

INTEGRAÇÃO DE AUTÓMATOS CELULARES E REDES NEURONAIS EM AMBIENTE SIG PARA PREVISÃO DE ALTERAÇÕES NOS PADRÕES DE USO E OCUPAÇÃO DO SOLO

Jorge Rocha¹, José António Tenedório², Sara Encarnação³ e Saudade Pontes⁴

¹ Centro de Estudos Geográficos, Faculdade de Letras – Universidade de Lisboa,
jrocha@fl.ul.pt.

^{2, 3 e 4} e-GEO, Centro de Estudos de Geografia e Planeamento Regional, Faculdade de Ciências Sociais e Humanas – Universidade Nova de Lisboa, ja.tenedorio@fcsh.unl.pt,
sara.encarnacao@fcsh.unl.pt, saudade.pontes@fcsh.unl.pt

Resumo

A geocomputação é um campo emergente de investigação que advoga o uso de métodos de análise espacial computacionalmente intensivos como os que recorrem a pesquisas heurísticas, redes neuronais e autómatos celulares. Este trabalho apresenta um método de simular a evolução do uso do solo numa realidade periurbana, com recurso à integração de redes neuronais e agentes autónomos em ambiente SIG. A simulação das alterações no uso do solo recorrendo unicamente a agentes autónomos é desaconselhada, em virtude destes modelos, na sua forma mais convencional, apresentarem limitações na definição dos parâmetros espaciais e das regras de transição. Neste trabalho é utilizada uma rede neuronal para aferir o grau de importância que cada variável de predição (probabilidade) tem nos constrangimentos geográficos. Estas variáveis são obtidas com recurso ao SIG, que também é utilizado num fim de todo o processo com o intuito de verificar a validade dos resultados obtidos.

1. Introdução

Nos últimos anos registou-se um acréscimo na disseminação das metodologias de análise espacial com base em suporte informático, dando origem a uma nova área de conhecimento, que combina abordagens recorrentes dos Sistemas de Informação Geográfica (SIG) com outras emergentes (com principal incidência no seio da inteligência artificial) como os algoritmos genéticos – Autómatos Celulares (AC), os sistemas periciais - pesquisas heurísticas e lógica de predicados, análise de dados com incerteza e Redes Neuronais (RN).

Da premente necessidade de catalogar esta nova área interdisciplinar, para que não fosse catalogada como uma simples extensão das técnicas estatísticas para análise de dados espaciais, nasceu a designação geocomputação. Proposta por Stan Openshaw, esta designação tem granjeado uma crescente aceitação no meio, visto permitir enquadrar no seio das tecnologias de informação geográfica, um certo tipo de investigação aplicada que recorre aos SIG como ferramenta, mas que, simultaneamente, se afasta das soluções comerciais (Goodchild e Longley, 1999).

Este termo descreve o uso de métodos que recorrem intensivamente a computadores tendo em vista a perscrutação de conhecimento em geografia, com especial destaque para aqueles que utilizam formas de *clustering* não convencionais e/ou técnicas de análise. Ultimamente, esta definição enquadrou-se num contexto bastante mais lato, que envolve a análise espacial, a modelização dinâmica e a visualização da dinâmica espaço-tempo. Efectivamente, a geocomputação enquadra-se num conjunto de modelos e métodos computacionais desvelados à solução de problemas geográficos de grande complexidade (Couclelis, 1997). Não se está face a uma simples utilização do computador como ferramenta, como era proposto nos primórdios, mas sim perante o reconhecimento de que o computador já não faz parte dos meios de pesquisa, é sim - ele próprio - o meio (ambiente) da pesquisa (Goodchild e Longley, 1999).

Mas, quais são as motivações que incentivam a investigação no domínio da geocomputação? Existem três razões principais para a difusão desta “nova ciência”: A primeira prende-se com o grande volume de informação que actualmente se encontra disponível e, principalmente, com a sua facilidade de circulação; a segunda materializa-se de duas formas: a emergência da tecnologia SIG e um conjunto de técnicas decorrentes da algorítmica, como as redes neuronais, a lógica de predicados e os agentes autónomos; Por fim, a terceira, relaciona-se com os processos de análise espacial baseados em estatísticas espaciais, os quais têm sido um tópico fundamental da investigação geográfica nas últimas décadas.

2 – Modelos de Alterações nos Padrões de Uso do Solo

As mudanças que se verificam ao nível das mudanças de uso do solo resultam de diversos factores, nomeadamente: políticos, de gestão, económicos, culturais, do comportamento humano e ambientais, entre outros (Houghton, 1994; Dale *et al*, 1993; Medley *et al*, 1995;

Richards, 1990; Wilder, 1985; Vesterby e Heimlich, 1991). A compreensão da forma como estes factores se interligam e como influenciam a ocorrência de alterações no uso do solo, afigura-se como crítica na medida em que este processo de índole antropogénica pode conduzir a um vasto leque de impactos ambientais, como a alteração de ciclos hidrológicos (Steiner e Osterman, 1988), da dimensão e organização dos habitats naturais (Dale *et al*, 1993) e da diversidade das espécies (Costanza, Kemp e Boyton, 1993). No entanto, apesar de preocupantes, estas consequências não são únicas pois os efeitos das alterações no uso do solo não se cingem meramente a questões ambientais, podendo mesmo afectar as economias locais e regionais (Burchell, 1996).

Os modelos de alterações de uso do solo encontram utilidade como poderosa ferramenta para (1) explorar os vários mecanismos que levam à ocorrência de alterações de uso do solo e identificar quais as variáveis económicas, sociais e espaciais que os condicionam (Batty e Longley, 1994; deKoning *et al*, 1999), (2) prever potenciais (futuros) impactos ambientais e socioeconómicos das alterações de uso do solo (Alig, 1986; Theobald, Miller e Hobbes, 1997), e (3) avaliar a influência de políticas de planeamento alternativas no uso do solo e nos padrões de desenvolvimento (Bockstael *et al*, 1995). A projecção dos padrões de uso do solo é passível de ser calculada através de dados empíricos, sendo os percursos genéricos de mudança identificados como sequências de alterações (Lambin, 1997).

Apesar de muitos modelos focarem essencialmente a modelação de percentagens de usos referentes a uma determinada área geográfica, como por exemplo um concelho (Alig e Healey, 1987), os que prevêem os padrões de alteração do uso do solo providenciam um maior manancial de informação, a partir do qual se podem avaliar os impactos da mudança. A simulação de alterações no uso do solo é importante para uma grande variedade de questões relacionadas com o planeamento e a gestão. Esta pode fornecer a linha de base de um cenário de crescimento para fornecer uma imagem do futuro desenvolvimento dos padrões de uso, assumindo que as premissas de desenvolvimento vigentes se prolongam no futuro. A obtenção desta linha de base pode ser determinante para identificar futuros problemas relacionados com o crescimento urbano e, paralelamente, verificar se os planos existentes estão de acordo com as necessidades previstas ou se por outro lado é necessário proceder a correcções.

A simulação da evolução do uso do solo fornece informação de vital importância sobre o tipo, a escala, a quantidade e a densidade das transformações que provavelmente ocorrerão. Os modelos de alteração de uso do solo podem pois constituir ferramentas de grande valor para: *i*) explorar os vários mecanismos que despoletam e/ou influenciam as alterações, bem como as variáveis sociais, económicas e espaciais que lhes servem de base; *ii*) prever futuros impactos económicos e ambientais derivados das alterações; *iii*) avaliar a influência de políticas alternativas de intervenção e/ou gestão, no desenvolvimento dos padrões espaciais *iv*) preparar planos reguladores do uso do solo e procurar obter padrões de uso óptimos. Em qualquer dos casos poderá ser necessário recorrer a zonamentos rígidos para prevenir os potenciais problemas identificados através da simulação. Esta, irá municiar os agentes de planeamento com a informação necessária para fornecer ao público os serviços, equipamentos e infra-estruturas necessários ao desenvolvimento sustentado (Hathout, 1998).

