

2. Metodologias para a Avaliação do Uso e Ocupação do Solo em Diferentes Épocas

Bruno M. Meneses; Maria José Vale; Rui Reis; Pedro Marrecas; Edgar Barreira

2.1. Processo de amostragem de LULC

As avaliações realizadas na identificação e caracterização das alterações de uso e ocupação do solo foram suportadas por unidades amostrais (UA) (Figura 1), distribuídas aleatoriamente pelo território de Portugal Continental.

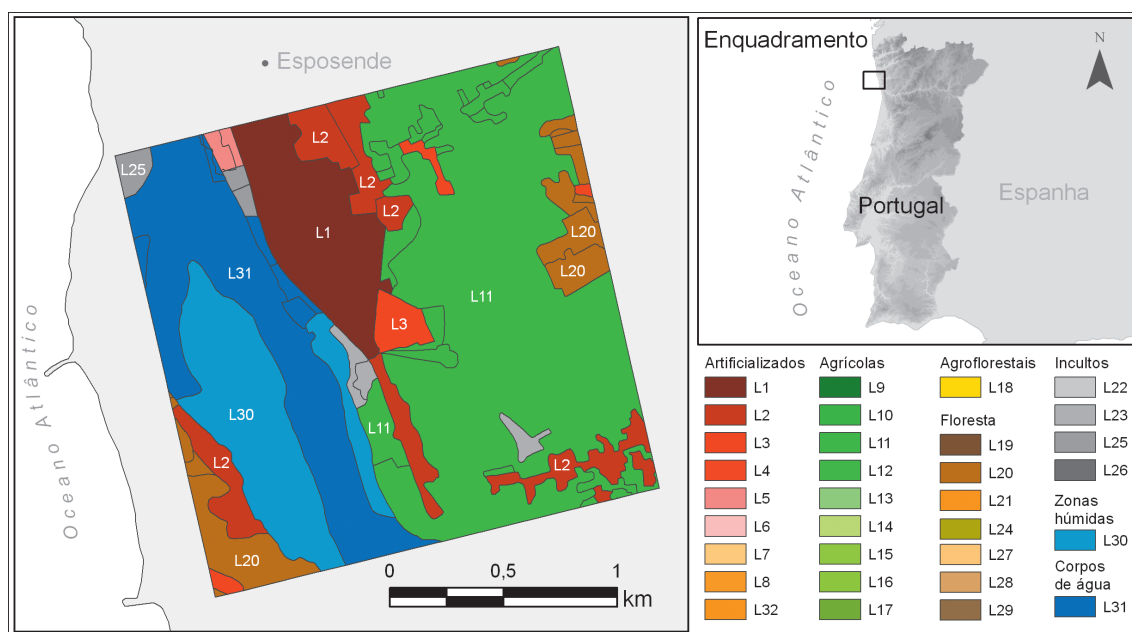


Figura 1 – Exemplo de uma unidade amostral LANDYN (ver legenda na Tabela 2).

Na definição da malha de amostragem, adotou-se a grelha de referência de 1 x 1 km (ETRS89-LAEA 52N 10E) da Agência Europeia do Ambiente (EEA), consistente com a malha EEA para Land cover accounts (LEAC). Para garantir a consistência espacial com os dados LUCAS 2009 do Eurostat, adotou-se uma amostragem por *clusters* constituídos por 2 x 2 elementos da grelha de referência, isto é, elementos ou UA com 4 km². Esta grelha é constituída por 1279 UA, distribuídas aleatoriamente pelo território continental (Figura 2).

Para a extrapolação dos dados das amostras ao território de Portugal Continental, utilizou-se como base a CAOP 2008.1 (tema auxiliar na elaboração da COS2007 utilizada na avaliação da amostragem do Projeto LANDYN). Com a sobreposição das amostras sobre este tema, verificou-se inconsistências entre os limites dos mesmos, nomeadamente no litoral, sendo necessário proceder à sua retificação, i.e. correção geométrica.

