



Universidade de Coimbra

Caracterização e análise de campos dunares marcianos
utilizando técnicas de reconhecimento de padrões

Mestrado em Astrofísica e Instrumentação para o Espaço

por: Pedro Ivo Kercadio Sarmento

Orientado por:

Doutor João Fernandes

Doutor David A. Vaz

Agradecimentos

Em primeiro lugar tenho de agradecer à minha mãe, por todo o apoio que me deu durante este ano e todos os anteriores, por sempre me ter impelido a avançar e a procurar novas oportunidades. Tenho também de dar o meu reconhecimento à minha namorada, Elisa, que me ajudou e que esteve do meu lado durante este Mestrado, e o Nuno, que foi sempre um bom irmão e companheiro. Uma palavra também para os meus avós, para a Janick e o Bernardo, que tantas vezes me receberam na sua casa e me deram palavras de apoio, e também a minha preocupada e carinhosa avó Branca, que tantas vezes ligou e ofereceu ajuda, e o meu avô Alberto, que faleceu em Fevereiro.

Agradeço também ao David, sempre disponível, e que me veio auxiliar e dar conselhos sempre que havia problemas com o programa, às vezes para prejuízo do seu próprio trabalho. Também ao Dr. João Fernandes, que me acolheu no Observatório e acompanhou a escrita da tese, e à Dr. Teresa Barata, que criou esta bolsa de mestrado e me aceitou para o cargo, e que sem ela nada disto teria acontecido, e ainda ao Dr. Nuno Peixinho, que, apesar de ter partido para o Chile, foi uma preciosa ajuda durante o Mestrado. Deixo uma nota para o resto do pessoal do observatório, que foram animando os dias com almoços e churrascos deliciosos. E não posso terminar sem mencionar o meu colega de Mestrado, o Marco, que partilhou as aulas comigo e sempre foi um bom amigo.

Abstract

The goal of this Master Thesis is the design and development of a Python computer program, taking advantage of the ArcGIS software to achieve a precise and effective automatic characterization of martian dune fields. The developed methodology uses neural networks to predict the different dominant transport directions from the dune patterns present on different regions of the planet. MRAMS data (a set of simulations of the wind directions on specific regions of Mars) is used for that, as well as the MGDDD (Mars Global Digital Dune Database). These two data sources are analyzed and compared to what the newly developed method can offer in the field.

The program also intends to be able to differentiate between different types of dunes, using clustering methods to distinguish areas where different patterns are present. The testing and validation of the developed methods have as a goal the application of these algorithms to new dune fields, or even to the whole planet.

It is concluded that the program cannot be applied blindly to an area, requiring previous study of the dominant wind directions indicated by the ground images. The application of the method or similar techniques to dune areas or fields of planet Earth is also suggested, where informations about the dominant wind directions would be easier to obtain and to study. The developed methodology is focused on the application of these techniques to three martian dune fields, selected for their characteristics, Aonia Terra's dune field, Gale crater, and the Herschel crater. Further work to improve the method and its capabilities is also suggested.

This thesis was written with the support of FCT, under the COMPETE program, project PTDC/CTE-SPA/117786/2010.

Resumo

O objetivo desta tese de mestrado é o desenvolvimento e teste de um programa em Python, utilizando o software ArcGIS, de forma a conseguir uma caracterização automática, precisa e eficaz dos campos de dunas marcianas. A metodologia desenvolvida utiliza redes neuronais de forma a conseguir prever as direções predominantes de transporte dos sedimentos a partir dos padrões das dunas em diversas regiões do planeta. São, para tal, utilizados dados obtidos pelo MRAMS, um conjunto de simulações das direções de transporte em regiões específicas do planeta Marte, e os dados da MGDDD, a base de dados global digital das dunas marcianas. Estas duas fontes de dados são comentadas e comparadas com aquilo que o novo método pode oferecer.

Além disso, o programa pretende ser também capaz de diferenciar entre diferentes tipos de dunas, utilizando métodos de *clustering* de forma a distinguir entre áreas onde existem predominâncias de diferentes padrões. O teste e a validação dos métodos desenvolvidos são feitos tendo em vista a aplicação destes algoritmos a novos campos de dunas, ou mesmo à totalidade do planeta.

Conclui-se que o programa não deve ser aplicado cegamente a uma área, sendo necessário um estudo prévio manual das direções gerais de transporte de sedimentos indicadas pelas marcas no terreno. É também sugerida a aplicação de técnicas idênticas a áreas dunares do planeta Terra, onde as informações de direções predominantes de transporte seriam mais fáceis de estudar/obter. O trabalho desenvolvido foca-se na aplicação destas técnicas a 3 campos de dunas marcianas selecionados para o efeito, o campo de dunas de Aonia Terra, a cratera Gale e a cratera Herschel. São sugeridas ideias e linhas de trabalho para melhorar o método desenvolvido.

Esta dissertação foi apoiada pela FCT, no âmbito do programa COMPETE, projecto PTDC/CTE-SPA/117786/2010 .

Lista de acrónimos

HiRISE - High-Resolution Imaging Science Experiment
MSL - Mars Science Laboratory
MGDDD - Mars Global Dune Database
MRAMS - Mars Regional Atmospheric Modeling System
THEMIS TIR - Thermal Emission Imaging System thermal infrared
NASA - National Aeronautics and Space Association
SRTM - Shuttle Radar Topography Mission
ESA - European Space Agency
MOLA - Mars Orbiter Laser Altimeter
GCM - General Circulation Model
MRO - Mars Reconnaissance Orbiter
CTX - Context Camera
ISIS - Integrated Software for Imagers and Spectrometers
DPC - Desvio Padrão Circular

Conteúdo

1	Objetivos e Estrutura Geral da Tese	1
2	Introdução e Estado da Arte	3
2.1	Marte - uma vista geral	3
2.1.1	Dados gerais do planeta	4
2.1.2	Características orbitais de Marte	5
2.1.3	Topografia Marciana	6
2.1.4	História Geológica do planeta Marte	8
	2.1.4.1 Período Pré-Noaquiano e Noaquiano	8
	2.1.4.2 Período Hesperiano	9
	2.1.4.3 Período Amazoniano	9
2.2	Dunas e processos eólicos em Marte	10
2.2.1	Efeitos do vento na superfície	10
2.2.2	Fenómenos eólicos: tempestades de pó e redemoinhos de poeira	11
2.2.3	Geomorfologia - Tipos de dunas e sua classificação	14
2.2.4	Análise visual do terreno e direções de transporte de se- dimentos	17
2.3	Dados científicos atuais - fontes em análise	20
2.3.1	Dunas - análise global (MGDDD)	20
2.3.2	Dados MRAMS	22
2.4	Áreas de estudo em Marte	25
2.4.1	Cratera Gale	25
2.4.2	Campo de dunas de Aonia	29
2.4.3	Cratera Herschel	31
3	Métodos e dados utilizados	35
3.1	Imagem de Marte - exemplo cratera Herschel	35
3.1.1	Construção do mosaico	36
3.1.2	Delimitação da área dunar	36
3.1.3	Cartografia e caracterização dos lineamentos dunares	37
3.1.4	Eliminação falsos positivos	39
3.1.5	Parâmetros de caracterização das dunas	40
3.1.6	Cálculos da estatística regional das dunas	40

3.2	Tratamento de dados e simulações dos ventos	44
3.2.1	<i>Clustering</i> como método de classificação das dunas	45
3.2.2	Redes neuronais	46
3.2.2.1	Dados do MRAMS e as Redes Neuronais	46
3.2.2.2	Arquiteturas de rede neuronal e campos de dunas	49
4	Apresentação e análise de resultados	51
4.1	<i>Clustering</i> - resultados obtidos	51
4.2	Redes Neuronais - Resultados obtidos	56
4.2.1	Arquitetura de redes - resultados	58
4.2.2	Comparação dos resultados obtidos com o <i>input</i> (MRAMS)	63
4.2.3	Comparação dos resultados obtidos com o estado da arte da análise de dunas - MGDDD	63
5	Conclusões e trabalhos futuros	71
5.1	<i>Clustering</i> - possibilidades e conclusões	71
5.2	Redes Neuronais - possibilidades e conclusões	71
5.2.1	MRAMS e as Redes Neuronais	71
5.2.2	MGDDD e as Redes Neuronais	72
5.3	Trabalhos futuros	73
6	Apêndices e Referências	75
6.1	Algoritmos e métodos de análise de dados utilizados - <i>Clustering</i> e Redes Neuronais	75
6.1.1	Redes Neuronais - o que são e como funcionam	75
6.1.1.1	Método de <i>Back-propagation</i>	77
6.1.2	<i>Clustering</i> - O que é e como funciona	79
6.1.2.1	Algoritmo <i>K-Means</i>	82
6.2	Programa criado - Esquema geral e estrutura	83
6.3	Referências e citações	85

Lista de Figuras

2.1	Imagem do planeta Marte, com o Valles Marineris no centro. Mosaico construído a partir dos dados da sonda orbital que acompanhou a Viking 1 na sua missão, em 1975.	4
2.2	Topografia marciana, a partir dos dados MOLA, com os campos de dunas e outras áreas do planeta mencionadas na tese assinadas. A cor indica a elevação dos locais, com a legenda no canto inferior direito.	7
2.3	Imagem de um <i>dust devil</i> em Marte, capturada pela Mars Global Surveyor em 2005. A marca escura na diagonal é a marca deixada por ele, e a mancha cinzenta na parte superior central da imagem é o próprio <i>dust devil</i>	12
2.4	Comparação do planeta Marte com e sem uma tempestade de pó global ativa. A imagem mostra como todo o planeta fica obscurecido, e existem até estudos sobre o potencial que estas tempestades têm para afetar instrumentos científicos no planeta Kinch et al. (2007).	13
2.5	Guia para os diversos tipos de dunas existentes, adaptado de McKee (1979). Do lado esquerdo temos os três tipos de dunas simples e mais comuns, que surgem quando existe apenas uma direção dominante do vento na região, enquanto que do lado direito se encontram dunas mais complexas e que representam diferentes direções do vento na mesma região.	14
2.6	Tipos de dunas evidenciados numa imagem do campo de dunas de Aonia. Pode ver-se claramente na parte inferior da imagem a forma das barcanas, com a face de deslizamento a indicar a direção dominante de transporte de sedimentos de Sul para Norte, e as dunas complexas que surgem no campo principal, com faces de deslizamento e dunas em todas as direções, indicando diferentes direções de transporte no local.	16

- 2.7 Direções de transporte de sedimentos na cratera Gale, deduzidas a partir da forma e da direção das faces de deslizamento das dunas barcanas/barcanóides na parte norte da cratera. As dunas, como se pode ver, não são muito uniformes, mas o conjunto do campo de dunas permite a dedução da direção preferencial do vento, traçada na imagem a verde. 18
- 2.8 Imagem das direções de transporte na cratera Gale, deduzidas a partir dos indicadores do vento na imagem e de algumas dunas na parte superior. Neste caso não existem tantas barcanas para indicar a direção preferencial de transporte, mas as dunas existentes na parte norte da imagem, assim como as marcas deixadas pelo vento na parte sul/sudeste, permitem uma dedução da direção preferencial de transporte de sedimentos, de sul/sudeste para norte/noroeste, com alguma margem de erro. 19
- 2.9 Mais um exemplo das direções dominantes de transporte traçadas pelas indicações no terreno das direções do mesmo, comparadas com os resultados do MRAMS no mesmo local. Pode ver-se na imagem dunas barcanas e outras marcas de direções do vento indicando uma direção razoavelmente constante de sul para norte, enquanto que as direções indicadas pelo MRAMS cobrem uma vasta gama de direções, entre o norte para sul e o noroeste para sudeste. Assim, torna-se evidente a necessidade de melhoramento destes dados e um novo trabalho de forma a criar uma base de dados mais precisa e exata nos campos de dunas do planeta Marte (um pequeno aparte; nem todos os dados do MRAMS discordam das evidências observacionais como este, o exemplo dado é apenas o mais evidente. Muitas direções por elas traçadas coincidem de uma maneira muito razoável com o aspeto do terreno e das dunas nele existentes). 21
- 2.10 Análise global da distribuição das dunas em Marte. Podemos ver que grande parte dos campos de dunas se situam nas zonas polares do planeta, e obedecem a uma distribuição bimodal, com máximos entre as longitudes 120°e 240°E e 270°e 330°E. Imagem retirada de Hayward et al. (2014) 23
- 2.11 Imagem da cratera Gale, comparando o mosaico criado no programa ISIS com a topografia obtida com dados do HRSC (resolução de 50 metros por pixel), com os campos de dunas assinalados a vermelho. De notar a elevação central da cratera e a depressão na sua parte norte. 26

- 2.12 Diagrama indicando as direções preferenciais de transporte de sedimento em cada zona da cratera Gale. Os diagramas de roseta indicam a direção de sotavento predominante e a segunda principal em cada um dos sub-campos de dunas estudados. BAR indica que os dados foram obtidos a partir de dunas barcanas e TRA indica que foram obtidos a partir de dunas transversas. Quando não existia a possibilidade de efetuar análises separadas dos dois tipos de duna, os diagramas apenas apresentam as direções preferenciais de transporte de sedimento. Imagem adaptada de Hobbs et al. (2010) 28
- 2.13 Imagem de mosaico representando o campo de dunas de Aonia e uma comparação com a topografia da mesma zona. A região mais escura no meio da imagem é a área onde se podem encontrar as dunas, com as áreas de interesse assinaladas a vermelho. Para além do grande campo de dunas representado, existem algumas dunas a Sudoeste, mas não foram consideradas nesta análise. Nesta zona o aspeto da topografia é pior do que o apresentado na cratera Gale porque tem uma resolução de apenas 463 metros por pixel. 30
- 2.14 Caracterização dos tipos de dunas no campo de Aonia, evidenciando as diferenças evidentes entre as dunas em cada área. De notar as barcanas isoladas a Sul, mostrando a direção de transporte Sul-Norte com a sua frente de avalanche, e as dunas complexas no centro do campo, ilustrando a presença de direções conflituosas do vento. Esta classificação manual em 3 tipos distintos de dunas será utilizada mais tarde na análise com *clusters*. 32
- 2.15 Topografia da cratera Herschel, obtida a partir dos dados do MOLA, e a imagem de mosaico da mesma, construída no programa ISIS. Os campos de dunas de interesse encontram-se assinalados a vermelho. De notar a grande área plana no centro da cratera, onde se acumulam as dunas, assim como a cratera menor no sudoeste, onde também existem vestígios de dunas. 33

- 3.1 Da esquerda para a direita, de cima para baixo: 1: Delimitação da área com dunas na cratera Herschel, close-up. As áreas interessantes (com areia e dunas) são seleccionadas para análise na etapa seguinte, enquanto que as áreas sem interesse são deixadas sem interferência, de forma a diminuir o tempo de processamento para o cálculo das possíveis linhas das dunas e dos seus parâmetros de análise. 2: Resultados iniciais da análise das linhas da cratera. Comparar resultado com imagem anterior. Estas linhas não representam ainda apenas as dunas verdadeiras, há muitas linhas que não correspondem a faces de deslizamento nem características da duna. Essas linhas serão ainda filtradas antes de termos uma imagem a partir da qual possamos começar a analisar os resultados. 3: Representação do método de eliminação de falsos positivos e linhas falsas criadas na imagem. No exemplo, as linhas marcadas a verde e a vermelho serão usadas como *input* num programa de redes neuronais utilizada para classificar as linhas como 'Duna' e 'Não_Duna'. Com uma amostra razoável (1000-3000 linhas para cada categoria) o programa vai aprender com a classificação manual e generalizá-la para o resto das linhas na imagem. A classificação final será uma compilação dos resultados obtidos pelas redes e da classificação manual, onde uma linha utilizará o resultado das redes a menos que tenha sido seleccionada manualmente para uma das categorias, e nesse caso utiliza a categoria seleccionada manualmente. 4: Resultados finais da classificação das linhas na cratera Herschel. De notar ainda a existência de alguns falsos positivos e falsos negativos, mas são em número reduzido e espalhados pela área. As linhas marcadas a vermelho serão descartadas para a análise posterior e a partir daí o foco do trabalho será na transformação destas linhas em pontos que conservem a sua informação de alguma forma. 38
- 3.2 Representação gráfica de alguns parâmetros calculados das dunas 42
- 3.3 Imagem retirada da cratera Herschel, com o intuito de mostrar o aspeto visual de alguns dos dados obtidos. Neste exemplo em concreto, podemos observar uma seção da cratera, contemplar o campo de dunas nela presente, e avaliar visualmente os dados apresentados. O campo mostrado na imagem representa o comprimento médio das linhas que representam o padrão das dunas mais próximas de cada ponto na imagem. Os pontos estão espalhados numa rede uniforme com 1000m de espaçamento entre cada ponto. Neste caso em particular, podemos constatar que, à primeira vista, o comprimento médio das dunas parece aumentar com o grau de complexidade das mesmas. Assim, podemos obter um parâmetro útil capaz de prever a complexidade dum dado campo de dunas, ou diferenciar entre dois tipos de dunas no mesmo campo, a partir do seu comprimento médio, por exemplo. 43

- 3.4 Close-up de uma zona da cratera Hershel onde as direções de transporte de sedimentos obtidas pelo MRAMS não coincidem com as deduzidas a partir da análise visual do terreno. Para corrigir este e outros problemas nas direções de transporte de sedimentos indicadas pelo MRAMS na cratera, foram introduzidas linhas do vento de forma 'manual', tomando o lugar dos dados do MRAMS quando estes discordavam das direções deduzidas pela forma das dunas no terreno. 47
- 3.5 Imagem de Herschel onde foram assinaladas algumas direções de transporte de sedimentos, a partir das faces de deslizamento das barcanas e da distribuição da areia no terreno. A direção preferencial do vento, nesta zona e na maior parte da cratera, é de Norte-Sul. 48
- 4.1 Classificação manual das dunas na zona de Aonia, tal como a exemplificada na imagem 2.14. A classificação foi feita manualmente e de acordo com o aspeto visual das dunas. O objetivo desta classificação rude dos tipos de dunas foi uma avaliação mais objetiva dos resultados da aplicação dos algoritmos de *clustering*. Esta classificação foi efetuada dividindo o campo de dunas de acordo com o tipo de duna dominante em cada área, de forma a utilizar um algoritmo de *clustering* com 3 classes. Com um algoritmo de *clustering* dividindo o campo de dunas em apenas duas classes, foram unidas as dunas complexas e barcanas, e apenas diferenciadas estas das lineares. 52
- 4.2 Resultados da aplicação do algoritmo de *clustering k-means* com duas classes no campo de dunas Aonia (*close-up*). De realçar as previsões corretas na maior parte das zonas onde existem dunas lineares, assim como a interpretação correta das zonas onde as barcanas são predominantes. Apenas se distinguem alguns erros na zona NW e SE da imagem, onde dunas claramente barcanas se encontram assinaladas como se de dunas lineares se tratassem. 53
- 4.3 Resultados da aplicação do algoritmo de *clustering k-means* com três classes no campo de dunas Aonia (*close-up*). As dunas lineares, tal como no exemplo acima, são corretamente identificadas na maior parte dos casos. Existe, contudo, uma confusão significativa entre as dunas complexas e as barcanas/barcanóides. Esse foi também um dos motivos porque se deu preferência à classificação das dunas com apenas duas classes, aglomerando esses dois tipos de dunas na mesma classe, em vez de os tentar diferenciar. 54

- 4.4 De cima para baixo: 1: Matriz de confusão depois de aplicado o algoritmo *K-means* com 2 classes ao campo de dunas de Aonia. De realçar a capacidade muito elevada de identificar corretamente as dunas complexas/barcanas, com uma precisão de 95%, e a capacidade inferior, mas positiva, de identificar dunas lineares (56%, aproximadamente). Estes resultados estão longe do perfeito, mas demonstram a possibilidade da utilização de métodos de *clustering* como uma ferramenta para classificar as dunas consoante o seu tipo; 2: Matriz de confusão depois de aplicado o algoritmo *K-means* com 3 classes ao campo de dunas de Aonia. De realçar a capacidade muito fraca para identificar corretamente dunas barcanas, e que a precisão nos outros dois casos mal ultrapassa os 60%. Uma vez que a mistura de resultados se deu principalmente entre as dunas barcanas e complexas (no caso das barcanas apenas 36% das dunas foram identificadas corretamente, percentagem similar a uma escolha ao acaso), esses foram os tipos de dunas aglomerados quando se correu o algoritmo com apenas duas classes. 55
- 4.5 Exemplo dum área da cratera Herschel onde foram aplicadas as redes neuronais de forma a reproduzir os dados das direções de transporte de sedimentos (compilação dos dados do MRAMS e das direções de transporte de sedimentos traçadas manualmente). De realçar a concordância geral nas direções do MRAMS e das redes neuronais na área assinalada. 57
- 4.6 Resultados das redes neuronais na cratera Gale. Mais uma vez, existe, à primeira vista, uma grande concordância entre os resultados das redes neuronais e as direções esperadas a partir dos dados do MRAMS e as direções de transporte de sedimentos traçadas manualmente. Estes resultados foram obtidos com uma rede neuronal com 3 camadas, com número de neurónios em cada uma igual a $t-2t-3$, respetivamente. 58
- 4.7 Representação dos resultados na cratera Gale da aplicação de redes neuronais para calcular o DPC (Desvio Padrão Circular) nos campos de dunas, a partir de dados do MRAMS. Os valores apresentados estão em graus e dão-nos indicações da variabilidade ao longo do ano das direções dos ventos na superfície. De salientar o efeito visual e grande semelhança nos padrões de cores entre os dados de entrada e de saída das redes neuronais. 59

- 4.8 Comparação inicial entre duas arquiteturas de rede neuronal na cratera Gale. Estão disponíveis um histograma polar das direções calculadas de transporte de sedimentos (lado direito) e uma representação dos valores obtidos pelas duas redes neuronais para o desvio das direções de transporte de sedimentos, comparando esses resultados com os valores do MRAMS (lado esquerdo). Neste caso, pode notar-se uma tendência linear, se bem que com grande variação, no gráfico que mostra os resultados do desvio das direções de transporte de sedimentos (valores de $r^2 = 0.425$ para rede com 4 camadas e $r^2 = 419$ para a rede com 3 camadas). No caso do gráfico polar, observa-se uma distribuição de valores do MRAMS com duas modas proeminentes, com direções aproximadas de 160° e 205° (coordenadas geográficas - Norte/Sul) e uma moda menos intensa, com valores entre 350° e 10° (Sul/Norte). Visualmente, a rede com 4 camadas não conseguiu reproduzir essas modas com as direções Sul/Norte, ao passo que a rede com 3 camadas apresentou alguns valores nesse intervalo. Dessa forma, e à primeira vista, a rede com 3 camadas será melhor a reproduzir os valores do que a rede com 4 camadas, que ficou 'viciada' na direção Norte/Sul. 61
- 4.9 Comparação dos resultados para duas redes neuronais. De notar o número muito superior de pontos visível na nuvem de pontos representativa desta cratera (a área é muito superior, e existem 3x mais pontos disponíveis para a análise). Nesta imagem podemos comparar duas redes neuronais com os mesmos parâmetros, mas treinadas em áreas diferentes. A rede a vermelho foi treinada com pontos tanto da cratera Gale como de Herschel, enquanto que a rede ilustrada a verde foi treinada apenas com pontos da cratera Herschel. O histograma a azul representando os dados iniciais de transporte de sedimentos utilizados como *output* esperado para a rede mostram que nesta cratera temos uma direção preferencial de transporte de sedimentos no sentido Norte/Sul ($\sim 180^\circ$), e uma direção secundária no sentido Este/Oeste ($\sim 250^\circ$). Podemos ver que o histograma é ligeiramente mais fiel aos dados do MRAMS no caso da rede treinada com as duas áreas (apesar de nenhum deles reproduzir fielmente a direção Este/Oeste), enquanto que os pontos do desvio da direção de transporte de sedimentos parecem melhor no caso da rede treinada apenas em Herschel. Com a tabela abaixo podem confirmar-se estas impressões. Conclui-se que a rede treinada nas duas áreas poderá ser mais fiel em termos de direção de transporte de sedimentos, mas não prevê com tanta precisão o desvio dessas direções. 62

- 4.10 Ilustração do campo de dunas de Aonia, com os vetores assinalados pertencendo a uma rede neuronal treinada nos campos de Gale e Herschel. Estão visíveis três áreas do campo de dunas. Como se pode ver, principalmente na imagem 1, este tipo de dados não corresponde exatamente à direção de transporte que se pode inferir a partir das barcanas (~Sul-Norte, enquanto que as direções calculadas apresentam Norte-Sul). 64
- 4.11 Dados disponíveis da MGDDD sobre a cratera Gale, comparar com a figura 2.11. De notar a fraca resolução e a indicação de apenas um vetor predominante em cada campo de dunas, bem como a ausência de dunas marcadas no lado Nordeste da cratera. É de notar o facto que a MGDDD possui alguma informação sobre os faces de deslizamento individuais das dunas que não está assinalada na imagem, devido ao seu tamanho reduzido, que ajudam a caracterizar os campos de dunas. Os vetores assinalados a verde traçam a direção desde o centro da cratera até ao centro do campo de dunas, e pretendem ser uma representação possível da direção de transporte de sedimentos dominante nesse campo de dunas. 66
- 4.12 Comparação de alguns dos resultados obtidos com as redes neuronais (neste caso, 3 camadas com 2t neurónios, ver seção 3.2.2.2 para mais informação) para campos de duna na cratera Gale com os dados da mesma área na MGDDD. O mosaico utilizado nas imagens foi criado especificamente para este trabalho, a resolução da imagem disponibilizada pela MGDDD era bastante pior (a exemplo da imagem 4.11), pelo que foi utilizado o nosso mosaico para maior clareza na apresentação dos dados. 67
- 4.13 Histograma polar comparando as direções de transporte de sedimentos previstas pelas redes neuronais em diferentes campos de dunas, as direções do MRAMS nesses mesmos pontos e as direções das frentes de avalanche apresentadas pela MGDDD (no caso de Aonia não existem dados do MRAMS). 69
- 6.1 Um esquema simples de uma rede neuronal '*Multilayer perceptron*', mostrando os neurónios de *input*, escondidos e de *output*. O número dos neurónios é uma das variáveis a modificar na arquitetura da rede. Neste caso, temos 3 camadas, com 4 neurónios de *input*, 5 escondidos e 1 de *output*, respetivamente. Imagem retirada de <http://www.codeproject.com/Articles/175777/Financial-predictor-via-neural-network>. 77
- 6.2 Imagem representativa de um *K-means clustering* num conjunto de pontos. Cada cor representa um grupo dentro do conjunto de pontos. Os '+' no centro de cada grupo indicam o ponto representativo de cada grupo individual. Imagem retirada de <http://www.mathworks.com/matlabcentral/fileexchange/screenshots/2240/original.jpg>. 80

6.3 Esquema programa Python 84

Lista de Tabelas

2.1	Tabela com os tipos de dunas, extensão, e tamanho dos grãos de areia em cada uma das seções da cratera Gale. Os tamanhos dos grãos foram calculados utilizando medidas de albedo das áreas dunares e representam uma estimativa grosseira dos tamanhos das partículas. Os campos de dunas encontram-se identificados na imagem 2.12. As direções preferenciais de transporte de sedimento estão indicadas a partir do Norte e em graus. Bar e Tra indicam as dunas barcanas e transversas, respetivamente.	27
3.1	Parâmetros calculados das linhas da imagem.	41
4.1	Resumo dos resultados obtidos pelo algoritmo <i>Kmeans</i> como método para diferenciação dos tipos de dunas, com duas e três classes disponíveis. De realçar o baixo valor de \hat{k} obtido para 3 classes, por comparação com duas classes, e os valores relativamente elevados da precisão e do <i>recall</i> nos dois casos.	51
4.2	Tabela contendo uma compilação dos resultados de várias redes neuronais, treinadas e testadas nos dois campos de dunas, Gale e Herschel. Os valores angulares estão todos medidos em graus. Nos campos dos resultados, 'Diferença' indica a diferença média entre os desvios padrões do MRAMS e da rede, para todos os pontos da área dunar, 'Med Dif Ang' indica a média das diferença entre as direções de transporte de sedimentos do MRAMS e da rede neuronal, para todos os pontos da área dunar, e 'Ang Tot Dif' indica a diferença entre o vetor médio da direção de transporte de sedimentos para toda a área duna do MRAMS e o medido pela rede neuronal. 'G+H' significa um conjunto de dados incluindo as dunas tanto de Gale como de Herschel.	60

- 4.3 Comparação direta das frentes de avalanche apresentadas pela MGDDD, as direções azimutais entre os centros das crateras e os campos de dunas, e os resultados obtidos pelas redes neuronais nas mesmas áreas. O valor apresentado v é uma medida da desvio padrão circular, apresentada no livro Jupp (2000), e corresponde a $v = \{-2 \log \bar{R}\}^{1/2}$, onde \bar{R} corresponde ao comprimento resultante médio, obtido pela equação $\bar{R} = (\bar{C}^2 + \bar{S}^2)^{1/2}$, onde \bar{C} e \bar{S} são os valores médios dos senos e cossenos, respectivamente. Esse valor é calculado em radianos e posteriormente convertido para graus. Deve ser evidenciado que os valores das 'Redes' indicados nas colunas do meio correspondem a valores completos dos campos de dunas, enquanto que os das colunas da direita apenas correspondem a pontos onde foram assinaladas as frentes de avalanche na MGDDD, podendo algumas discrepâncias entre os resultados estar associadas a essa diferença nos dados. Os campos apresentados na cratera Herschel foram marcados na MGDDD e representam quatro campos de dunas, ordenados de Sul para Norte. 68

Capítulo 1

Objetivos e Estrutura Geral da Tese

Esta tese tem como motivação e principal objetivo a construção de um programa em Python, utilizando software ArcGIS, capaz de, utilizando dados relativos a dunas marcianas, prever as direções de transporte na área e de identificar corretamente o tipo de dunas dominante em cada zona. Para tal, será utilizada uma análise prévia do campo de dunas que vai centrar-se na identificação das dunas aí existentes e do cálculo de uma série de parâmetros relativos às mesmas (ver seção 3.1.6). O programa construído pretende, desta forma, surgir como um complemento e melhoramento ao estado atual do estudo de direções de transporte e de frentes de avalanche no planeta Marte. Assim, a tese começa por explicar e detalhar a história geológica do planeta (ver seção 2.1), referenciando as características gerais orbitais e da superfície do mesmo, assim como os períodos mais relevantes na sua história. Serão ainda explicados os efeitos mais relevantes do vento na superfície marciana (ver seção 2.2), dando particular destaque à formação de dunas. O estado da arte, com um foco principal nos aspetos do MRAMS e do MGDDD (ver seção 2.3), será posto em perspetiva e análise, e serão exploradas as características e motivos de escolha dos três campos de dunas utilizados nesta tese (ver seção 2.4). passando depois para uma explicação dos métodos utilizados na construção das imagens (ver seção 3.1.1) e na obtenção de resultados. O estado da arte disponível neste ramo do estudo dos campos de dunas será também explorado, assim como os pontos em que pode ser feita uma contribuição importante e relevante. Os resultados obtidos pelo programa serão então analisados e comparados com os estudos anteriores (ver capítulo 4), assinalando pontos fortes e fracos de cada um. Finalmente, serão explicitados possíveis trabalhos futuros e avanços no tema (ver seção 5.3). Será ainda apresentado um esquema geral dos programas desenvolvidos (ver seção 6.2) de forma a ilustrar o trabalho criado e produzido na construção desta tese de Mestrado, e detalhando qual foi feito por mim e qual foi aproveitado de dados já existentes/inicialmente do Dr. David A. Vaz.

O trabalho realizado pretende, desta forma, mostrar uma perspectiva e tipo de análise diferentes dos tradicionais, explorando novas técnicas de análise de dados e reconhecimento de padrões (clustering e redes neuronais) para obter resultados em relação aos tipos de dunas, possibilitando ainda análises futuras dos dados e trabalhos de desenvolvimento na área. Espera-se que o futuro venha a trazer novos avanços nesta área, que mantenham este trabalho relevante e útil para a posteridade.

Capítulo 2

Introdução e Estado da Arte

Marte é um dos planetas do sistema solar mais interessantes do ponto de vista científico, sendo um dos poucos locais no sistema solar onde ainda se admite a possibilidade de encontrar indícios de vida, atualmente, na sub-superfície ou no registo fóssil. Nos últimos anos tem havido muito interesse científico no planeta marciano, com as várias missões da NASA, ESA, Rússia e China a adquirirem dados que nos permitem estudar, com cada vez mais detalhe, a sua superfície, história e interior. Há diversos motivos de interesse para este tipo de missões a Marte, entre elas o estudo dos processos eólicos no planeta. Esta tese irá começar com um resumo do conhecimento científico atual sobre Marte, para depois se focar na análise das dunas, bem como dos ventos que lhes dão forma e as modelam e as suas direções de transporte. Depois, será explicado o trabalho realizado, incluindo os métodos utilizados para obter os mosaicos do terreno e as bases de dados digitais das formas das dunas, bem como a forma como se pretende chegar aos objetivos propostos, que são uma simulação, através de redes neuronais, das direções de transporte e suas distribuições nos campos dunares marcianos, utilizando para isso uma análise computadorizada das formas das dunas que os constituem. Todos os valores angulares apresentados nesta tese seguem a notação geográfica, em que 0° corresponde ao norte geográfico, e em que os valores aumentam no sentido dos ponteiros do relógio.

2.1 Marte - uma vista geral

Começamos por apresentar as características gerais do planeta, assim como os motivos de interesse para o seu estudo.