A Maior parte da modelação económica das alterações nos padrões de uso deriva das teorias do custo do solo de Von Thünen e Ricardo (Mertens e Lambim, 2000). Nestes casos, qualquer parcela de território, tendo em conta os seus atributos, é assumida como estando alocada ao uso que mais a rentabiliza, verificando-se uma competição entre os usos pelas localizações mais favoráveis. Theobald e Hobbs (1998) delineiam dois tipos essenciais de modelos espaciais de alteração dos padrões de uso do solo: os modelos do tipo regressão linear múltipla (Mertens e Lambim, 2000) e os modelos baseados em transições espaciais. O objectivo primordial dos primeiros é estabelecer relações funcionais entre um conjunto de variáveis espaciais condicionantes (ex. distância à rede viária) que são utilizadas para localizar em antecipação as presumíveis alterações. As variáveis espaciais e as tendências de alterações no uso do solo são avaliadas com recurso a informação multi-temporal, permitindo estabelecer relações funcionais entre todos os parâmetros considerados, de forma a calcular as probabilidades futuras de mudanças no uso do solo. Por outro lado, os modelos de transição espacial são uma extensão do modelo (não espacial) de Markov (Hathout, 1988) e uma forma de autómato celular estocástico (Theobald e Hobbs, 1998).

O modelo de Markov propriamente dito entrou em desuso em virtude de apenas prever as alterações nas categorias de uso, não procedendo à espacialização dos resultados. Por seu lado, os autómatos celulares incorporam regras simples sobre os efeitos da adjacência espacial que condicionam as dinâmicas de um sistema de evolução do uso do solo e dão relevância a comportamentos e padrões emergentes, os quais são geralmente mais complexos do que

aqueles gerados pelos modelos de equilíbrio simples (Batty e Longley, 1994; Clarke *et al*, 1997). Ultimamente, tem-se assistido a um incremento das simulações do crescimento urbano com recurso a agentes autónomos (Wu, 1998; White e Engelen, 1993; Batty e Xie, 1994). A aplicação de AC em modelos urbanos pode dar destaque a uma grande variedade de fenómenos urbanos. Os modelos de AC urbanos têm melhor desempenho que os modelos tradicionais na simulação do crescimento urbano, principalmente devido ao facto de serem bastante mais simples que as complexas equações matemáticas que suportam esses modelos (tradicionais). No entanto, os seus resultados têm provado ser mais úteis e significativos quando observados numa perspectiva intuitiva (Deadman *et al*, 1993; White e Engelen, 1993; Wu, 1998).

As complexidades espaciais e temporais dos sistemas urbanos podem ser modeladas através da correcta e apropriada definição de regras de transição nos modelos de AC. Efectivamente, os modelos de evolução de uso do solo baseados em autómatos celulares são bastante válidos para descrever a dinâmica e o comportamento desses sistemas (Theobald e Hobbs, 1998). A simulação com AC fornece informação importante para a compreensão das teorias urbanas, tais como a evolução da forma e da estrutura (Li e Yeh, 2002). Por outro lado, a simulação de alterações em múltiplos usos do solo com recurso a AC é consideravelmente mais complexa que a simulação do crescimento urbano, a qual é normalmente efectuada numa base binária, i.e. o território ou é (1) ou não é (0) anexado a uso urbano. Quando estão presentes vários usos do solo, as regras de transição dos modelos urbanos com base em AC tornam-se substancialmente mais complexas porque a simulação envolve um mais vasto conjunto de variáveis espaciais e parâmetros, e modelos com estrutura deveras mais complexa.

Recentemente, foram propostas algumas alterações ao conceito original de AC, para que este permitisse incorporar constrangimentos geográficos. Neste processo, a característica mais importante a descartar é a de espaço isotrópico, substituído por um espaço em que cada célula tem inerente um conjunto próprio de atributos que representa as suas características relevantes. Estes avanços, de que a ligação entre os AC e o modelo de Markov é um exemplo, têm sido acompanhados de um crescimento da complexidade dos modelos (Couclelis, 1997). Estas modificações permitiram a integração, não só conceptual mas também prática, dos agentes autónomos em ambientes SIG. Uma vez aplicados a um espaço anisotrópico (idêntico na sua essência ao encontrado num SIG de base matricial) os agentes autónomos podem ser entendidos como uma espécie de SIG dinâmico (Batty e Xie, 1994).

3. Autómatos Celulares com Base em Redes Neurais

A integração entre SIG e Detecção Remota (tanto na vertente da fotografia aérea como das imagens de satélite) tem servido de base a inúmeros estudos sobre a detecção, identificação e quantificação de alterações no uso do solo (Howarth, 1986; Jensen *et al*, 1995; Li e Yeh, 1998). Em geral, estes modelos prevêm as alterações nos usos de acordo com as variáveis espaciais independentes geradas através das tradicionais ferramentas SIG de análise espacial (Mertens e Lambin, 2000). Estas variáveis costumam ser determinísticas e imutáveis durante o processo de modelação (Landis, 1995; Landis e Zhang, 1998). A modelação em SIG também pode ser utilizada para modelar as alterações aplicando vários tipos de constrangimentos. No entanto, muito embora seja possível prever futuras ocupações com base em dados empíricos, existe uma nítida carência de estudos nessa área devido à complexidade que tal tarefa acarreta.

No entanto, e como já foi referido, a crescente proeminência dos computadores conduziu a uma nova forma de ver o mundo. Esta nova perspectiva encara a natureza como uma forma de (geo)computação. Efectivamente, a representação em ambiente SIG do espaço geográfico é essencialmente estática, logo, um foco de investigação com elevada relevância na geocomputação é a elaboração de modelos que combinem os elementos estruturais do espaço (objectos geográficos) com os processos que o modificam (acções humanas e a forma como se processam ao longo do tempo). Estes modelos visam libertar o analista da visão estática do espaço, incutida após séculos de produção cartográfica tradicional e, ressaltar a componente dinâmica como uma parte essencial do espaço geográfico.

Esta motivação conduziu à utilização dos AC como método de simulação do crescimento urbano e regional. Os AC estendem esta analogia de forma a providenciar a visualização de todo um conjunto de células (pixéis) em interacção, sendo cada uma delas(es) um computador (automatização). Estes modelos adoptam uma abordagem de baixo-para-cima (*bottom-up*) na medida em que as interacções locais (vizinhança) dão relevo à formação de complexos padrões globais. Desta forma os AC podem ser entendidos como um sistema espacial dinâmico e relativamente simples, no qual o estado de cada célula da matriz depende do estado prévio das células que se encontram dentro de uma determinada vizinhança, de acordo com um conjunto de regras de transição. Nestes modelos o resultado da iteração anterior tem um papel determinante no desfecho da seguinte podendo, ao fim das diversas iterações que

compõem a simulação, formarem-se complexos padrões globais de uso do solo (Batty e Xie, 1994).