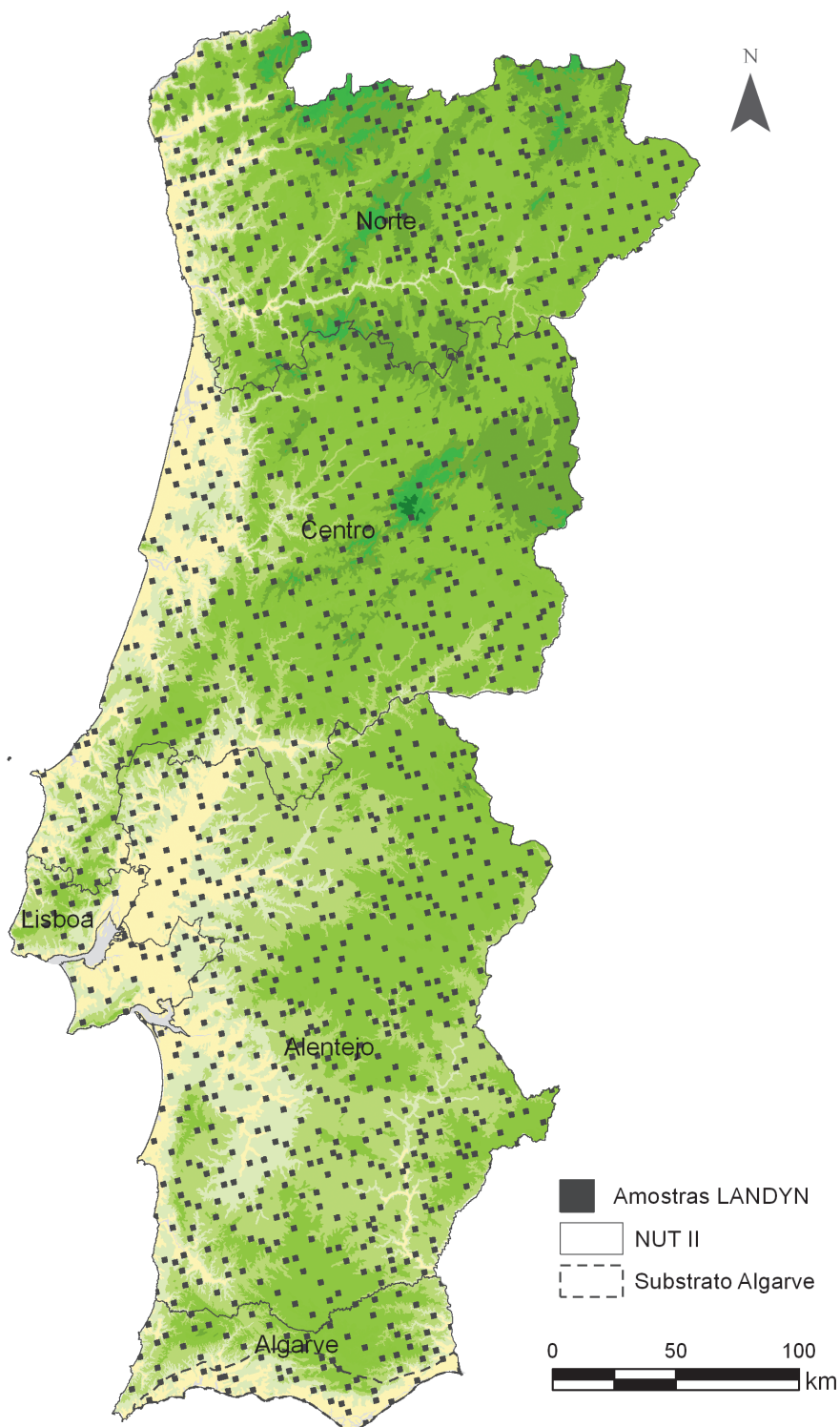


Figura 2 – Distribuição espacial das unidades amostrais LANDYN em Portugal Continental.

2.2. Definição da nomenclatura do uso e ocupação do solo

Para avaliar o uso e ocupação do solo é necessário estabelecer à partida uma legenda que descreva os vários tipos de ocupação que se pretendem integrar na análise. Existindo já um vasto trabalho sobre esta temática na DGT, optou-se pela adaptação da nomenclatura

da COS na definição inicial da nomenclatura a usar para cumprir os objetivos descritos no ponto 1.3 e a que se fará doravante referência como nomenclatura LANDYN.