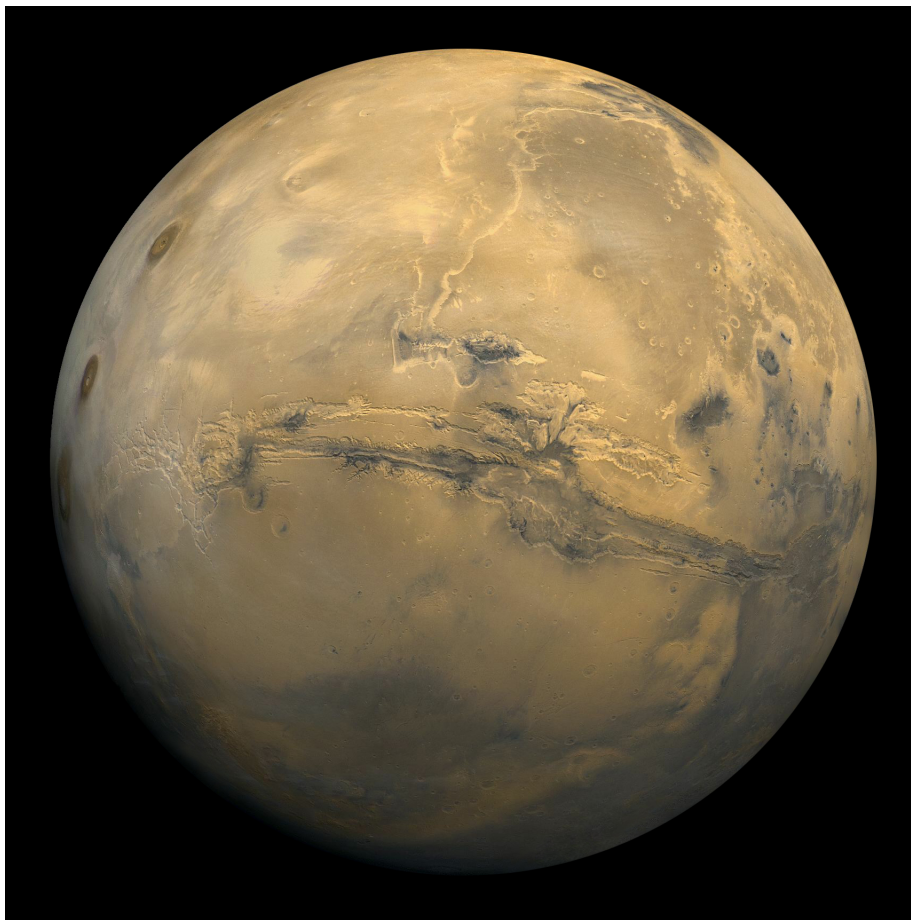


Figura 2.1: Imagem do planeta Marte, com o Valles Marineris no centro. Mosaico construído a partir dos dados da sonda orbital que acompanhou a Viking 1 na sua missão, em 1975.

2.1.1 Dados gerais do planeta

Marte é o quarto planeta do sistema solar a contar do Sol. É um planeta rochoso, tal como Mercúrio, Vénus e a Terra. O período de rotação do planeta é de cerca de 24h37min22s, colocando-o bastante próximo do terrestre. Tem um raio médio de 3390 km, cerca de 53% do terrestre, o que resulta num volume total de aproximadamente 14.8% do da Terra. Marte tem, portanto, uma massa de apenas $6.4 \times 10^{23} \text{kg}$ (10.7% da terrestre), uma vez que apresenta uma densidade ligeiramente inferior à da Terra ($\rho_{\oplus} = 5.52 \text{g} \cdot \text{cm}^{-3}$; $\rho_{\text{M}} = 3.933 \text{g} \cdot \text{cm}^{-3}$). Marte apresenta temperaturas na superfície que variam entre 130 e 310 Kelvin (-143

a 35 graus centígrados), com uma temperatura média de 210K (-63°C). Estas temperaturas apresentam grandes oscilações diárias e anuais, devido à reduzida atmosfera do planeta. O planeta tem uma atmosfera com apenas 2% da densidade da terrestre, constituída, principalmente, por dióxido de carbono (95.3%), azoto (2.7%) e Árgon (1.6%) (Carr (2006)). Devido ao seu pequeno volume (relativamente à Terra e a Vénus), o núcleo de Marte já há muito que arrefeceu, e a visão atual é a de que o planeta está moribundo geologicamente, sem movimento tectónico ou geológico dignos desse nome, ao contrário da Terra ou até de Vénus, planetas com tamanhos muito superiores, e que, ainda hoje, apresentam registos de atividade vulcânica a partir do seu interior. Contudo, alguns indícios recentes têm colocado dúvidas nesta afirmação, havendo resultados que indicam que Marte não está tão morto como seria de pensar. Existem, por exemplo, trabalhos que indicam vestígios de atividade recente, incluindo possíveis movimentos tectónicos (Spagnuolo et al. (2011)). Assim sendo, permanecem ainda muitas dúvidas sobre qual o grau de atividade geológica atualmente no planeta.

2.1.2 Características orbitais de Marte

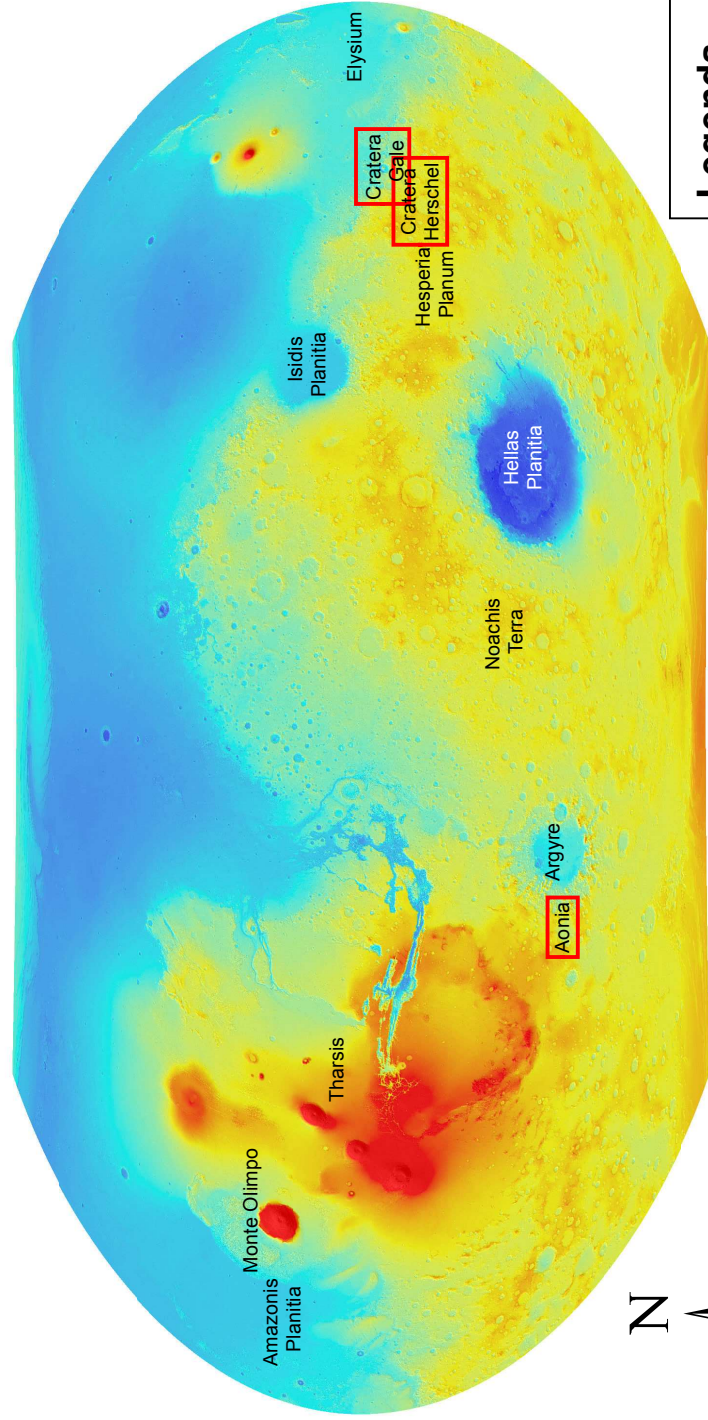
Sobre as características da órbita marciana há ainda alguns fatores a considerar. O planeta demora 669 dias marcianos a percorrer a sua órbita, que apresenta uma excentricidade de 0.093, valor que resulta numa órbita um pouco excêntrica, mas ainda razoavelmente circular. O planeta está a uma distância de apenas 1.38AU do Sol no periélio, comparando com a distância máxima de 1.67AU no afélio (Carr (2006)). Devido a esta excentricidade, e devido à diminuição da luz solar com o quadrado da distância ao mesmo, Marte receberá mais 45% de luz solar no periélio do que no afélio. O periélio ocorre no fim da primavera/início do verão no hemisfério sul marciano, resultando em estações mais quentes neste hemisfério do que no norte. Além disso, devido à segunda lei de Kepler, a velocidade do planeta na sua órbita será superior no periélio, resultando em estações mais curtas. Marte tem, ainda, valores de precessão mais elevados que sentidos na Terra, devido à ausência de uma lua de tamanho comparável com a da Terra. Desta forma, simulações da órbita marciana indicam que o planeta tem uma propensão elevada para grandes oscilações nos valores de inclinação (valor atual de cerca de 25°, próximo dos valores terrestres de 23.7°). Assim, a órbita marciana será caracterizada por oscilações de 100 000 anos (Ward (00/1992)), governadas por diferentes eixos de rotação e taxas de precessão do plano da órbita durante as quais a obliquidade pode mudar 20°. Além disso, simulações numéricas das equações de movimento da órbita marciana indicam que o planeta pode ter passado por um evento de ressonância há 5 milhões de anos atrás, atingindo obliquidades até 45°(Ward (00/1992)). Por isso, comprova-se que a órbita de Marte é mais instável que a terrestre e mais propensa a extremos. Todos estes estudos sugerem um forte controle externo no clima marciano ao longo da sua história geológica. Sendo que as estruturas sedimentares eólicas podem registar indícios climáticos a diversas escalas temporais (nomeadamente as direções de ventos predominantes na superfície) o estudo deste tipo de estruturas geológicas a uma escala local e global pode ajudar na compreensão das

mudanças climáticas que ocorreram no passado.

2.1.3 Topografia Marciana

Marte apresenta uma dicotomia muito forte entre os seus dois hemisférios, norte e sul, sendo que o hemisfério norte tem uma altitude média de cerca de 5.5km inferior à do sul, apresentando, ainda, menos crateras e 'acidentes' topográficos. Pensa-se, portanto, que este hemisfério seja mais novo que o sul, e esta dicotomia está entre as mais velhas características marcianas. Além da dicotomia entre os hemisférios, Marte apresenta, ainda, a província de Tharsis, uma zona com 30 milhões de quilómetros quadrados de extensão que se eleva $\sim 10km$ acima da elevação média da superfície. Tharsis é, também, uma zona de crista antiga, com cerca de 3.5-3.8Ga (Wenzel et al. (2004)). Esta zona apresenta alguns dos maiores vulcões do sistema, solar, incluindo o Monte Olimpo, a maior montanha do sistema solar, não só em altitude mas também em extensão. Com 600km de diâmetro no seu sopé e uma altitude de 21km se medida a partir da base (27km se medida a partir dos pontos mais baixos da superfície), o Monte Olimpo ergue-se, imponente, na superfície do planeta. Para além do Monte Olimpo, existem outros vulcões nesta zona, e pelo menos um deles (Arsia Mons) esteve ativo nos últimos 40 milhões de anos (Carr (2006)). Existem algumas teorias que podem explicar esta dicotomia e a formação da zona Tharsis, apesar de ainda não haver consenso (Carr (2006)). Entre as possibilidades, apresenta-se a teoria de que a diferença entre os hemisférios se explique por um impacto de um corpo gigantesco, que nivelou todo o hemisfério norte do planeta. Simulações recentes indicam que pode ter mesmo sido este o caso, uma vez que um impacto de energias entre $\sim (3-6) \times 10^{29} J$, a velocidade baixa ($6-10 km \cdot s^{-1}$) e um ângulo de impacto oblíquo ($30-60^\circ$), a fronteira de remoção da crista resultante é similar em tamanho e elipticidade com as características observadas no hemisfério norte marciano (Marinova et al. (2008)). Por outro lado, surge a teoria de que a dicotomia é um vestígio de um enorme oceano primordial, coberto de gelo, que ocupava todo o hemisfério norte do planeta, pelo menos durante os primeiros mil milhões de anos de evolução (Clifford & Parker (2001)). Este oceano seria, contudo, principalmente uma consequência da dicotomia global, e não a sua causa. A terceira teoria afirma que as diferenças observadas à superfície têm origem no próprio interior marciano, e resultam numa dicotomia no próprio núcleo do planeta. Estes efeitos podem resultar numa subida da crista em todo o hemisfério sul do planeta e nas plumas de magma que podem explicar os vulcões observados na província de Tharsis (Wenzel et al. (2004)). Para além de ser mais elevado do que o hemisfério norte, o hemisfério sul, por seu lado, está coberto de crateras e de escarpas (Wenzel et al. (2004)), incluindo Hellas Planitia, que, com os seus 2300km de diâmetro, é a maior cratera de impacto visível no sistema solar. Outros locais de interesse no planeta são Valles Marineris, conjunto de ravinas e penhascos que se situa na fronteira entre os hemisférios. Com mais de 4000km de extensão e até 7km de profundidade, é um

Topografia de Marte



Legenda

MOLA_Marte
High : 21128
Low : -7810



Figura 2.2: Topografia marciana, a partir dos dados MOLA, com os campos de dunas e outras áreas do planeta mencionadas na tese assinaladas. A cor indica a elevação dos locais, com a legenda no canto inferior direito.

dos maiores canyons do sistema solar, apenas superado pelos vales do rifte na Terra e (apenas em comprimento) pelo Baltis Vallis, em Vénus (Carr (2006)).

Para além de todos estes acidentes geográficos interessantes e curiosos, Marte tem, também, um enorme e complexo sistema de dunas espalhado por todo o planeta, principalmente no hemisfério norte (próximo do polo) (Hayward et al. (2014)) e em crateras e outras depressões no resto do planeta. Vamos agora focar as nossas atenções nestes conjuntos de dunas, nas suas características, e no que nos podem ensinar sobre o presente e o passado do planeta.

2.1.4 História Geológica do planeta Marte

A história geológica do planeta Marte divide-se em três períodos: o Noaquiano, o Hesperiano, e o Amazoniano, do mais antigo para o mais recente. Vamos distinguir entre os três, apontando as características mais importantes de cada um, e o seu significado na situação atual do planeta.

2.1.4.1 Período Pré-Noaquiano e Noaquiano

O período Pré-Noaquiano contém o início e a formação do planeta Marte, representando o intervalo desde a acreção e diferenciação do planeta (há cerca de 4.5Ga, juntamente com o resto do sistema solar) até à formação da cratera de impacto de Hellas (entre 4.1 e 3.8 Ga). A maior parte do registo geológico deste período foi, entretanto, apagado, devido à erosão subsequente e às grandes taxas de impacto de meteoritos. Pensa-se que a dicotomia crustal do planeta se tenha formado já nesta altura, assim como as crateras de impacto de Argyre e de Isidis.

Depois deste período, surgiu o período Noaquiano (nome atribuído devido à zona de Noaquis Terra), o mais interessante dos períodos da história do planeta. Este período é caracterizado pela grande taxa de impacto de meteoritos no planeta, assim como a muito provável presença de água abundante na superfície. Apesar de não existirem limites definidos para este período da história do planeta, o intervalo aceite pela comunidade científica está entre 4.1 e 3.7Ga, durante o Intenso Bombardeamento Tardio (época durante a qual uma quantidade superior de meteoritos e asteroides colidiram com os planetas do sistema solar interno). Estes impactos geraram muitas das crateras visíveis em Marte e na Lua, e este período será o equivalente aos éons Hádico e Arcaico da Terra, onde as primeiras formas de vida começaram a surgir.

Devido a estes fatores, as zonas e terrenos em Marte onde ainda existem vestígios deste período são de interesse superior em termos de buscas por vestígios fósseis da vida em Marte. Existem, também, indícios que apontam para um clima mais quente no planeta durante este período (Carr (2006)), sendo que muito do terreno marciano datado desta altura se apresenta dissecado por redes de vales que se parecem muito com sistemas de rios terrestres. Esse é mais um dos indícios que aponta para a presença de água líquida de forma abundante no planeta, principalmente em grandes lagos e rios no hemisfério Sul do planeta. Além disso, a extensa quantidade de vulcanismo no planeta (principalmente na

zona de Tharsis, que se pensa ter-se formado durante este período) e a enorme diversidade de minerais e argilas disponíveis parecem fornecer as condições químicas ideais para a criação de vida microbial no planeta. Vestígios de vida em particular é que ainda não foram encontrados.

2.1.4.2 Período Hesperiano

O Hesperiano (nome com origem na região de Hesperia Planum, uma zona moderadamente coberta de crateras a noroeste da cratera de impacto de Hellas) corresponde a um período da história do planeta Marte caracterizado por vulcanismo intenso e, ainda, por cheias catastróficas, que traçaram enormes canais na zona de Chryse Planitia e noutros locais do planeta. Este período terá decorrido entre há 3.7 e 3.0Ga, e está associado à formação de extensas planícies de lava. Pensa-se, também, que o Monte Olimpo se terá começado a formar durante este período. Durante esta fase da história do planeta existiram, também, lagos e mares efêmeros nas terras mais baixas do hemisfério Norte do planeta.

Este período distingue-se do Noaquiiano porque o motor principal dos processos geológicos no planeta deixa de ser a atividade meteorítica e passa a ser a atividade vulcânica. O vulcanismo neste período produziu vastas planícies de basaltos e deixou grandes marcas vulcânicas (*paterae* nas terras altas do planeta). Além disso, as emissões de gases devido aos vulcões (principalmente SO_2 e H_2S) causaram alterações no clima. A água líquida à superfície tornou-se mais localizada e mais acídica, à medida que interagiu com SO_2 e H_2S para formar ácido sulfúrico. Estas alterações fizeram com que os minerais que se formaram mais abundantemente fossem sulfatos e não filosilicatos (argilas) como ocorreu durante o Noaquiiano. A atmosfera continuou, também, a ficar mais fina e o planeta arrefeceu, levando ao eventual desaparecimento da água líquida à superfície, que se foi acumulando nos polos e formando calotas de gelo.

No fundo, o período Hesperiano representa a transição entre o período Noaquiiano ancestral, coberto de água, rios e lagos e atividade geológica, e o planeta Marte atual, mais seco e parado no tempo (mas não completamente).

2.1.4.3 Período Amazoniano

O Amazoniano, assim denominado devido à região de Amazonis Planitia, que tem uma baixa densidade de crateras de impacto, é o período atual do planeta Marte, e dura desde há cerca de 3.0Ga. Este período é caracterizado pela baixa taxa de impacto de meteoritos no planeta, e pelas condições frias e áridas que se verificam hoje em dia em Marte. Uma vez que é o período mais recente entre aqueles que caracterizam a história do planeta, o Amazoniano é o mais bem conhecido. O facto de a erosão em Marte ser praticamente inexistente neste período permite uma reconstrução mais bem detalhada da sua história e clima, Marte foi um planeta com atividade geológica muito reduzida, com poucas marcas de eventos catastróficos capazes de alterar a superfície como em períodos mais antigos. O vulcanismo abrandou também bastante, apesar de existirem, ainda, vestígios de possível atividade vulcânica recente no planeta (Hartmann

et al. (1999)). Ainda assim, existem diversos motivos de interesse para o estudo deste período, tais como possíveis mudanças nas mecânicas orbitais do planeta, a dinâmica de glaciares nos polos, os movimentos do vento e das dunas, a variação nas propriedades da atmosfera, a dinâmica das calotas polares, e a presença e as características dos resquícios de água na superfície, todos eles temas com grande avanço na atualidade e com a promessa de mais surpresas que nos podem ajudar a compreender melhor o planeta.

2.2 Dunas e processos eólicos em Marte

Desde as missões Viking a Marte, em 1975, que existem trabalhos sobre as distribuições das dunas no planeta e sobre os ventos que lhes deram origem (Ward et al. (1985)). Desde então, e devido ao melhoramento progressivo na capacidade das missões a Marte, foi-se obtendo um mapa cada vez melhor e mais detalhado da distribuição de dunas e outras características eólicas no planeta.

2.2.1 Efeitos do vento na superfície

O vento é um dos poucos agentes de mudança da superfície que podemos ver em ação em Marte. Foram observadas no planeta tempestades de poeira, tanto locais como globais. A formação de '*dust devils*' (redemoinhos de poeira) é periodicamente observada tanto a partir da superfície como da órbita do planeta (Carr (2006)). São observadas, ano após ano, mudanças na distribuição de poeiras e material mais fino e solto no planeta, por ação do vento. Há, até, registos de movimento regular de *ripples* (estruturas sedimentares que consistem em pequenas ondulações) na superfície, deslocando-se cerca 1.7 metros em apenas 4 meses, ou seja, até 4.1 metros por ano (Silvestro et al. (2010b)). Nesse artigo, são mencionados resultados recentes que indicam que em Nili Patera existem *ripples* em migração, e que continua a existir saltação e mudanças nas faces de deslizamento e na forma das dunas hoje em dia. Para além disso, outro artigo mais recente revela atividade eólica no percurso do rover *Curiosity*, na cratera Gale (Silvestro et al. (2013)), sendo observados movimentos de 0.66m por ano marciano (0.4 por ano terrestre). Estes resultados foram obtidos utilizando dados da câmara HiRISE, uma vez que apenas estes dados atingem a resolução espacial necessária para visualizar as estruturas de pequenas dimensões ($\sim 0.25m/\text{pixel}$). Portanto, pode-se então concluir que existem dunas ativas no planeta, apesar de outros estudos mais antigos sugerirem que a eficiência do vento para remodelar a superfície seria limitada. As taxas de migração observadas nos *ripples* mostram que existe atividade no planeta ainda no presente, e só estudos posteriores podem revelar toda a extensão desta atividade, confirmando o que está realmente vivo e o que está fossilizado no panorama eólico do planeta.

O ciclo dos sedimentos que podem ser movidos pelo vento é muito diferente na Terra e em Marte. Na Terra, a maior parte dos detritos rochosos de tamanho capaz de ser movido pelo vento é areia de quartzo. Esta areia é criada, principalmente, devido ao enfraquecimento e erosão de rocha granítica a partir

de processos mecânicos e químicos, passando um curto período de tempo (em termos geológicos) participando nos processos eólicos à superfície. O seu destino será mergulhar num oceano, onde será enterrada e tornada parte do registo sedimentar. Movimentos tectónicos podem trazer o material de volta à superfície, onde o ciclo se renova continuamente. Existe, assim, um depósito de material em constante movimento, e que é reciclado a escalas de tempo geológicas. Partículas mais finas, argilas, por exemplo, são mais facilmente mobilizáveis pelo que são mais rapidamente removidas do ciclo eólico.

Em Marte, não é isto que se verifica. Os sedimentos mobilizáveis são constituídos maioritariamente por olivinas e piroxenas provenientes da alteração de rochas basálticas. Para além disso, embora possam ter ocorrido depósitos de material em oceanos, foram processos muito limitados, pelo menos a partir do Noaquiano, e a reciclagem crustal por processos tectónicos não ocorreu. Este fator implica que, uma vez que o material se encontra solto à superfície do planeta, ele vai continuar disponível para ser movimentado e erodido pelo vento indefinidamente. Desta forma, os materiais livres existentes à superfície do planeta devem ser bastante antigos, pelo menos do Noaquiano, quando as taxas de vulcanismo, impactos e erosão, processos capazes de produzir partículas, eram muito mais elevados do que hoje em dia. Adições posteriores de sedimentos eólicos terão sido muito reduzidas, excetuando detritos criados em impactos de crateras na superfície marciana. Concluindo, enquanto que na Terra o material fragmentado é rapidamente removido da superfície e reincorporado no registo sedimentar, em Marte mantém-se na superfície durante milhares de milhões de anos para ser arrastado e fustigado pelo vento. Apesar de se terem formado em Marte espessos depósitos de rochas sedimentares, estes mantêm-se à superfície, onde podem ainda ser afetados pelo vento, resultando em alguns dos mais espetaculares exemplos de erosão em Marte.

2.2.2 Fenómenos eólicos: tempestades de pó e redemoinhos de poeira

Para além de movimentarem a areia e formarem as dunas, os ventos em Marte também produzem fenómenos únicos e curiosos. Entre eles, estão as tempestades de pó globais e os *dust devils* (redemoinhos de poeira), que deixam as suas marcas na superfície marciana.

Os redemoinhos de poeira (ver imagem 2.3) são comuns em diversas regiões do planeta, incluindo a zona de Amazonis Planitia, Cimmerium, Sinai e Solis (Carr (2006)). Os dados experimentais mostram que a criação de redemoinhos de poeira é, fundamentalmente, diferente da deslocação de partículas maiores, como as que constituem as dunas, pelo vento, uma vez que o pó (que forma os redemoinhos de poeira) pode ser levantado do chão por ventos unidirecionais relativamente pouco intensos. Um fator chave na criação de redemoinhos de poeira pode ser a diferença nas pressões atmosféricas sentidas pelas partes superior e inferior das partículas à medida que o redemoinho de poeira avança sobre a superfície. Isto resulta na elevação da partícula nos ventos do vórtice do redemoinho de poeira. Apesar de a maior parte dos redemoinhos de poeira

ter apenas algumas centenas de metros de altura (Ferri et al. (2003)), o maior observado até ao momento tinha 70 metros de diâmetro e 20 quilómetros de altura (NASA (2012)), e foi observado a 14 de Março de 2012, pela câmara HiRISE .

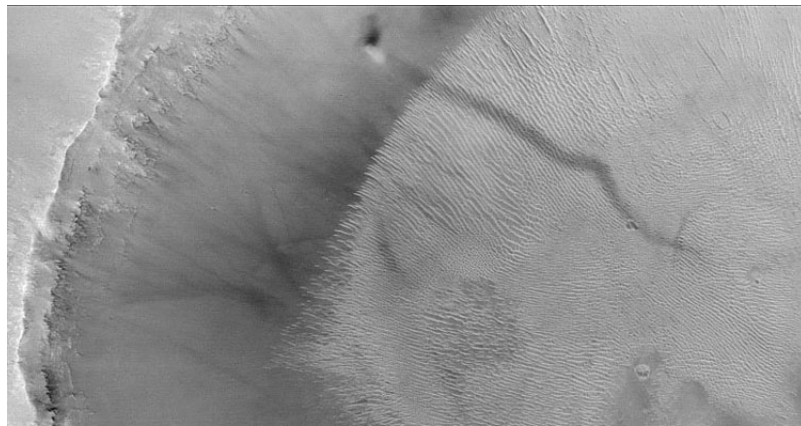


Figura 2.3: Imagem de um *dust devil* em Marte, capturada pela Mars Global Surveyor em 2005. A marca escura na diagonal é a marca deixada por ele, e a mancha cinzenta na parte superior central da imagem é o próprio *dust devil*.

Quanto às tempestades de pó, estas são observadas em Marte desde há séculos atrás, e têm uma tendência para surgirem na primavera ou verão locais. Muitas têm uma extensão regional, mas em alguns anos, como em 1971 ou em 2001, tornaram-se verdadeiramente globais. Alguns locais com mais tendência para o surgimento destas tempestades são os limites das calotas polares e a faixa entre as latitudes 10°-30°S, assim como a Hellas, as terras altas de Noaquis, Argyre, entre outras.

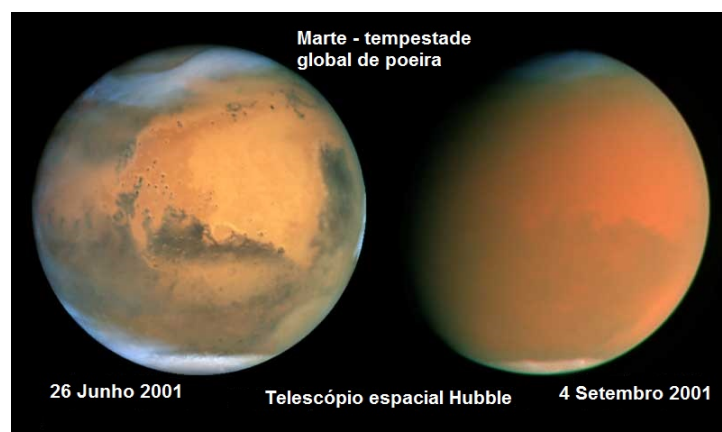


Figura 2.4: Comparação do planeta Marte com e sem uma tempestade de pó global ativa. A imagem mostra como todo o planeta fica obscurecido, e existem até estudos sobre o potencial que estas tempestades têm para afetar instrumentos científicos no planeta Kinch et al. (2007).

O significado geológico das tempestades de pó está ainda pouco claro. Os efeitos observacionais, como o escurecer da superfície pelo pó na atmosfera, são bastante intensos (ver imagem 2.4). Mas, apesar disso, a quantidade de pó envolvida nestas tempestades é bastante reduzida, sendo equivalente a uma camada superficial com apenas alguns micrómetros de espessura, mesmo para as maiores tempestades, e a maioria deste material deve ser reciclado a cada nova tempestade.

2.2.3 Geomorfologia - Tipos de dunas e sua classificação

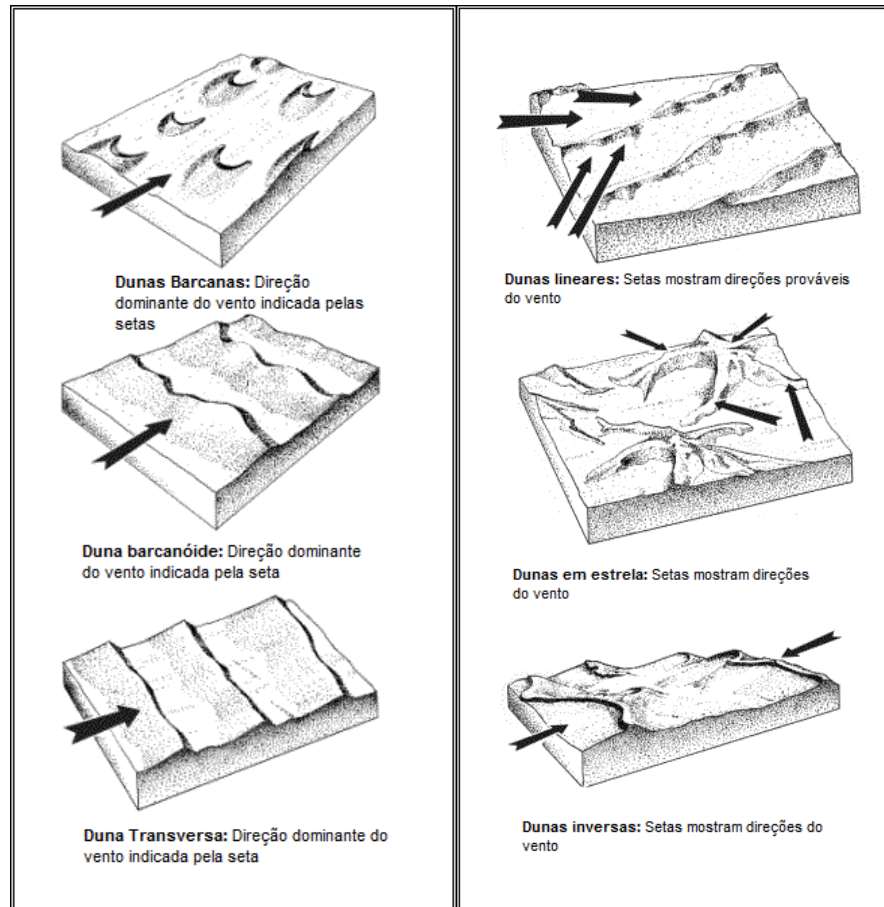


Figura 2.5: Guia para os diversos tipos de dunas existentes, adaptado de McKee (1979). Do lado esquerdo temos os três tipos de dunas simples e mais comuns, que surgem quando existe apenas uma direção dominante do vento na região, enquanto que do lado direito se encontram dunas mais complexas e que representam diferentes direções do vento na mesma região.

Existem diversos tipos de dunas, consoante a topografia da região, a disponibilidade de sedimentos, a idade do campo de dunas e as direções de transporte predominantes que se fazem sentir no local. Estes diversos tipos seguem uma classificação consoante a forma da duna e dos ângulos verificados entre a 'face de deslizamento' (face de deslizamento ou frente de avalanche) da duna e a(s) direção(ões) do vento que as forma. Começando pelas dunas barcanas, o tipo mais simples e dos mais comuns de duna, vamos ver o que cada uma destas classificações significa e como a podemos usar para deduzir a direção de transporte

dominante do vento que lhe deu origem.

As dunas barcanas apresentam uma inclinação suave (5° a 10°) na direção do vento, por onde as partículas por este arrastadas se deslocam por saltação, e uma face mais inclinada (30° a 32°) no lado oposto onde as partículas descem pelo efeito de avalanche. Além desta característica, a barcana apresenta, ainda, dois prolongamentos na direção oposta à do vento, que lhe dão o seu aspeto característico de crescente. Se a areia for muito abundante no local onde a duna se forma, as barcanas podem crescer e unir-se, formando cristas de areia numa direção transversa à do vento. Se as cristas destas dunas ondularem, referimo-nos a elas como barcanóides. Se a duna for linear, denomina-se duna transversa (ver a imagem 2.5 com exemplos). Estes três tipos de dunas, barcanas, barcanóides e transversas, são denominadas de 'Dunas Simples', uma vez que têm apenas uma face de deslizamento, ou seja, são o produto de uma direção única e constante do vento no local onde a duna se forma.

A partir da forma básica da duna barcana pode desenvolver-se uma variedade de tipos diferentes de dunas. Por exemplo, os braços exteriores da barcana podem estender-se, de forma a que um conjunto de barcanas se torne em dunas longitudinais lineares. Alternativamente, os dois braços podem unir-se e formar uma entidade única a apontar na direção do vento. Este tipo de variantes de dunas surge, tipicamente, devido a direções ou intensidades variáveis do vento no local de formação das dunas.

Para além dos dois exemplos acima mencionados, podem existir outros tipos de duna quando as direções de transporte de sedimentos não são uniformes. Se existem duas direções predominantes do vento com uma diferença inferior a 90° , podem formar-se longas dunas lineares, com frentes de avalanche nos dois lados da duna (ver imagem 2.5, lado direito, parte superior). Nesse caso, a duna bissectará o ângulo formado entre as duas direções do vento. Podem, ainda, formar-se dunas em forma de estrela, com tantos braços e faces de deslizamento quantas as direções do vento presentes no local, e dunas invertidas, que se formam quando as direções do vento são opostas. Todos estes tipos diferentes de dunas podem unir-se e formar padrões mais complexos, como podemos ver na imagem 2.6 localizada na zona de Aonia Terra.