Os modelos de AC têm uma capacidade de modelação superior à dos SIG, quer ao nível do crescimento urbano, quer ao nível das alterações de usos do solo. As variáveis espaciais incluídas nos AC são actualizadas de forma dinâmica durante o ciclo iterativo o que se traduz em resultados não determinísticos. Em contraste, a generalidade dos modelos SIG encontra dificuldades em simular a evolução de uso do solo sem recorrer a regras locais e ciclos iterativos, usando quase sempre variáveis espaciais estáticas. Por outro lado, também é difícil capturar os elementos não-lineares que estão presentes em muitos dos fenómenos geográficos. Não é fácil explicar o significado teórico e intuitivo dos fenómenos quando a simulação é puramente baseada em SIG. Também os algoritmos utilizados na modelação em SIG são mais complexos que os utilizados pelos AC, tornando o processo computacionalmente mais exigente e alongando o tempo de simulação. A eficiência computacional dos AC deve-se ao facto de serem sistemas discretos e iterativos, que envolvem unicamente iterações entre regiões ao invés de um par de células. O facto de permitirem trabalhar com grandes resoluções espaciais confere aos AC uma importante vantagem em termos de modelação das dinâmicas de uso do solo (White e Engelen, 1997) e a correcta definição das regras de transição pode até permitir, durante o processo de simulação, o advento de variáveis não previstas (Wu, 1998) como por exemplo a criação de novos centros de agregação (Wu, 1998) ou as propriedades fractais das parcelas (White e Engelen, 1997).

Os modelos de AC tornaram-se bastante atractivos para as simulações em ambiente urbano porque permitem gerar resultados bastante interessantes (Xia e Yeh, 2002), constituindo uma poderosa ferramenta para compreender a cidade, vista como um sistema complexo e evolucionário. Numa cidade auto-organizada a evolução do uso é um processo intimamente ligado à história, em que a evolução passada condiciona a futura através de interacções locais entre as parcelas de terreno (Wu e Webster, 1998). Ao construírem-se regras apropriadas dentro de um AC, pode-se simular um extenso conjunto de comportamentos complexos. Os AC incorporam regras simples sobre os efeitos da adjacência espacial que condicionam a dinâmica dos sistemas e dão importância a comportamentos e padrões emergentes normalmente mais complexos do que os gerados pelos modelos de equilíbrio simples.

Um AC convencional consiste em: *i*) um espaço euclidiano dividido numa matriz regular; *ii*) uma janela móvel com uma função de vizinhança associada; *iii*) um conjunto de *layers* discretos que estabelecem o estado das células; *iv*) um conjunto de regras de transição; *v*) um número de iterações predefinido. Precisamente devido à natureza espacial de muitas das variáveis iniciais (principalmente utilizadas no ponto *iv*), a integração dos AC com os SIG é essencial, tendo a aplicação dos agentes autónomos em sistemas de informação sido inicialmente proposto por Tobler (1979). Os SIG permitem ao utilizador manipular e analisar dados espaciais explicitamente associados aos modelos. Por exemplo, o SIG permite ao utilizador elaborar variáveis de entrada para o modelo (constrangimentos ou probabilidades), identificando a existência de heterogeneidades ou padrões nos dados (Openshaw e Clarke, 1996), quantificando as alterações observadas e/ou previstas nos padrões espaciais (deKoning *et al*, 1999), e acedendo a factores que operam através de uma grande variedade de escalas (Qi e Wu, 1996). A grande maioria dos modelos de alteração de uso do solo com base em tecnologia SIG está organizada numa estrutura matricial (Clarke *et al*, 1997; Landis, 1994; Veldkamp e Fresco, 1996), porque este formato simplifica a representação do espaço através da sua partição em unidades de forma e tamanho iguais. Esta estrutura tem ainda a vantagem de permitir a fácil integração de dados provenientes de detecção remota, que são por natureza matriciais.

Muito embora os AC tenham muitas vantagens têm um problema que é como definir as regras de transição e a estrutura do modelo. Estas são normalmente dependentes da aplicação em causa, pois muito embora existam diversos modelos de AC de natureza genérica (Wu, 1998; Batty *et al*, 1999), eles apresentam formas substancialmente diferentes: As variações devem-se à existência de diversas formas de definir as regras de transição e as estruturas dos modelos. Por exemplo, Batty e Xie (1994) utilizaram a concentração num espaço de vizinhança e uma função decrescente de distância relativamente aos centros de crescimento para determinar as probabilidades de transição, Wu e Webster (1998) definiram as regras de transição com base em métodos de análise multi-critério, White e Engelen (1993) utilizaram para o mesmo efeito uma matriz de parâmetros predefinidos e Li e Yeh (2000) propuseram um modelo baseado numa matriz (imagem) em tons de cinzento para acomodar o processo gradual de conversão para urbano. Estes modelos também podem incluir constrangimentos para gerar formas urbanas idealizadas (Li e Yeh, 2000), opções e objectivos de planeamento para produzir cenários alternativos e teorias urbanas neoclássicas (Wu e Webster, 2000). Nestes modelos, têm sido propostas estruturas e regras de transição substancialmente

diferentes para responder a vários objectivos e especificações. O dilema da escolha do modelo apropriado está sempre presente na medida em que existe um variado leque de opções.

Outro problema dos modelos de AC, e talvez o maior, é o da determinação das ponderações a atribuir a cada factor. No passado estes modelos apenas eram utilizados para simular o crescimento urbano na perspectiva da transição rural-urbano. A simulação deste tipo de crescimento, que apenas lida com estados binários – urbanizado ou não, é relativamente fácil, mas os modelos AC tornam-se consideravelmente mais complexos quando são introduzidos múltiplos usos, como residencial, comercial e industrial (Batty *et al*, 1999). Quando se lida com diversos usos do solo em competição entre si pelo território o número de factores de ponderação aumenta consideravelmente e os modelos tornam-se mais complexos. Existem numerosos parâmetros que precisam de ser determinados para que uma simulação reflecta um sistema urbano particular e o naipe de possíveis modelos a empregar é enorme (Batty *et al*, 1999).

A simulação envolvendo múltiplos usos do solo implica a utilização de bastantes variáveis espaciais. A contribuição de cada uma destas variáveis para a simulação é quantificada pelo peso, ou parâmetro, que lhe está associado: e existem numerosos parâmetros que têm de ser quantificados antes de se dar início à simulação. O valor destes parâmetros tem um grande peso (efeito) nos resultados da simulação, verificando-se que diferentes combinações de valores conduzem a formas urbanas totalmente diferentes (Batty *et al*, 1999).

Na maioria das situações é necessário calibrar o modelo AC de modo a assegurar que a simulação gere resultados próximos da realidade, sendo que esta calibração é extremamente difícil de realizar quando a conversão ocorre entre diversas classes de uso do solo. Existem dois tipos principais de processos de calibração: os baseados em métodos estatísticos e os baseados em abordagens tipo tentativa e erro. No primeiro caso pode-se utilizar a regressão logística. Este tipo de modelos preocupa-se apenas com as conversões binárias (urbanizado ou não) e podem apresentar algumas limitações quando os factores espaciais e as estruturas dos modelos se apresentam complicadas. Paralelamente, não são válidos para trabalhar com factores espaciais que apresentem um elevado grau de correlação entre si e têm dificuldades em operar com dados de fraca qualidade e/ou não tratados. No segundo caso, não é necessária a utilização de métodos matemáticos rígidos. Um método simples é comparando visualmente os resultados de simulações efectuadas de acordo com diferentes combinações de pesos

(Clarke *et al*, 1997; Ward *et al*, 2000). Contudo, quando existem muitas variáveis é difícil definir as combinações e aceder aos resultados visualmente, visto que os padrões gerados são muito complexos. White *et al* (1997) propuseram um modelo alternativo que se traduzia numa matriz intuitiva de pesos. Este modelo utilizava 378 parâmetros (21 x 18) e não se baseava em modelos matemáticos rígidos o que se acabava por traduzir num grande consumo de tempo para efectuar a calibração necessária à obtenção da matriz. Para obviar ao problema da subjectividade por detrás da calibração visual, Clarke e Gaydos (1998) desenvolveram um método relativamente robusto, que calcula os ajustes entre dados reais (históricos) e vários resultados de simulações. O melhor conjunto de parâmetros será aquele que proporcionar o melhor desempenho na simulação. Este tipo de calibração é computacionalmente intensivo (no caso, centenas de horas até encontrar o resultado ideal), mas esse facto deve-se essencialmente aos algoritmos de ajuste. Assim, uma forma de proceder a este tipo de calibração mas encurtando ao mesmo tempo o tempo de processamento poderá ser através da utilização de redes neuronais.