Assim, definiu-se que a nomenclatura LANDYN é composta por 32 classes (Tabela 2), acompanhando de perto o nível 3 da nomenclatura da COS (DGT, 2013a), com algumas adaptações. Estas 32 classes estão agrupadas em 7 classes mais agregadas de uso e ocupação do solo, de forma a permitir realizar avaliações mais simples das diversas transições de uso e ocupação do solo entre os vários momentos considerados.

Tabela 2 – Classes de uso e ocupação do solo do Projeto LANDYN.

| Código | Classificação | Descrição | Simplificada |
|--------|---------------|---|------------------|
| L1 | TUC | Tecido urbano contínuo | Artificializados |
| L2 | TUD | Tecido urbano descontínuo | |
| L3 | ICE | Indústria, comércio e equipamentos gerais | |
| L4 | RVF | Redes viárias e ferroviárias e espaços associados | |
| L5 | APO | Áreas portuárias | |
| L6 | AER | Aeropostos e aeródromos | |
| L7 | AEI | Áreas de extração de inertes | |
| L8 | ADR | Áreas de deposição de resíduos | |
| L9 | ACO | Áreas em construção | |
| L10 | CTS | Culturas temporárias de sequeiro | |
| L11 | CTR | Culturas temporárias de regadio | |
| L12 | ARR | Arrozais | |
| L13 | VIN | Vinhas | |
| L14 | POM | Pomares | |
| L15 | OLI | Olivais | |
| L16 | PAP | Pastagens permanentes | |
| L17 | AAH | Áreas agrícolas heterogéneas | Agroflorestais |
| L18 | SAF | Sistemas agroflorestais | |
| L19 | FFO | Florestas de folhosas (excluindo o eucalipto e espécies invasoras) | Florestas |
| L20 | FRE | Florestas de resinosas | |
| L21 | FEE | Florestas de eucalipto e espécies invasoras | |
| L22 | VHN | Vegetação herbácea natural | Incultos |
| L23 | MAT | Matos | |
| 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 ardidas | |
| L27 | FAF | Florestas abertas de folhosas (excluindo o eucalipto e espécies invasoras) | Florestas |
| L28 | FAR | Florestas abertas de resinosas | |
| L29 | FAE | Florestas abertas de eucalipto e espécies invasoras | |
| L30 | ZHU | Zonas húmidas | Zonas húmidas |
| L31 | H20 | Corpos de água | Corpos de água |
| L32 | GLF | Campos de Golfe | Artificializados |

2.3. Processo de fotointerpretação e vectorização

Para perceber do rigor dos dados apresentados convirá explicar com algum detalhe o processo de produção da informação.

Com a definição da nomenclatura LANDYN, e a definição da área a levantar através da delimitação das unidades amostrais de 2 x 2 km que incide a produção, e uma

vez seleccionadas as diversas coberturas aerofotográficas a utilizar relativas à cobertura temporal definida, inicia-se o processo de interpretação visual, seguida de digitalização em ecrã das coberturas ortofotocartográficas disponíveis na DGT.

O processo de interpretação visual e digitalização recorreu a informação auxiliar disponível relacionada com ocupação e uso do solo existente na DGT e também proveniente de outras instituições. As imagens foram interpretadas por técnicos especializados, com recurso a software de Sistemas de Informação Geográfica (SIG). Toda a informação foi validada e armazenada de forma a serem criados produtos com as características técnicas requeridas.

A produção inicia-se com a validação da COS2007, a partir da informação auxiliar disponível relativa às unidades amostrais. Posteriormente utilizou-se esta base de trabalho para a produção relativa ao ano de 2010 (fazendo atualização), 1995, 1980 e, finalmente, para o ano de 1970 (apenas um substrato da Região do Algarve), fazendo a “desatualização” da ocupação para cada instante temporal considerado (Figura 3).