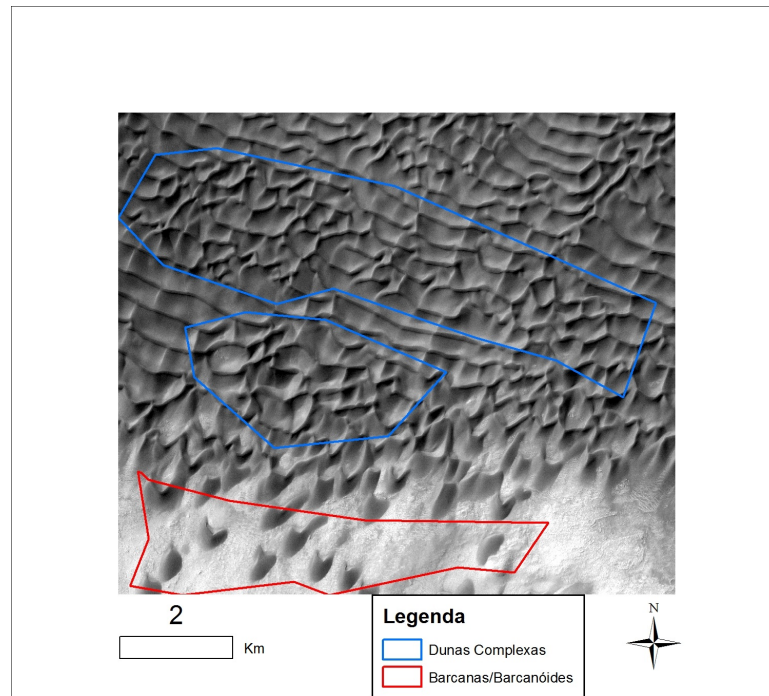


Figura 2.6: Tipos de dunas evidenciados numa imagem do campo de dunas de Aonia. Pode ver-se claramente na parte inferior da imagem a forma das barcanas, com a face de deslizamento a indicar a direção dominante de transporte de sedimentos de Sul para Norte, e as dunas complexas que surgem no campo principal, com faces de deslizamento e dunas em todas as direções, indicando diferentes direções de transporte no local.

Assim, e como podemos deduzir a partir da classificação de dunas apresentadas acima, o único tipo de dunas onde podemos ter a certeza de que só existe uma direção de transporte e qual esta é será a duna barcana/barcanóide, uma vez que a própria duna longitudinal se pode confundir com a duna transversal em termos de aspeto visual, apesar das duas terem origem em direções de transporte de sedimentos distintas. Por isso, para a análise das direções de transporte do vento numa região, o tipo de duna que queremos estudar e analisar será, principalmente, a duna barcana. Só depois de verificadas as direções das faces de deslizamento das barcanas dum campo de dunas podemos começar a deduzir as direções de transporte que dão origem às outras dunas desse campo.

Existe, ainda, um tipo de dunas típico do planeta Marte e que não se verificou na Terra, as dunas tipo 'alvo' (Carr (2006)). Estas curiosas formações dunares surgem quando existe areia acumulada no fundo de uma cratera, sendo o padrão de ventos único a esta formação essencial para a disposição resultante das dunas em anéis concêntricos. O seu interesse neste trabalho é, contudo, limitado, uma vez que os campos de dunas em estudo não apresentam este tipo de duna. Fica

o registo, no entanto, a título de curiosidade.

2.2.4 Análise visual do terreno e direções de transporte de sedimentos

Uma parte importante do trabalho desenvolvido nesta tese implica a dedução das direções do vento a partir do aspeto visual da superfície e das dunas nela presentes. Para isso precisamos de saber quais os melhores indicadores de direção dominante de transporte de sedimentos na superfície, e como os interpretar da maneira correta. Um dos melhores indicadores, como foi mencionado acima, são as dunas barcanas. A partir da observação da sua frente de avalanche, podemos deduzir a direção do vento que lhe deu origem (ver imagem 2.7). Contudo, para além dessas dunas, existem outras marcas na superfície que nos podem ajudar a assinalar as direções de transporte de sedimentos no terreno.

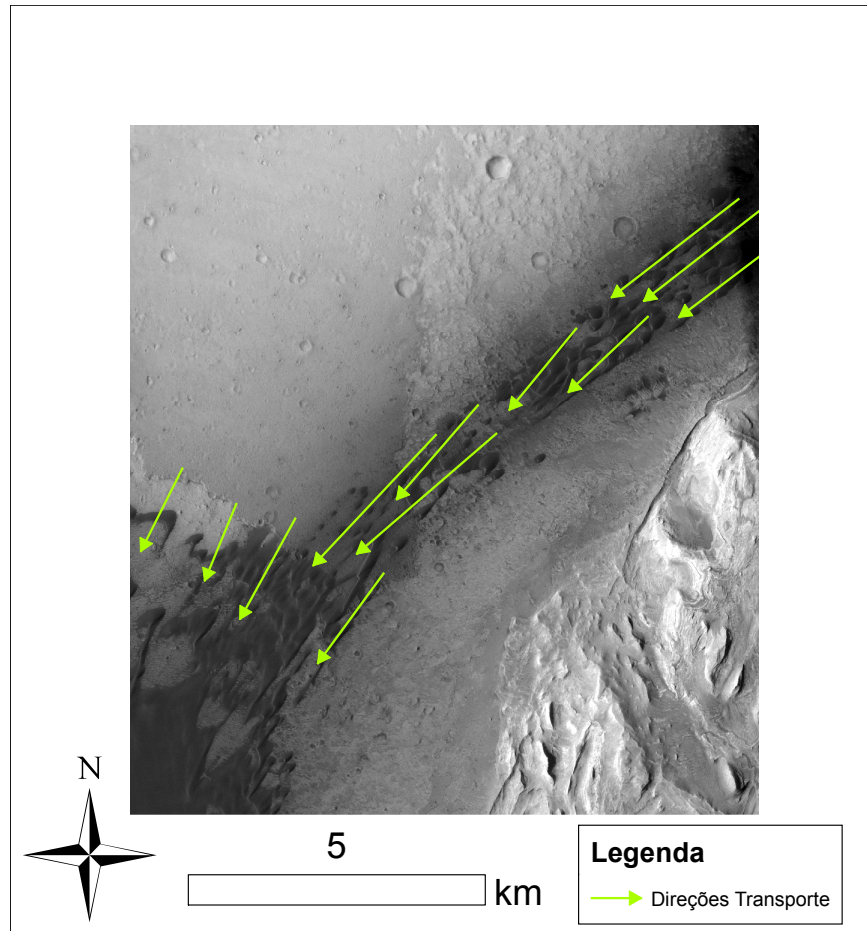


Figura 2.7: Direções de transporte de sedimentos na cratera Gale, deduzidas a partir da forma e da direção das faces de deslizamento das dunas barcanas/barcanóides na parte norte da cratera. As dunas, como se pode ver, não são muito uniformes, mas o conjunto do campo de dunas permite a dedução da direção preferencial do vento, traçada na imagem a verde.

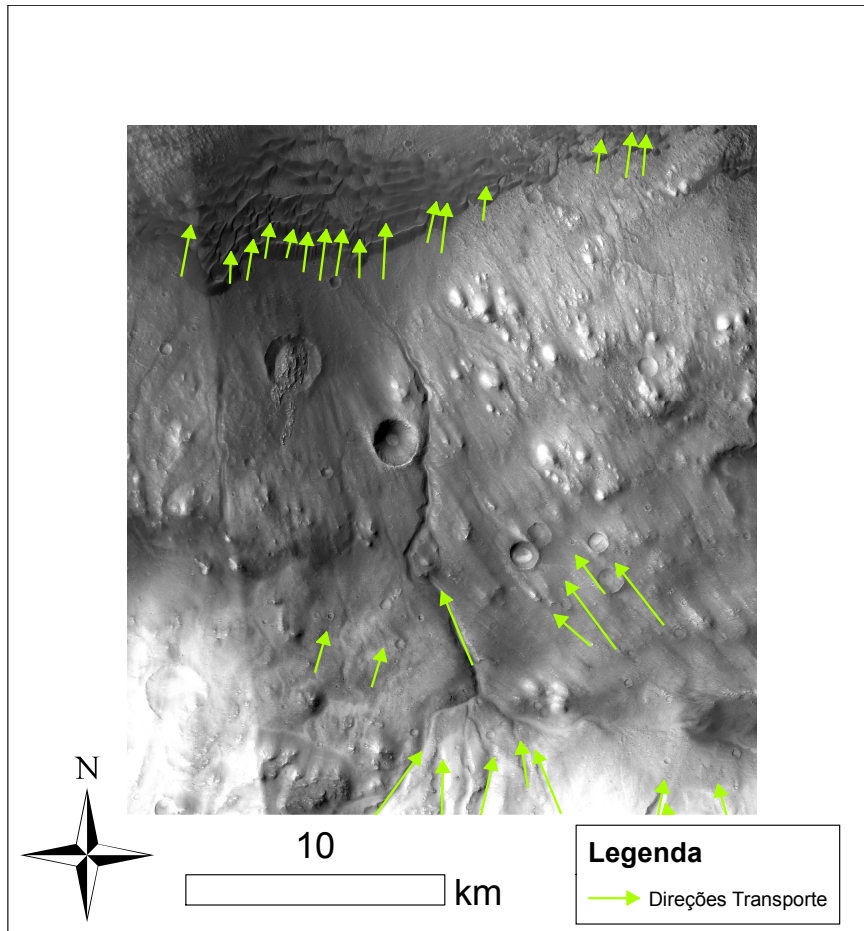


Figura 2.8: Imagem das direções de transporte na cratera Gale, deduzidas a partir dos indicadores do vento na imagem e de algumas dunas na parte superior. Neste caso não existem tantas barcanas para indicar a direção preferencial de transporte, mas as dunas existentes na parte norte da imagem, assim como as marcas deixadas pelo vento na parte sul/sudeste, permitem uma dedução da direção preferencial de transporte de sedimentos, de sul/sudeste para norte/noroeste, com alguma margem de erro.

Para além das direções das dunas barcanas, que são o nosso indicador principal e mais útil na análise das direções de transporte, podemos, ainda, analisar as *wind streaks* (marcas deixadas no terreno devido à ação do vento, como rastos de areia, manchas de pó, etc., que ajudam a caracterizar as direções de transporte numa região), como se pode observar na imagem 2.8. As marcas observadas são *streaks* deposicionais claros, e são o tipo mais comum de *streaks* encontrados na superfície marciana. Estas marcas formam-se na direção do

vento partir de depósitos de material sedimentar. A maior parte destes *streaks* mantém uma direção estável, apesar de alguns nas zonas de Tharsis e de Elysium apresentarem mudanças sazonais, que não vão ser consideradas nesta análise das direções eólicas. As suas direções apresentam boas correlações com as direções predominantes de transporte de sedimentos, e é por esse motivo que podem ser utilizados na nossa análise de marcas do vento. Formam-se, provavelmente, devido à deposição de poeiras no acalmar das tempestades de pó em Marte (ver seção 2.2.2, sobre tempestades de pó e redemoinhos de poeira). Se isso se verificar, estes *streaks* podem não passar de uma camada muito fina, com poucas dezenas de micrómetros de espessura. Podemos, ainda, encontrar alguns *streaks* mais espessos, capazes de afetar a inércia térmica da superfície, tendo espessuras na ordem de grandeza do centímetro/metro, e, neste caso, serão mais parecidos com 'caudas de vento', observadas em alguns locais de aterragem.

Estas direções preliminares de transporte vão ser muito importantes neste trabalho, uma vez que nos ajudam a confirmar ou desacreditar resultados das nossas simulações e da MGDDD (ver seção 2.3.1) e do MRAMS (ver seção 2.3.2) (por vezes os dados obtidos a partir das bases de dados não coincidem com as observações diretas da superfície, como podemos ver na imagem 2.9. Nessa imagem torna-se evidente a importância de uma análise cuidada das marcas no solo e das direções de transporte do vento que estas implicam, uma vez que nem todos os dados que temos disponíveis serão corretos.

2.3 Dados científicos atuais - fontes em análise

Nesta seção vão ser explorados os conjuntos de dados científicos disponíveis atualmente sobre os campos de dunas marcianas, e que vão ser utilizados como ponto de comparação e avaliação dos nossos próprios resultados. As duas fontes de dados utilizadas neste trabalho são a MGDDD (Base de dados global das dunas marcianas, ver seção 2.3.1) e os dados do MRAMS (Modelos regionais da atmosfera marciana, ver seção 2.3.2).

2.3.1 Dunas - análise global (MGDDD)

As dunas de areia formam-se quando existem sedimentos com a dimensão adequada disponíveis e ventos suficientemente fortes para promover o movimento das partículas por saltação. Assim, os depósitos de areia e dunas existentes num dado local são ótimos indicadores das condições de vento e dos sedimentos aí existentes. Mesmo campos de dunas completamente fossilizados, onde as dunas já não se deslocam, uma vez que não se verificam regularmente ventos com energia suficiente para o fazer, podem ser um indicador sobre as características do local no momento de formação das mesmas. Uma vez que as dunas podem existir numa grande variedade de terrenos e altitudes, um registo global do panorama dunar em Marte foi sempre considerado um passo fundamental no estudo do planeta. Infelizmente, as investigações sobre as dunas são necessariamente limitada às áreas onde existem imagens de qualidade suficiente para

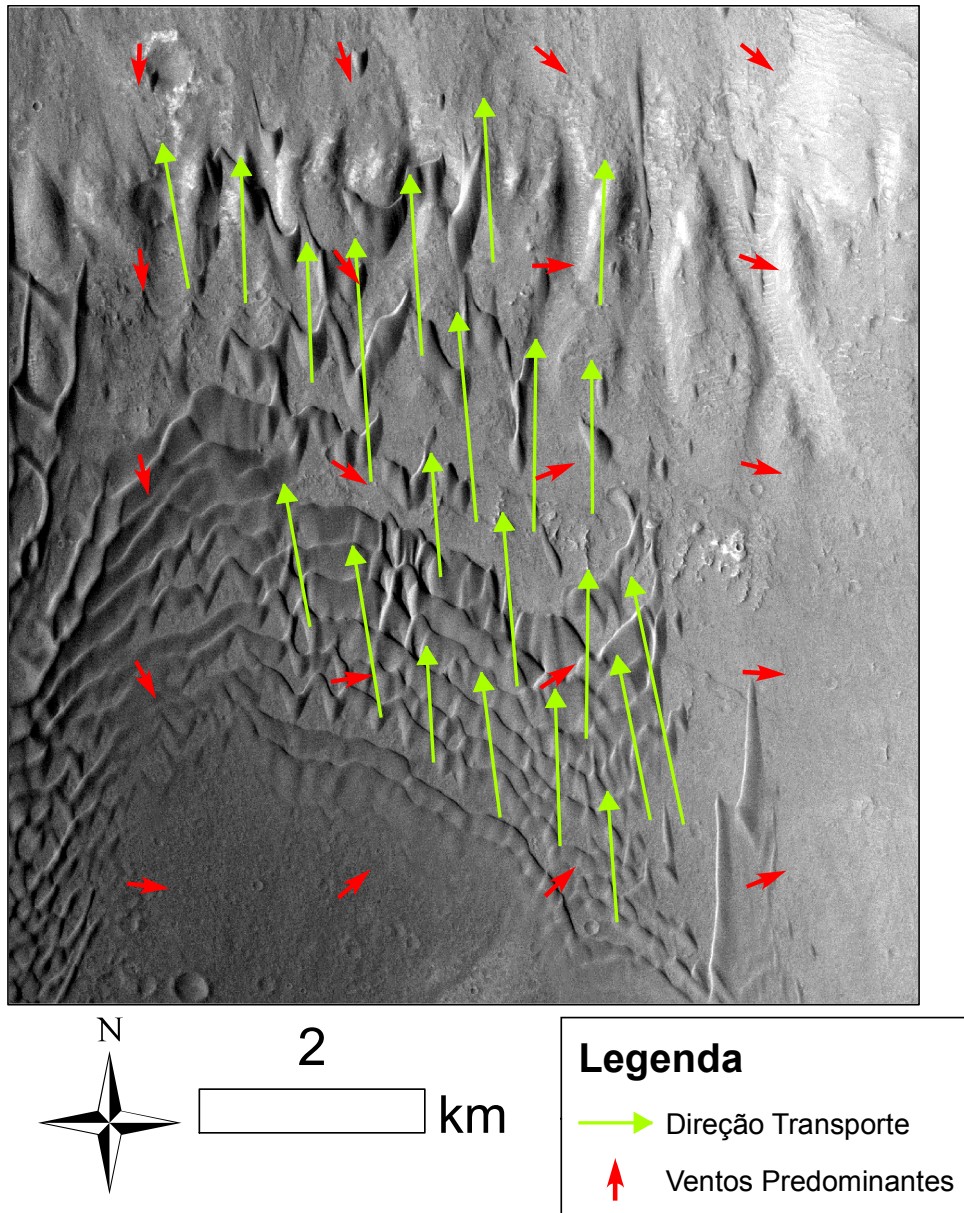


Figura 2.9: Mais um exemplo das direções dominantes de transporte traçadas pelas indicações no terreno das direções do mesmo, comparadas com os resultados do MRAMS no mesmo local. Pode ver-se na imagem dunas barcanas e outras marcas de direções do vento indicando uma direção razoavelmente constante de sul para norte, enquanto que as direções indicadas pelo MRAMS cobrem uma vasta gama de direções, entre o norte para sul e o noroeste para sudeste. Assim, torna-se evidente a necessidade de melhoramento destes dados e um novo trabalho de forma a criar uma base de dados mais precisa e exata nos campos de dunas do planeta Marte (um pequeno aparte; nem todos os dados do MRAMS discordam das evidências observacionais como este, o exemplo dado é apenas o mais evidente. Muitas direções por elas traçadas coincidem de uma maneira muito razoável com o aspeto do terreno e das dunas nele existentes).

identificar e caracterizar as dunas aí existentes. Por isso, quando as imagens da THEMIS TIR conseguiram uma cobertura completa de Marte em 2003, com uma resolução de 100m/pixel, um estudo global das dunas marcianas tornou-se finalmente possível. Vamos aqui fazer apenas um apanhado geral dos resultados obtidos recentemente neste aspeto, e sobre aquilo que podemos melhorar em relação a este tipo de mapas.

Mapas globais das dunas marcianas já estão disponíveis, como a MGDDD (Hayward et al. (2014)), que se estende por todo o planeta, desde os 90°N aos 90°S (ver imagem 2.10). As áreas observadas dividem-se entre polo Norte (dunas entre 65°e 90°N, área total $\sim 845\,000\text{km}^2$), equatorial (65°S e 65°N, área $\sim 70\,000\text{km}^2$) e polo Sul (65°e 90°S, área de $\sim 60\,000\text{km}^2$ de campos de dunas médios-grandes e uns $\sim 15\,000\text{km}^2$ de depósitos de areias e de pequenos campos de dunas). Entre as dunas do polo Norte, quase todas pertencem a grandes campos de dunas, ao contrário das equatoriais e do polo Sul, que pertencem a pequenos campos de dunas localizados nas crateras existentes. A partir das faces de deslizamento das dunas e do azimute do centróide dos campos de dunas (uma medida da localização dos campos de dunas em relação ao centro da cratera respetiva) foi possível começar a efetuar estimativas das direções e intensidades locais do vento que gera as dunas. Existem, contudo, problemas vários relacionados com esta metodologia, que serão explorados em mais detalhe aquando da comparação destes resultados com as simulações obtidas pelos nossos programas (ver seção 4.2.3). Com os novos dados que surgem diariamente da superfície marciana com resoluções melhoradas, seria necessária uma atualização completa e permanente desta base de dados, utilizando os dados disponíveis com maior resolução. Contudo, um empreendimento desse nível com o objetivo de traçar manualmente todas as dunas visíveis no planeta seria um projeto impossível, pelo que a necessidade de um algoritmo automatizado com esse fim parece evidente. A verificação da precisão destas estimativas e a sua comparação com as direções de transporte induzidas a partir das faces de deslizamento das dunas é uma questão crucial que deve ser explorada de forma a melhorar o catálogo das dunas já existente. Esse será, portanto, um dos objetivos deste trabalho, utilizando, para isso, os algoritmos criados para o efeito. Desta forma, pretendemos melhorar e automatizar os métodos de dedução das direções dominantes de transporte de sedimentos.

2.3.2 Dados MRAMS

O MRAMS consiste numa série de simulações da atmosfera de Marte, a diferentes escalas e em diferentes pontos do planeta (Michaels & Rafkin (2008)). Estas simulações usam como pano de fundo trabalhos no âmbito de GCM (Modelo Geral de Circulação), (Haberle et al. (1993)) para simular o comportamento da atmosfera a grandes escalas espaciais e temporais, traduzindo depois os resultados para um modelo de mesoescala que simula o comportamento local da atmosfera em três dimensões, assim como uma simulação de *eddies* (vórtices) para resolver os problemas gerados pela turbulência a três dimensões. Apesar de todas estas simulações apresentarem graus elevados de precisão (teoricamente)

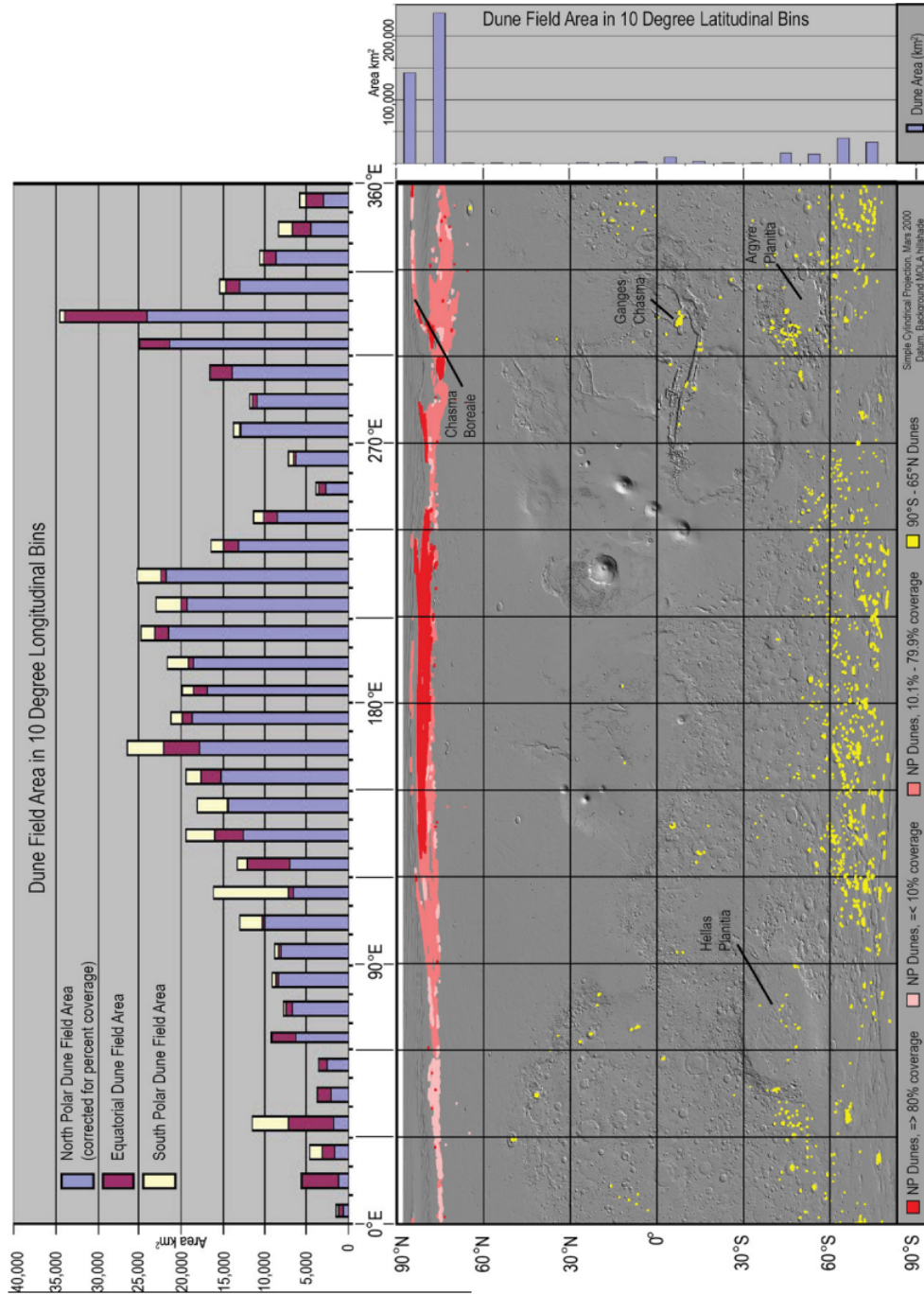


Figura 2.10: Análise global da distribuição das dunas em Marte. Podemos ver que grande parte dos campos de dunas se situam nas zonas polares do planeta, e obedecem a uma distribuição bimodal, com máximos entre as longitudes 120°E e 240°E e 270°E e 330°E. Imagem retirada de Hayward et al. (2014)

estas simulações devem ser interpretadas como modelos climáticos, e não de previsão meteorológica. Apenas se podem prever tendências gerais da atmosfera e dos ventos no local, e não tempestades esporádicas ou os ventos numa dada data, uma vez que estes comportamentos apresentam uma tendência extremamente caótica. Assim, devem interpretar-se estes resultados como os de uma simulação de uma atmosfera independente que se comporta de maneira similar à marciana.

Mais concretamente, os modelos de mesoescala (resolução média), que são aqueles que nos interessam, foram desenvolvidos em 2001 (Rafkin et al. (2001)), e foram já utilizados de forma a prever as condições locais da atmosfera nos locais de aterragem da Phoenix (Michaels & Rafkin (2008)) e de outras sondas na superfície marciana. Estas simulações estão distribuídas numa rede de até 200km de lado, e com um espaçamento entre os pontos que pode variar de acordo com o local, mas que se encontra aproximadamente entre 1 e 7km. A simulação cobre uma altitude até 70km, mas só nos vão interessar os vetores do vento na superfície do planeta, uma vez que estes é que ditam as direções de transporte dos sedimentos nos campos de dunas. Para gerar este tipo de resultados à superfície do planeta, é levada em conta toda a topografia local do terreno, e esses dados são o que nos permite obter resultados fiáveis mesmo a altitudes baixas/ junto à superfície. Em termos de intensidade, os valores não são medidos de forma direta, mas sim como um rácio entre a velocidade média do vento e o limite mínimo de stress de saltação das partículas de areia. Ou seja, os valores de magnitude correspondem ao stress aerodinâmico do vento/ stress mínimo de saltação das partículas de areia. Estas simulações produzem um resultado correspondente a um Sol (dia marciano) em cada estação do ano. O programa simula a atmosfera a cada 20 minutos (72 pontos por dia), durante um dia inteiro, em quatro dias diferentes, um em cada estação do ano marciano. e o resultado que vamos utilizar em cada ponto será uma média das direções médias de cada dia. Assim, o nosso objetivo pretendido não é o de utilizar resultados que correspondam a uma simulação do vento, mas sim das direções médias de transporte na superfície marciana. Evidentemente, este procedimento pode levar a erros, uma vez que a direção predominante do vento não tem que ser constante ao longo do ano. Mesmo assim, este é o melhor método disponível na comunidade científica para prever e simular as direções do vento no planeta Marte, existindo já trabalhos que utilizam estes dados como fontes das direções de transporte de sedimentos em campos de dunas marcianos (Fenton et al. (2014)) Além disso, quando os resultados do MRAMS entrarem em conflito com as direções induzidas a partir dos padrões de dunas na superfície, serão utilizadas as direções indicadas por estas últimas. Ou seja, se as direções indicadas pelo MRAMS diferirem da nossa análise do terreno a partir das indicações visuais das dunas, serão utilizadas essas indicações visuais de direções de transporte de sedimentos e não as direções fornecidas pelo MRAMS. Finalmente, a direção de transporte considerada será uma média ponderada, de acordo com os valores do rácio entre o stress aerodinâmico/limite de saltação. Este passo é efetuado uma vez que os resultados apresentam rácios stress aerodinâmico/limite de saltação inferiores a 0.5, e desta forma a saltação de facto nunca ocorreria. É assumido

que o transporte de partículas ocorra em regimes de extremos de velocidade do vento no local, pelo que o momento exato em que isto acontece está fora do nosso alcance e conhecimento. É importante compreender bem este parágrafo e a maneira como estes dados são utilizados no programa de forma a perceber os objetivos e o método utilizado posteriormente.

Estes dados serão passados para um ficheiro *shapefile* do ArcGIS e anexados aos nossos ficheiros das dunas existentes utilizando os métodos desse programa. A resolução utilizada será superior à dos resultados do MRAMS, pelo que terão de ser efetuadas médias e aproximações entre os pontos da simulação para ser possível a anexação pretendida. Para trabalhos futuros, seria importante e até necessária a obtenção de simulações com resoluções superiores (1km, ou melhor) das direções do vento, de forma a eliminar estes cálculos de médias que podem levar a erros nos resultados obtidos.

2.4 Áreas de estudo em Marte

Além das características gerais do planeta Marte e da sua caracterização eólica e morfológica, será necessária uma análise em particular de cada um dos campos de dunas que vamos analisar durante esta tese. Uma vez que este trabalho se foca na caracterização de alguns campos de dunas em particular e na indução destas análises para todo o planeta, será muito importante perceber quais as características exatas de cada um destes campos de dunas, e porque são importantes no contexto geral do planeta. Será ainda importante perceber o que a bibliografia disponível sobre cada um destes campos de dunas nos pode indicar, assim como as análises já existentes sobre as direções preferenciais de transporte em cada um dos locais, ou os tipos de duna predominantes e a sua história geológica. Assim, vamos tomar em consideração cada um destes aspetos na caracterização dos campos de dunas a estudar, e mostrar uma vista geral, a partir do ponto de vista geomorfológico e eólico, de cada um dos campos.

2.4.1 Cratera Gale

A cratera Gale mede aproximadamente 150km de diâmetro, como se pode ver na imagem 2.11, e está situada na fronteira entre as terras a sul, cobertas de crateras, e as terras a norte da Elysium Planitia (5.4°S e 222.2°W) (ver imagem 2.2). A cratera Gale é dominada por uma montanha central, com cerca de 5km de altitude, denominada Monte Sharp. Pensa-se que este monte consista numa sequência de estratos de origem sedimentar. Existem também vestígios de atividade fluvial na cratera, como canais no monte central e nas paredes da cratera. Estes indícios mostram que a cratera terá tido água líquida no seu interior durante o período Noaquiano (ver seção 2.1.4.1) Existe ainda uma grande variedade de vestígios eólicos na cratera (Hobbs et al. (2010)), tornando-a um local excelente para estudos de atividades eólicas. São visíveis campos de dunas que seguem toda a face norte da montanha central, assim como um grande mar de areia situado entre a face ocidental da cratera e o Monte Sharp, e campos de

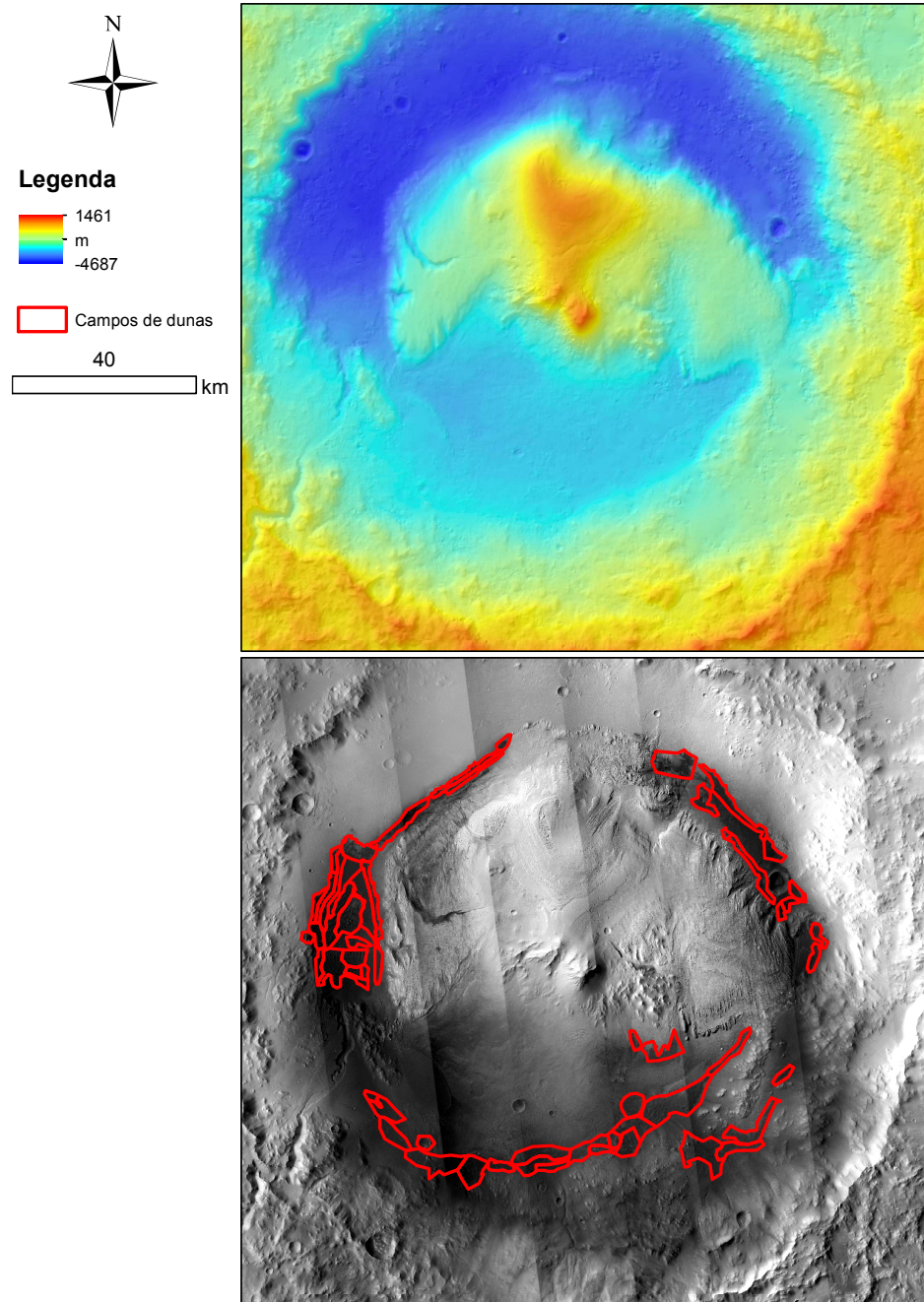


Figura 2.11: Imagem da cratera Gale, comparando o mosaico criado no programa ISIS com a topografia obtida com dados do HRSC (resolução de 50 metros por pixel), com os campos de dunas assinalados a vermelho. De notar a elevação central da cratera e a depressão na sua parte norte.

