,5	9	38%
115	U110	644	12	575	12%	U120	954	9	287	232%	U130	17	98,41	12	34%	U340	24	102,93	3	644%
116	U110	128	nr	92	39%	U120	436	12	52	732%	U130	--	--	15	--	U340	--	--	2	--
117	U110	1806	6	1235	46%	U120	834	11	473	76%	U130	--	--	128	--	U340	--	--	1	--
118	U110	3513	5	2700	30%	U120	1833	8	938	95%	U130	--	--	35	--	U340	--	--	3	--
150	U110	1428	8	743	92%	U120	888	12	466	91%	U130	48	55,22	12	284%	U340	33	70,52	23	44%
161	U110	476	nr	233	105%	U120	967	6	123	688%	U130	--	--	9	--	U340	30	70,16	5	467%
162	U110	659	9	332	98%	U120	1042	7	169	515%	U130	13	101,39	1	1279%	U340	16	100,37	5	221%
163	U110	471	13	135	249%	U120	888	8	74	1097%	U130	--	--	28	--	U340	--	--	2	--
164	U110	386	18	137	182%	U120	1655	6	177	835%	U130	--	--	5	--	U340	--	--	1	--
165	U110	741	11	383	93%	U120	1513	7	328	362%	U130	--	--	15	--	U340	15	99,8	4	238%
166	U110	287	20	98	192%	U120	1254	7	312	302%	U130	--	--	0	--	U340	--	--	0	--
167	U110	225	25	145	56%	U120	301	21	58	419%	U130	--	--	1	--	U340	--	--	0	--
168	U110	1565	7	1218	28%	U120	957	11	254	276%	U130	21	100,88	50	-59%	U340	17	98,58	2	851%
169	U110	1272	8	1339	-5%	U120	1853	6	568	226%	U130	16	98,3	0	--	U340	14	99,96	2	654%
16A	U110	527	nr	395	33%	U120	596	10	144	314%	U130	--	--	18	--	U340	--	--	1	--
16B	U110	986	8	590	67%	U120	485	13	114	327%	U130	40	56,94	4	1001%	U340	--	--	6	--
16C	U110	679	12	390	74%	U120	1087	8	287	279%	U130	--	--	0	--	U340	16	98,23	4	339%
171	U110	446	12	312	43%	U120	66	48	43	53%	U130	--	--	0	--	U340	40	54,77	17	138%
172	U110	565	12	512	10%	U120	363	16	146	149%	U130	38	72,66	0	--	U340	28	70,92	9	196%
181	U110	1659	6	2324	-29%	U120	2667	5	556	380%	U130	--	--	3	--	U340	50	57,44	4	1200%
182	U110	2546	5	4348	-41%	U120	3000	4	500	501%	U130	--	--	1	--	U340	--	--	4	--
183	U110	3482	4	5340	-35%	U120	3247	4	425	665%	U130	27	101,8	47	-44%	U340	--	--	5	--
184	U110	4936	3	5351	-8%	U120	2728	6	561	386%	U130	--	--	35	--	--	--	--	3	--
185	U110	1743	6	1828	-5%	U120	2005	5	748	168%	U130	18	98,55	5	249%	U340	33	70,35	7	351%

Anexo 15: Comparação das Estimativas para as classes de uso do solo U110, U120, U130, U340 e U612, dadas através dos dados LUCAS 2012, ao nível das NUTS 3, com os resultados obtidos de outras fontes de dados

NUTS3	Classes uso do solo	LUCAS	CV	NDS	DIF. %	Classes uso do solo	LUCAS	CV	NDS	DIF. %	Classes uso do solo	LUCAS	CV	NDS	DIF. %	Classes uso do solo	LUCAS	CV	NDS	DIF. %	LU	LUCAS	CV	NDS	DIF. %
111	U110	516	13	721	-28%	U120	646	10	253	155%	U130	--	--	36	--	U340	--	--	3	--	U612	647	12	355	82%
112	U110	461	10	310	49%	U120	495	10	145	241%	U130	--	--	7	--	U340	--	--	2	--	U612	109	31	167	-35%