Cada uma destas etapas englobou fases intermédias de correção de erros geométricos e/ou topológicos e de validação temática de transições entre os diferentes anos, de modo a garantir uniformidade e uma maior qualidade da informação em todo o processo.

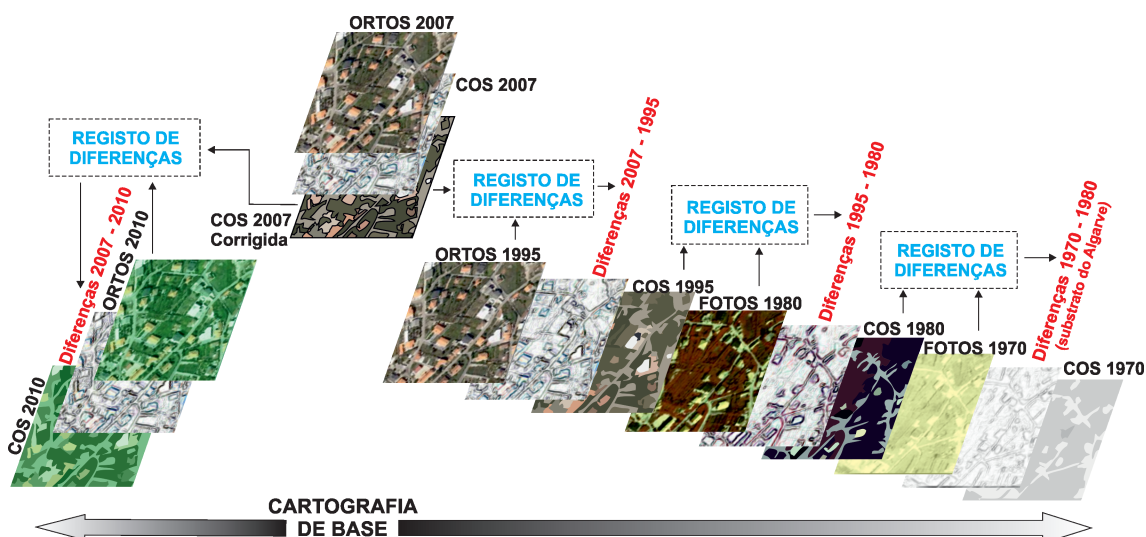


Figura 3 – Processo de fotointerpretação e cartografia resultante.

No processo de reclassificação da COS2007, realizou-se a correspondência das categorias da COS2007 (nomenclatura organizada em cinco níveis hierárquicos, com uma desagregação máxima em 193 categorias) para a nomenclatura Quioto, tendo-se em consideração que todas as ocupações florestais são discriminadas com base na espécie dominante detentora de um coberto arbóreo superior às restantes.

Esta estratégia levou à identificação de um conjunto de classes que careciam de reclassificação, dada a necessidade de desagregação por tipo de espécie florestal, como é o caso dos cortes rasos, das novas plantações, das áreas ardidas e dos sistemas agroflorestais (SAF) de outras espécies, onde foi necessário desagregar a espécie de pinheiro manso (*Pinus pinea*).

Com a realização desta etapa, pretendeu-se produzir cartografia com o detalhe necessário para monitorizar as áreas abrangidas pelos artigos 3.3 (desflorestação e florestação) e 3.4 (gestão florestal, gestão agrícola e gestão de pastagens) relacionadas com a análise e balanço de carbono, e, conseqüentemente, ao cumprimento do protocolo de Quioto.

Para o território de Portugal Continental foram reclassificados 3569 polígonos correspondendo a uma área de 28012 ha, tendo como suporte a cobertura ortofotocartográfica de 2007. Concluída esta etapa obteve-se um novo conjunto de dados COS2007.1, contemplando este 225 classes.

A COS2010 produziu-se por atualização da informação geométrica e temática da COS2007.1, com base na fotointerpretação da cobertura ortofotocartográfica adquirida pelo ex-IGP entre Agosto e Outubro de 2010, informação sujeita a correções de erros (geometria e atributos), do qual resultou a COS2007.2.