Campo de dunas	Descrição	Tipo de dunas	Area (km ²)	Tamanho Partículas	Dir. pref. vento (princ/sec)
A	Campo dunas Noroeste	Barcanas, barcanóides, transversas	17	Areias muito grossas (1-2 mm)	Bar - 351/70; Tra - 326/94
B	A Sudoeste de A	'Dome dune', barcana, barcanóide	5	Areias grossas - muito grossas (1/2-2 mm)	55/77
C	Campo dunas Ocidental	Lençol de areia, Barcanas, barcanóides, transversas	303	Areia média - cascalho	Bar - 297/156; Tra -297/124
D	Campo dunas Nordeste	Streak, 'Dome dune', barcana, barcanóide	171	Areia grossa - muito grossa, algum cascalho	275/50
E	Campo dunas Sudeste	Barcanóides, transversas	203	Areia média - grossa	236/195
F	Campo dunas Sul	Barcanas modificadas, transversas	78	Sem dados	142/244
G	Campo dunas Sudoeste	Streaks, Barcanas	50	Sem dados	324/255

Tabela 2.1: Tabela com os tipos de dunas, extensão, e tamanho dos grãos de areia em cada uma das seções da cratera Gale. Os tamanhos dos grãos foram calculados utilizando medidas de albedo das áreas dunares e representam uma estimativa grosseira dos tamanhos das partículas. Os campos de dunas encontram-se identificados na imagem 2.12. As direções preferenciais de transporte de sedimento estão indicadas a partir do Norte e em graus. Bar e Tra indicam as dunas barcanas e transversas, respetivamente.

dunas dos lados Sul e Este desse monte. A extensão destas características indica que a atividade eólica relativamente recente tem influenciado significativamente a geomorfologia da cratera Gale, e formou muitas das características vistas hoje em dia. Estão agora a surgir vestígios que indicam atividade eólica no presente, tais como deslocamentos de dunas (Silvestro et al. (2013)). Com a chegada da sonda MRO (Mars Reconnaissance Orbiter), foi possível obter imagens com enorme resolução da superfície marciana e clarificar processos eólicos e a morfologia de dunas da cratera Gale. Todos estes motivos de interesse motivaram inclusivamente a escolha desta cratera como local de aterragem da MSL (Mars Science Laboratory), ou Curiosity, num passado recente. Esta missão veio trazer nova importância à cratera, sendo agora local de muito interesse no panorama geral do planeta, e os estudos nesta zona adquirem um cariz fundamental.

As direções preferenciais de transporte de sedimentos e o tipo de dunas predominante na cratera já foram identificadas (Hobbs et al. (2010)), e denotam uma predominância das dunas barcanas. As crestas das dunas barcanóides e transversas estão localizadas principalmente a sotavento em relação aos campos de barcanas. Encontraram-se também '*wind streaks*' e dunas em domo (dunas sem face de deslizamento, apenas um domo de areia) na cratera. As morfologias de tipos de duna encontrados em cada um dos campos de duna específicos estão explicitadas na tabela 2.1.

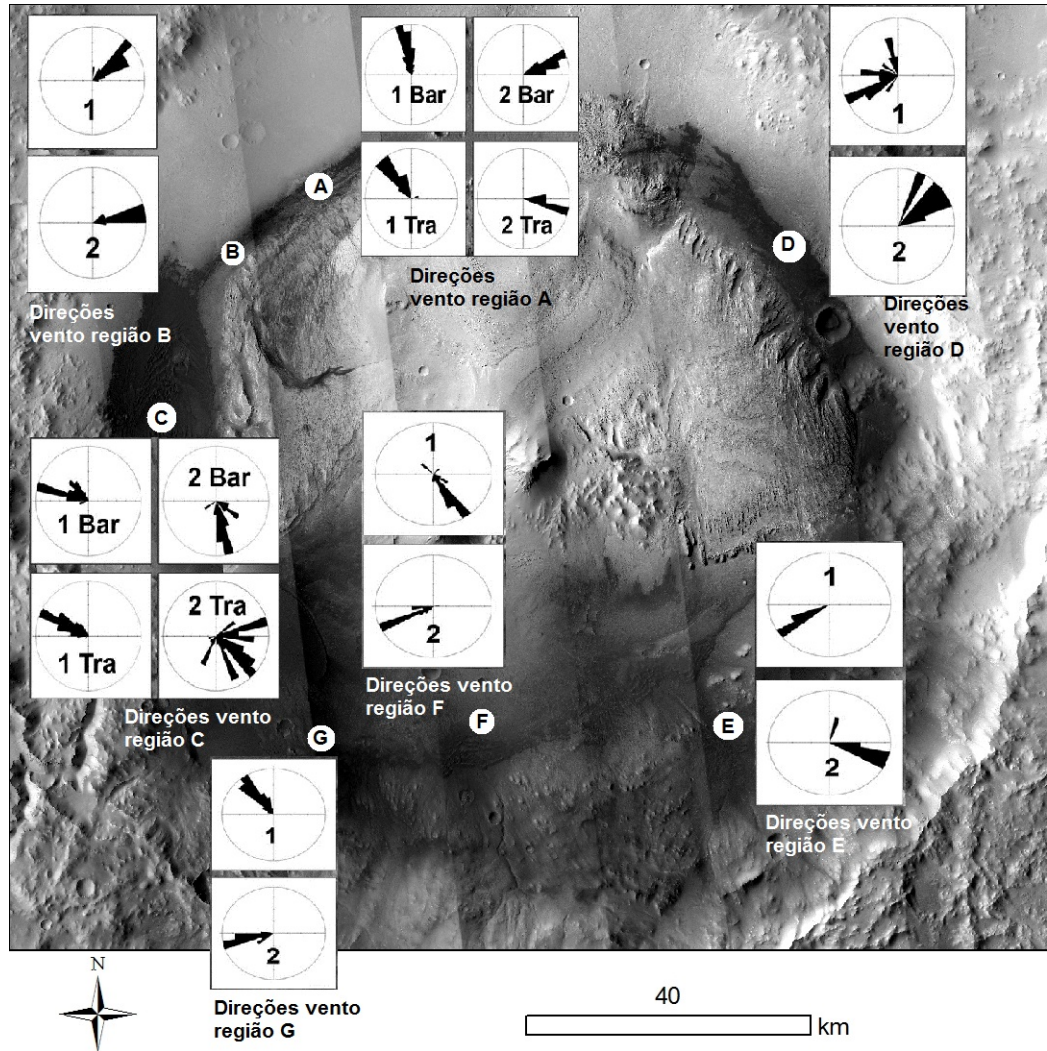


Figura 2.12: Diagrama indicando as direções preferenciais de transporte de sedimento em cada zona da cratera Gale. Os diagramas de roseta indicam a direção de sotavento predominante e a segunda principal em cada um dos subcampos de dunas estudados. BAR indica que os dados foram obtidos a partir de dunas barcanas e TRA indica que foram obtidos a partir de dunas transversas. Quando não existia a possibilidade de efetuar análises separadas dos dois tipos de duna, os diagramas apenas apresentam as direções preferenciais de transporte de sedimento. Imagem adaptada de Hobbs et al. (2010)

Na imagem 2.12 estão explicitadas as direções preferenciais de transporte de sedimento em cada um dos campos de dunas secundários da cratera Gale (imagem adaptada de Hobbs et al. (2010), direções assinaladas apenas como estimativa e fornecidas nesse artigo). Assim, podemos constatar que existe um domínio dos ventos da direção Norte-Noroeste na parte norte da cratera, e da direção Sul-Sudoeste na parte mais a sul da mesma. De notar ainda que estas direções preferenciais são consistentes entre os dois tipos de duna estudados, existindo apenas disparidade entre as direções principais e secundárias em cada zona. Estes dados foram obtidos através da cartografia das faces de deslizamento em cada zona da cratera e prevendo a partir das mesmas a direção preferencial de transporte de sedimentos.

A topografia da cratera Gale influenciou obviamente a localização e morfologia dos seus campos de dunas, como se pode verificar pelas direções preferenciais de transporte de sedimento indicadas na imagem 2.12 e na tabela 2.1. As dunas identificadas como barcanas são o tipo de morfologia mais comum na cratera Gale, unindo-se em dunas barcanóides e transversas a sotavento. Provavelmente devido a uma maior acumulação de sedimentos nessas áreas. Estas acumulações de sedimento são influenciadas pelas direções de transporte de sedimentos predominantes na cratera, fazendo com que estes sedimentos se acumulem nas depressões que rodeiam o Monte Sharp, e levando à disponibilidade de material necessário para formar as dunas barcanóides e transversas. Desta forma, o fluxo de vento local na cratera é dominado/controlado pelo Monte Sharp, contrastando dessa forma a cratera com outras onde não existe este tipo de elevações no centro. Existem mais análises sobre os efeitos da topografia do Monte Sharp no regime de ventos na região, mas o que se encontra exposto acima será o suficiente para os nossos propósitos.

2.4.2 Campo de dunas de Aonia

Aonia Terra representa um dos mais extensos campos de dunas na região de Thaumasia, a Nordeste de Aonia Planum (49°S e 65°W) (ver imagem 2.2 e 2.13). Nesta área temos disponíveis seis ergs (campos de dunas), na base de diversas crateras, e interconectados por um sistema extenso de transporte de sedimentos (Silvestro et al. (2010a)). Esta área apresenta bastantes características análogas a desertos terrestres de Mojave, Sahara, ou da Arábia. A identificação dos regimes locais de vento e das áreas fontes dos materiais para as diferentes zonas permitem uma explicação detalhada da morfologia local, que apresenta um grau de complexidade capaz de rivalizar com os análogos terrestres mencionados. Desta forma, vamos apresentar então uma vista geral do campo de dunas de Aonia.

Uma análise à região da Thaumasia revela que toda a área parece ter sofrido uma série de processos geológicos, envolvendo atividades tectónicas e vulcânicas, acompanhadas de atividade hidrológica esporádica Silvestro et al. (2010a). Já em 1986 (Scott (1986)) foram identificadas duas unidades geológicas, Npl1 e Nplh, que representam zonas com crateras e com colinas, respetivamente. Estas formaram-se durante o período de forte bombardeamento e são, provavelmente,

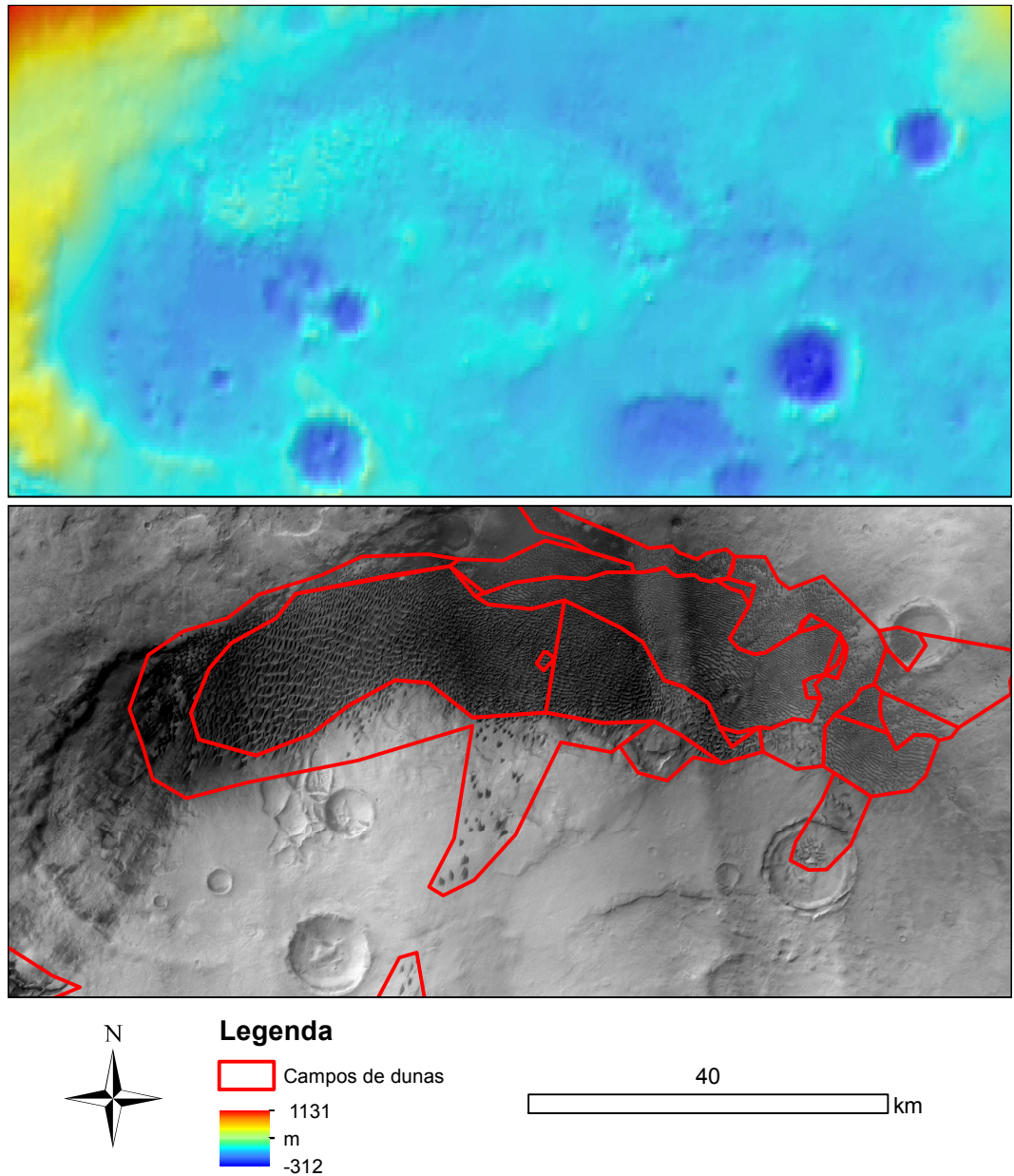


Figura 2.13: Imagem de mosaico representando o campo de dunas de Aonia e uma comparação com a topografia da mesma zona. A região mais escura no meio da imagem é a área onde se podem encontrar as dunas, com as áreas de interesse assinaladas a vermelho. Para além do grande campo de dunas representado, existem algumas dunas a Sudoeste, mas não foram consideradas nesta análise. Nesta zona o aspeto da topografia é pior do que o apresentado na cratera Gale porque tem uma resolução de apenas 463 metros por pixel.

constituídas de fluxos de lava e brechas de impacto. Estas unidades encontram-se bem cobertas por uma rede dendrítica de drenagem. A zona que estamos a estudar é, portanto, mais uniforme, do período Hesperiano (3700-3000 Ma), e interpretada como sendo composta por espessos fluxos de lava cobertos por depósitos eólicos. Encontram-se espalhadas “*wind streaks*” e barcanas, indicando a presença de areias e as suas direções de transporte. Entre as dunas no local encontram-se vários tipos, tais como barcanas, complexas e longitudinais.

A forma e as direções das frentes de avalanche no campo de dunas indica que estamos na presença de um regime de dunas complexo e multi-direcional (Silvestro et al. (2010a)). Em locais onde se verifica a convergência de ventos de diferentes direções surgem extensos padrões de dunas complexas (como se pode ver na imagem 2.14). As direções preferenciais destes ventos foram calculadas, depois de uma análise da distribuição das dunas e da sua forma, e são de Sudoeste e de Nordeste, soprando simultaneamente. Depois disso, uma alteração nos regimes de vento terá levado a uma modificação morfológica dos campos de dunas, sendo esculpido por fluxos mais recentes.

2.4.3 Cratera Herschel

A cratera Herschel (coordenadas 11.5°-17°Sul e 127°-132.5°Este, ver imagem 2.2) está localizada na região do Mare Tyrrenium. Esta cratera, de cerca 300km de diâmetro, está situada nos terrenos cobertas de crateras do hemisfério Sul marciano, a cerca de 100km a Leste dos terrenos vulcânicos de Hesperia Planum. Os bordos da cratera situam-se cerca de 1.5km acima da base, tendo o bordo sul sido coberto de várias outras crateras de impacto menores, com cerca de 10km de diâmetro (como se pode ver na imagem 2.15).

Encontram-se dentro desta cratera vários tipos distintos de dunas, tais como barcanas, barcanóides, dunas de domo, campos de areia, dunas transversas e complexas (ver imagens 3.13.3, onde se podem ver algumas imagens detalhadas das dunas da cratera Herschel), e foi por essa razão que foi escolhida para este trabalho, servindo como um laboratório de estudos de faces de deslizamento, direções das dunas e os seus tipos, etc. Está, além disso, relativamente próximo da cratera Gale, onde está situada a MSL (ver seção 2.4.1) e pode servir como ponto de referência/contraste com essa cratera. As direções das frentes de avalanche nas dunas indica uma direção preferencial de transporte de sedimentos no sentido Norte-Sul, com ventos de outras direções contribuindo para a acumulação de dunas mais complexas em algumas regiões da cratera e no interior das crateras menores. Este regime de ventos é, portanto, fortemente dependente da topografia local (Cardinale et al. (2012a)).

Os ventos no local continuam ativos no presente (Cardinale et al. (2012b); Runyon et al. (2014)), existindo vestígios e imagens que ilustram o movimento de *ripples* e dunas mesmo hoje em dia. Desta forma, podemos concluir que a cratera Herschel é também um importante foco do trabalho sobre os campos de dunas e os movimentos eólicos na superfície marciana.

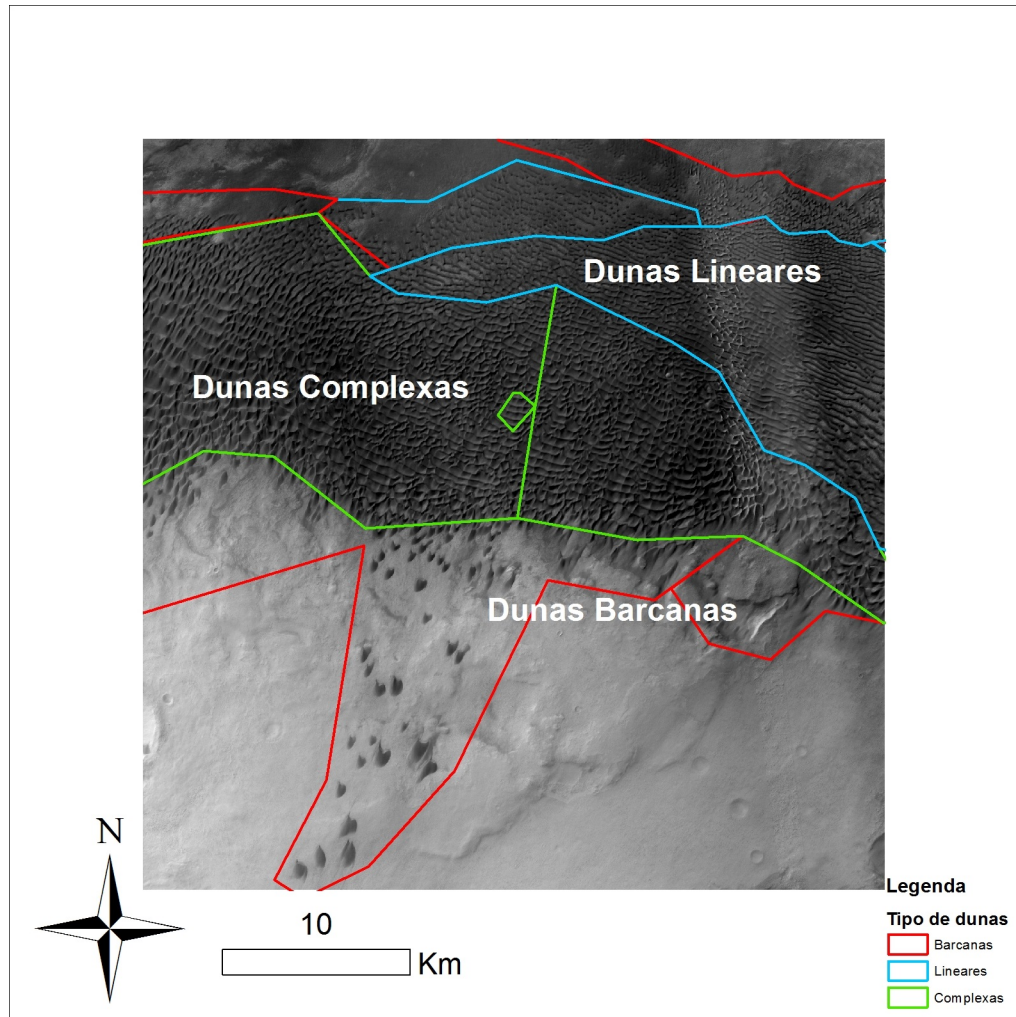


Figura 2.14: Caracterização dos tipos de dunas no campo de Aonia, evidenciando as diferenças evidentes entre as dunas em cada área. De notar as barcanas isoladas a Sul, mostrando a direção de transporte Sul-Norte com a sua frente de avalanche, e as dunas complexas no centro do campo, ilustrando a presença de direções conflituosas do vento. Esta classificação manual em 3 tipos distintos de dunas será utilizada mais tarde na análise com *clusters*.

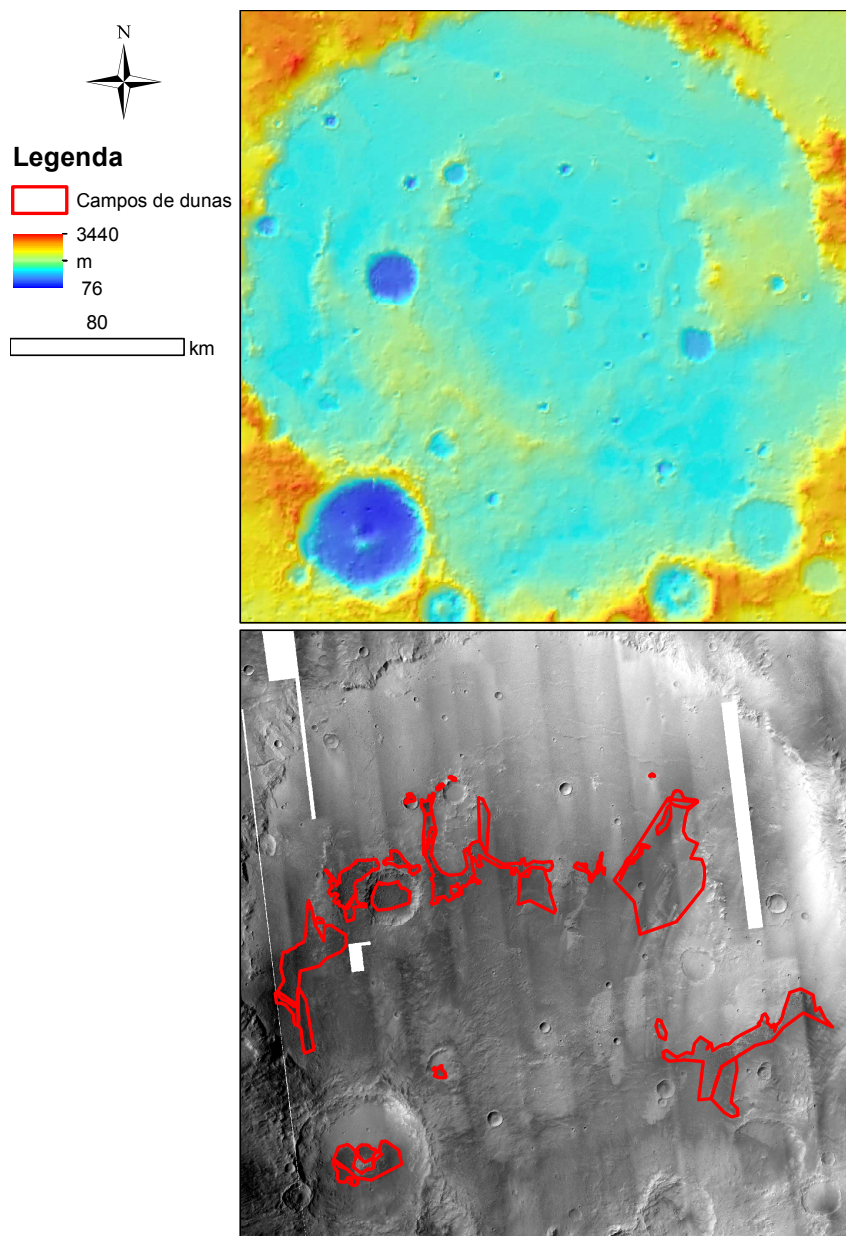


Figura 2.15: Topografia da cratera Herschel, obtida a partir dos dados do MOLA, e a imagem de mosaico da mesma, construída no programa ISIS. Os campos de dunas de interesse encontram-se assinalados a vermelho. De notar a grande área plana no centro da cratera, onde se acumulam as dunas, assim como a cratera menor no sudoeste, onde também existem vestígios de dunas.

Capítulo 3

Métodos e dados utilizados

Neste capítulo será apresentado um resumo das metodologias utilizadas neste trabalho para chegar aos resultados de análise das dunas e dos ventos que lhes dão origem. O trabalho divide-se em duas partes, a criação de mosaicos de imagens dos campos de dunas marcianas com as dunas evidenciadas, e a construção do programa capaz de condensar estes resultados e extrapolar a partir deles os parâmetros do vento e do tipo de duna constituintes. Vamos, por isso, dar um exemplo de como se pode construir um mosaico de imagens de um campo de dunas e traçar digitalmente as linhas que vão representar as dunas que o constituem e, depois, vamos ver como podemos utilizar essa informação calculada para reconstituir virtualmente as direções de transporte de sedimentos que lhe deram origem e testar alguns métodos de classificação digital das dunas que constituem o campo. O campo de dunas que vamos utilizar como exemplo neste caso é o da cratera Herschel, que contém vários campos dunares na sua extensão e é portanto um bom alvo para este tipo de análise.

3.1 Imagem de Marte - exemplo cratera Herschel

O processo de construção de um mosaico de imagens obtidas pelo instrumento CTX (Context Camera) será explicado utilizando como exemplo as imagens da cratera Herschel (ver seção 2.4.3). Para começarmos a análise das dunas em si, primeiro temos de conseguir encontrar um campo de dunas com as características procuradas (boa distribuição de dunas, indicadores da direção de transporte de sedimentos, grande área a analisar, etc), e a cratera Herschel é um bom exemplo; e obter uma imagem de boa resolução ($<10\text{m}/\text{pixel}$) dessa área. Com as diversas sondas que se encontram a orbitar o planeta Marte, encontrar as imagens neste momento não será uma tarefa particularmente difícil. A nossa maior dificuldade será condensar e juntar as diversas imagens que podemos obter dos campos de dunas e criar um mosaico onde possamos visualizar as dunas e o meio envolvente de uma forma integrada. Para além desta cratera temos, como já foi dito anteriormente, outras áreas disponíveis para análise, como a

cratera Gale (ver seção 2.4.1) , o campo de dunas Aonia (ver seção 2.4.2). Todos estes campos de dunas foram utilizados no desenvolvimento deste programa, e os resultados em cada um deles serão posteriormente analisados.

3.1.1 Construção do mosaico

Utilizando o programa ISIS (Integrated Software for Imagers and Spectrometers) criámos o mosaico que se pode observar na imagem 2.15. Para tal foram descarregadas as imagens em bruto disponíveis *online* através do interface Mars Orbiter Data Explorer (<http://ode.rsl.wustl.edu/mars/>). Foram usadas as imagens da câmara CTX que se encontra a bordo da sonda MRO, com resolução aproximada de 5-6m/pixel (não se utilizaram as imagens da câmara HiRISE porque, apesar de possuírem uma melhor resolução espacial, esta não seria necessária para a análise que pretendíamos efetuar, e não existe uma cobertura completa dos campos de dunas que pretendemos estudar). Assim, usaram-se as imagens com resolução mais baixa, mas que permitem uma cobertura completa dos campos de dunas. As principais etapas de processamento implicam: 1) calibração radiométrica e remoção de ruído; 2) equalização das várias imagens; 3) criação de uma rede de pontos de controle de forma a permitir o refinamento das correções geométricas; 4) aplicação das correções geométricas e posterior projeção para um sistema de coordenadas centrado em cada área de estudo; 5) criação do mosaico e filtragem das áreas de sobreposição das várias imagens, de forma a reduzir as diferenças de iluminação nas zonas de transição entre as imagens que compõem o mosaico. Após todas estas etapas obtêm-se mosaicos geometricamente controlados com uma resolução de 7m/pixel. Para cobrir algumas áreas onde não existiam imagens CTX disponíveis, foram utilizadas imagens com resolução mais baixa (obtidas pela câmara HRSC - High Resolution Stereo Camera). Mesmo com esse método, continuaram a existir 'buracos' no nosso mosaico, onde não se encontrou nenhuma imagem disponível para os preencher.

O método utilizado é bastante 'manual', uma vez que envolve várias operações diretamente pela linha de comandos, e, dependendo dos processadores do computador utilizado para criar o mosaico, pode ser demorado (2-3 dias, pelo menos, para criar um mosaico). Desta forma é importante dominar o programa ISIS e utilizar todas as ferramentas à nossa disposição para criar um mosaico de uma maneira direta, sem perder tempo precioso com passos inúteis. Depois dos conhecimentos adquiridos com o processamento desta cratera, as imagens seguintes foram certamente mais fáceis de processar. De qualquer forma, a criação destes mosaicos permite obter uma imagem homogénea e continua de grandes extensões da superfície marciana, e são ideais para cartografar e estudar a superfície do planeta.

3.1.2 Delimitação da área dunar

Depois da construção do mosaico da nossa cratera no ISIS passamos a trabalhar com o sistema de informação geográfica ArcGIS, e o resto do processamento será

feito aí, com o construção do algoritmo na linguagem de programação python. Será necessária, portanto, a delimitação das zonas dunares entre toda a área da cratera no mosaico. De forma a simplificar o procedimento, esta seleção é feita manualmente, assinalando as regiões de interesse para análise, no entanto existem já trabalhos que sugerem a possibilidade de automatização desta etapa de processamento Bandeira et al. (2011, 2013). Utilizando métodos de extração de informação local das imagens, tal como características dos gradientes, será possível distinguir zonas onde as dunas estão presentes de onde estas se encontram ausentes com elevada precisão. Infelizmente, os trabalhos atuais apenas permitem a distinção de dunas barcanas e barcanóides. Desta forma, o método não será utilizado neste trabalho, mas fica a nota de interesse. Tal como na criação do mosaico, nesta fase é preferível selecionar terreno a mais do que a menos. A filtragem das dunas para o cálculo dos parâmetros que as caracteriza será feita no passo seguinte, nesta altura o importante é reunir toda a área importante disponível no mosaico.

As dunas são visualmente identificadas pela cor escura do terreno que as constitui, evidenciando a areia aí existente, por contraste com a 'cama de rocha' que se encontra subjacente, e pelos complexos padrões nelas observados, que foram analisados anteriormente nesta tese (seção 2.2.3 - Geomorfologia). Com estes dois critérios, podemos identificar visualmente as áreas onde existem dunas e o seu tipo com relativa facilidade. A análise posterior a estas áreas é efetuada com métodos computacionais.

Para além de campos de dunas densos, com dunas a preencher toda a superfície, observamos, ainda, zonas com apenas uma linha de dunas evidente, e o resto da areia dispersa pelo terreno. Apesar da menor relevância destas dunas solitárias para a análise do terreno, vamos incluí-las na nossa seleção, não só porque são dunas e podem servir como pontos importantes posteriormente, como são dos pontos onde as direções de transporte de sedimentos são mais evidentes e facilmente observadas. A partir das dunas visíveis e dos rastros de areia deixados no terreno, podemos concluir com relativa facilidade as direções de transporte de sedimentos que dão origem aos padrões observados.