As redes neuronais foram desenvolvidas à imagem do sistema de neurónios interconectados que constituem o cérebro, para que os computadores pudessem, observando os relacionamentos entre os dados, imitar a sua capacidade de discriminar padrões e aprender através da experiência. Rosenblatt (1958) é creditado com o desenvolvimento de uma das primeiras redes neuronais devido a ter criado o “perceptrão”. O perceptrão consiste num único nó que recebe informação ponderada e limita os resultados de acordo com uma regra definida. Este tipo de máquina neuronal simples é capaz de classificar dados linearmente separáveis e operacionalizar funções lineares.

O perceptrão multi-camada (Rumelhart, Hinton e Williams, 1986) constitui um dos tipos de RN mais utilizados. É constituído por 3 níveis: *i*) nível de entrada, *ii*) nível escondido e *iii*) nível de saída, e possui a vantagem de poder identificar relacionamentos de natureza não-linear. Os algoritmos que compõem as RN calculam os pesos das variáveis de entrada e dos nós dos níveis de entrada, escondido(s) e saída recorrendo à introdução dos dados iniciais (de entrada) através de um processo designado de alimentação para a frente, que os propaga através dos níveis escondido(s) e de saída. Os sinais propagam-se de nó para nó, sendo modificados pelo peso associado a cada ligação. O nó receptor procede à soma ponderada dos valores de todos os nós do nível anterior que lhe estão conectados. A saída deste nó é então calculada em função dos valores ponderados de entrada, sendo designada de função de

activação. Deste modo, os dados movem-se para a frente, de nó para nó, com a ocorrência de múltiplos somatórios ponderados, até atingirem o nível de saída.

Numa RN os pesos são determinados através da utilização de um algoritmo de treino, dos quais o mais popular é o de retropropagação (RP). O algoritmo RP selecciona ao acaso os pesos iniciais e depois compara o resultado obtido com o esperado. A diferença entre os valores obtidos e os esperados para todos os usos é sumariado através do erro médio quadrático. Depois da rede ter testado todos os usos os pesos são modificados de acordo com a regra *delta* modificada, para que o erro total seja distribuído pelos vários nós da rede. Este processo de alimentação para a frente e retropropagação dos erros é repetido iterativamente (nalguns casos, milhares de vezes) até que o erro estabilize num nível baixo.

4. Autómatos Celulares com Base em Redes Neurais na Previsão de Alterações nos Padrões de Uso do Solo

4.1. Área de Estudo e Evolução do Uso do Solo

O primeiro passo na persecução deste objectivo centrou-se na escolha da área teste, onde a metodologia vai ser aplicada. Esta deve ser bastante heterogénea relativamente aos usos, na medida em que o método proposto visa uma posterior extrapolação a toda a Grande Área Metropolitana de Lisboa (GAML) e, portanto, como acontece sempre nestas situações, é vital escolher um espaço onde pontifiquem, na medida do possível, todas as variáveis existentes.

Esta opção é tanto mais laboriosa quanto a GAML, em virtude de ser o pólo dinamizador da economia nacional, detendo mais de 60% do total do emprego do Continente, nos sectores secundário e terciário, ser o paradigma da consequente concentração da população e crescimento exponencial das áreas urbanas, constituindo um fenómeno territorial complexo. A urbanização ocorreu, neste território, de forma casuística, provocando graves desequilíbrios, cuja principal causa, de ordem estrutural e funcional, se deve ao facto de a evolução da economia não ter sido acompanhada pelas necessárias medidas de ordenamento do território (Machado *et al*, 1997). O concelho de Almada, como espaço integrado da Grande Área Metropolitana de Lisboa, apresentou-se desde o início como um espaço privilegiado para a aplicação da metodologia desenvolvida.

A escolha deste concelho prende-se com diversos factores de grande importância, se não veja-se: bastante próximo da cidade-centro (Lisboa), gozando de acessibilidade acrescida pela construção da Ponte sobre o Tejo, da possibilidade de expansão urbana e de potencialidades naturais raras, o território concelhio assistiu, entre 1967 e 1986, a uma pressão urbanística desmedida que se traduziu numa ocupação irreversível do solo mesmo em áreas de fortes restrições físicas. Entre 1960 e 1991 (datas correspondentes ao recenseamentos gerais da população) a sua população não deixou de crescer ainda que em ritmos diferenciados; a dependência em termos de emprego é uma realidade que os movimentos pendulares entre Lisboa e Almada corroboram; as suas praias impuseram-se como alternativa às da linha do Estoril; a segunda residência assume no concelho uma forma de habitação com grande peso percentual.

A envergadura e rapidez destes processos não podia deixar de ter expressão cartográfica (figura 1) significativa. O período de análise escolhido (1967 a 1991) parece suficientemente lato para evidenciar, recensear em tipo e área, avaliar e interpretar as grandes modificações de uso do solo entretanto ocorridas. A sua escolha prende-se também com o facto de a primeira cobertura aérea utilizada (1967) ter sido realizada um ano após a abertura da Ponte sobre o Tejo. A quantificação da informação contida nas cartas de uso solo (figura 2) permite corroborar tudo o que até agora foi afirmado. Avaliando a situação, em 1986, e a dinâmica do uso do solo, entre 1967 e 1986, verifica-se que a carta de uso do solo de 1986 representa um território diferenciadamente edificado. De um total de 7156,51 ha, 35% dizem respeito a áreas construídas e deste valor 15,7% referem-se unicamente à área de habitação essencialmente unifamiliar: a nível concelhio é este o uso que, depois da área agrícola (2115,67 ha) e da mata (1548,95 ha), ocupa maior superfície, cerca de 1125 ha.

Tomando em consideração a classificação utilizada pode-se concluir que dos 2115,67 ha de área agrícola calculada (cerca de 30% em relação ao total de usos) a maior parte não tem qualquer significado produtivo. De facto, algumas áreas, embora apresentando potencialidades agro-florestais, estão em situação expectante face a previsíveis usos urbanos pelo que, melhor se deveriam denominar “incultos urbanos” dado o abandono e dispersão que frequentemente apresentam entre conjuntos de lotes e edifícios, do foro legal, onde se tornou difícil ou mesmo impossível, qualquer prática agrícola de rendimento. Deste modo decorre o empolamento do peso da área agrícola em todo o concelho.