113	U110	375	11	248	51%	U120	371	13	137	170%	U130	--	--	1	--	U340	--	--	3	--	U612	184	NaN	2	9676%
114	U110	202	15	151	34%	U120	240	14	36	567%	U130	--	--	1	--	U340	4	143	9	-50%	U612	4	129	4	-9%
115	U110	681	11	575	19%	U120	956	8	287	233%	U130	7	132	12	-41%	U340	--	--	3	--	U612	435	15	30	1368%
116	U110	279	13	92	202%	U120	358	12	52	583%	U130	--	--	15	--	U340	--	--	2	--	U612	90	33	0	--
117	U110	1431	7	1235	16%	U120	775	10	473	64%	U130	--	--	128	--	U340	--	--	1	--	U612	593	13	264	125%
118	U110	3564	4	2700	32%	U120	2051	6	938	119%	U130	--	--	35	--	U340	16	84	3	514%	U612	760	12	1451	-48%
150	U110	1997	5	743	169%	U120	1408	7	466	202%	U130	29	62	12	135%	U340	24	71	23	3%	U612	282	19	471	-40%
161	U110	493	10	233	112%	U120	842	6	123	586%	U130	27	64	9	198%	U340	31	67	5	474%	U612	199	20	7	2619%
162	U110	543	10	332	64%	U120	932	6	169	450%	U130	4	146	1	381%	U340	--	--	5	--	U612	174	22	7	2567%
163	U110	496	10	135	268%	U120	844	7	74	1037%	U130	--	--	28	--	U340	4	141	2	87%	U612	15	NaN	176	-91%
164	U110	381	14	137	178%	U120	1365	6	177	671%	U130	--	--	5	--	U340	23	73	1	1564%	U612	458	NaN	4	12087%
165	U110	912	9	383	138%	U120	1253	7	328	282%	U130	--	--	15	--	U340	--	--	4	--	U612	708	12	0	--
166	U110	227	19	98	131%	U120	1378	5	312	341%	U130	--	--	0	--	U340	--	--	0	--	U612	81	38	0	--
167	U110	171	nr	145	18%	U120	211	21	58	264%	U130	--	--	1	--	U340	--	--	0	--	U612	182	24	386	-53%
168	U110	1379	7	1218	13%	U120	765	11	254	201%	U130	29	76	50	-41%	U340	7	139	2	269%	U612	520	14	588	-12%
169	U110	1563	6	1339	17%	U120	1968	5	568	247%	U130	9	130	0	--	U340	--	--	2	--	U612	18	92	389	-95%
16A	U110	538	nr	395	36%	U120	509	11	144	254%	U130	--	--	18	--	U340	--	--	1	--	U612	93	34	142	-35%
16B	U110	1043	6	590	77%	U120	540	11	114	375%	U130	29	58	4	696%	U340	--	--	6	--	U612	92	34	97	-6%
16C	U110	765	9	390	96%	U120	1049	7	287	266%	U130	--	--	0	--	U340	--	--	4	--	U612	30	56	94	-68%
171	U110	587	11	312	89%	U120	131	27	43	205%	U130	16	104	0	--	U340	29	61	17	71%	U612	120	35	219	-45%
172	U110	648	9	512	27%	U120	410	15	146	181%	U130	--	--	0	--	U340	28	66	9	195%	U612	153	31	229	-33%
181	U110	3128	4	2324	35%	U120	1675	7	556	201%	U130	--	--	3	--	U340	5	179	4	32%	U612	292	16	545	-46%
182	U110	3767	3	4348	-13%	U120	2046	6	500	310%	U130	--	--	1	--	U340	--	--	4	--	U612	189	24	565	-66%
183	U110	4248	3	5340	-20%	U120	2486	5	425	486%	U130	34	58	47	-29%	U340	--	--	5	--	U612	141	28	0	--
184	U110	5665	2	5351	6%	U120	2116	6	561	277%	U130	--	--	35	--	U340	--	--	3	--	U612	122	27	697	-82%
185	U110	2138	5	1828	17%	U120	1599	7	748	114%	U130	--	--	5	--	U340	21	73	7	183%	U612	48	50	111	-57%