3.1.3 Cartografia e caracterização dos lineamentos dunares

Uma vez concluída a seleção das zonas de interesse do mosaico entre toda a área disponível na imagem, foram aplicadas o mesmo conjunto de técnicas desenvolvidas anteriormente para análise paleotectónica de Marte (Vaz (2010); Vaz & Silvestro (2014)) mas que se revelaram igualmente úteis para a cartografia e caracterização de *ripples* eólicos em Marte. São aplicadas um conjunto de técnicas de análise de imagem que possibilitam a segmentação e vectorização das estruturas lineares existentes nas imagens, no presente caso as estruturas lineares que formam os padrões de dunas. Este procedimento possibilita uma análise de imagem orientada a objetos, o que permite uma simples integração da informação textural, morfológica, espectral e geométrica dos objetos que constituem uma imagem Blaschke (2010). Esta análise baseada em objetos é apenas possível

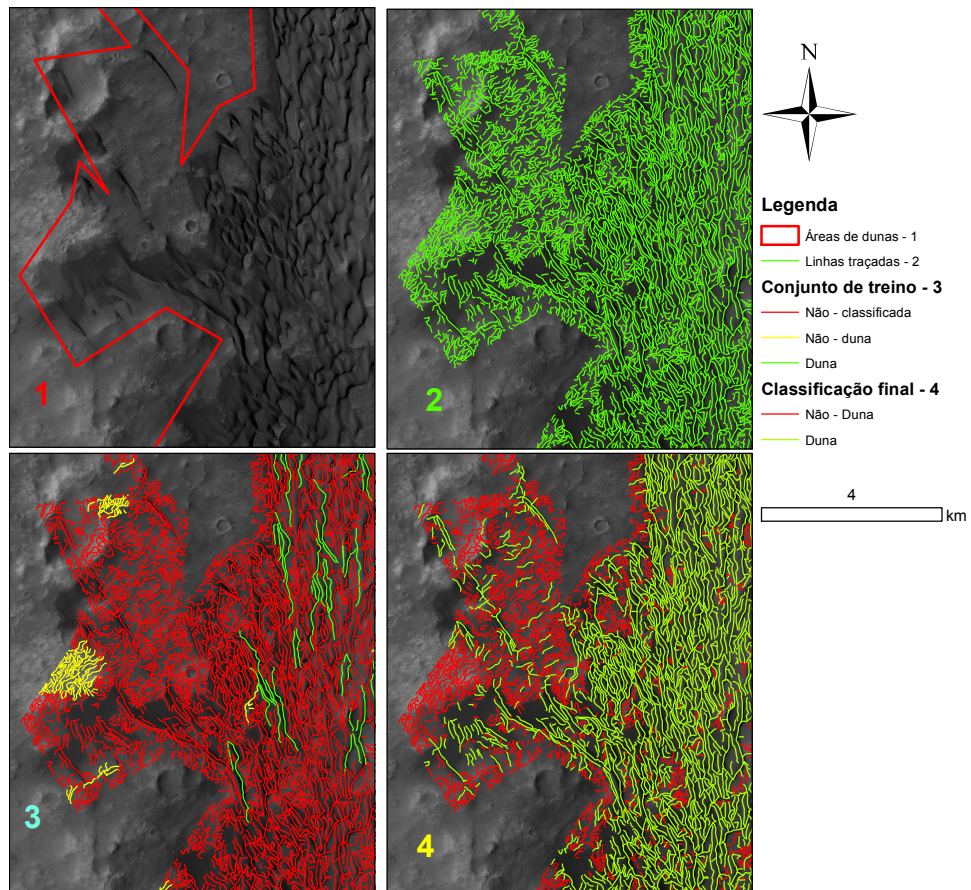


Figura 3.1: Da esquerda para a direita, de cima para baixo: 1: Delimitação da área com dunas na cratera Herschel, close-up. As áreas interessantes (com areia e dunas) são seleccionadas para análise na etapa seguinte, enquanto que as áreas sem interesse são deixadas sem interferência, de forma a diminuir o tempo de processamento para o cálculo das possíveis linhas das dunas e dos seus parâmetros de análise. 2: Resultados iniciais da análise das linhas da cratera. Comparar resultado com imagem anterior. Estas linhas não representam ainda apenas as dunas verdadeiras, há muitas linhas que não correspondem a faces de deslizamento nem características da duna. Essas linhas serão ainda filtradas antes de termos uma imagem a partir da qual possamos começar a analisar os resultados. 3: Representação do método de eliminação de falsos positivos e linhas falsas criadas na imagem. No exemplo, as linhas marcadas a verde e a vermelho serão usadas como *input* num programa de redes neuronais utilizada para classificar as linhas como 'Duna' e 'Não_Duna'. Com uma amostra razoável (1000-3000 linhas para cada categoria) o programa vai aprender com a classificação manual e generalizá-la para o resto das linhas na imagem. A classificação final será uma compilação dos resultados obtidos pelas redes e da classificação manual, onde uma linha utilizará o resultado das redes a menos que tenha sido seleccionada manualmente para uma das categorias, e nesse caso utiliza a categoria seleccionada manualmente. 4: Resultados finais da classificação das linhas na cratera Herschel. De notar ainda a existência de alguns falsos positivos e falsos negativos, mas são em número reduzido e espalhados pela área. As linhas marcadas a vermelho serão descartadas para a análise posterior e a partir daí o foco do trabalho será na transformação destas linhas em pontos que conservem a sua informação de alguma forma.

uma vez que temos resoluções que nos permitem distinguir claramente as dunas (pixeis individuais pequenos em relação a elas). Contudo, este processo gera um enorme número de falsos positivos, e, antes de proceder à análise destes resultados, será necessário filtrar as linhas obtidas. Na imagem 3.1, parte 2, estão evidenciadas as linhas obtidas na imagem depois da aplicação deste algoritmo, e é visível a quantidade de linhas produzidas que não correspondem a estruturas dunares na imagem subjacente. Para além da criação destas linhas, o programa utilizado permite também o cálculo de vários parâmetros através da análise do mosaico da cratera, nomeadamente o albedo da linha, o seu comprimento, o seu azimute médio, o seu pendor, entre outros. Estes parâmetros serão os utilizados numa fase posterior da análise para, primeiro, separar as dunas reais dos falsos positivos e, em segundo lugar, distinguir, possivelmente, as direções de transporte de sedimentos que geram cada área de dunas.

3.1.4 Eliminação falsos positivos

Depois da seleção automatizada das potenciais dunas entre as zonas selecionadas no mosaico, é então aplicado um novo filtro, através de uma análise utilizando redes neuronais, para distinguir as linhas que representam dunas ou as suas faces de deslizamento e os falsos positivos gerados automaticamente. Este filtro funciona através de uma seleção manual de algumas linhas na imagem que pertencem a um dos dois grupos (dunas ou falsos positivos) e a classificação automatizada do resto do grupo, utilizando como *input* as linhas selecionadas. Este método de seleção é iterativo, uma vez que podemos ter de selecionar as linhas como verdadeiras/falsas mais do que uma vez e repetir o processo até ao máximo de sucesso ser atingido. Nas imagens que ilustram esta seção podem ver-se alguns dos resultados obtidos pelo método, destacando o sucesso por ele conseguido.

As redes neuronais funcionam utilizando as linhas selecionadas como treino, e o resto das linhas como teste. Todos os parâmetros que acompanham as linhas, mencionados na seção anterior (3.1.3, são aqui utilizados como auxiliares na seleção das dunas entre as linhas a analisar. Para classificar as dunas foram criados campos na tabela de conteúdos das mesmas, e os valores das mesmas preenchidos durante a seleção. Como uma medida extra para a obtenção de uma boa classificação das dunas, foi selecionado como o campo a utilizar aquele da classificação manual, em caso de discórdia entre esse campo e o calculado pelas redes neuronais. Assim, evita-se a existência de possíveis falsos negativos/falsos positivos em que a classificação das redes diferia da classificação visual e manual da linha.

Este método de seleção de dunas/falsos positivos foi também desenvolvido pelo Doutor David A. Vaz e está baseado na linguagem de Matlab, assim como o programa de identificação de linhas original mencionado acima. As dunas obtidas na filtragem final serão utilizadas finalmente para a construção de uma rede de parâmetros regionais do campo de dunas, como se pode comprovar na seção seguinte (3.1.6) da tese.

3.1.5 Parâmetros de caracterização das dunas

Os parâmetros individuais utilizados e armazenados na rede de pontos estão entre os calculados e processados pelo programa de identificação de linhas, e, em particular, os campos utilizados foram os visíveis na tabela 3.1. Para além dos parâmetros utilizados nas redes neuronais e no *clustering* existiram outros parâmetros calculados por esse programa que não foram utilizados posteriormente. Esta seleção foi feita de acordo com o conjunto de parâmetros que apresentava melhores resultados após a aplicação das redes neuronais e do *clustering*, se bem que um estudo mais aprofundado pudesse levar a uma seleção mais otimizada nesse aspeto, levando a um programa capaz de obter o melhor conjunto de resultados com os parâmetros de input. As categorias aplicadas na tabela referem-se aos métodos de obtenção e cálculo destes parâmetros. Em particular, os parâmetros texturais utilizam informação da cor e da textura da imagem, os parâmetros Geométricos calculam resultados referentes à distribuição de linhas e às suas características geométricas, os parâmetros morfométricos utilizam dados referentes à topografia e ao modelo digital de terreno, e os parâmetros mistos utilizam dois ou mais destes tipos de informação. Em particular, e no caso de parâmetros direcionais referentes à duna, estes não serão calculados diretamente, de modo a diminuir possíveis enviesamentos devido a tendências direcionais no campo de dunas. Em vez disso, são medidos os ângulos agudos entre pares de parâmetros ("AZIM" e "DF1", por exemplo) e os valores utilizados são dos senos e cosenos desses ângulos, e não dos campos originais. Na imagem 3.2 podem ver-se alguns padrões da imagem que permitem a aplicação posterior das técnicas de processamento criadas. A representação gráfica destes parâmetros permite a compreensão do tipo de padrões que os algoritmos das redes neuronais e do *clustering* vão procurar de forma a poderem extrair informação importante da imagem.

3.1.6 Cálculos da estatística regional das dunas

Com a eliminação dos falsos positivos concluída, o passo seguinte consiste na criação de uma rede uniformemente espaçada de pontos na área onde existem dunas na imagem, onde serão condensados os dados das dunas individuais, de forma a possibilitar um processamento mais rápido e mais geral do campo de dunas do que uma análise das dunas individuais permitiria. Assim, vamos formar uma rede de pontos, com um espaçamento de 1km, onde vamos armazenar alguns dos campos disponíveis mencionados acima no processamento das dunas individuais. Serão armazenados em cada ponto a média e o desvio padrão dos campos selecionados, para permitir uma descrição completa dos campos dunares, evitando ao máximo as perdas de informação ao mudar de paradigma de informação (passagem de dunas individuais para pontos que representam várias dunas simultaneamente). Serão também armazenadas informações relativas à densidade de linhas em redor dos pontos nos quais estamos a calcular os nossos valores, uma vez que é necessário descartar da nossa análise pontos nos quais o número de dunas disponível para análise não é estatisticamente significativo.

Parâmetros calculados					
Nome	Descrição	Tipo	Utilizado?	Escalar/Angular?	
DN	Medida do albedo/cor/nível de cinzento na duna;	Textural	Sim	Escalar	
ENTRP	Medida da entropia da duna, corresponde a variação do parâmetro DN ao longo da duna	Textural	Sim	Escalar	
BRG_D	Cálculo local do nível de brilho da duna	Textural	Sim	Escalar	
DRK	Cálculo local do nível de escuridão da duna	Textural	Não	Escalar	
DF1, DF2	Força do campo direcional da duna, (1 e 2 indicam diferentes escalas a que este é calculado)	Textural	Sim	Angular	
DFS1, DFS2	Força do campo direcional da duna, (S representa a força do mesmo)	Textural	Sim	Escalar	
LENGTH	Comprimento da frente de avalanche da duna (em metros)	Geométrico	Sim	Escalar	
IM_DIR	Direção Longitudinal, calculada a partir da refletância da imagem	Textural	Não	Angular	
IM_DIP	Pendor vertical da duna, calculado a partir da refletância da imagem	Textural	Sim	Escalar	
AZIM	Direção da frente de avalanche (2D)	Geométrico	Sim	Angular	
LINE_COUNT	Quantidade/densidade de linhas presente na vizinhança do ponto da rede	Geométrico	Sim	Escalar	
DF1AZ, DF2AZ, DF12	Ângulos agudos calculados entre os parâmetros DF1 e DF2 e AZIM	Misto	Sim	Angular	
SINU	Curvatura da linha (0-100)	Geométrico	Não	Escalar	
LIN_DIR	Direção do vetor longitudinal da linha (a partir do modelo digital de terreno)	Morfométricos	Não	Angular	
LIN_DIP	Pendor do vetor da linha (a partir do modelo digital de terreno)	Morfométricos	Não	Escalar	
SLOPE	Inclinação da topografia, ângulo de pendor	Morfométricos	Não	Angular	
TOPO	Elevação da linha	Morfométricos	Não	Escalar	

Tabela 3.1: Parâmetros calculados das linhas da imagem.

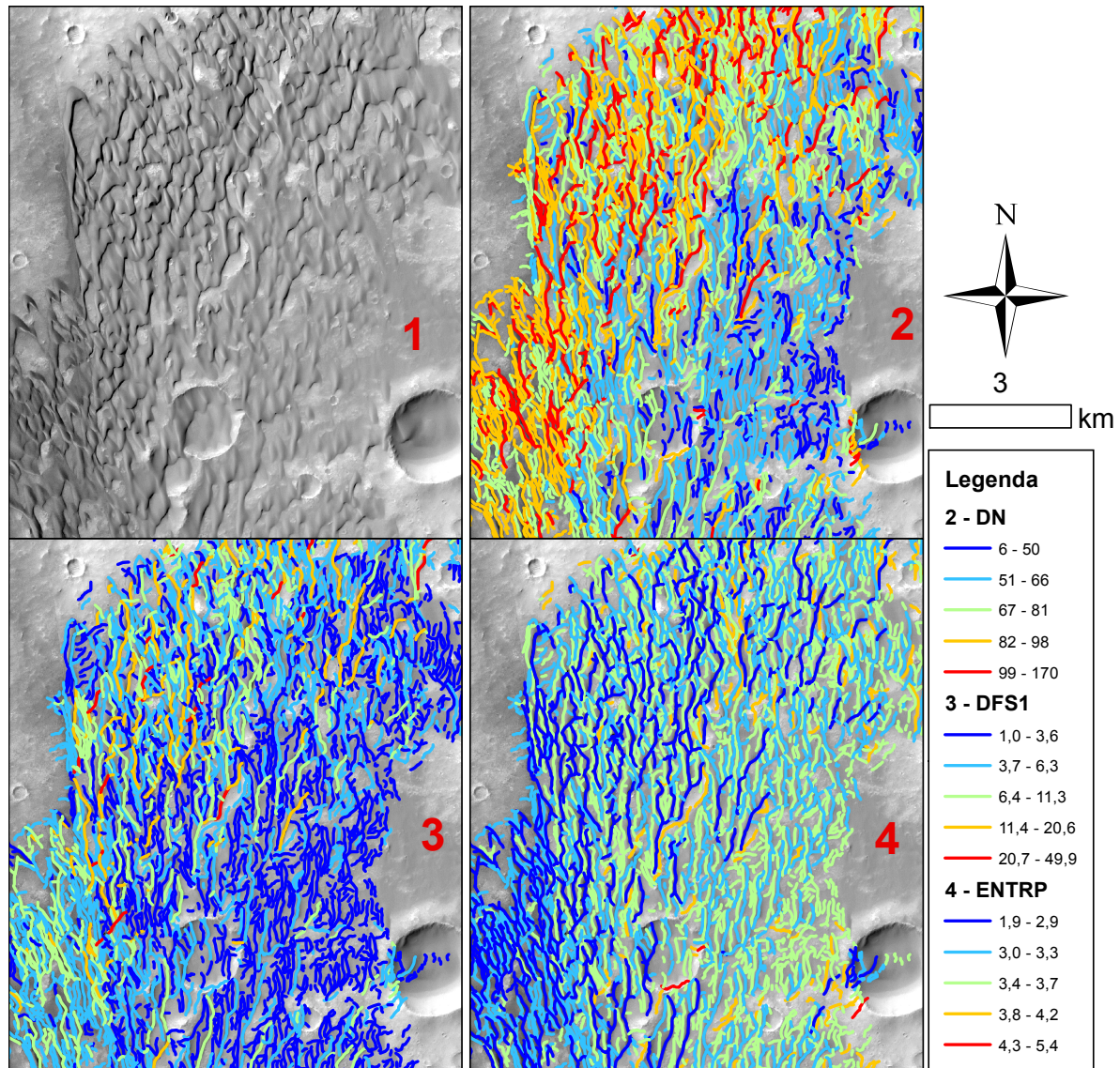


Figura 3.2: Representação gráfica de alguns parâmetros calculados das dunas

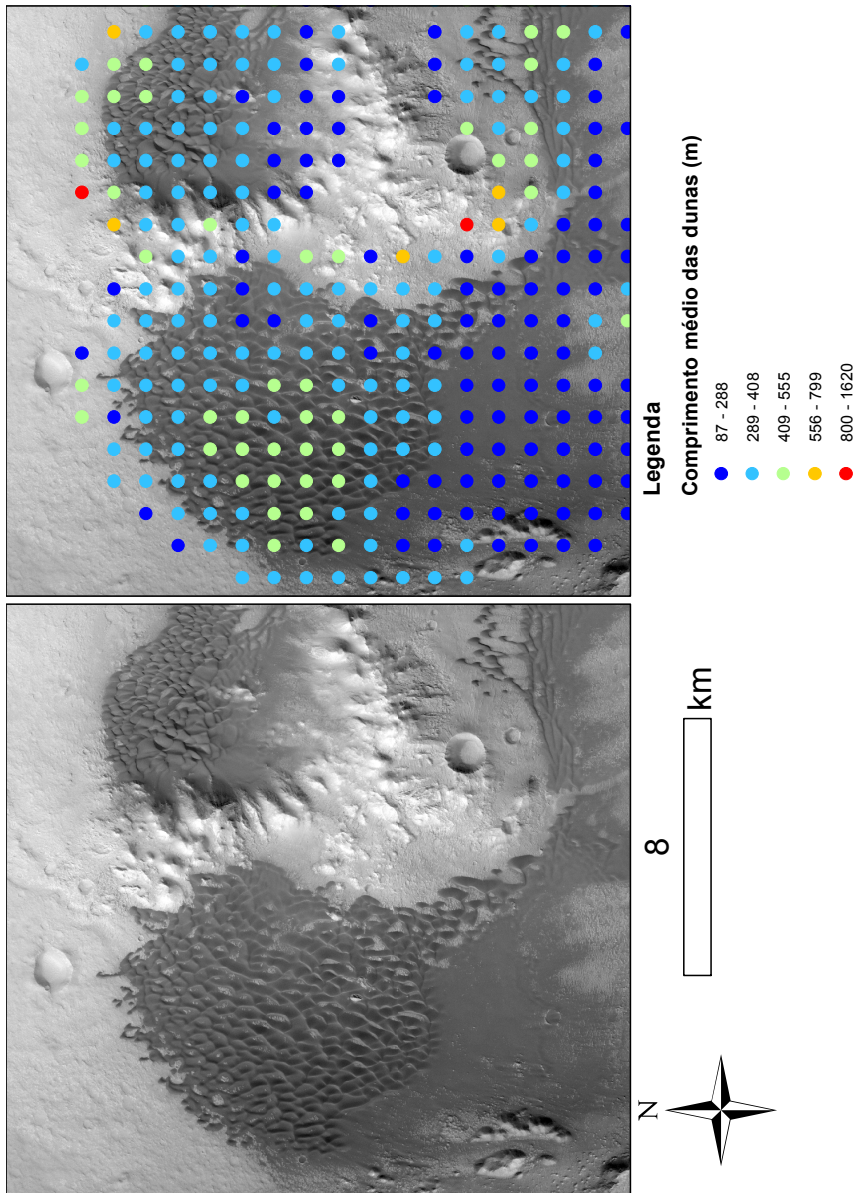


Figura 3.3: Imagem retirada da cratera Herschel, com o intuito de mostrar o aspecto visual de alguns dos dados obtidos. Neste exemplo em concreto, podemos observar uma seção da cratera, contemplar o campo de dunas nela presente, e avaliar visualmente os dados apresentados. O campo mostrado na imagem representa o comprimento médio das linhas que representam o padrão das dunas mais próximas de cada ponto na imagem. Os pontos estão espalhados numa rede uniforme com 1000m de espaçamento entre cada ponto. Neste caso em particular, podemos constatar que, à primeira vista, o comprimento médio das dunas parece aumentar com o grau de complexidade das mesmas. Assim, podemos obter um parâmetro útil capaz de prever a complexidade dum dado campo de dunas, ou diferenciar entre dois tipos de dunas no mesmo campo, a partir do seu comprimento médio, por exemplo.

Os dados direcionais requerem, obviamente, um tipo de processamento que será ligeiramente diferente do utilizado nos dados escalares, e essa diferença é aplicada aqui. Quando é criada a estatística da média e do desvio padrão de cada um dos campos, são tomados os devidos cuidados com esse tipo de dados, utilizando as indicações na literatura disponíveis sobre estatística direcional (Jupp (2000)). Assim, é assegurada a integridade dos campos direcionais nos passos seguintes do algoritmo. A direção média entre vários vetores unitários $\mathbf{x}_1, \dots, \mathbf{x}_n$, com os seus ângulos correspondentes $\theta_i, i = 1, \dots, n$, é calculada de acordo com o seguinte método indicado no livro:

Para calcular a direção média $\bar{\theta}$ de todos os valores $\theta_1, \dots, \theta_n$ temos de calcular a resultante da soma dos vetores $\mathbf{x}_1 + \dots + \mathbf{x}_n$, correspondente também à direção do centro de massa $\bar{\mathbf{x}}$ dos vetores $\mathbf{x}_1, \dots, \mathbf{x}_n$. Uma vez que as coordenadas cartesianas dos pontos \mathbf{x}_j são dadas por $(\cos \theta_j, \sin \theta_j)$, para $j = 1, \dots, n$, as coordenadas dos centros de massa serão (\bar{C}, \bar{S}) , onde $\bar{C} = \frac{1}{n} \sum_{j=1}^n \cos \theta_j$ e $\bar{S} = \frac{1}{n} \sum_{j=1}^n \sin \theta_j$. Desta forma $\bar{\theta}$ é a solução das equações $\bar{C} = \bar{R} \cos \bar{\theta}$ e $\bar{S} = \bar{R} \sin \bar{\theta}$ (com $\bar{R} > 0$) e onde o comprimento médio resultante \bar{R} é dado por $\bar{R} = (\bar{C}^2 + \bar{S}^2)^{1/2}$. De notar que $\bar{\theta}$ não está definido quando $\bar{R} = 0$. Quando $\bar{R} > 0$, $\bar{\theta}$ é dado explicitamente por:

$$\bar{\theta} = \begin{cases} \tan^{-1}(\bar{S}/\bar{C}); & \bar{C} \geq 0 \\ \tan^{-1}(\bar{S}/\bar{C}) + \pi; & \bar{C} < 0 \end{cases}$$

onde o inverso da tangente toma valores entre $[-\pi/2; \pi/2]$. Outros cálculos utilizados, como por exemplo, para o desvio padrão circular destes valores, encontram-se no livro referenciado (Jupp (2000)).

O passo seguinte na nossa análise será uma renormalização dos dados disponíveis para uma escala padrão, entre -1 e 1, utilizando o valor 0 como a média e fazendo um *scaling* com o valor do desvio padrão de cada campo. Esta escala será aplicada a todos os campos não direcionais a analisar, e esta renormalização é efetuada de forma a facilitar o trabalho das redes neuronais que vão ser utilizadas num passo seguinte. Com os valores todos condensados entre -1 e 1, é minimizado o risco de divergência de extremos dos valores utilizados quando aplicamos as redes neuronais aos campos de treino e de teste. Por exemplo, num dado campo com valores entre 10 e 10000, podemos ter um conjunto de treino com um mínimo de 10 e máximo 5000, e um conjunto de teste com valores entre 50 e 10000. Com a renormalização, estas diferenças surgem a uma escala muito menor, eliminando este problema. Além disso, as redes neuronais funcionam normalmente com valores entre -1 e 1, usando esses valores nos pesos de cada neurónio. A renormalização e conseqüente uso desses valores também nos *inputs* facilita o trabalho das redes neuronais, possibilitando assim a obtenção de resultados melhores e mais fiáveis nessa parte do algoritmo.

3.2 Tratamento de dados e simulações dos ventos

Depois de termos condensado os dados disponíveis sobre o campo de dunas numa série de pontos representativos das áreas em redor, vamos anexar à nossa base de dados as informações disponíveis sobre os ventos na região. Para tal pode-

mos agir de duas maneiras: compilar a informação já existente nos modelos do MRAMS (ver seção 2.3.2) , e desenhar os nossos próprios vetores das direções de transporte de sedimentos. A técnica utilizada depende da disponibilidade dos dados e da topografia do terreno de região para região, mas, no nosso caso, na zona da cratera Gale, onde testamos inicialmente este método, incluímos as informações obtidas pelos dois processos. Combinamos assim os dados direcionais do vento fornecidos pelo MRAMS com as nossas próprias observações dos efeitos eólicos no terreno (*wind streaks*, direções das barcanas, etc) e anexamos estes dados à nossa base de dados. O passo seguinte seria utilizar as redes neurais para, utilizando os dados das dunas como *input*, atingir as direções de transporte de sedimentos que definimos como *output*. Se se verificasse o sucesso das redes, podíamos usar a sua 'inteligência' para prever as direções de transporte de sedimentos de outras áreas, e comparar esses resultados com aqueles que seriam previstos utilizando os efeitos eólicos no terreno. Só um teste numa área diferente poderia ser conclusivo para o sucesso ou insucesso do método, uma vez que o treino e teste na mesma área serão sempre mais fáceis do que os testes das mesmas redes em áreas diferentes.

3.2.1 *Clustering* como método de classificação das dunas

Para além das redes neurais, e como outro dos objetivos do trabalho, tentou-se estimar qual o tipo de duna predominante em cada área a partir de uma análise com *clusters*. Para tal, foram utilizadas ferramentas próprias para o python com as características pretendidas para a classificação (Pedregosa et al. (2011)). Este conjunto de ferramentas, o *sklearn*, permitia diversos tipos de *clustering*, assim como o ajuste dos diversos parâmetros descritos no apêndice 6.1.2.1. Tal como foi explicado acima, será utilizada uma classificação gradual, não distinguindo os pontos das dunas apenas entre as duas ou mais classes existentes, mas sim apresentando uma pequena medida de proximidade do ponto ao centro da classe. Esta classificação gradual apenas pode ser feita corretamente com duas classes existentes, uma vez que, com três ou mais classes, se torna mais difícil quantificar as distâncias relativas entre as classes com apenas um valor numérico (seriam precisos dois ou mais, um para cada uma das outras classes que não aquela atribuída ao ponto a analisar).

A classificação de tipos de dunas pode ser feita não só com o algoritmo *K-means*, mas sim com outros métodos de classificação disponibilizados pelo software *sklearn* (Pedregosa et al. (2011)). Como tal, e para distinguir o melhor tipo de algoritmo de *clustering* para o nosso caso, foram testados métodos alternativos ao algoritmo *K-means*, mas todos apresentaram problemas, como limites de memória, ou resultados não correspondentes com a classificação esperada *à priori*, pelo que apenas se discutirão os resultados do algoritmo *K-means*. Os resultados do *clustering* apresentados são ainda preliminares, existindo ainda espaço para testar outros parâmetros e tipos de *clustering*, de forma a obter uma melhor classificação dos tipos de dunas (ver classificação geomorfológica das dunas na seção 2.2.3). Além disso, temos de efetuar uma filtragem maior nos campos que podemos utilizar como *inputs* para determinar as classes, especial-

mente se os mosaicos apresentarem diferenças significativas na textura devido a diferenças entre as imagens utilizadas para os criar. Tentou fazer-se o *clustering* utilizando parâmetros relacionados com aspetos texturais da imagem, mas esta abordagem foi abandonada posteriormente, sendo o algoritmo utilizado apenas com parâmetros geométricos das dunas, e não com medidas da textura da imagem. De facto, a informação geométrica do padrão das dunas é aquela que é fundamental para a foto-interpretação do seu significado geomorfológico.

3.2.2 Redes neuronais

As redes neuronais foram utilizadas nesta tese como uma ferramenta capaz de analisar as informações disponíveis dos campos de dunas e utilizá-las para prever as direções de transporte de sedimentos dominantes em cada campo de dunas. Esta análise foi efetuada utilizando a ferramenta pybrain para a linguagem de programação python, disponível online (Schaul et al. (2010)). O programa permite a criação de redes neuronais com diferentes tipos de camadas, com diferentes métodos para calcular os pesos relativos dos neurónios da rede neuronal. As redes inicialmente usadas para testar o código e verificar os resultados tinham uma camada sigmoide, com um número de neurónios escondido igual a duas vezes o número de neurónios de *inputs* disponível. O programa disponibiliza outros tipos de cálculo, estes foram utilizados inicialmente para treino. Foram testadas também redes com 4+ vezes o número de *inputs* como neurónios escondidos, assim como redes neuronais com mais do que três camadas. Mais informação sobre as arquiteturas de redes neuronais utilizadas nas seções 3.2.2.2 e 4.2.1. Além disso, é utilizado o algoritmo de '*Back Propagation*', como se pode ver na imagem 6.1 (explicado em mais detalhe no apêndice 6.1.1.1).

3.2.2.1 Dados do MRAMS e as Redes Neuronais

Uma vez que havia disponibilidade de dois conjuntos de dados do MRAMS (para as áreas de Gale e de Herschel), as redes neuronais foram primariamente treinadas singularmente em cada um desses dois campos, e depois aplicadas na área oposta e noutros campos de dunas disponíveis, mas sem dados do vento. O objetivo neste momento seria a construção da melhor rede neuronal possível, controlando os *inputs* e *outputs* ao nosso alcance. Obter resultados positivos numa zona diferente daquela onde a rede tinha sido treinada é o objetivo do método. Contudo, nem todas as direções de transporte de sedimentos no MRAMS coincidiam com as direções deduzidas visualmente a partir dos padrões dunares da imagem (a partir das dunas indicados na imagem 2.5). Isto significava que havia a necessidade de traçar manualmente as direções de transporte de sedimentos (ver imagem 3.5, onde se mostra um *close-up* de uma das zonas onde foram traçadas estas direções) observadas nas zonas em que havia a possibilidade de as distinguir, seja a partir das dunas ou a partir de marcas no terreno que mostram as direções preferenciais de transporte de sedimentos. Os dados do MRAMS, para além das direções de transporte de sedimentos predominantes numa dada área, indicam também o desvio padrão circular do vento nessa

mesma área. Infelizmente, era impossível corrigir ou melhorar de alguma forma esse parâmetro manualmente, pelo que os valores tomados como *output* serão, por definição, os valores do MRAMS nesse local.

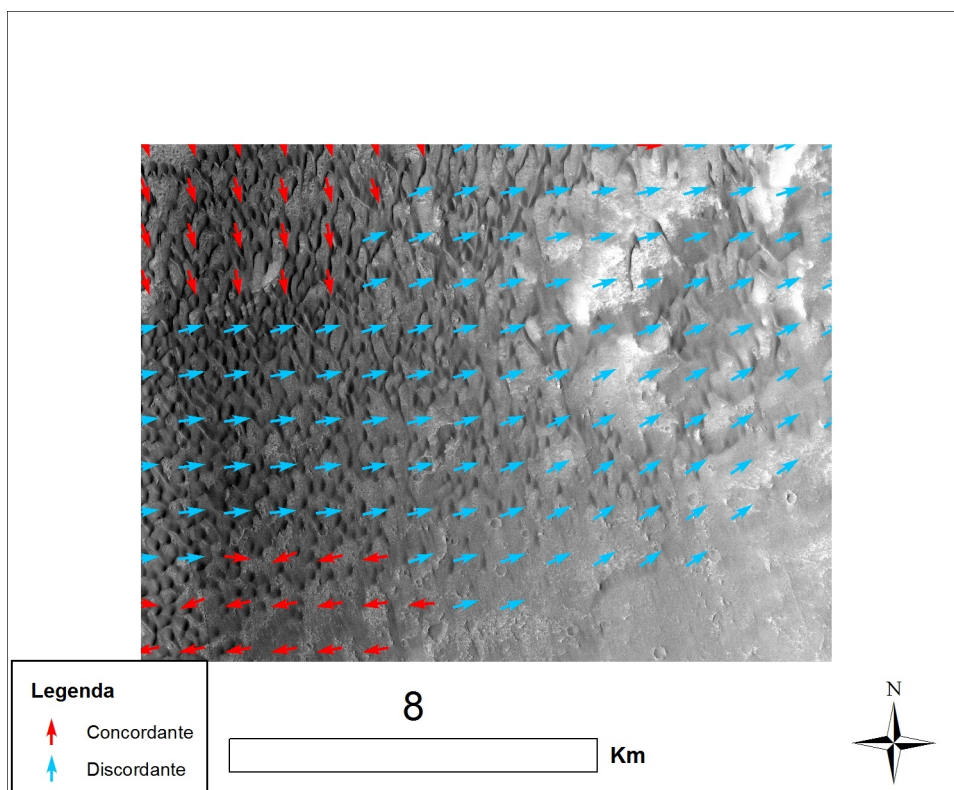


Figura 3.4: Close-up de uma zona da cratera Hershel onde as direções de transporte de sedimentos obtidas pelo MRAMS não coincidem com as deduzidas a partir da análise visual do terreno. Para corrigir este e outros problemas nas direções de transporte de sedimentos indicadas pelo MRAMS na cratera, foram introduzidas linhas do vento de forma 'manual', tomando o lugar dos dados do MRAMS quando estes discordavam das direções deduzidas pela forma das dunas no terreno.

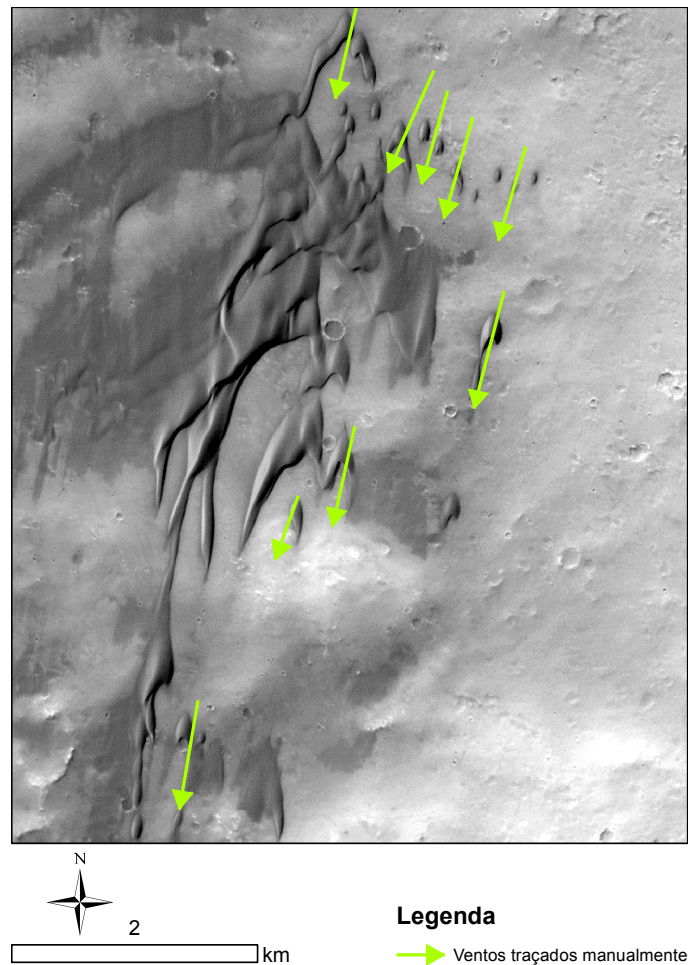


Figura 3.5: Imagem de Herschel onde foram assinaladas algumas direções de transporte de sedimentos, a partir das faces de deslizamento das barcanas e da distribuição da areia no terreno. A direção preferencial do vento, nesta zona e na maior parte da cratera, é de Norte-Sul.