Já a COS1995 foi produzida por “desatualização” da informação proveniente da COS2007.2, com base na fotointerpretação das fotografias aéreas de falsa-cor ortorretificadas referentes ao ano de 1995. Neste processo de produção recorreu-se, ainda, às fotografias aéreas ortorretificadas com referência a 2007 e 2010, assim como a outros elementos auxiliares. No decurso desta recolha de informação para produção da COS deste ano, detetaram-se incongruências geométricas e temáticas na informação armazenada da COS2007.2 e da COS2010, obrigando às respetivas correções.

Para a produção dos dados LANDYN 1970 e 1980, elaborou-se um amplo trabalho de levantamento da informação existente na fototeca da DGT relativamente à cobertura aerofotográfica relativa às unidades amostrais. Este trabalho envolveu a sistematização de dados e digitalização dos planos de voo, o levantamento dos fotopontos (pontos centrais das fotografias integrados no plano de voo) adequados para cada amostra, sua digitalização em scanner de alta resolução, e a georreferenciação das imagens obtidas. No processo de produção das amostras LANDYN encontraram-se algumas dificuldades que se prendem com os planos de voo das fotografias aéreas das décadas de 1970 e 1980 representados em suporte analógico que possuíam algumas imprecisões, tornando impossível selecionar diretamente as fotos necessárias para cobrir na totalidade as unidades amostrais, sendo necessário, após a primeira seleção de fotos, fazer uma segunda seleção de modo a conseguir cobrir a totalidade das UA.

No total foram digitalizadas e georreferenciadas 1876 imagens da década de 1980, e 46 da década de 1970 (substrato do Algarve). Esta informação foi produzida por “desatualização” da informação proveniente da COS1995, com base na fotointerpretação das fotografias aéreas pancromáticas georreferenciadas referentes à década de 1980. Neste procedimento optou-se por coberturas aéreas do centro da década (1985), sempre que possível, prevalecendo a opção por coberturas aéreas de datas próximas do centro da década, e com nível de detalhe equivalente.

Assim, consideraram-se os limites verificados nas amostras e excluíram-se destas todos os polígonos com referência ao oceano (classe 5.2.3.01.1 da COS), evitando-se desta forma a integração destas áreas na classe LANDYN Corpos de água e, conseqüentemente, erros na extrapolação dos resultados de LULC obtidos nas amostras para o território continental. Este ajustamento também se realizou no Estuário do Sado,

devido aos limites da CAOP não abrangerem o mesmo (exclusão dos polígonos com referência à classe 5.2.2.01.1 da COS). Todos os polígonos das amostras referentes ao território de Espanha também foram eliminados.

Com a informação dos dois temas corrigida, apurou-se para as 1279 UA a área total de 499651 ha e, para o território continental 8897135 ha, constituindo esta última a base das diversas avaliações realizadas à escala continental.

Também se avaliou a variação de LULC por região (NUT II) nas três décadas, mas no caso do Algarve foi possível determinar o tipo de LULC na década de 1970 num substrato a sul da região (172244 ha), conforme demarcado no mapa da Figura 2.

2.4. Correção e validação temática

A correção temática realizou-se em simultâneo com a deteção de alterações e em presença de dados auxiliares relevantes, nomeadamente o cadastro olivícola, cartografia de áreas ardidadas, os pontos de campo utilizados no controlo dos dados para a COS2007 e o inventário florestal de 1995 e de 2005.

A existência de informação de anos diferentes (e.g. Ortofotos de 1995, 2004, 2007 e 2010; imagens de satélite de 2005) permitiu reconstituir a sequência lógica de eventos de transformação de paisagem e, assim, efetuar uma melhor validação temática.

As classes que mais se prestam à dedução através da identificação do tipo de alteração são: áreas de cortes e novas plantações (identificadas quando numa das datas se identifica floresta e na outra solo descoberto) e áreas em construção, que por vezes podem ser confundidas com outras zonas sem vegetação.

De uma forma geral as classes identificadas com maior facilidade são: os territórios artificializados; algumas plantações puras como o pinheiro e o eucalipto, cortes no meio de zonas florestais, onde foi possível identificar o tipo de floresta adulta numa das datas; zonas descobertas e sem vegetação; zonas húmidas costeiras e os corpos de água.