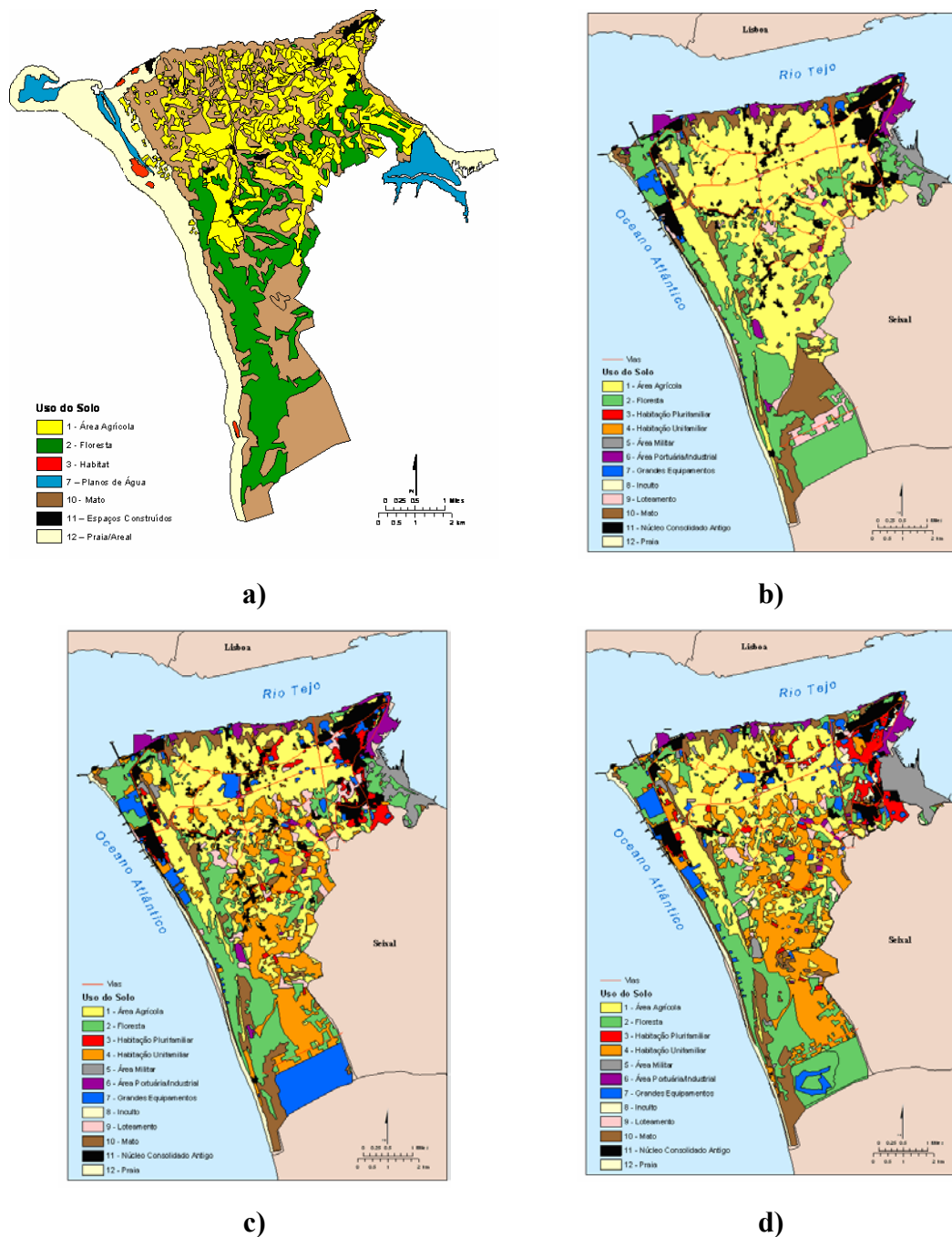


Figura 1 – Uso do solo em Almada nos anos de (a) 1869, (b) 1967, (c) 1981 e (d) 1991.

Entre 1980 e 1991 mantiveram-se as tendências dos anos anteriores embora a um ritmo não tão acentuado. As classes “naturais” continuaram a perder importância relativa face às classes “artificiais”, com especial incidência para a agrícola que perdeu 4%, cifrando-se em 1991 em cerca de 26 % da área total do concelho. O grande crescimento verificou-se ao nível dos usos urbanos como a habitação plurifamiliar e principalmente a unifamiliar (grande consumidora de espaço e bastante influenciada pelo mercado de segunda habitação). O facto mais grave é

que este crescimento não foi acompanhado pelo crescimento das áreas de equipamentos e infra-estruturas que se mantiveram relativamente estáveis durante o período.

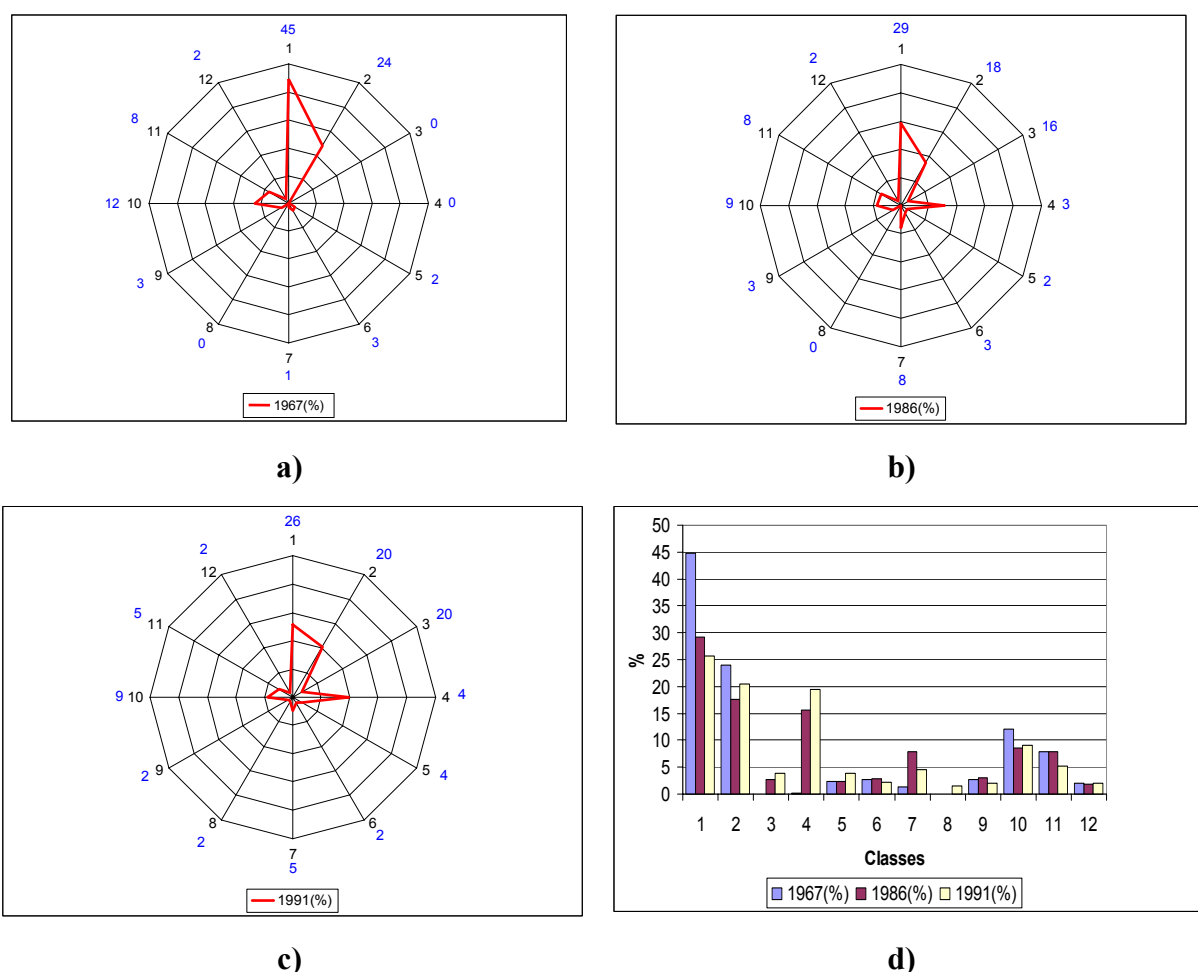


Figura 2 – Diagramas de uso do solo para 1967 (a), 1980 (b) e 1991 (c), e gráfico de evolução entre as três datas (d). (Nota: nos diagramas o primeiro anel de valores corresponde ao número da classe e o segundo às percentagens)

4.2. Implementação do modelo

O modelo utilizado é baseado em cinco fases sequenciais: *i)* processamento/codificação da informação de forma a criar níveis espaciais a partir das variáveis de previsão; *ii)* aplicação de regras espaciais que relacionem as variáveis de previsão com a transformação de uso do solo em cada local da área; *iii)* integração de todos os níveis de informação utilizando uma de três técnicas (regressão logística, análise multi-critério ou redes neuronais) e *iv)* ordenar os dados de maneira a criar uma série temporal que permita prever futuros usos.