Primariamente, depois um cálculo da direção média de transporte de sedimentos em cada ponto, condensando as informações visuais das dunas e os dados fornecidos pelo MRAMS, foi efetuado um teste geral à área a analisar, descartando as zonas onde a direção de transporte de sedimentos indicada não coincidia com a direção de transporte de sedimentos deduzida a partir de uma análise visual do terreno. Esta filtragem inicial está indicada na imagem 3.4. Os resultados obtidos foram então analisados, e elaboraram-se gráficos indicando a direção principal de transporte de sedimentos predominante na região, tanto a inicial fornecida pelo MRAMS como a calculada pelo modelo.

3.2.2.2 Arquiteturas de rede neuronal e campos de dunas

O objetivo, quando se utilizam redes neuronais, é a utilização da rede com arquitetura mais simples que traz melhores resultados. Contudo, a rede mais simples disponível pode nem sempre ser a rede ideal para o nosso trabalho. Desta forma, para melhorar os resultados obtidos, tiveram de se realizar testes a várias arquiteturas de redes neuronais disponíveis, e a outros métodos de filtragem de *inputs* e *outputs*, de forma a concluir qual seria o melhor conjunto de parâmetros de redes neuronais disponíveis para o trabalho a efetuar. Assim, foram escolhidos alguns tipos de redes neuronais, variando parâmetros como o número de camadas, número de neurónios por camada, e também desenvolvidos alguns critérios de seleção de áreas no campo de dunas, de forma a obter os melhores resultados que as redes neuronais são capazes de obter numa determinada zona. Foram, ainda, testados campos treinados apenas numa das áreas ou num conjunto de dados representativo das duas áreas em conjunto. Os testes foram efetuados através da construção de histogramas e testes estatísticos aos resultados obtidos pela arquitetura de rede neuronal testada. As arquiteturas mais promissoras foram então aplicadas a outros campos de dunas e avaliado o seu desempenho nessas áreas. Contudo, este esforço pode ser desenvolvido com mais tempo e disponibilidade, uma vez que a variedade de arquiteturas de rede neuronal e parâmetros ajustáveis é bastante extensa. Aqui apenas foi testada uma pequena variedade, uma vez que o objetivo desta tese era apenas a demonstração da possibilidade do método e não a sua exploração completa.

Capítulo 4

Apresentação e análise de resultados

4.1 *Clustering* - resultados obtidos

Vamos agora observar os resultados que foram obtidos pelos algoritmos. Serão mostrados os resultados do campo de dunas de Aonia (ver seção 2.4.2), uma vez que esta zona apresenta maior diversidade no tipo de dunas e permite uma classificação mais abrangente das mesmas. Estes algoritmos foram corridos com duas e com três classes possíveis, e os resultados foram verificados efetuando uma análise visual dos tipos de dunas e comparando a triagem efetuada pelo algoritmo com uma efetuada manualmente. Como *input* (posições centrais a testar com o algoritmo *k-means*, ver seção 6.1.2.1) foram utilizadas as coordenadas médias dos pontos de cada uma das classes de dunas. Os resultados foram mais positivos para a distinção entre apenas duas classes de dunas.

Número de classes	\hat{k}	Sum(Diagonal)/ Total (precisão)	Dunas corretas/ classificação manual (recall)
2	0.74	0.79	0.79
3	0.13	0.54	0.55

Tabela 4.1: Resumo dos resultados obtidos pelo algoritmo *Kmeans* como método para diferenciação dos tipos de dunas, com duas e três classes disponíveis. De realçar o baixo valor de \hat{k} obtido para 3 classes, por comparação com duas classes, e os valores relativamente elevados da precisão e do *recall* nos dois casos.

Podemos ver pelos resultados obtidos (imagens 4.1 a 4.4 e tabela 4.1) que a segmentação com *clusters*, realizada com o intuito a diferenciar tipos de dunas

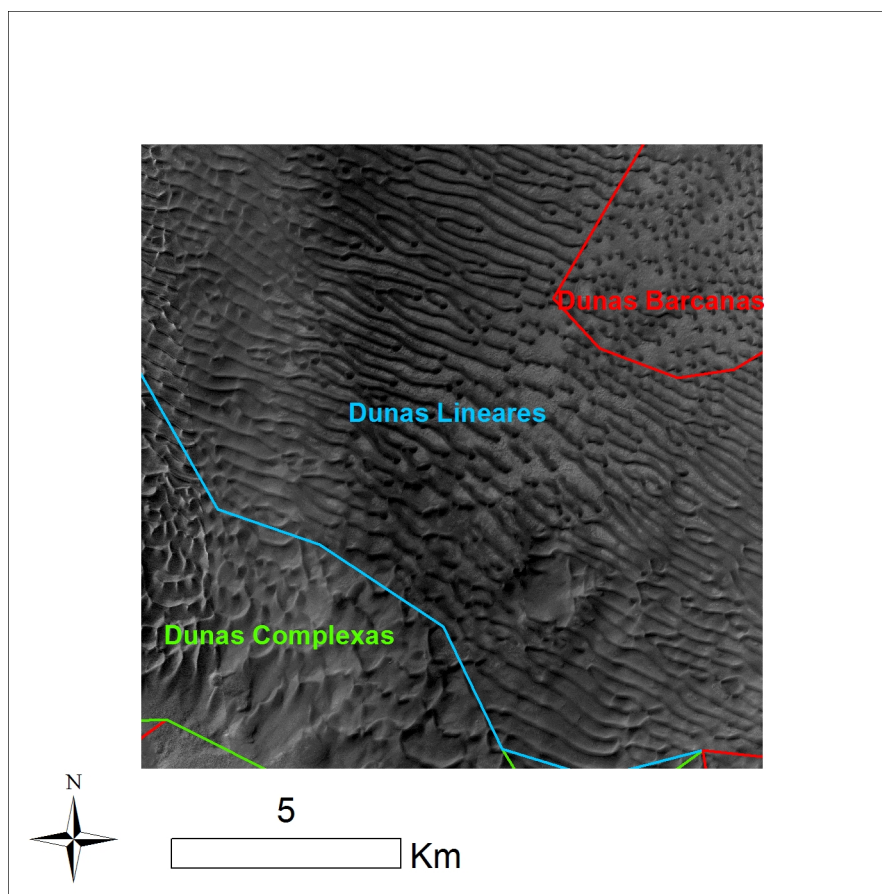


Figura 4.1: Classificação manual das dunas na zona de Aonia, tal como a exemplificada na imagem 2.14. A classificação foi feita manualmente e de acordo com o aspeto visual das dunas. O objetivo desta classificação rude dos tipos de dunas foi uma avaliação mais objetiva dos resultados da aplicação dos algoritmos de *clustering*. Esta classificação foi efetuada dividindo o campo de dunas de acordo com o tipo de duna dominante em cada área, de forma a utilizar um algoritmo de *clustering* com 3 classes. Com um algoritmo de *clustering* dividindo o campo de dunas em apenas duas classes, foram unidas as dunas complexas e barcanas, e apenas diferenciadas estas das lineares.

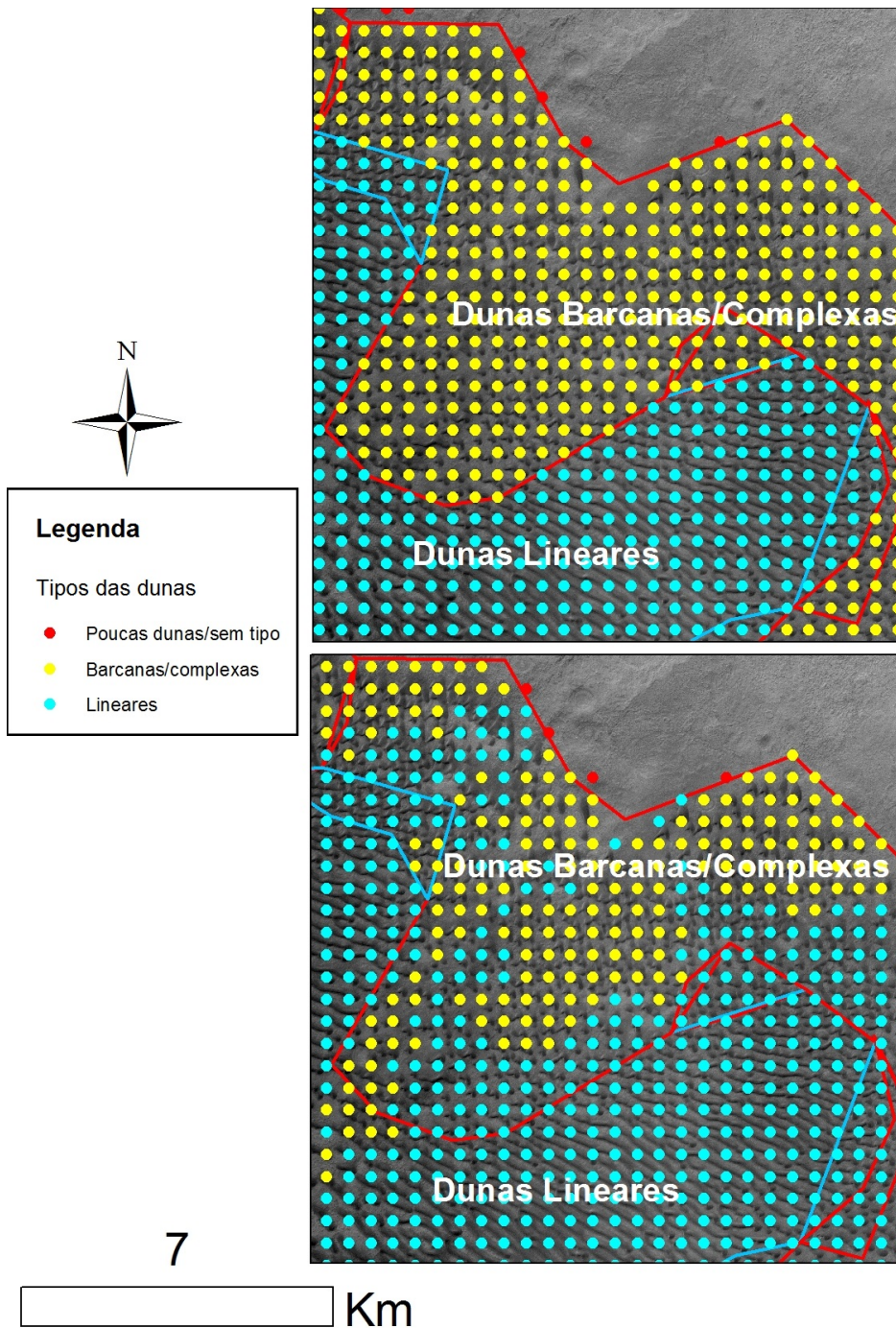


Figura 4.2: Resultados da aplicação do algoritmo de *clustering k-means* com duas classes no campo de dunas Aonia (*close-up*). De realçar as previsões corretas na maior parte das zonas onde existem dunas lineares, assim como a interpretação correta das zonas onde as barcanas são predominantes. Apenas se distinguem alguns erros na zona NW e SE da imagem, onde dunas claramente barcanas se encontram assinaladas como se de dunas lineares se tratassem.

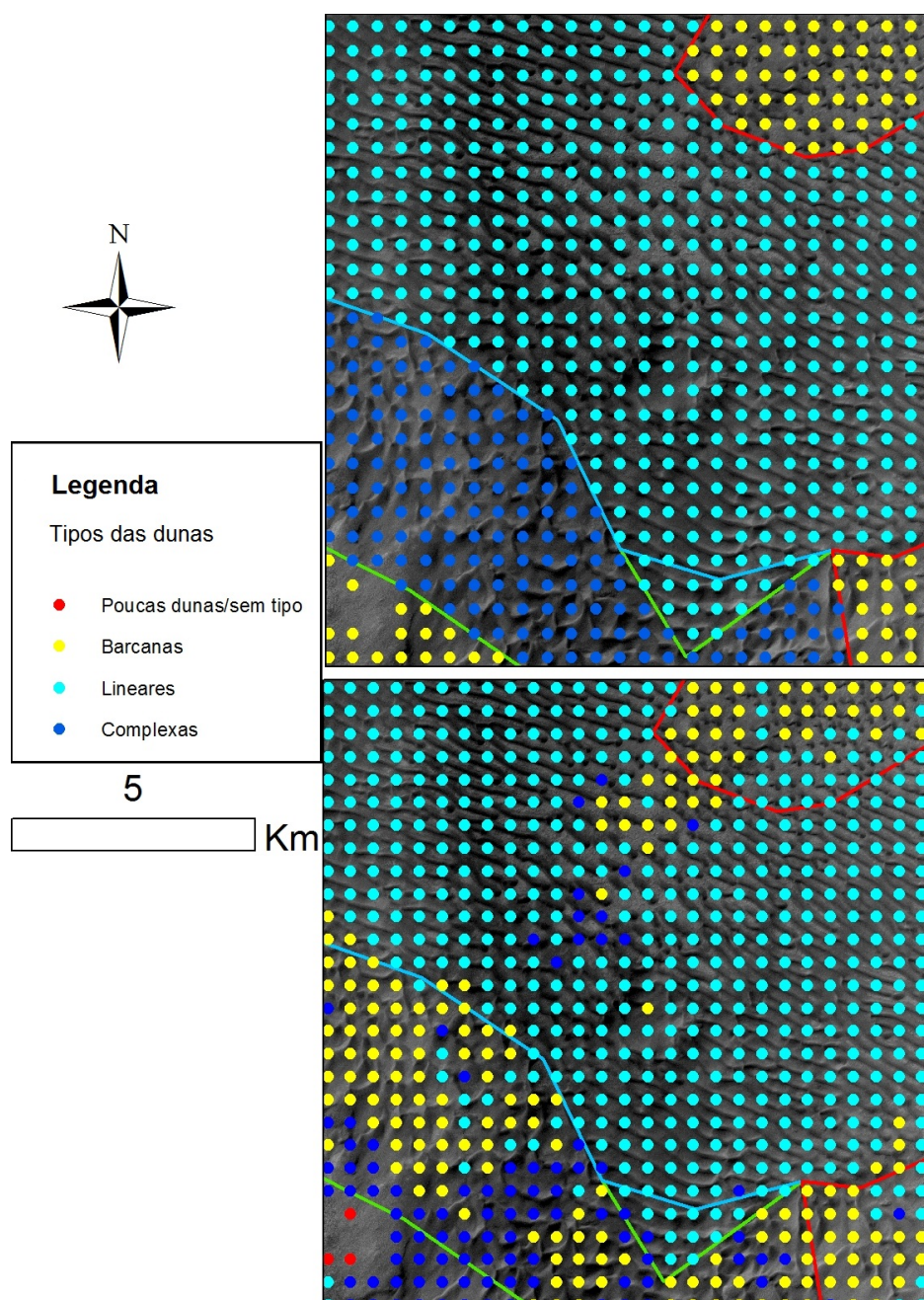


Figura 4.3: Resultados da aplicação do algoritmo de *clustering k-means* com três classes no campo de dunas Aonia (*close-up*). As dunas lineares, tal como no exemplo acima, são corretamente identificadas na maior parte dos casos. Existe, contudo, uma confusão significativa entre as dunas complexas e as barcanas/barcanóides. Esse foi também um dos motivos porque se deu preferência à classificação das dunas com apenas duas classes, aglomerando esses dois tipos de dunas na mesma classe, em vez de os tentar diferenciar.

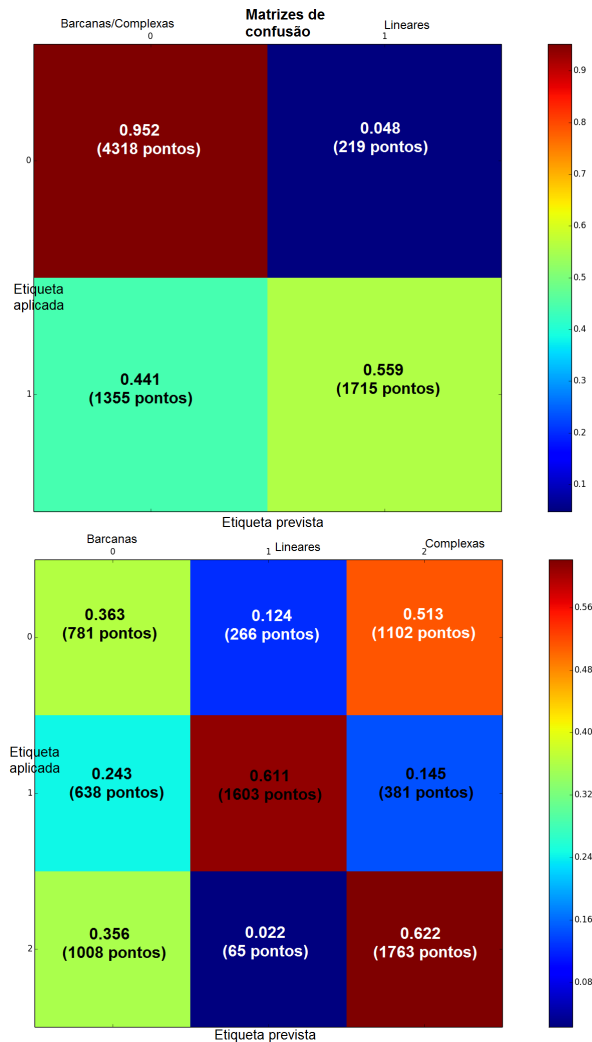


Figura 4.4: De cima para baixo: 1: Matriz de confusão depois de aplicado o algoritmo *K-means* com 2 classes ao campo de dunas de Aonia. De realçar a capacidade muito elevada de identificar corretamente as dunas complexas/barcanas, com uma precisão de 95%, e a capacidade inferior, mas positiva, de identificar dunas lineares (56%, aproximadamente). Estes resultados estão longe do perfeito, mas demonstram a possibilidade da utilização de métodos de *clustering* como uma ferramenta para classificar as dunas consoante o seu tipo; 2: Matriz de confusão depois de aplicado o algoritmo *K-means* com 3 classes ao campo de dunas de Aonia. De realçar a capacidade muito fraca para identificar corretamente dunas barcanas, e que a precisão nos outros dois casos mal ultrapassa os 60%. Uma vez que a mistura de resultados se deu principalmente entre as dunas barcanas e complexas (no caso das barcanas apenas 36% das dunas foram identificadas corretamente, percentagem similar a uma escolha ao acaso), esses foram os tipos de dunas aglomerados quando se correu o algoritmo com apenas duas classes.

é difícil e pode resultar em erros de classificação significativos. Contudo, se diminuirmos o número de classes em jogo, podemos obter resultados satisfatórios e retirar uma análise útil do conjunto de dunas a analisar. Apesar disso, este tipo de análise mantém-se como uma ferramenta a utilizar cuidadosamente, e apenas nos campos de dunas adequados para o efeito, uma vez que um número de classes exagerado pode levar a resultados diferentes dos esperados.

4.2 Redes Neurais - Resultados obtidos

Depois de construídas as redes neuronais, estava na altura de avaliar os seus resultados iniciais. Como se pode ver nas imagens 4.5 e 4.6, onde temos dois exemplos de resultados calculados pelas redes neuronais, para duas áreas disponíveis, existe uma concordância geral com as direções de transporte calculadas pelas redes neuronais e aquelas utilizadas como *input* nessas redes. Além disso, como se confirma pela imagem 4.7, o padrão geral dos desvios padrão circulares é mantido depois de aplicadas as redes neuronais. Contudo, e como já foi mencionado na seção 3.2.2.2, é importante realizar um estudo sobre os tipos de arquiteturas de redes disponíveis, pelo que se compararam resultados obtidos para diferentes campos e com diferentes arquiteturas de redes. Esta comparação será efetuada na seção 4.2.1.

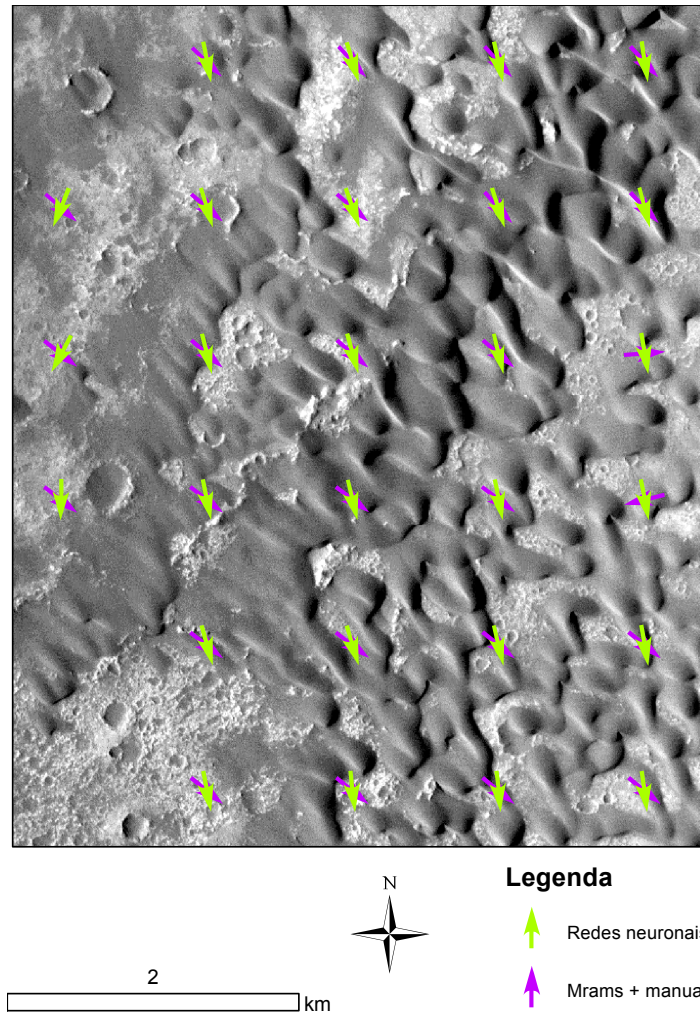


Figura 4.5: Exemplo duma área da cratera Herschel onde foram aplicadas as redes neuronais de forma a reproduzir os dados das direções de transporte de sedimentos (compilação dos dados do MRAMS e das direções de transporte de sedimentos traçadas manualmente). De realçar a concordância geral nas direções do MRAMS e das redes neuronais na área assinalada.

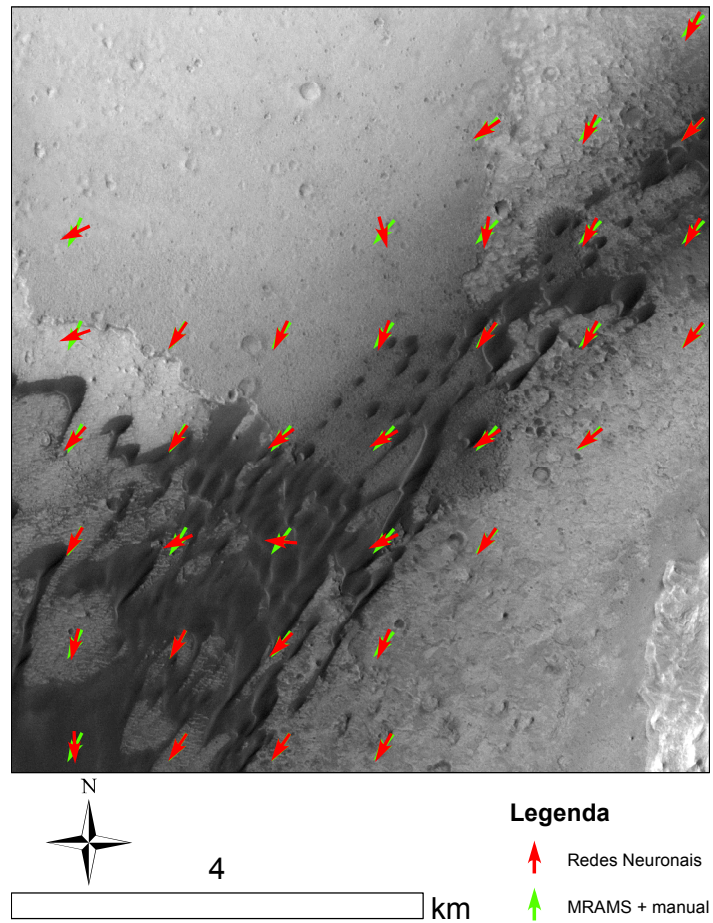


Figura 4.6: Resultados das redes neuronais na cratera Gale. Mais uma vez, existe, à primeira vista, uma grande concordância entre os resultados das redes neuronais e as direções esperadas a partir dos dados do MRAMS e as direções de transporte de sedimentos traçadas manualmente. Estes resultados foram obtidos com uma rede neuronal com 3 camadas, com número de neurónios em cada uma igual a $t-2t-3$, respetivamente.

4.2.1 Arquitetura de redes - resultados

Foram então testadas as diferentes redes neuronais com o programa criado. Os parâmetros usados para avaliar a qualidade de uma dada rede neuronal são:

- r^2 - Coeficiente de determinação, entre -1 e 1, avalia a correlação entre os valores do desvio padrão da direção média de transporte do MRAMS e os da rede neuronal, medindo a soma dos quadrados das distâncias entre os pontos dos dois conjuntos

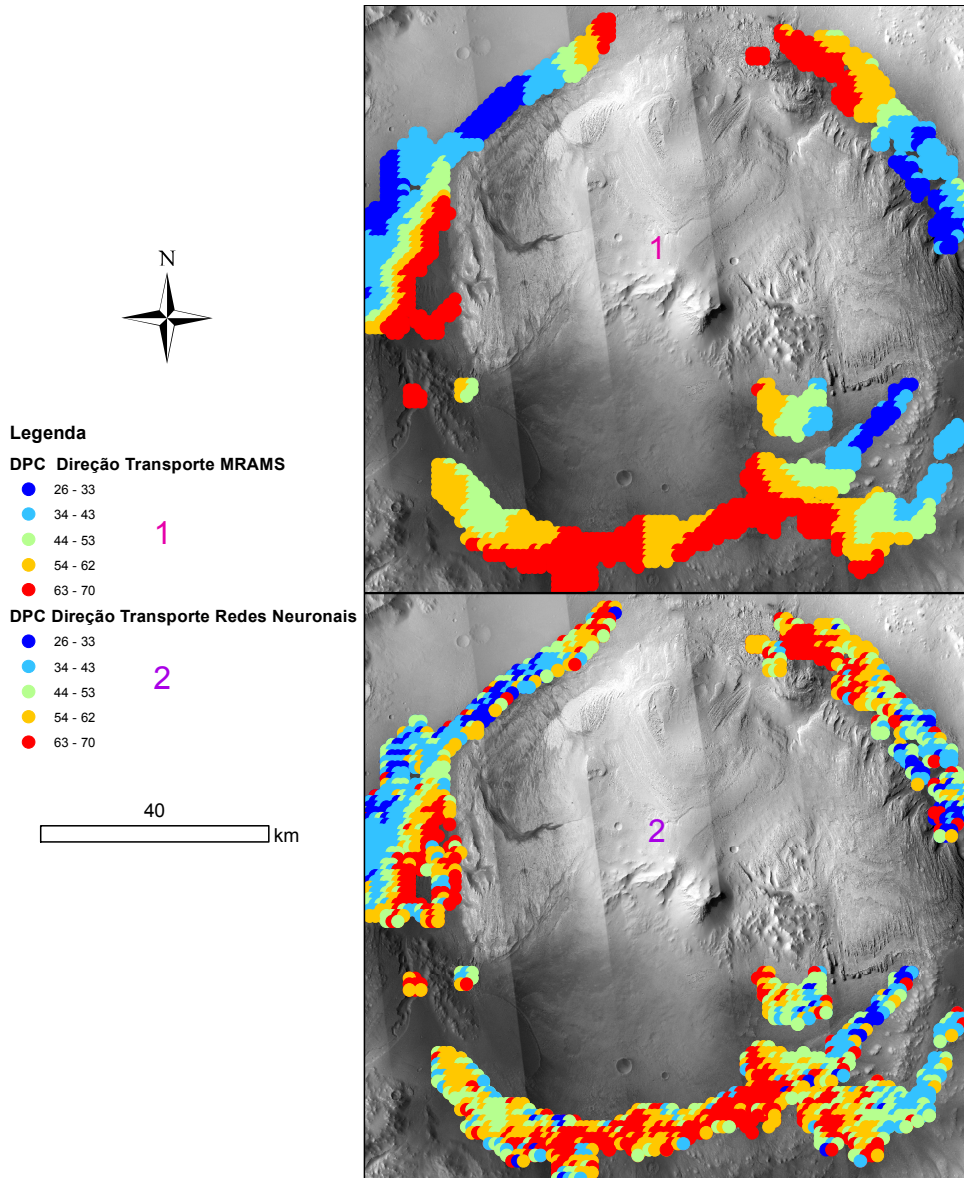


Figura 4.7: Representação dos resultados na cratera Gale da aplicação de redes neurais para calcular o DPC (Desvio Padrão Circular) nos campos de dunas, a partir de dados do MRAMS. Os valores apresentados estão em graus e dão-nps indicações da variabilidade ao longo do ano das direções dos ventos na superfície. De salientar o efeito visual e grande semelhança nos padrões de cores entre os dados de entrada e de saída das redes neurais.

- A diferença média, em graus, entre o desvio padrão de transporte de sedimentos de cada um dos pontos do MRAMS e os da rede neuronal
- Uma avaliação visual do histograma com as direções angulares de transporte previstas pela rede neuronal
- A diferença média, em graus, entre a direção de transporte de sedimentos de cada um dos pontos do MRAMS e os previstos pela rede neuronal
- A diferença entre a direção média (de todos os pontos no campo de dunas) prevista pelo MRAMS e pela rede neuronal

Os resultados encontram-se apresentados na tabela seguinte:

Campo		Rede		Desvio Padrão		Direções	
Treinada	Testada	Camadas	Neuronios	r ²	Diferença	Med Dif Ang	Ang Tot Dif
Gale	Gale	4	t-t/2	0,425	6,63	39,3	0,1
Gale	Gale	3	2t	0,419	6,68	36,6	10,51
Gale	Gale	4	t-t/3	0,428	6,67	43,06	9,4
Gale	Gale	4	t-t	0,42	6,52	43,33	16,37
G+H	Gale	4	t-t	0,359	7,6	45,34	4,01
G+H	Herschel	4	t-t	0,34	3,6	26,92	0,6
Gale	Herschel	4	t-t/2	0,0005	19,33	36,24	15,68
Herschel	Herschel	4	t-t/2	0,62	2,08	26,19	2,94
Herschel	Herschel	3	t	0,54	2,47	27,96	7,94
Herschel	Herschel	4	t-t	0,64	2,03	25,79	14,18

Tabela 4.2: Tabela contendo uma compilação dos resultados de várias redes neuronais, treinadas e testadas nos dois campos de dunas, Gale e Herschel. Os valores angulares estão todos medidos em graus. Nos campos dos resultados, 'Diferença' indica a diferença média entre os desvios padrões do MRAMS e da rede, para todos os pontos da área dunar, 'Med Dif Ang' indica a média das diferença entre as direções de transporte de sedimentos do MRAMS e da rede neuronal, para todos os pontos da área dunar, e 'Ang Tot Dif' indica a diferença entre o vetor médio da direção de transporte de sedimentos para toda a área duna do MRAMS e o medido pela rede neuronal. 'G+H' significa um conjunto de dados incluindo as dunas tanto de Gale como de Herschel.

Como se pode ver na tabela 4.2, a arquitetura da rede escolhida para uma dada simulação e o campo onde esta é treinada pode influenciar de forma significativa os resultados obtidos na simulação. Podemos ver que a diferença média entre as direções de transporte de sedimentos nos pontos individuais do MRAMS e as previstas pelas redes neuronais não ultrapassa os 45°, mesmo na pior das redes utilizadas, pelo que os resultados estão no nível esperado. Em termos de diferença entre o ângulo médio da direção de transporte de sedimentos para todo o campo de dunas, esta diferença é ainda mais pequena, não ultrapassando os 20° no pior dos casos. Este tipo de resultados é bastante promissor, se se mantiver para outros campos de dunas. No que toca ao desvio padrão, temos uma

correlação sempre positiva (> 0.35) entre os dados das redes e do MRAMS, exceto no caso da rede treinada em Gale e testada em Herschel. Os resultados são melhores para os dados de Herschel, tanto para a correlação como para a diferença média entre os pontos do MRAMS e da rede neuronal, o que será explicado não só pela muito maior quantidade de dados de *input* disponíveis (4800 vs 1900), mas também pela maior uniformidade nas direções preferenciais de transporte de sedimentos nessa zona.

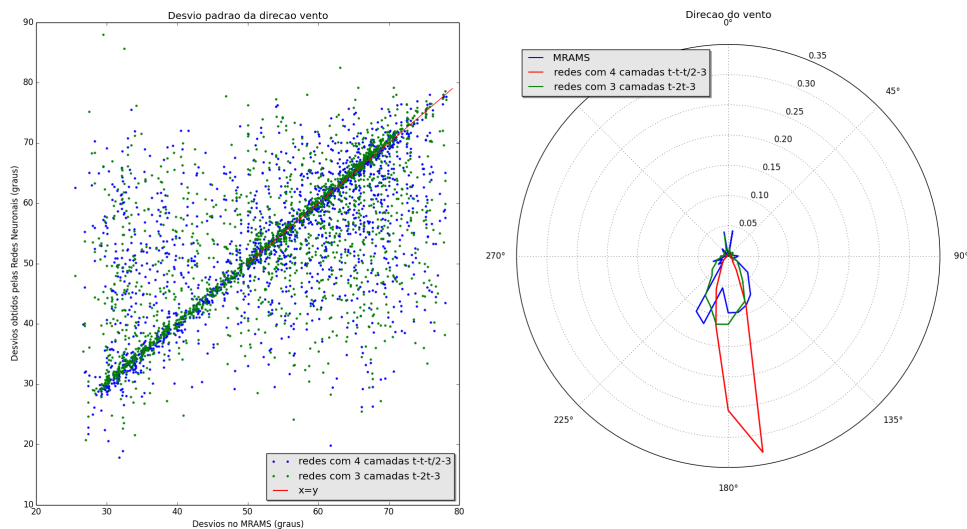


Figura 4.8: Comparação inicial entre duas arquiteturas de rede neuronal na cratera Gale. Estão disponíveis um histograma polar das direções calculadas de transporte de sedimentos (lado direito) e uma representação dos valores obtidos pelas duas redes neuronais para o desvio das direções de transporte de sedimentos, comparando esses resultados com os valores do MRAMS (lado esquerdo). Neste caso, pode notar-se uma tendência linear, se bem que com grande variação, no gráfico que mostra os resultados do desvio das direções de transporte de sedimentos (valores de $r^2 = 0.425$ para rede com 4 camadas e $r^2 = 0.419$ para a rede com 3 camadas). No caso do gráfico polar, observa-se uma distribuição de valores do MRAMS com duas modas proeminentes, com direções aproximadas de 160° e 205° (coordenadas geográficas - Norte/Sul) e uma moda menos intensa, com valores entre 350° e 10° (Sul/Norte). Visualmente, a rede com 4 camadas não conseguiu reproduzir essas modas com as direções Sul/Norte, ao passo que a rede com 3 camadas apresentou alguns valores nesse intervalo. Dessa forma, e à primeira vista, a rede com 3 camadas será melhor a reproduzir os valores do que a rede com 4 camadas, que ficou 'viciada' na direção Norte/Sul.