As classes que apresentaram maior dificuldade de identificação, com base apenas nas coberturas ortofotocartográficas, são: as pastagens, as culturas permanentes de sequeiro e de regadio e as classes florestais (com exceção das classes puras).

A avaliação da exatidão temática da cartografia foi feita com base num teste de hipóteses, que permitiu decidir se existiam evidências estatísticas para aceitar ou rejeitar a hipótese dos dados que compõe as Cartas de Ocupação e Uso do Solo, para os anos de 1995, 2007 e 2010 possuírem uma exatidão global inferior à desejada.

O processo de validação efetuado recorreu a uma amostragem aleatória simples. A dimensão da amostra, assim como a sua modalidade, resultou do compromisso entre o erro máximo admissível, na estimativa da exatidão global temática, e a exigência operacional para concretizar a amostragem.

No processo de amostragem simples, procurou-se manter as seguintes condições:

- A cada unidade amostral é associado um e um só dos seguintes valores, “correto” ou “incorreto”, i.e., cada unidade amostral terá um comportamento binário

(independentemente da regra de concordância definida; neste caso a cada unidade amostral pode-se atribuir 4 classificações alternativas);

- As unidades amostrais têm igual probabilidade de estarem corretas;
- As unidades amostrais são independentes entre si, i.e., a probabilidade de uma unidade amostral estar correta não influencia a probabilidade de outra unidade estar correta;
- O número de unidades amostrais é fixado à partida. Nestas condições, o modelo matemático mais apropriado é o modelo da distribuição binomial (Ginevan, 1979; Aronoff, 1982; Aronoff, 1985).

Para elaborar este teste de hipótese foi necessário fixar um valor de exatidão global mínima para aceitação, o risco do utilizador e a dimensão da amostra, ou seja, o número de unidades amostrais que foram lançadas no mapa.

A exatidão global de um mapa é a estimativa pontual da probabilidade de um qualquer ponto no mapa se encontrar corretamente classificado, de acordo com a nomenclatura em utilização. Neste processo, a exatidão mínima global é uma condição preliminar e é a definida nas especificações associadas à produção da COS, fixada em 85%, já o risco do utilizador é fixo em 5%, o que implica que o nível de confiança do teste seja de 95% (Ginevan, 1979; Aronoff, 1982; Aronoff, 1985).

Para a seleção da dimensão da amostra de teste é aconselhado por Aronoff (1982) a consideração do risco associado ao custo necessário para o desenvolvimento da amostra de validação. Nestas condições, o compromisso entre risco e custo da elaboração da amostra de validação aponta para uma amostra com 300 unidades amostrais.

O teste de hipótese foi, então, definido do seguinte modo (Aronoff, 1982):

- H0 – a cartografia não possui a exatidão mínima de aceitação;
- H1 – a cartografia possui uma exatidão global superior à mínima para aceitação.

Nestas condições, a regra de decisão do teste foi a seguinte: se o mapa apresentasse um número de unidades amostrais incorretas superior ao valor crítico, então concluíam-se que, segundo este teste de hipóteses, não existiam evidências estatísticas suficientes para rejeitar a H0, portanto, concluíam-se que o mapa não possuía a exatidão mínima necessária para aceitação, com um nível de confiança de 95%. Por outro lado, se o mapa apresentasse um número de unidades amostrais incorretas inferior ou igual ao valor crítico, então existiam evidências para rejeitar a H0 e, portanto, concluíam-se que o mapa possuía uma exatidão global superior à exatidão mínima para aceitação (Ginevan, 1979).

Neste processo de avaliação da exatidão temática consideraram-se os dados originais da COS2007, admitindo por um lado a extensão do território continental e, por outro, a extensão da amostra. Na Tabela 3 apresenta-se a superfície ocupada por cada uma das classes LANDYN e a sua contribuição para a área da amostra, bem como para a área total do território (área total da COS2007, sem as alterações referidas anteriormente).