Na fase 1 - processamento dos dados espaciais - os dados de entrada (figura 3) são gerados a partir de uma série de níveis de base integrados e geridos em ambiente SIG. Esta base de dados contém informação tanto em formato vectorial como em formato matricial e foi constituída com o propósito de fornecer a informação espacial básica para a simulação. Os níveis de informação representam temáticas diversas como o histórico dos usos do solo (e.g. floresta, urbano), a topografia ou elementos da paisagem (e.g. estradas). Muito embora a base de dados contenha informação em formato matricial e vectorial, esta última teve que ser convertida em matricial (pixel 10 x 10 m), de modo a viabilizar a simulação. Como a grande maioria dos modelos de AC, o modelo proposto assenta sobre uma estrutura celular. Por fim, os pixéis são codificados de forma a representarem constrangimentos ou probabilidades de ocorrência. No primeiro caso representam *layers* binários, onde o “0” representa ausência e o “1” presença, e no segundo caso constituem variáveis contínuas.

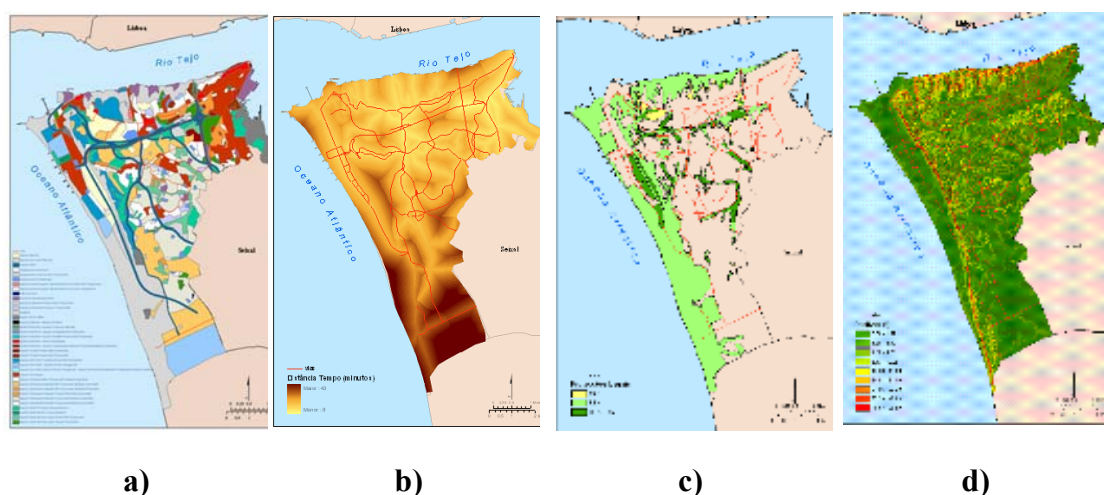


Figura 3 – Dados de entrada: PDM (a), acessibilidade a Lisboa (b), RAN e REN (c) e declives (d)

Para a fase 2 - aplicação de regras de transição – os dados de entrada são desenvolvidos recorrendo a um conjunto de regras de transição que quantificam os efeitos espaciais que as células de previsão detêm nas mudanças de uso do solo (Pijanowski *et al.*, 2000). Foram utilizados duas classes de regras de transição: *i*) vizinhanças ou densidades e *ii*) distância às células de previsão. Os efeitos de vizinhança são baseados na premissa que a composição das células vizinhas (i.e. janela de vizinhança) tem efeito na tendência de uma célula central mudar de uso. Por seu lado, as regras de transição espacial baseadas na distância relacionam a distância euclidiana entre cada célula e a variável de previsão mais próxima.

Certas localizações são codificadas de forma a inviabilizarem quaisquer mudanças. Esta acção torna-se necessária em zonas, dentro das quais, a expansão urbana é interdita (e.g. Reserva Ecológica Nacional – REN). Nestes casos utilizou-se o “0” para codificar todas as células onde a mudança não se pode verificar, atribuindo-se o “1” a todas as outras. Posteriormente, foi calculado o produto de todos estes níveis, obtendo-se um único nível correspondente às “zonas de exclusão”.

A fase 3 - integração das variáveis de previsão – a integração das variáveis de previsão pode ser feita através de regressão logística, análise multi-critério (Pijanowski *et al*, 2000) ou redes neuronais (Pijanowski *et al*, 2002). Cada um destes processos requer uma forma diferente de normalização dos dados. Neste caso optou-se pela utilização das RN, dimensionando-se todos os ficheiros relativamente a um *layer* de referência, neste caso o limite administrativo do concelho de Almada (célula de 10 x 10 m). O resultado desta fase é um mapa de probabilidade de mudança para cada célula, obtido por integração através da RN de todos os valores de mudança derivados de todas as variáveis de predição.

A fase 4 – indexação temporal – a quantidade de células (território) que se prevê que transite para um tipo uso diferente (tabela 1) pode ser calculada através de vários processos como a análise de componentes principais (Li e Yeh, 1998), é calculado recorrendo aos campos de *Markov*. Neste caso a matriz da probabilidade da transição é o resultado do cruzamento das duas imagens de treino (uso do solo inicial e final) ajustadas pelo erro proporcional. A matriz da área da transição é produzida pela multiplicação de cada coluna na matriz da probabilidade da transição pelo número das células do uso do solo correspondente na imagem mais antiga. O erro proporcional expressa a probabilidade das classes uso do solo no mapas da entrada estarem incorrectas (isto é, 0.0 indicaria um mapa perfeitamente exacto). As probabilidades condicionais da saída são multiplicadas por “1 - erro proporcional” para produzir os valores finais da probabilidade condicional.

Esta comunicação demonstra que as redes neuronais podem ser convenientemente incorporadas com autómatos celulares para simular a evolução de múltiplos usos do solo. O método proposto pode obviar a algumas das dificuldades impostas pelos modelos de autómatos celulares tradicionais, na simulação de sistemas urbanos complexos e múltiplas alterações de uso e ocupação do solo, através da redução significativa do tempo necessário para definir os valores dos parâmetros, as regras de transição e as estruturas do modelo. O

SIG permite a fácil obtenção de dados de treino que permitem calibrar modelo e obter, com facilidade, o valor dos referidos parâmetros. Esta abordagem tem como principal vantagem o facto de conseguir lidar com dados de entrada incompletos e erróneos e da superfície de previsão gerada ser claramente não linear, a qual abre um leque de probabilidades bastante superior ao das superfícies obtidas pelos modelos de regressão (linear) tradicionais. Em grande parte dos fenómenos geográficos as variáveis encontram-se correlacionadas e os métodos tradicionais, como as técnicas de análise multi-critério, são inadequados para avaliar os pesos correctos das variáveis (correlacionadas). No modelo de autómatos celulares com base em redes neuronais, as variáveis espaciais não têm necessariamente que ser independentes umas das outras.