Pela tabela, a melhor rede a utilizar será com certeza a treinada no próprio campo, mas a rede treinada nos dois campos em conjunto não fica muito

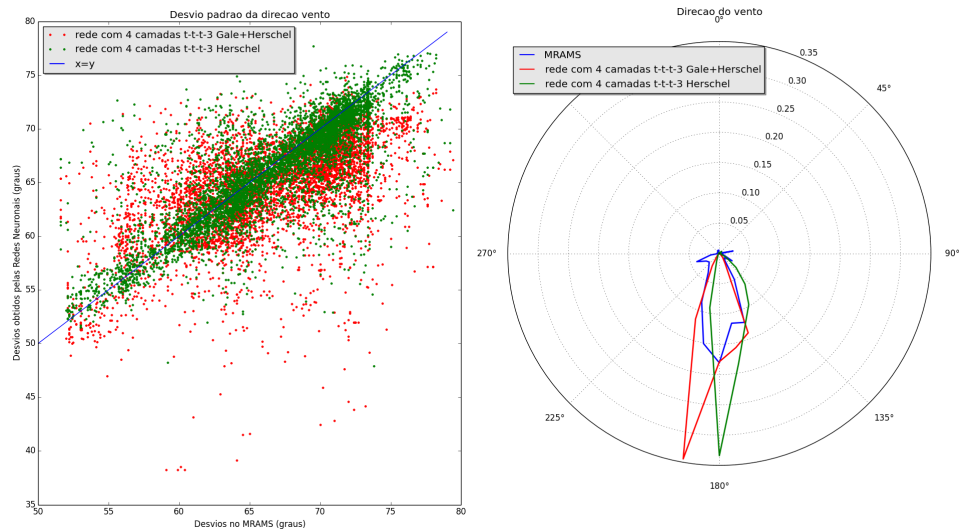


Figura 4.9: Comparação dos resultados para duas redes neurais. De notar o número muito superior de pontos visível na nuvem de pontos representativa desta cratera (a área é muito superior, e existem 3x mais pontos disponíveis para a análise). Nesta imagem podemos comparar duas redes neurais com os mesmos parâmetros, mas treinadas em áreas diferentes. A rede a vermelho foi treinada com pontos tanto da cratera Gale como de Herschel, enquanto que a rede ilustrada a verde foi treinada apenas com pontos da cratera Herschel. O histograma a azul representando os dados iniciais de transporte de sedimentos utilizados como *output* esperado para a rede mostram que nesta cratera temos uma direção preferencial de transporte de sedimentos no sentido Norte/Sul ($\sim 180^\circ$), e uma direção secundária no sentido Este/Oeste ($\sim 250^\circ$). Podemos ver que o histograma é ligeiramente mais fiel aos dados do MRAMS no caso da rede treinada com as duas áreas (apesar de nenhum deles reproduzir fielmente a direção Este/Oeste), enquanto que os pontos do desvio da direção de transporte de sedimentos parecem melhor no caso da rede treinada apenas em Herschel. Com a tabela abaixo podem confirmar-se estas impressões. Conclui-se que a rede treinada nas duas áreas poderá ser mais fiel em termos de direção de transporte de sedimentos, mas não prevê com tanta precisão o desvio dessas direções.

atrás. A pior, como se seria de certa forma de esperar, é a rede treinada no campo oposto. Como o nosso objetivo é a aplicação destas redes a outros campos de dunas, utilizar-se-á uma rede neuronal com o maior número possível de pontos de treino, ou seja, utilizando os dois campos como *inputs*. Entre as arquiteturas testadas, chegou-se à conclusão que a melhor rede será uma com 4 camadas, incluindo a de *input* e a de *output*, e com duas camadas intermédias com t neurónios, em que t será o nosso número de *inputs*. Esta rede pode não ser a mais otimizada para estas medidas, mas é seguramente a melhor entre aquelas testadas. Para mais desenvolvimentos na área sugere-se um estudo com uma maior variedade de redes neuronais, de forma a concluir qual será, de facto, a rede 'perfeita' para este objetivo. Além disso, seriam necessários mais campos de dunas a introduzir como *input*, de forma a obter amostras mais representativas da população de dunas marciana, com direções preferenciais de transporte de sedimentos espalhadas igualmente pelas diferentes direções cardiais. Só dessa forma seria possível obter resultados capazes de prever as direções de transporte de sedimentos em todas as áreas de teste existentes no planeta, mesmo que a qualidade das previsões individuais para um dado campo de dunas diminuísse.

Resumindo, e de acordo com os resultados obtidos, a rede final e que servirá para testar outros campos de dunas (neste caso, o de Aonia) será a rede com *inputs* dos dois campos de dunas ('G+H') e com 4 camadas.

4.2.2 Comparação dos resultados obtidos com o *input* (MRAMS)

Foram já explorados e apresentados alguns resultados obtidos com as redes neuronais na seção 3.2.2.2. Contudo, todos esses resultados refletem a utilização de redes neuronais em campos onde estas próprias foram, de alguma forma, treinadas. Abaixo, na imagem 4.10, está ainda um resultado da aplicação destas redes a campos onde elas não foram treinadas, de forma a fornecer uma ideia dos resultados por elas produzidas sem treinos específicos com direções preferenciais de transporte de sedimentos.

A partir destes resultados, e em linha com aquilo que se pôde observar em relação às redes neuronais nos outros campos de treino (Gale, Herschel, ver imagens 4.5 e 4.6, ou ainda os histogramas das direções de transporte de sedimentos nesses campos, nas imagens 4.8 e 4.9) podemos concluir que é provável que exista um enviesamento natural das direções de transporte de sedimentos calculadas pelas redes neuronais. Ou seja, se o campo de treino apresenta direções preferenciais (neste caso, maioria dos ventos na direção Norte-Sul) os resultados nos campos de teste vão sofrer esse mesmo enviesamento, mesmo que as dunas nesse caso não indiquem essas direções de transporte de sedimentos.

4.2.3 Comparação dos resultados obtidos com o estado da arte da análise de dunas - MGDDD

Para termos uma apreciação e compreensão deste trabalho, é necessária uma comparação dos resultados obtidos, em termos de resolução e de precisão das direções indicadas de transporte de sedimentos e das faces de deslizamento, com

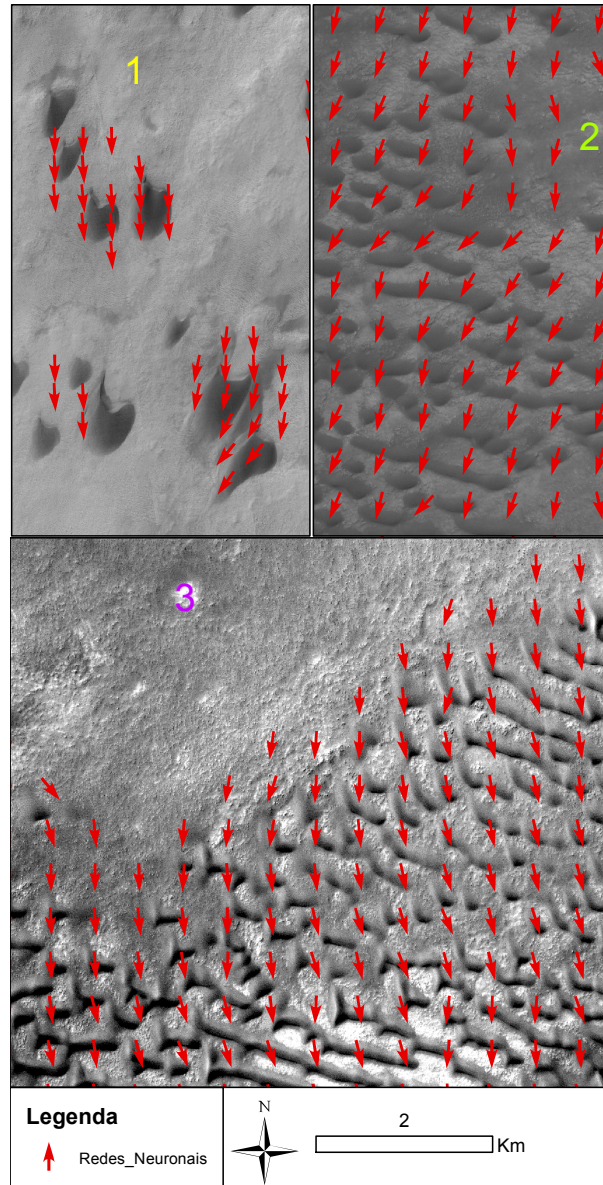


Figura 4.10: Ilustração do campo de dunas de Aonia, com os vetores assinalados pertencendo a uma rede neuronal treinada nos campos de Gale e Herschel. Estão visíveis três áreas do campo de dunas. Como se pode ver, principalmente na imagem 1, este tipo de dados não corresponde exatamente à direção de transporte que se pode inferir a partir das barcanas (~Sul-Norte, enquanto que as direções calculadas apresentam Norte-Sul).

o estado da arte do tema. Ou seja, é importante concluir em que aspetos pode este tipo de análises melhorar o trabalho já existente da MGDDD em termos de direções de transporte de sedimentos e de caracterizações de campos de dunas. Estes dados estão disponíveis online, e procedeu-se então à comparação de resultados.

Como se pode confirmar na figura 4.11, os dados da MGDDD têm uma resolução muito inferior aos nossos uma vez que resultam da foto-interpretação dos mosaicos THEMIS com uma resolução espacial de 100 m/pix. Em relação às direções de transporte de sedimentos propriamente ditas, podemos confirmar que existem algumas discrepâncias entre os dois casos, uma vez que os dados disponíveis na MGDDD apenas ilustram a direção entre o centro da cratera e o centro do campo de dunas. Essa aproximação é muito básica e não consegue ilustrar corretamente o comportamento do vento na região, nem tem em consideração as direções de transporte ou das faces de deslizamento das dunas. Ainda assim, podemos comparar os dados das frentes de avalanche do campo de dunas (vetores a roxo) com os dados das redes neuronais, e testar dessa forma os nossos resultados com análises independentes do mesmo campo de dunas. Estas frentes de avalanche digitalizadas manualmente infelizmente não coincidem exatamente com os locais onde estão traçadas as dunas. Este problema surge muito provavelmente devido ao facto de não serem utilizados mosaicos como fonte da imagem onde são traçadas as frentes de avalanche, e sim apenas imagens individuais. Além disso, as imagens onde foram traçadas estas frentes de avalanche têm uma resolução inferior à disponível no nosso mosaico da zona, pelo que existe também uma incerteza na direção traçada associada a esse facto. Podem ainda existir interpretações diferentes das direções de transporte numa mesma duna por examinadores independentes, pelo que alguma da discrepância pode também provir desse facto. Finalmente, todos estes dados são recolhidos remotamente a partir de satélites, sem análises diretas no terreno, e sem uma medição em primeira mão das direções de transporte de sedimentos, apenas induções a partir da forma das dunas, o que pode induzir em erro. Portanto, os resultados que apresentamos aqui devem ser analisados com alguma cautela e atenção, uma vez que uma diferença entre os dados das redes neuronais e os da MGDDD não indica necessariamente que os primeiros estejam errados e os segundos corretos.

Analisando a tabela 4.3 e a imagem 4.13, podemos chegar a algumas conclusões relativamente aos dados publicados pela MGDDD e sobre os dados obtidos pelas redes neuronais nos diferentes campos. O primeiro ponto que salta à vista é a enorme disparidade nos dados de Gale, com uma diferença entre as direções médias da MGDDD e das redes neuronais de $133,72^\circ$, muito superior ao esperado. Esta diferença surge de um problema que se pode notar na imagem 4.12. Nessa imagem, nota-se que as frentes de avalanche fornecidas pela MGDDD apresentam uma tendência Sul-Norte, verificada nessa zona do campo de dunas, mas que não está presente nos nossos dados provenientes das redes neuronais. Esta discrepância e enviesamento na direção Sul-Norte nos dados da MGDDD, que não surge nos dados utilizados como *input* do MRAMS para essa área, provocam a diferença visível entre esses dados e os das redes neuronais.

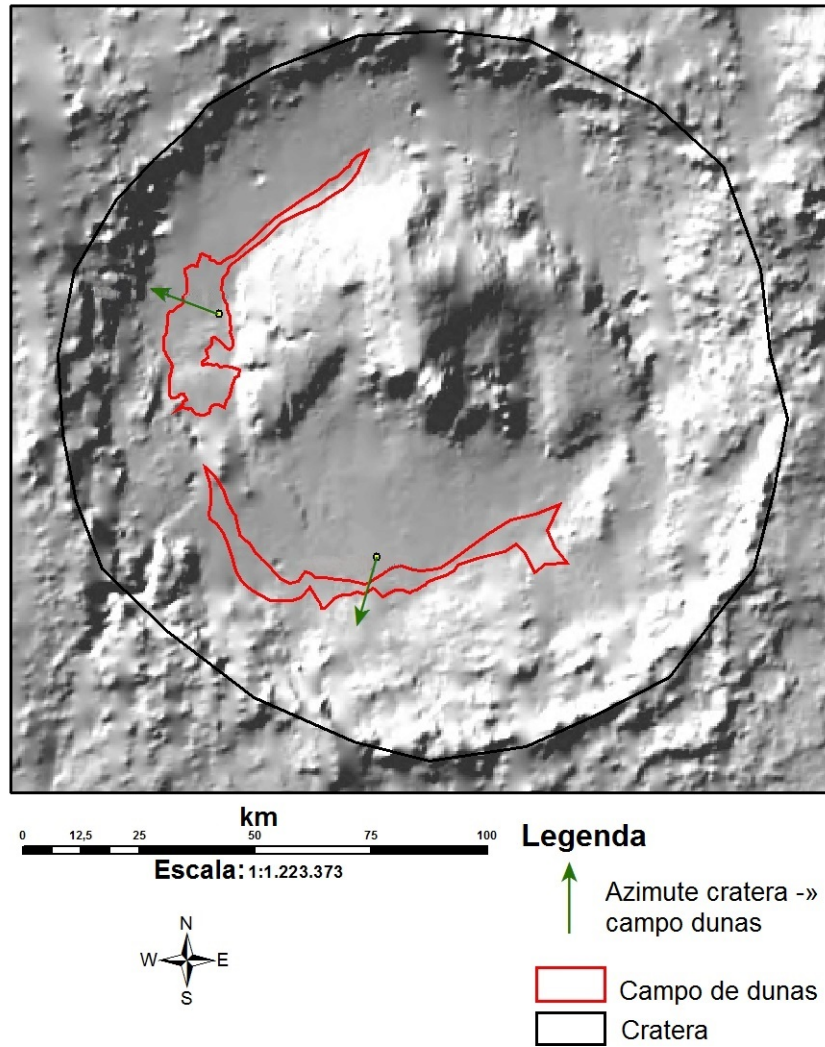


Figura 4.11: Dados disponíveis da MGDDD sobre a cratera Gale, comparar com a figura 2.11. De notar a fraca resolução e a indicação de apenas um vetor predominante em cada campo de dunas, bem como a ausência de dunas marcadas no lado Nordeste da cratera. É de notar o facto que a MGDDD possui alguma informação sobre os faces de deslizamento individuais das dunas que não está assinalada na imagem, devido ao seu tamanho reduzido, que ajudam a caracterizar os campos de dunas. Os vetores assinalados a verde traçam a direção desde o centro da cratera até ao centro do campo de dunas, e pretendem ser uma representação possível da direção de transporte de sedimentos dominante nesse campo de dunas.

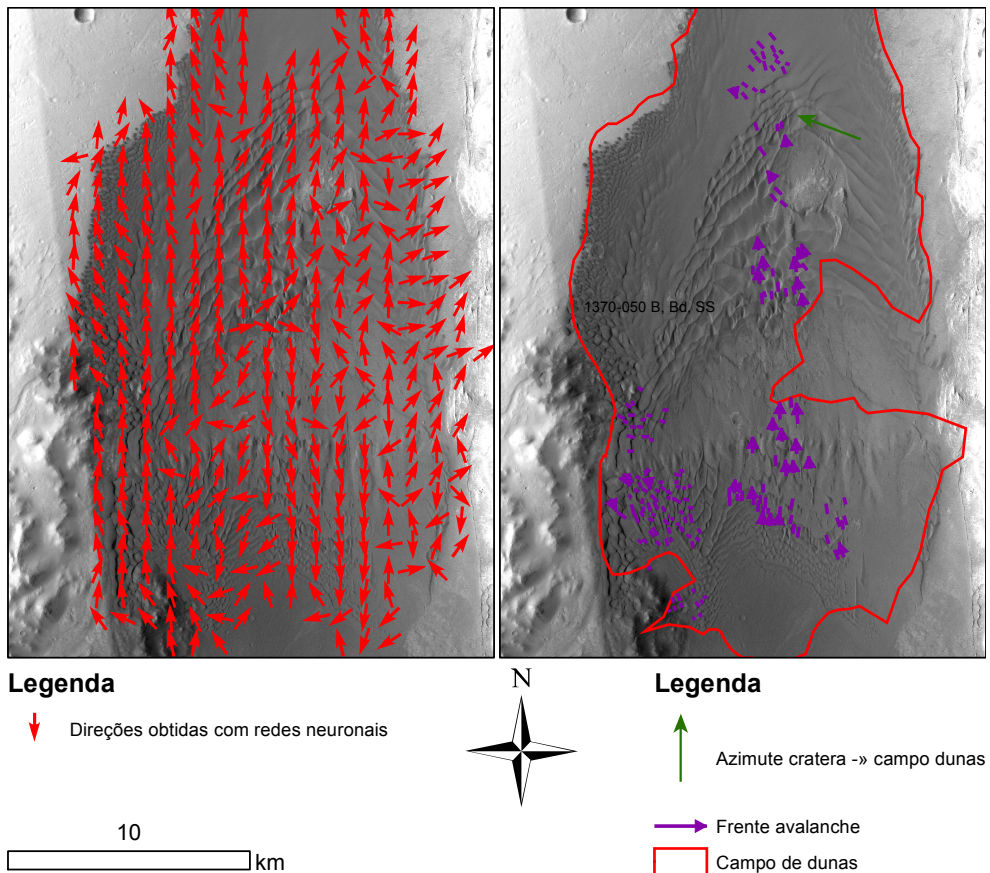


Figura 4.12: Comparação de alguns dos resultados obtidos com as redes neuronais (neste caso, 3 camadas com 2t neurónios, ver seção 3.2.2.2 para mais informação) para campos de duna na cratera Gale com os dados da mesma área na MGGDD. O mosaico utilizado nas imagens foi criado especificamente para este trabalho, a resolução da imagem disponibilizada pela MGGDD era bastante pior (a exemplo da imagem 4.11), pelo que foi utilizado o nosso mosaico para maior clareza na apresentação dos dados.

Área	Campo Dunas	MGDD			Rede vs MGDD				
		Azim Campo	Frentes Avalanche		Redes		Dif Direções cada ponto		
		Média	Média	v	Média	v	Média	v	Dif Dir Médias
Gale	Norte	290,64	82,45	94,8	200,32	71,0	90,03	76,2	133,72
	Sul	195,81	sem dados	sem dados	190,46	68,2	sem dados	sem dados	sem dados
Aeonia	1	336,87	226,74	57,5	174,24	51,0	72,46	45,3	60,4
	todos	186,2	170,01	39,5	183,12	27,9	27,18	25,8	6,51
	1	45,5	235,9	18,72	167,4	50,7	63,84	20,1	63,81
	2	262,6	151,6	23,95	170,6	31,1	16,66	12,0	6,1
	3	284,5	166,1	9,21	170,83	34,0	7,2	4,6	4,29
Herschel	4	294,4	149,8	20,22	170,2	49,2	31,67	16,0	28,8

Tabela 4.3: Comparação direta das frentes de avalanche apresentadas pela MGDD, as direções azimutais entre os centros das crateras e os campos de dunas, e os resultados obtidos pelas redes neuronais nas mesmas áreas. O valor apresentado v é uma medida da desvio padrão circular, apresentada no livro Jupp (2000), e corresponde a $v = \{-2 \log \bar{R}\}^{1/2}$, onde \bar{R} corresponde ao comprimento resultante médio, obtido pela equação $\bar{R} = (\bar{C}^2 + \bar{S}^2)^{1/2}$, onde \bar{C} e \bar{S} são os valores médios dos senos e cossenos, respectivamente. Esse valor é calculado em radianos e posteriormente convertido para graus. Deve ser evidenciado que os valores das 'Redes' indicados nas colunas do meio correspondem a valores completos dos campos de dunas, enquanto que os das colunas da direita apenas correspondem a pontos onde foram assinaladas as frentes de avalanche na MGDD, podendo algumas discrepâncias entre os resultados estar associadas a essa diferença nos dados. Os campos apresentados na cratera Herschel foram marcados na MGDD e representam quatro campos de dunas, ordenados de Sul para Norte.

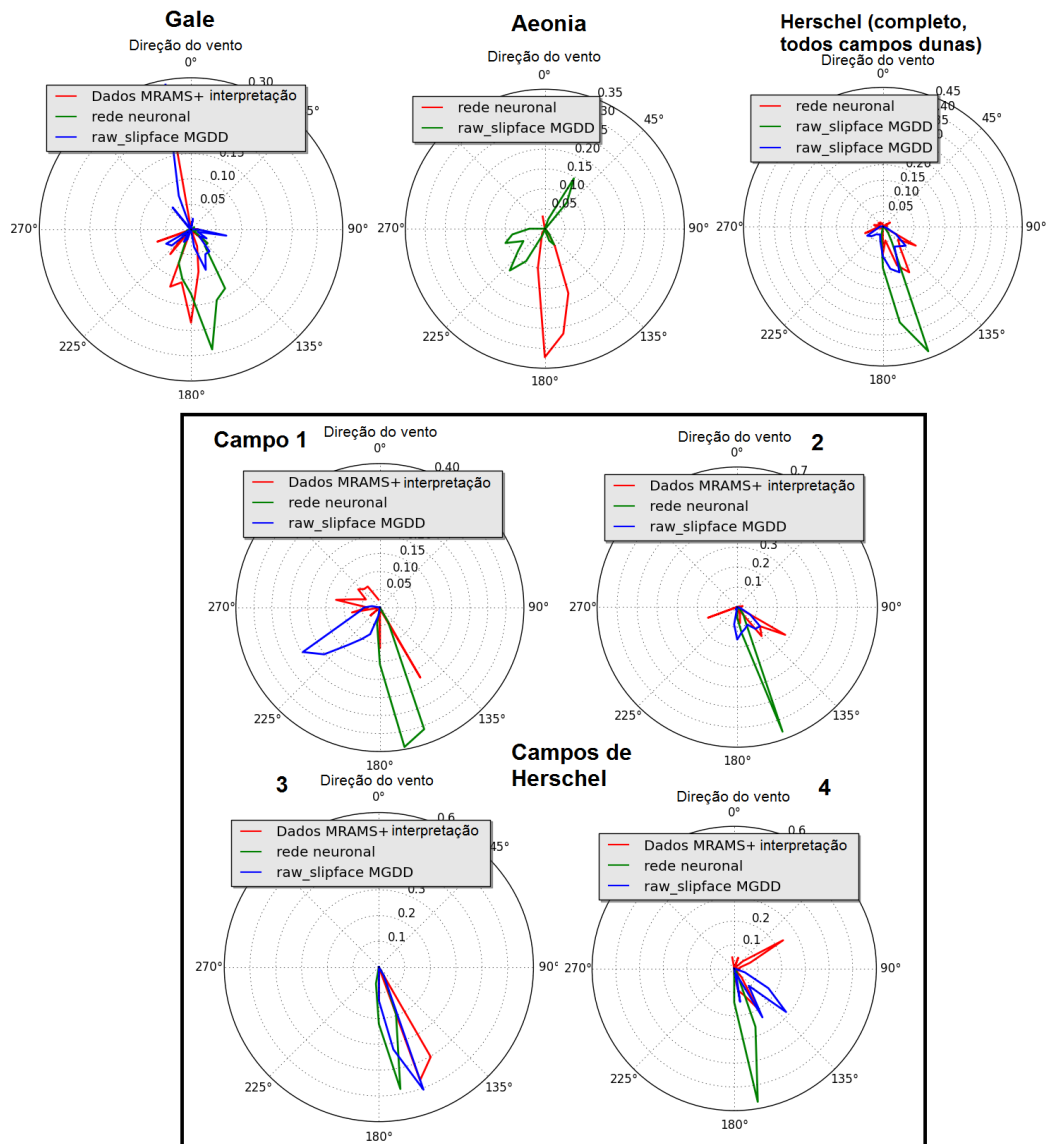


Figura 4.13: Histograma polar comparando as direções de transporte de sedimentos previstas pelas redes neurais em diferentes campos de dunas, as direções do MRAMS nesses mesmos pontos e as direções das frentes de avalanche apresentadas pela MGDD (no caso de Aonia não existem dados do MRAMS).

Com um conjunto novo de dados de *input* poder-se-ia melhorar os dados nessa área em particular. Nos outros casos, e até no caso do campo de Aonia, as diferenças são muito menores, sendo sempre da ordem dos 60° ou inferior. No caso do campo 1 da cratera Herschel, onde a diferença é elevada, esta explica-se com algumas discrepâncias entre as direções traçadas na MGDDD e no MRAMS, muito provavelmente causada por características da topografia local (este campo de dunas em particular situa-se numa depressão), e os resultados das frentes de avanço também diferem dos do azimute do campo apresentado na MGDDD. É de assinalar ainda a grande precisão / pequena diferença nos resultados nos casos dos campos de dunas 2,3 e 4 da cratera Herschel, e numa análise conjunta aos campos disponíveis, onde a diferença nas direções médias ronda os 6.5°. Nota-se ainda a predominância e enviesamento nas direções Norte-Sul nos dados calculados pelas redes neuronais, estando as direções médias dos campos sempre nos intervalos entre 150° e 200°. Apesar de estes valores coincidirem com os das direções de transporte em Herschel e Gale, isso já não acontece no campo de Aonia, e é um dos problemas que devem ser resolvidos. Em relação aos valores da desvio padrão, é de realçar o facto destes valores serem superiores nos campos de dunas individuais em Herschel, mas inferiores no panorama geral da cratera. Esta característica pode explicar-se a partir do facto de existirem algumas direções opostas nestes campos, e dos campos utilizados para obter estes resultados terem mais dunas do que aqueles utilizados na MGDDD. É, portanto, natural que estes valores apresentem uma variância superior aos da MGDDD. É importante ainda apontar o facto de que os valores da variância são inferiores (nos campos de Gale e de Aonia) nas simulações das redes neuronais em relação aos dados da MGDDD. Este facto pode ser uma consequência de um enviesamento dos dados das redes neuronais, que faz as direções de transporte de sedimentos apontarem sensivelmente numa mesma direção. Este tipo de enviesamento estaria ausente de uma interpretação manual, dando azo a essa diferença assinalável.

Capítulo 5

Conclusões e trabalhos futuros

5.1 *Clustering* - possibilidades e conclusões

Como se viu acima, os métodos de clustering como forma de classificação de dunas têm ainda de avançar de forma a tornar o método numa ferramenta capaz e útil nesse aspeto, capaz de contribuir significativamente para o avanço da compreensão dos padrões das dunas marcianas, e para facilitar e simplificar o estudo imediato de grandes campos de dunas no planeta. Ainda assim, os resultados sugerem que a possibilidade existe, sendo a distinção entre dunas lineares e barcanas/complexas a mais promissora neste estudo. Sugere-se a aplicação dos algoritmos a mais parâmetros do campo de dunas, como os topográficos, por exemplo, e testando exaustivamente as combinações de dados de forma a obter a melhor classificação possível para os tipos de dunas presentes na área. Apesar de neste trabalho ter sido concluído que o melhor algoritmo de classificação seria o *Mini-batch K-means*, podem existir outros algoritmos não testados que possibilitem melhores resultados. Mais testes nesse aspeto são também sugeridos, de forma a melhorar o método para futuras aplicações.

5.2 Redes Neurais - possibilidades e conclusões

5.2.1 MRAMS e as Redes Neurais

Um dos problemas que sobressaiu nos dados do MRAMS no início da sua utilização como fonte para as direções de transporte de sedimentos foi a sua resolução nitidamente fraca. Muitos pontos do MRAMS estão distanciados de 5-7km, e surgem inexatidões quando os tentamos passar para a resolução utilizada nos campos de dunas (500m-1km, dependendo da área). Existem também algumas áreas sem informação disponível, como é o caso do campo de dunas de Aonia e outros campos de dunas no planeta, que tivemos de eliminar da análise

de resultados precisamente por esse facto. Além disso, é evidente em algumas áreas as diferenças entre os dados apresentados pelo MRAMS e as direções de transporte visíveis nas dunas no terreno. Em muitos destes casos foi necessária uma correção destes mesmos dados, de forma a aumentar a coincidência entre aquilo que o terreno indica e aquilo que os nossos dados nos dizem. Todas estas limitações do MRAMS valorizam a metodologia proposta, uma vez que esta permitirá a obtenção de dados com uma melhor resolução espacial e que resultam diretamente das estruturas eólicas cartografáveis na superfície.

Por outro lado, esta ausência de dados imediatos sobre as direções de transporte de sedimentos em determinados campos de dunas é precisamente o problema que nos propomos resolver no início deste trabalho, e neste campo não se pode dizer que o sucesso seja estrondoso. No caso do campo de dunas de Aonia, que é o verdadeiro teste deste método, uma vez que é um campo de dunas diferente daqueles onde foram treinadas as redes neuronais, os resultados não foram os melhores, apresentando uma tendência Norte-Sul quando a interpretação visual das frentes de avalanche das dunas não sugeriria esta direção. Estes resultados sugerem que será necessário treinar as redes neuronais com mais campos de dunas, que apresentem diferentes direções preferenciais de transporte de sedimentos, de forma a possibilitar uma previsão mais certa das direções de transporte. Alternativamente, pode ser necessária a criação de uma rede neuronal específica para cada campo de dunas individual, onde são levados em conta os enviesamentos naturais do vento nesse mesmo campo, evitando assim erros que surgem quando se passa de um dado campo de dunas para outro. Esta última alternativa é obviamente a que colocará mais obstáculos à aplicação deste tipo de métodos a uma escala global, no entanto, mesmo esta abordagem será mais objetiva e implicará menos trabalho que a clássica cartografia manual das dunas marcianas.

5.2.2 MGDDD e as Redes Neuronais

Como se pôde ver na seção 4.2.3, existem aspetos da base de dados global das dunas marcianas que podem ser melhoradas. Com novas imagens, mosaicos bem feitos dos campos de dunas, e uma utilização de ferramentas informáticas para ajudar a interpretação humana, podem ser construídas novas bases de dados, mais completas e precisas, sobre os campos de dunas do planeta. Este método, utilizando as redes neuronais, pode ajudar a lá chegar, mas é preciso ainda muito trabalho para ultrapassar algumas das limitações apontadas, e que se centram na passagem das interpretações humanas para uma interpretação mecânica e automatizada, com elevada precisão, das direções de transporte dos campos de dunas. Ainda assim, os nossos resultados, se aplicados a nível global em Marte, poderiam melhorar a qualidade dos dados existentes na base de dados existente, uma vez que as informações disponíveis em termos de direções de transporte são vagas e, em certos casos, contraditórias. Uma aplicação cuidada das redes neuronais em cada um dos campos de dunas do planeta seria um método capaz de dar uma nova perspetiva sobre o tema, e simultaneamente capaz de nos fazer compreender melhor a história eólica de Marte, ao mesmo

tempo que nos fornece pistas com mais detalhe sobre os padrões de circulação atmosférica a uma escala global.

5.3 Trabalhos futuros

O método apresentado, como se pode ver na seção 4.2.3 pode melhorar certos aspetos da base de dados atual das dunas marcianas, fornecendo dados referentes às frentes de avalanche e direções de transporte que de outra maneira não estariam disponíveis para a comunidade científica. As redes neuronais surgem naturalmente como uma ferramenta capaz de prever as direções indicadas por um *input*, e, até certo ponto, com capacidades de generalização e aplicação a áreas não treinadas.

Uma das mudanças possíveis numa futura aplicação do programa a outros campos de dunas, passará pela integração de parâmetros topográficos na informação disponível como *input* das redes neuronais. O motivo desta sugestão foi a apreciação, depois de estudadas as direções de transporte de sedimentos e das dunas nas diferentes áreas, da influência que este fator tem, tantas vezes, na direção de transporte dos sedimentos. A integração deste tipo de informação localmente nos pontos de amostragem das redes neuronais poderá, portanto, resultar numa maior precisão dos resultados.

Seria também útil melhorar os mosaicos desenvolvidos neste trabalho. Cobrir os espaços em branco, devido à falta de imagens do planeta, e melhorar as diferenças entre imagens de modo a diminuir os problemas de transições bruscas poderiam ser passos importantes de modo a melhorar os dados disponíveis e possíveis de calcular sobre os campos de dunas. Quanto melhores e mais precisos os dados dos campos de dunas, melhores resultados se podem obter.

Outra ideia que surgiu no desenvolvimento deste trabalho foi a necessidade de melhorar as simulações atuais sobre a direção de transporte de sedimentos no planeta, nomeadamente as do MRAMS. Muitas das direções de transporte dos sedimentos previstas por este modelo não estão de acordo com as direções inferidas através das dunas, e em alguns casos estão simplesmente erradas. É entendido o facto de que estes dados não pretendem simular as direções de transporte de sedimentos, trata-se apenas de um modelo de circulação atmosférica, no entanto a análise das estruturas sedimentares eólicas é até ao momento a única forma de aferir a validade destes modelos.