Tabela 3 – Área das amostras e de Portugal Continental por cada classe LANDYN.

| Legenda LANDYN | Área das Amostras (ha) | Área das Amostras (%) | Área de Portugal (ha) | Área de Portugal (%) | DELTA |
|----------------|------------------------|-----------------------|-----------------------|----------------------|-------|
| L1 | 8656 | 1,7 | 149300 | 1,7 | 0,1 |
| L2 | 8352 | 1,7 | 152446 | 1,7 | 0,0 |
| L3 | 2875 | 0,6 | 56629 | 0,6 | -0,1 |
| L4 | 1802 | 0,4 | 30014 | 0,3 | 0,0 |
| L5 | 25 | 0,0 | 1620 | 0,0 | 0,0 |
| L6 | 52 | 0,0 | 2667 | 0,0 | 0,0 |
| L7 | 978 | 0,2 | 14827 | 0,2 | 0,0 |
| L8 | 58 | 0,0 | 1137 | 0,0 | 0,0 |
| L9 | 856 | 0,2 | 14907 | 0,2 | 0,0 |
| L10 | 51295 | 10,3 | 779455 | 8,7 | 1,5 |
| L11 | 25432 | 5,1 | 418793 | 4,7 | 0,4 |
| L12 | 1261 | 0,3 | 32887 | 0,4 | -0,1 |
| L13 | 11875 | 2,4 | 196625 | 2,2 | 0,2 |
| L14 | 7098 | 1,4 | 117935 | 1,3 | 0,1 |
| L15 | 24279 | 4,9 | 367165 | 4,1 | 0,7 |
| L16 | 27927 | 5,6 | 441441 | 4,9 | 0,6 |
| L17 | 24795 | 5,0 | 432493 | 4,8 | 0,1 |
| L18 | 42867 | 8,6 | 703965 | 7,9 | 0,7 |
| L19 | 37155 | 7,4 | 703948 | 7,9 | -0,5 |
| L20 | 45510 | 9,1 | 839775 | 9,4 | -0,3 |
| L21 | 30353 | 6,1 | 577708 | 6,5 | -0,4 |
| L22 | 12785 | 2,6 | 245767 | 2,8 | -0,2 |
| L23 | 62742 | 12,6 | 1218055 | 13,6 | -1,1 |
| L24 | 29312 | 5,9 | 546839 | 6,1 | -0,3 |
| L25 | 6546 | 1,3 | 129692 | 1,5 | -0,1 |
| L26 | 306 | 0,1 | 14390 | 0,2 | -0,1 |
| L27 | 14163 | 2,8 | 281861 | 3,2 | -0,3 |
| L28 | 11202 | 2,2 | 226976 | 2,5 | -0,3 |
| L29 | 2195 | 0,4 | 39647 | 0,4 | 0,0 |
| L30 | 1089 | 0,2 | 28966 | 0,3 | -0,1 |
| L31 | 6020 | 1,2 | 156293 | 1,8 | -0,6 |
| L32 | 80 | 0,0 | 3177 | 0,0 | 0,0 |



Pela aplicação do teste de hipóteses verificou-se que é legítimo aceitar a hipótese nula para o nível de confiança de 95% (Tabela 4) , o que nos indica que as variações observadas nas amostras são resultantes de variações de natureza aleatória introduzidas nos dados de entrada e, no processo de avaliação da variável delta.

Tabela 4 – Descrição estatística da variável delta e respetivos resultados da aplicação do teste de hipóteses para a avaliação da exatidão temática da cartografia.

| Descrição | Resultados |
|------------------------|------------------------|
| Valor Médio | $-5,7 \times 10^{-16}$ |
| Desvio Padrão | 0,4441 |
| Estatística de teste | $-7,3 \times 10^{-15}$ |
| Significância do teste | 0,05 |
| Valor normal tabelado | 1,96 |
| Resultado do teste | Aceitar hipótese nula |

2.5. Harmonização e partilha de dados

As infraestruturas de dados geográficos constituem, na atualidade, a principal estratégia utilizada para o armazenamento e disponibilização de dados. A ocupação e o uso do solo têm sido tratados de forma diferenciada ao longo da Europa, exceto na obtenção da *Corine Land Cover* (CLC). Nesse sentido, a Diretiva INSPIRE vem ampliar a discussão da normalização da informação geográfica, incluindo os temas tratados no Projeto LANDYN.