Classes		1	2	3	4	5	6	7	8	9	10	11	12
1	Área Agrícola	0.74	0.01	0.07	0.07	0.00	0.00	0.02	0.00	0.10	0.00	0.00	0.00
2	Floresta	0.00	0.75	0.00	0.00	0.00	0.00	0.16	0.00	0.06	0.04	0.00	0.00
3	Habitação Plurifamiliar	0.00	0.00	0.58	0.42	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	Habitação Unifamiliar	0.00	0.00	0.17	0.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	Área Militar	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	Área Portuária/Industrial	0.00	0.00	0.00	0.00	0.00	0.91	0.04	0.00	0.00	0.00	0.00	0.05
7	Grandes Equipamentos	0.00	0.00	0.00	0.00	0.00	0.15	0.85	0.00	0.00	0.00	0.00	0.00
8	Inculto	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	Loteamento	0.06	0.00	0.05	0.61	0.00	0.00	0.04	0.02	0.20	0.01	0.02	0.00
10	Mato	0.00	0.10	0.00	0.15	0.00	0.00	0.00	0.00	0.02	0.73	0.00	0.00
11	Núcleo Consolidado Antigo	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.85	0.00
12	Praia	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.00	0.00	0.84

Tabela 1. Probabilidades de transição de usos entre 1991 e 2015

5. Resultados e comentários finais

Quando a análise incide sobre múltiplos usos do solo, a calibração dos autómatos celulares afigura-se bastante complexa. Os modelos de calibração tradicionais não são robustos porque assentam principalmente em abordagens do tipo tentativa e erro. Estas abordagens envolvem o teste de um conjunto variado de possíveis combinações de parâmetros de forma a tentar atingir o melhor ajustamento possível e este processo, além de ser computacionalmente intensivo, pois existem inúmeras combinações possíveis, também assenta em algoritmos muito dependentes da aplicação. As redes neuronais são bastante robustas e funcionam satisfatoriamente na calibração de modelos de simulação, com recurso a algoritmos de

retropropagação. Neste estudo, o processo de treino da rede neuronal determina automaticamente o valor dos parâmetros (pesos), os quais são posteriormente importados para o modelo de autómatos celulares de forma a simulas as múltiplas alterações de uso e ocupação do solo. O método desenvolvido torna o processo de previsão mais célere que nos modelos AC tradicionais, embora ainda possa ser considerado computacionalmente pesado. Relativamente aos valores apurados, e.g. mapa de uso do solo para 2015 (figura 4), pode-se observar que: i) o mapa resultante é coerente com as recentes tendências de uso do solo observadas no concelho; ii) prevê-se que a agricultura continue a perder importância percentual para a habitação; iii) observa-se uma tendência para a colmatação dos espaços-ilha.

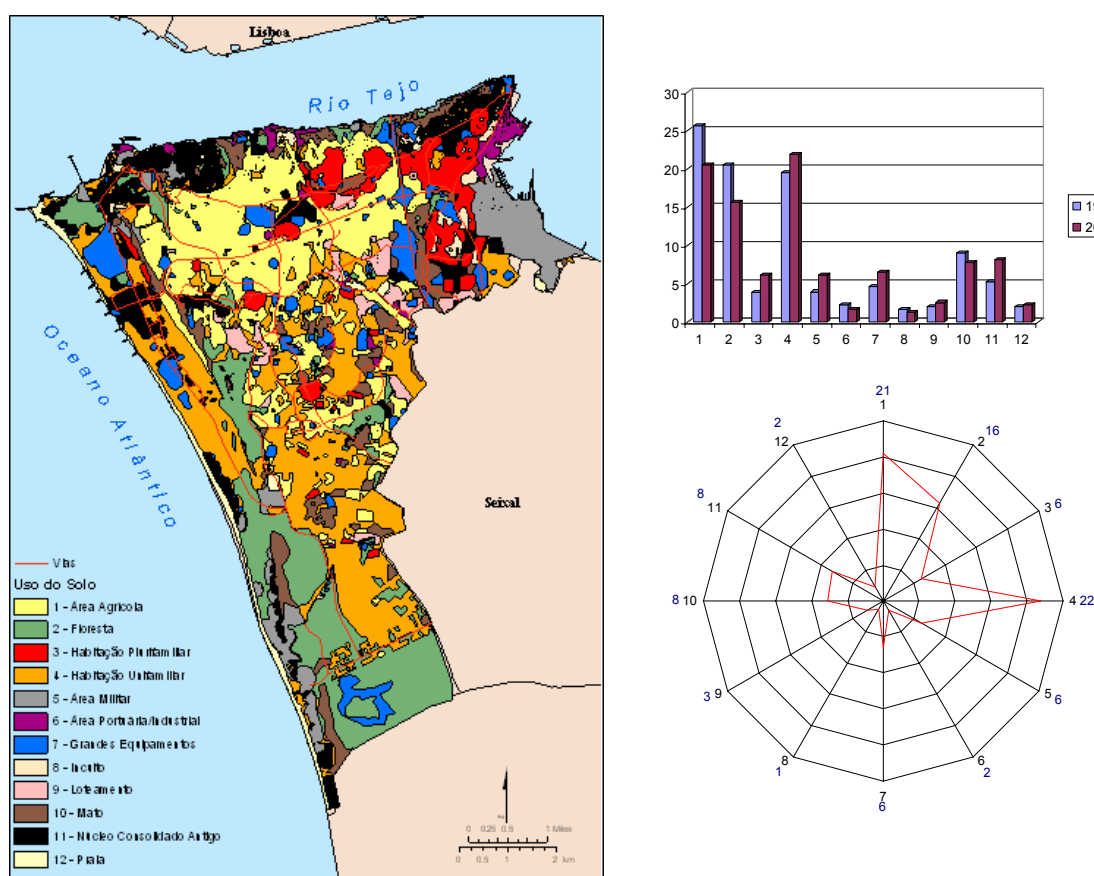


Figura 4 – Dados apurados através da previsão de evolução do uso do solo para 2015

6. Bibliografia

- Alig, R. J. (1986): Econometric analysis of the factors influencing forest acreage trends in the southeast, *Forest Science*, 32, pp. 119-134.
- Alig, R. J. e Healey, R. G. (1987): Urban and built-up land area changes in the United States: an empirical investigation of determinants, *Land Economics*, 63(3), pp. 216-226.
- Batty e Xie (1994): *From Cells to Cities*, Environment and Planning B, 21:pp. 531-548.
- Batty *et al*, (1999): Modelling urban dynamics through GIS-based cellular automata *Computer, Environment and Urban Systems*, 23, pp. 205 – 233.
- Batty, M. e Longley, P. (1994): Urban modeling in computer-graphic and geographic information system environments, *Environment and Planning B*, 19, pp. 663-688.
- Bockstael *et al* (1995): Ecological economic modeling and valuation of ecosystems, *Ecological Economics*, 14, pp. 73-90.
- Burchell, R. W. (1996): Economic and fiscal impacts of alternative land-use patterns, In S. Batie (Ed.), *The land use decision making process: its role in a sustainable future for Michigan*, January 9-10, Michigan State University Extension, East Lansing, Michigan.
- Clarke e Gaydos (1998): Loose-coupling a cellular automata model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science*, 12, pp. 699 – 714.
- Clarke *et al* (1997): A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B: Planning and Design*, 24, pp. 247 – 261.
- Costanza, R., Kemp, W. M. e Boyton, W. R. (1993): Predicability, scale and biodiversity in coastal and estuarine ecosystems: implications for management, *Ambio*, 22, pp. 88-96.
- Couclelis (1997): From cellular automata to urban models: new principles for model development and implementation. *Environment and Planning B: Planning e Design*, 24, pp. 165 – 174
- Dale *et al* (1993): Causes and effects of land use change in Central Rondonia, Brazil, *Photogrammetric Engineering and Remote Sensing*, 59(6), pp. 97-105.
- Deadman *et al* (1993): Modelling urban dynamics through GIS-based cellular automata, *Computer Environment and Urban Systems*, 23, pp. 205-233.
- deKoning *et al* (1999): Multi-scale modelling of land use change dynamics in Ecuador. *Agricultural Systems*, 61, pp. 77 – 93.

- Goodchild e Longley (1999): *Geographical Information Systems – Volume 1*, 2^a ed., John Wiley e Sons, INC., New York, p. 580.
- Hathout, S. (1988): Land use change analysis and prediction of the suburban corridor of Winnipeg, Manitoba, *Journal of Environmental Management*, 27, pp. 325-335.
- Houghton, R. A. (1994): The world-wide extend of land use change, *Bioscience*, 44(5), pp. 305-313.
- Howarth (1986): Landsat digital enhancements for change detection in urban environment. *Remote Sensing of Enviornment*, 13, pp. 149 – 160.
- Jensen *et al* (1995): Inland wetland change detection in the Everglades water conservation area 2A using a time series of normalized remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 61, pp. 199 – 209.
- Lambin, E. (1997): Modelling and monitoring land cover change processes in tropical regions, *Progress in Physical Geography*, 21, pp. 375-393.
- Landis e Zhang (1998): The second generation of the California urban features model. Part 1: model logic and theory. *Environment and Planning B: Planning and Design*, 25, pp. 447 – 461.
- Landis, J. D. (1994): The California urban features model: a new-generation of metropolitan simulation-models. *Environment and Planning B*, 21, pp. 399 – 420.
- Landis, J.D. (1995): Imagining land use features: applying the California urban futures model. *Journal of American Planning Association*, 6, pp. 438 – 457.
- Li, X. e Yeh, A. (2002): Neural-network-based cellular automata for simulating multiple land use changes using GIS, *International Journal of Geographical Information Science*, 16(4), pp. 323-343.
- Li, X. e Yeh, A. G.O. (1998): Principal component analysis of stacked multi-temporal images for the monitoring of rapid urban expansion in the Pearl River Delta. *International Journal of Remote Sensing*, 19, pp. 1501 – 1518.
- Li, X. e Yeh, A. G.O. (2000): Modelling sustainable urban development by the integration of constrained cellular automata and GIS. *International Journal of Geographical Information Science*, 14, pp. 131 – 152.
- Machado *et al* (1997): Os Planos Directores Municipais da Área Metropolitana de Lisboa. Emergência de uma Futura Base de Dados para um novo Planeamento Regional, *Proceedings do ESIG/97*, Junho, Lisboa.
- Medley *et al* (1995): Landscape Change with agricultural intensification in a rural watershed, Southwest Ohio, *USA Landscape Ecology*, 10(3), pp. 161-176.

- Mertens, B. e Lambin, E. F. (2000): Land cover-change trajectories in Southern Cameroon. *Annals of the Association of American Geographers*, 93, pp. 467 – 494.
- Openshaw, S. e Clarke, G. (1996): Developing spatial analysis functions relevant to GIS environments. *In: M. Fischer, H. J. Scholten. e D. Unwin (Eds.). Spatial analytical perspectives on GIS*, Tayler and Francis, London, pp. 21 – 37.
- Pijanowski *et al* (2002) : Using neural networks and GIS to forecast land use changes : a Land Transformation Model, *Computers, Environment and Urban Systems*, n° 26 (2002), pp. 553-575.
- Pijanowski, B. C. *et al.* (2000): A land transformation model :integrating policy, socioeconomics and environmental drivers using a geographic information system. *In: L. Harris e J. Sanderson (Eds.). Landscape ecology: a top down approach*. Lewis Publishers, Boca Raton
- Qi, Y.e Wu, J. (1996): Effects of changing spatial resolution on the results of landscape pattern analysis using autocorrelation indices. *Landscape Ecology*, 11(1), pp. 39 – 49.
- Richards, J. (1990): Land transformation, *In B. L. Turner (Ed.)*, Earth as transformed by human action: global and regional changes in the biosphere over the past 300 years (pp. 713), Cambridge University Press with Clark University Press, Cambridge, New York.
- Rosenblatt F. (1958): The perceptron: a probabilistic model for information storage and organization I the brain. *Psychological Review*, 65, pp. 386 – 408.
- Rumelhart, D., Hinton, G. e Williams, R. (1986): Learning internal representations by error propagation. *In: D. E. Rumelhart e J. L. McClelland (Eds.) Parallel distributed processing: explorations in the microstructures of cognition*, Cambridge, MIT Press, Vol. 1, pp. 318 – 362.
- Steiner, F. R. e Osterman, D. A. (1988): Landscape planning: a working method applied to a case study of soil conservation, *Landscape Ecology*, 1(4), pp. 213-226.
- Theobald, D. M. e Hobbs, N. T. (1998): Forecasting rural land-use change: a comparison of regression and spatial transition-based models, *Geographical and Environmental Modeling*, 21(1), pp. 65-82.
- Theobald, D. M., Miller, J. M. e Hobbes, N. T. (1997): Estimating the cumulative effects of development on wildlife habitat, *Landscape and Urban Planning*, 39, pp. 25-36.
- Tobler, W. R. (1979): Cellular geography. *In: Gale, S. e Olsson, G. (Eds.), Philosophy in Geography*, Dordrecht, Holland, D Reidel Publishing Company, pp. 379 – 386.
- Veldkamp, A. e Fresco, L. O. (1996): CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica. *J.D. Ecological Modelling*, 91, pp. 231 – 248.

- Vesterby, M. e Heimlich, R. (1991): Land use and demographic change: results from fast-growing counties, *Land Economics*, *Land Economics*, 67(3), pp. 271-291.
- Ward, D. P. *et al* (2000): A stochastically constrained cellular model of urban growth. *Computers, Environment and Urban Systems*, 24, 539 – 558.
- White, R. e Engelen, G. (1993): Cellular automata and fractal urban form: a cellular modeling approach to the evolution of urban land use patterns. *Environment and Planning A*, 25, pp. 1175 – 1199.
- White, R. e Engelen, G. (1997): Cellular Automata as the Basis of Integrated Dynamic Regional Modelling. *Environment and Planning B*: 24, pp. 235 – 246.
- White, R. *et al* (1997): The use of constrained cellular automata for high-resolution modeling of urban land use dynamics. *Environment and Planning B: Planning and Design* 24, pp. 323 – 343.
- Wilder, M. G. (1985): Site and situation determinations of land use change: an empirical example, *Economic Geography*, 61, pp. 332-344.
- Wu, F. (1998): An experiment on the generic polycentricity of urban growth in a cellular automatic city. *Environment and Planning B: Planning and Design* 25, pp. 103 – 126.
- Wu, F. e Webster, C. J. (1998): Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and Planning B: Planning and Design* 25, pp. 103 – 126.
- Wu, F. e Webster, C. J. (2000): Simulating artificial cities in a GIS environment: urban growth under alternative regulation regimes. *International Journal of Geographical Information Science*, 14, pp. 625 – 648.
- Xia e Yeh (2002): Neural-network-based cellular automata for simulating multiple land use changes using GIS, *International Journal of Geographical Information Science*, Vol. 14, nº 6, pp. 323-343.