A Diretiva INSPIRE, (Diretiva 2007/02/EC, de 14 de março), estabelece a criação da Infraestrutura Europeia de Informação Geográfica. A sua criação, que partiu de algumas diretivas ambientais precedentes, pretende promover a disponibilização de informação geográfica, normalizada para todos os Estados Membros aderentes (Figura 4). Portugal inclui-se no lote de países que têm responsabilidades de implementação da Diretiva INSPIRE, que foi transposta para a lei portuguesa através do Decreto-Lei nº 180/2009, publicado a 7 de agosto de 2009.

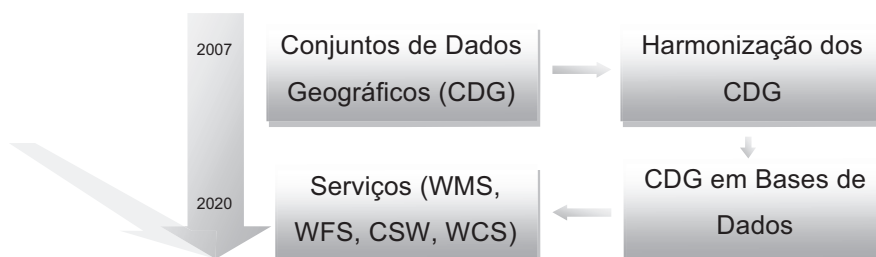


Figura 4 – Fluxo de procedimentos para tornar os CDG conforme com a Diretiva INSPIRE até ao ano de 2020. Adaptado de Barreira *et al.* (2012).

No âmbito da Diretiva INSPIRE foram criados três anexos de temas que englobam a grande maioria dos conjuntos de dados geográficos (CDG) produzidos no continente europeu (Rizone e Fonseca, 2010) e o estágio da transposição desta Diretiva em Portugal para esses conjuntos, onde se incluem os dados de ocupação do solo, podendo estes ser consultados em Vale *et al.* (2013).

Consideram-se, nesta abordagem, apenas os temas Ocupação do Solo (Anexo II) e Uso do Solo (Anexo III) pelo facto de serem os mais relevantes no âmbito do projeto e antecipando o esforço associado à sua disponibilização em Serviços web.

No caso da Ocupação do Solo, onde se inclui a Carta de Uso e Ocupação de Solo (COS), utilizada como dados de base, a especificação de dados contempla dois modelos de dados: vetor e raster.

No caso presente os dados considerados têm características vetoriais, e têm a particularidade da legenda LANDYN estar também direcionada para a monitorização das emissões de carbono (LULUCF). Esta foi uma das áreas que a Agência Europeia do Ambiente considerou como área de estudo para a elaboração das especificações de dados relativas ao tema Ocupação do Solo.

A nomenclatura usada por cada Estado Membro pode ser considerada no âmbito INSPIRE, havendo normalização apenas para os dados CORINE, cuja nomenclatura

é transversal para toda a Europa. Para esse fim são utilizadas nos modelos de dados as *code list* referentes à nomenclatura que se pretende utilizar. A classificação que é atribuída pelo LUCAS é também considerada na especificação de dados.

O tema do Uso do Solo está dividido em duas áreas fundamentais: uso do solo existente e uso do solo planeado. Os produtos finais do Projeto LANDYN permitirão apoiar as decisões de planeamento do território português no futuro, pelo que se deve considerar este tema.

A adequação dos CDG produzidos, em conformidade com o modelo de dados que se considere, no âmbito da Diretiva INSPIRE, é considerada como harmonização dos dados. Com a informação geográfica harmonizada considera-se a publicação dos CDG em Serviços web, servindo estes o propósito da partilha da informação geográfica através da Internet para a comunidade de utilizadores (Barreira *et al.*, 2014). Devem considerar-se no âmbito da Diretiva serviços WMS (Web Map Service), WMTS (Web Map Tile Service), WFS (Web Feature Service), CSW (Catalogue Services for the Web), WPS (Web Processing Services) e WCS (Web Coverage Services).