Existem também diversos outros tipos de análise de dados e de reconhecimento de padrões. Um estudo futuro no tema poderia, com os mesmos dados, explorar novas técnicas e métodos para classificar os tipos de dunas ou prever as direções de transporte de sedimentos. Aqueles explicados e utilizados aqui podem servir apenas como exemplo ou inspiração a outros tipos de análise, com mais e melhores resultados.

Uma possibilidade que surgiu também no desenvolvimento deste trabalho está relacionada com a sua aplicabilidade no planeta Terra e noutros planetas onde existam dados sobre os campos dunares e sobre os padrões das direções de transporte dos sedimentos. A vantagem dos dados terrestres seria a possibili-

dade interessante de recolha direta de dados sobre as direções de transporte dos sedimentos, sem ser necessária a utilização de dados provenientes de simulações imperfeitas e, por vezes, defeituosas, como é o caso do MRAMS. Infelizmente, os dados visuais do terreno terrestre não se encontram tão bem distribuídos e disponíveis como os marcianos, sendo necessário, na maior parte dos casos, pagar para obter os mesmos. Desta forma, a aplicação ao caso terrestre terá de ser adiada.

Quanto ao *clustering* como método de classificação de dunas, a sua utilização e capacidades serão sempre limitadas pela qualidade de informação disponível sobre um determinado campo de dunas, mas surge também como um apoio ou complemento à interpretação humana na diferenciação rápida de diferentes áreas de dunas. Os seus limites principais estão no número de *clusters*/tipos de dunas que é capaz de diferenciar, mas dentro desse limite pode obter resultados até interessantes. Mais testes com novos tipos de *clustering* e testes noutras áreas são necessários para se poder desenvolver este método como uma ferramenta capaz de classificar rápida e eficazmente grandes campos de dunas do planeta.

Capítulo 6

Apêndices e Referências

6.1 Algoritmos e métodos de análise de dados utilizados - *Clustering* e Redes Neurais

6.1.1 Redes Neurais - o que são e como funcionam

Uma rede neuronal (Haykin (1994)) é um paradigma de processamento de informação inspirado pela maneira como os sistemas nervosos biológicos, tal como o nosso cérebro, processam a informação. O elemento chave deste paradigma é a estrutura dos sistemas de processamento da informação. As redes neuronais são compostas por um grande número de elementos de processamento interconectados (neurónios) que trabalham em conjunto para resolver um determinado problema. As redes neuronais, tal como as pessoas, aprendem por exemplo. Para configurar uma rede neuronal para uma determinada aplicação, tal como reconhecimento de padrões ou classificação de dados, é utilizado um processo de aprendizagem, similar aos ajustamentos sinápticos existentes nos processos de aprendizagem nos sistemas biológicos. As redes neuronais, com a sua capacidade notável de decifrar significado em dados complicados ou pouco precisos, podem ser usadas para encontrar padrões ou tendências em conjuntos de dados que são demasiado complexos para serem detetados por humanos ou por outras técnicas computacionais.

Apesar das redes neuronais parecerem ser um desenvolvimento recente da tecnologia, na realidade são mais antigas que o computador, sendo produzidos pela primeira vez em 1943, pelo neurofisiologista Warren McCulloch e pelo lógico Walter Pitts. Infelizmente, a tecnologia disponível na altura não lhes possibilitou muito desenvolvimento. Em 1969 foi publicado um livro por Minsky e Papert, onde era resumido o sentimento geral de frustração de muitos investigadores em relação às redes neuronais. Desde então, com o aumento da velocidade de processamento nos computadores, foi aumentando a popularidade das redes neuronais, e são agora usadas em muitos campos e com diversos fins.

As redes neuronais utilizam uma abordagem diferente da dos computadores

convencionais no que toca à resolução de problemas. Um computador convencional utiliza uma abordagem algorítmica, isto é, o computador segue um conjunto detalhado de instruções de forma a conseguir resolver um dado problema. A menos que conheça os passos específicos a seguir, o computador não vai ser capaz de resolver o problema. Isto significa que a capacidade de resolução de problemas de um computador convencional está limitada a problemas que já compreendemos e somos capazes de resolver. Desta forma, um computador capaz de resolver problemas que não sabemos como resolver será, com certeza, um computador muito mais útil.

Por outro lado, as redes neuronais processam informação de uma maneira similar à de um cérebro humano. Isto significa que a rede é composta por um grande número de elementos de processamento altamente conectados, que vão trabalhar em paralelo para resolver um determinado problema. As redes neuronais aprendem por exemplo, não podendo ser programadas especificamente para resolverem uma tarefa específica. Os exemplos devem ser cuidadosamente escolhidos, de forma a evitar perdas de tempo precioso, ou prevenir funcionamentos incorretos da rede. A vantagem está no facto de que, uma vez que a rede aprende a resolver o problema 'sozinha', a operação pode ser imprevisível, e gerar resultados diferentes em cada treino.

Assim, as redes neuronais e os computadores algorítmicos convencionais não estão em oposição, mas sim complementam-se mutuamente. Existem tarefas mais apropriadas para cada um dos métodos, e podem ser ambos úteis nos seus próprios domínios. Além disso, um grande número de tarefas necessita a combinação das duas abordagens (tal como um computador convencional é utilizado, por exemplo, para supervisionar uma rede neuronal) de forma a conseguir uma operação com a eficiência máxima.

As redes neuronais são normalmente descritas como 'não-paramétricas', isto é, a utilização de uma rede neuronal não necessita de qualquer informação sobre a distribuição estatística dos dados subjacentes (Tso (2001)). A qualidade da performance de uma rede neuronal depende significativamente da qualidade do seu treino, e não da qualidade dos fatores assumidos sobre a distribuição estatística dos dados. Durante a fase de treino, a rede neuronal vai 'aprender' regularidades existentes nos dados de treino e, baseado nesses padrões, construir regras que possa estender para o resto dos dados desconhecidos. Esta é a habilidade que define uma rede neuronal. Contudo, o utilizador tem de definir a arquitetura da rede, o número de neurónios que esta contém, o tipo de treino a utilizar, e a capacidade de aprendizagem, que vão afetar o tempo de treino, a performance, e a velocidade de convergência de uma determinada rede neuronal. Não existem regras definidas para ajudar na criação de uma rede neuronal, apenas alguns 'guias' que podem ajudar os utilizadores na sua escolha de parâmetros na rede neuronal.

Neste trabalho em particular, as redes neuronais vão ser usadas para prever as direções de transporte de sedimentos em vários campos de dunas do planeta Marte. Em concreto, elas vão utilizar os neurónios para chegar do *input* (conjunto de atributos e características regionais do campo de dunas) para chegar a um *output* (conjunto de valores sobre as direções e variâncias de transporte

6.1. ALGORITMOS E MÉTODOS DE ANÁLISE DE DADOS UTILIZADOS - CLUSTERING E REDES NEURON

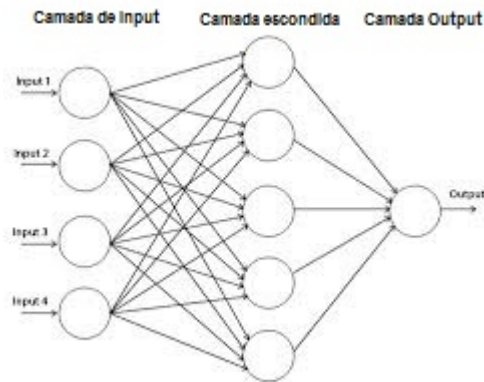


Figura 6.1: Um esquema simples de uma rede neuronal 'Multilayer perceptron', mostrando os neurónios de *input*, escondidos e de *output*. O número dos neurónios é uma das variáveis a modificar na arquitetura da rede. Neste caso, temos 3 camadas, com 4 neurónios de *input*, 5 escondidos e 1 de *output*, respetivamente. Imagem retirada de <http://www.codeproject.com/Articles/175777/Financial-predictor-via-neural-network>.

de sedimentos em determinada área). Mais detalhes serão fornecidos na seção 3.2.2, onde será analisado o trabalho desenvolvido com as redes neuronais nesta tese.

6.1.1.1 Método de *Back-propagation*

O 'Multilayer perceptron' usando algoritmos de aprendizagem com *back-propagation* é um dos modelos de redes neuronais mais utilizados. Na imagem 6.1 temos o aspeto da arquitetura típica deste tipo de rede neuronal. Do lado esquerdo temos a camada de *input*, que contém o conjunto de neurónios capaz de receber impulsos do exterior (na forma de parâmetros de caracterização de campos de dunas, neste caso, por exemplo). Esta camada de neurónios não efetua qualquer tipo de computações, ao contrário dos elementos das outras duas camadas. A camada central é a camada escondida (pode haver mais do que uma camada escondida nos tipos mais complexos de redes neuronais, mas neste caso vamos cingir o nosso estudo a uma rede neuronal com apenas uma camada escondida, uma vez que é esse tipo de rede neuronal que vamos utilizar como base no nosso algoritmo. Camadas escondidas suplementares fazem, de certa forma, um trabalho idêntico ao de uma única camada), e a camada da direita é a camada de *output*, que produz os resultados do processamento efetuado na camada do meio. Não existem interações entre neurónios da mesma camada, mas todos os neurónios de uma dada camada estão completamente ligados a todos os neurónios das camadas adjacentes. Estas conexões terão diferentes pesos numéricos, ajustados durante a fase de aprendizagem. O valor tomado por cada neurónio

é denominado a sua 'atividade'.

O algoritmo de *back-propagation* é o mais comum no que toca a atualizar as atividades neuronais e os pesos da interconectividade numa dada rede neuronal '*multilayer perceptron*'. A *back-propagation* envolve dois passos principais, a propagação para a frente e para trás, de forma a atingir o seu objetivo de modificar o estado neural. Inicialmente, o conjunto completo de dados a analisar é dividido em dois subconjuntos, agora denominados de 'Treino' e 'Teste', sendo o conjunto de treino maior do que o de teste. O conjunto de treino servirá para condicionar a rede neuronal, enquanto que o conjunto de teste servirá como grupo de controlo. Durante o treino, cada amostra (por exemplo, um dado pixel no campo de dunas e os parâmetros que lhe estão associados) é alimentada na camada de *input*, e as atividades dos neurónios são sequencialmente atualizadas entre as camadas de *input* e de *output* através de uma função de *mapping*. Assim que as passagens para a frente estão completas, as atividades dos neurónios de *output* são comparadas com as suas atividades esperadas, de acordo com os valores de *output* calibrados. A não ser em circunstâncias muito peculiares, os resultados do *output* da rede neuronal vão diferir em certa medida dos resultados expectados à partida. Estas diferenças são denominadas o 'erro da rede' e este é então novamente distribuído para as camadas anteriores através da rede, modificando os pesos relativos de cada neurónio. No fim de cada iteração de aprendizagem da rede, esta será aplicada ao conjunto de teste, de forma a verificar a qualidade da aprendizagem da mesma, e a confirmar que a rede não está a ficar 'viciada' no conjunto de treino. Estas passagens dos erros e dos resultados para trás e para diante continua até à rede neuronal ter 'aprendido' todas as características das classes de *input* e *output*, ou o utilizador achar que o erro da rede já está suficientemente reduzido. O momento de paragem deve, também, ser escolhido cuidadosamente, uma vez que treinos a mais podem diminuir os erros no conjunto de dados de treino mas aumentá-los nos conjuntos de testes. Este processo é denominado o 'treino da rede neuronal'.

Durante o processo de propagação para a frente, as atividades dos neurónios são atualizadas sequencialmente camada a camada, entre as camadas de *input* e *output*, de modo a gerar um *output* na forma de ativações dos neurónios da camada de *output*. Se x_j denominar o *input* total recebido pelo neurónio j , este será igual a algo do género: $x_j = \sum_i a_i w_{ji}$, onde a_i é a atividade do neurónio i e w_{ji} é o peso da conexão entre os neurónios i e j . Depois do valor de x_i ser calculado é convertido num valor de *output* (para transmissão para a camada seguinte se o neurónio estiver numa camada intermédia) utilizando uma função *mapping*. O método de função de *mapping* utilizado neste caso é a função de sigmoide, uma das funções mais utilizadas neste tipo de rede neuronal. A sigmoide é uma função monotonamente crescente não-linear, definida pela expressão: $a_j = f(x_j) = \frac{1}{1 + \frac{1}{\exp(\frac{x_j}{T})}}$, onde T é um parâmetro chamado 'Temperatura', e que torna a curvatura da sigmoide mais abrupta ($T > 1$) ou mais gradual ($T < 1$). No nosso caso, o parâmetro vai estar definido para o valor $T = 1$, um valor central e padrão, uma vez que não nos interessa alterar esse tipo de propriedades das redes neuronais, e o programa utilizado não permite

6.1. ALGORITMOS E MÉTODOS DE ANÁLISE DE DADOS UTILIZADOS - CLUSTERING E REDES NEURON

um acesso fácil a esses parâmetros.

Depois de computada a atividade de cada neurônio numa dada camada, é efetuado um processo semelhante na camada adjacente. De realçar que a camada de *input* é um caso à parte, uma vez que os neurônios nesta camada tomam os valores fornecidos pela amostra de treino. Para os nodos da camada de treino, a atividade do neurônio j é diretamente definida como a componente j do vetor com o padrão de *input*. No caso da *back-propagation*, os pesos inter-neuronais são modificados começando da camada de *output* e movendo-se na direção da camada de *input* (da direita para a esquerda, segundo o esquema da imagem 6.1). O objetivo da atualização dos pesos é a redução dos erros de identificação da rede (definidos acima como a diferença entre o *output* obtido pela rede e o *output* esperado). Para tal, é aplicado o critério de menor erro quadrado médio, definido pela expressão: $E(w) = \frac{1}{2} \sum_{j,k} (a_{j,k} - o_{j,k})^2$, onde w é um conjunto dos pesos da rede, $a_{j,k}$ é o neurônio j na camada de *output* obtido pelo valor k na amostra de *input*, e $o_{j,k}$ é o valor esperado no neurônio j na camada de *output* para a amostra de *input* k . O objetivo da nossa rede neuronal passa, obviamente, pela iteração dos processos de treino de modo a obter um valor mínimo para o parâmetro $E(w)$.

6.1.2 *Clustering* - O que é e como funciona

Para além da utilização das redes neuronais como método para prever as direções de transporte de sedimentos numa dada região a partir da forma das dunas e das direções das faces de deslizamento, este trabalho tinha também como um objetivo a classificação automatizada das dunas em diferentes categorias. Para tal, foram utilizadas ferramentas de *clustering*. O seu modo de funcionamento e as suas características encontram-se explicadas abaixo.

Os avanços na tecnologia em termos de deteção e armazenamento de dados, e um crescimento dramático em aplicações como as pesquisas na Internet, imagens digitais e vigilância de vídeo, criaram grupos enormes de dados, com múltiplas dimensões. A maior parte destes dados estão gravados digitalmente em meios eletrónicos, oferecendo um enorme potencial para o desenvolvimento de análises de dados automatizadas, bem como classificação e pesquisa. Além do aumento da quantidade de dados, a variedade destes dados (texto, imagens, vídeo, etc.) também tem aumentado. Muitos destes dados estão soltos, sem uma estrutura que os organize. Desta forma, a criação de uma ferramenta ou algoritmo capaz de classificar os dados automaticamente torna-se um imperativo.

A distribuição da informação em grupos é uma ferramenta essencial da compreensão e da aprendizagem, bem como uma etapa fundamental para organizar os dados obtidos. Por exemplo, um método comum de classificação científica envolve a distribuição de organismos segundo uma distribuição hierárquica e lógica (Domínio, Reino, Filo, Classe, Ordem, etc.). A análise com *clusters* é o estudo formal dos métodos e algoritmos para agrupar objetos segundo determinadas características ou semelhanças (Jain (2010)). A análise com *clusters* não utiliza etiquetas de categoria que identificam os objetos com observações *à priori*, isto é, etiquetas de classe. A ausência de categorias da informação distingue o

clustering de dados (aprendizagem não-supervisionada) de classificação ou análise discriminatória (aprendizagem supervisionada). Nesse aspeto, o *clustering* é similar às redes neuronais, no sentido em que o algoritmo consiste nas informações sobre como construir um método de classificação, não nas classificações em si. O objetivo do *clustering* é de encontrar padrões e estruturas nos dados, e é portanto exploratório na sua natureza. O *clustering* e a análise com *clusters* tem uma rica e longa história em diversos campos de estudo científico. Um dos mais populares, e aquele que vai ser analisado com mais pormenor nesta seção, dos métodos de *clustering*, é o método *K-Means*. Publicado inicialmente em 1955 (Jain (2010)), é o melhor exemplo de algoritmos de *clustering*, e é também um dos mais simples e fáceis de aplicar. Apesar do facto de ter sido proposto há mais de 50 anos, e milhares de algoritmos de *clustering* terem sido construídos e publicados desde então, o método *K-Means* é ainda bastante utilizado. Isto demonstra a dificuldade em desenvolver um algoritmo de propósito geral e os problemas que apresenta o *clustering*.

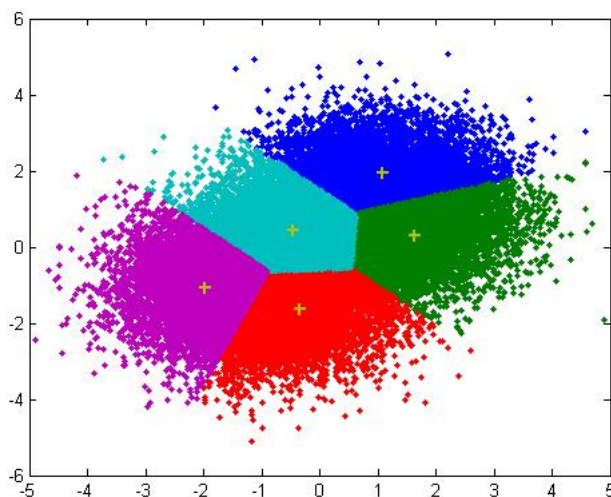


Figura 6.2: Imagem representativa de um *K-means clustering* num conjunto de pontos. Cada cor representa um grupo dentro do conjunto de pontos. Os '+' no centro de cada grupo indicam o ponto representativo de cada grupo individual. Imagem retirada de <http://www.mathworks.com/matlabcentral/fileexchange/screenshots/2240/original.jpg>.

O *clustering* funciona agrupando os dados disponíveis segundo características semelhantes. O objetivo é sempre o de descobrir os grupos naturais de um conjunto de padrões, pontos ou objetos. Esta classificação é estatística na sua natureza e envolve comparações quantitativas de múltiplas características. Na imagem 6.2 podemos ver o exemplo de um *clustering* aplicado numa

6.1. ALGORITMOS E MÉTODOS DE ANÁLISE DE DADOS UTILIZADOS - CLUSTERING E REDES NEURON

imagem, e os grupos dela resultantes. Uma definição operacional de *clustering* pode ser explicitada da seguinte forma: Dada uma representação de n objetos, encontram-se K grupos baseados numa medida de semelhança entre os objetos, de tal forma que a semelhança entre objetos do mesmo grupo seja elevada, e as diferenças entre objetos de grupos diferentes seja elevada. Mas, qual é a definição, neste caso, de semelhança? O que é um *cluster*? Os *clusters* podem diferir entre si em densidade, forma e tamanho. O objetivo de um algoritmo de *clustering* passa também pela descoberta do melhor número e tamanho dos *clusters* presentes num conjunto de objetos. O *cluster* ideal pode ser definido como um conjunto de pontos que é compacto e isolado. Na realidade, o *cluster* é uma identidade subjetiva, e pode diferir de acordo com os objetivos ou pretensões do observador. Mas, apesar de o cérebro humano ser um excelente reconhecedor e organizador de *clusters* em uma, duas, ou mesmo três dimensões, serão necessários algoritmos para organizar dados com um número de dimensões mais elevado. É esse desafio, bem como a determinação de um número desconhecido de *clusters* para organizar um determinado conjunto de dados, que deu origem aos inúmeros algoritmos de *clustering* que já foram publicados e que continuam a surgir, ano após ano.

O *clustering* de dados tem sido utilizado com os três objetivos abaixo em mente:

- Estrutura subjacente: para ganhar conhecimento sobre os dados, gerar hipóteses, detetar anomalias, e identificar características salientes ou subjacentes à estrutura dos dados
- Classificação natural: para identificar o grau de semelhança entre dois objetos, por exemplo
- Compressão: como um método para organizar os dados e os resumir segundo os protótipos de *cluster* gerados.

Por exemplo, a organização de dados por classes é utilizada de forma a reconhecer caracteres em letra escrita. Diferentes utilizadores podem utilizar diferentes métodos para escrever o mesmo dígito ou letra, pelo que utilizar o *clustering* para classificar esses métodos pode ser útil de forma a reconhecer a letra escrita pelo utilizador. Em vez de utilizar um único modelo para cada carácter, múltiplos modelos baseados no número de subclasses existentes podem ser utilizados de forma a aumentar a eficácia do reconhecimento.

No nosso caso em particular, o *clustering* é utilizado como um método para reproduzir a tradicional análise e classificação manual dos campos de dunas. A análise por *clusters* poderia reproduzir essa classificação de uma maneira automática e objetiva, aumentando a velocidade e a eficiência de estudos que dependem da análise dos campos de dunas. Uma aplicação de *clustering* nesse campo poderia servir como uma ferramenta útil e preciosa nesse aspeto.

6.1.2.1 Algoritmo *K-Means*

Deixemos $X = \{x_i\}, i = 1, \dots, n$ ser o conjunto de n pontos d -dimensionais que queremos agrupar num conjunto de K *clusters*, $C = \{c_k\}, k = 1, \dots, K$. O algoritmo *K-means* encontra uma partição tal que o erro quadrado entre a média empírica de um *cluster* e os pontos desse mesmo cluster será minimizada. Deixemos σ_k ser a média dum dado *cluster* c_k . O quadrado do erro entre o ponto σ_k e os pontos nesse *cluster* c_k é definido como: $J(c_k) = \sum_{x_j \in c_k} \|x_j - \sigma_k\|^2$. O objetivo de um algoritmo *K-means* é, portanto, a minimização da soma deste erro ao longo de todos os K *clusters*:

$$J(C) = \sum_{k=1}^K \sum_{x_j \in c_k} \|x_j - \sigma_k\|^2$$

Somatório de todos os erros ao longo dos diversos *clusters*

É sabido que a minimização desta função é um problema NP (mesmo para $K=2$) (Drineas et al. (2004)), ou seja, o tempo de computação do problema varia exponencialmente com o número de pontos a analisar. Logo, o algoritmo *K-means* pode apenas convergir para um mínimo local (apesar de poder convergir para um mínimo global, se os *clusters* estiverem suficientemente bem separados). Ou seja, se aplicado repetidas vezes a um mesmo conjunto de dados, o algoritmo de *clustering* pode encontrar diferentes *clusters*. O algoritmo começa com uma partição inicial escolhida aleatoriamente com K *clusters* e depois vai atribuir padrões a esses *clusters* de forma a diminuir o valor $J(C)$ calculado na equação 6.1.2.1. Uma vez que o erro diminui sempre com um aumento do número de *clusters* K (com o valor $J(C) = 0$ quando $K = n$), isto significa que o valor pode ser minimizado apenas para um número fixo de *clusters*. Desta forma, os principais passos do algoritmo *K-Means* são os seguintes:

1. Selecionar uma partição inicial com K *clusters*; repetir passos 2-3 até a partição dos objetos pelos clusters estabilizar
2. Gerar uma nova partição atribuindo a cada padrão o centro de cluster mais próximo
3. Computar novos centros de *cluster*

Este algoritmo *K-means* necessita de 3 parâmetros especificados pelo utilizador, que vão ajudar a definir a atribuição dos *clusters* pelos dados. Estes parâmetros são o número de *clusters* K , as coordenadas iniciais de cada *cluster*, e a métrica da distância. De entre estes 3 parâmetros, o mais importante e mais crítico para o algoritmo é a escolha do valor K . Embora não exista um critério matemático perfeito para esta escolha, existe sim um conjunto de guias para a escolha deste valor K . Tipicamente, corre-se o algoritmo para diferentes valores de K e o utilizador terá de escolher qual partição é melhor ou mais adequada para o objetivo pretendido. Além disso, diferentes inicializações dos *clusters* podem significar resultados finais completamente distintos, uma vez que o algoritmo apenas converge para os mínimos locais do padrão de dados, não garantindo a obtenção dum mínimo absoluto. Uma maneira de ultrapassar este problema

poderá ser correndo o algoritmo múltiplas vezes para o mesmo conjunto de dados, e escolhendo a partição com o menor valor $J(C)$.

Os algoritmos *K-Means* são tipicamente utilizados com a métrica Euclidiana para calcular as distâncias entre pontos e o centro dos *clusters* correspondentes. Consequentemente, o algoritmo costuma encontrar *clusters* com forma esférica nos dados. Contudo, existem outros métodos para medir a distância que podem ser utilizados, detetando *clusters* hiper-elipsoidais, ou outro tipo de agrupamento dos dados, mas esses casos não vão ser explorados nesta utilização, uma vez que nos vamos cingir à métrica euclidiana.

O algoritmo *K-means* permite ainda, devido à maneira como é calculado, a computação da distância de cada ponto ao centro do *cluster*. Isto permite uma medida extra da 'pertença' de cada ponto ao *cluster*, possibilitando a distinção entre um ponto próximo do *cluster* de um ponto que pertence à mesma classe, mas que está mais distante. Com essa medida, podemos fazer uma classificação mais subtil entre os diversos *clusters* existentes no conjunto de dados, tendo uma classificação gradual dos pontos a analisar.

Para este trabalho em concreto, foi utilizada a variante '*Mini Batch K-Means*' do algoritmo *K-Means*. Esta variante está disponível no software *sklearn* (Pedregosa et al. (2011)), que foi utilizado nos testes de *clustering*. Esta variante do algoritmo utiliza *mini-batches*, ou seja, subconjuntos dos dados de *input*, escolhidos aleatoriamente em cada iteração do treino. Isto é feito de forma a diminuir o tempo de computação, mantendo como objetivo a minimização do erro $J(C)$. Estas *mini-batches* reduzem significativamente o tempo de computação, produzindo resultados que são apenas ligeiramente piores do que no algoritmo padrão.

Para avaliar se um algoritmo de *clustering* é preciso, foi aplicada uma equação. Tomando os valores dos pontos avaliados corretamente, podemos obter uma medida da precisão do método de classificação do conjunto. Para tal, vamos utilizar o coeficiente kappa (Cohen (1960)). Este coeficiente tem em conta todas as informações na matriz de confusão criada, comparando os resultados do *clustering* com uma classificação manual dos dados, de forma a ter em conta a possibilidade de uma classificação aleatória dos pontos. O coeficiente kappa é definido como:

$$\hat{k} = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} \cdot x_{+i})}$$

Nesta equação, \hat{k} é o coeficiente kappa estimado, r é o número de linhas e colunas na matriz de confusão, x_{ii} é a entrada com as coordenadas (i, i) da matriz, x_{i+} e x_{+i} são os totais parciais da fila i e da coluna j , respetivamente, e N é o número total de observações.

6.2 Programa criado - Esquema geral e estrutura

Para ajudar a perceber e entender o programa desenvolvido no decorrer desta tese, foi criado um esquema geral com a sua estrutura. Nos capítulos e seções anteriores os passos estão explicados, mas uma síntese deste género auxilia

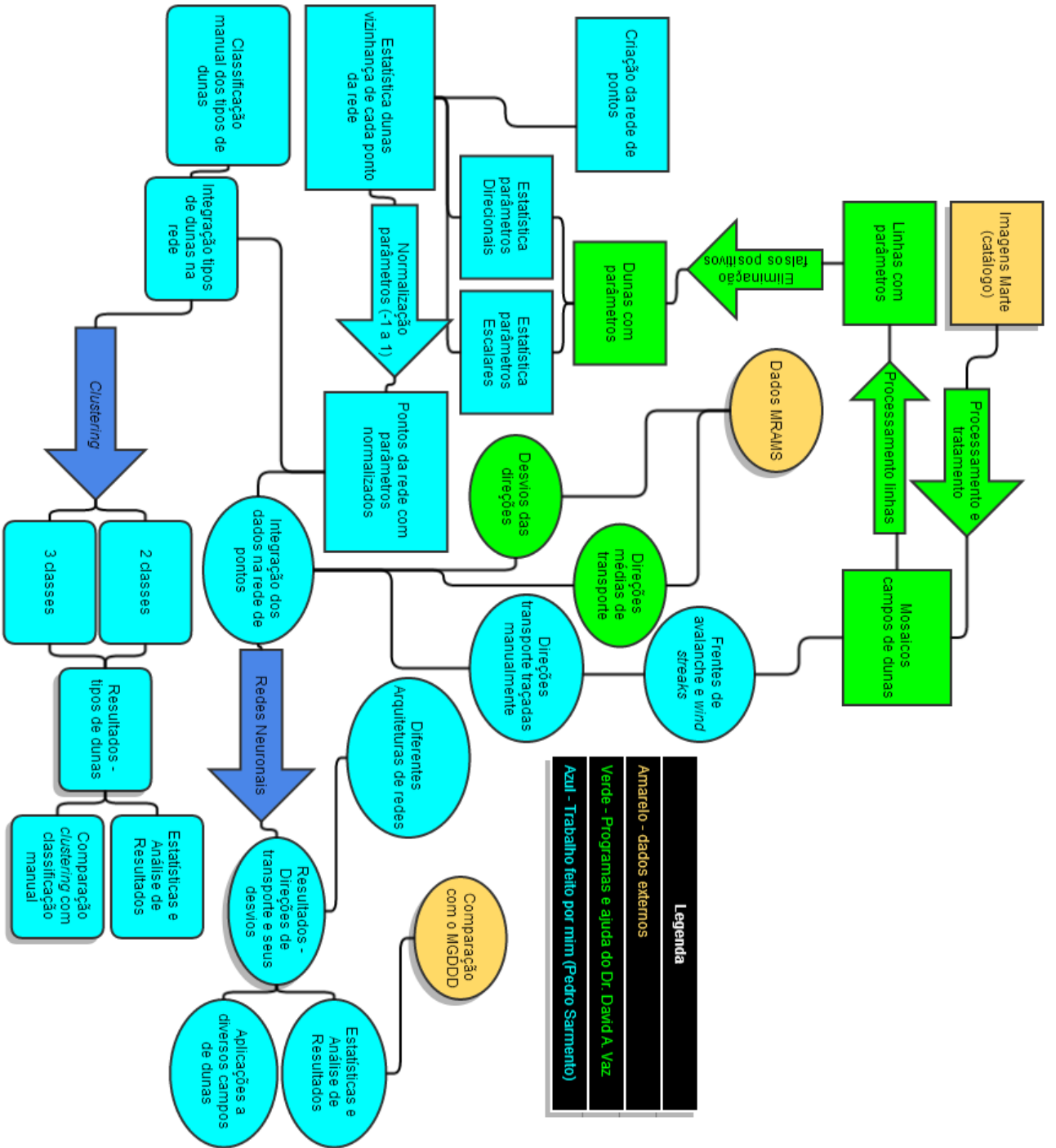


Figura 6.3: Esquema programa Python

à compreensão do método desenvolvido. A linguagem utilizada foi sempre o Python, com o auxílio do software ArcGIS.

Na figura 6.3 está esquematizado o programa escrito e desenvolvido, detalhando exatamente o que foi criado de raiz e o que foi aproveitado e auxiliado pelo Dr. David A. Vaz, assim como a posição onde se encontram os dados externos (imagens marcianas, MRAMS e MGD3D) utilizados na obtenção e avaliação de resultados. Encontra-se detalhado o paralelismo entre os dois métodos (Redes Neurais e *Clustering*), assim como os métodos indiretos de obtenção das direções de transporte de sedimentos (a partir de dados da imagem). No esquema podemos ainda verificar a ordem de processamento das imagens e das dunas, para referência futura, e a origem dos parâmetros de processamento das linhas (a partir de programas desenvolvidos previamente pelo Dr. David A. Vaz).

6.3 Referências e citações

- Bandeira, L., Marques, J. S., Saraiva, J., & Pina, P. (2011). Automated detection of martian dune fields. *Geoscience and Remote Sensing Letters, IEEE*, 8(4), 626–630.
- Bandeira, L., Marques, J. S., Saraiva, J., & Pina, P. (2013). Advances in automated detection of sand dunes on mars. *Earth Surface Processes and Landforms*, 38(3), 275–283.
URL <http://dx.doi.org/10.1002/esp.3323>
- Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), 2 – 16.
URL <http://www.sciencedirect.com/science/article/pii/S0924271609000884>
- Cardinale, M., Komatsu, G., Silvestro, S., & Tirsch, D. (2012a). The influence of local topography for wind direction on mars: two examples of dune fields in crater basins. *Earth Surface Processes and Landforms*, 37(13), 1437–1443.
URL <http://dx.doi.org/10.1002/esp.3289>
- Cardinale, M., Silvestro, S., Komatsu, G., Vaz, D., & Michaels, T. (2012b). Evidences for sand motion in the equatorial region of mars. In *Lunar and Planetary Institute Science Conference Abstracts*, vol. 43, (p. 2452).
- Carr, M. H. (2006). *The surface of Mars*, vol. 6. Cambridge University Press.
- Clifford, S. M., & Parker, T. J. (2001). The evolution of the martian hydro-sphere: Implications for the fate of a primordial ocean and the current state of the northern plains. *Icarus*, 154(1), 40–79.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educ. psychol. Meas Mt.*,

- Drineas, P., Frieze, A., Kannan, R., Vempala, S., & Vinay, V. (2004). Clustering large graphs via the singular value decomposition. *Machine Learning*, 56(1-3), 9–33.
URL <http://dx.doi.org/10.1023/B%3AMACH.0000033113.59016.96>
- Fenton, L. K., Michaels, T. I., Chojnacki, M., & Beyer, R. A. (2014). Inverse maximum gross bedform-normal transport 2: Application to a dune field in ganges chasma, mars and comparison with hirise repeat imagery and {MRAMS}. *Icarus*, 230(0), 47 – 63. Third Planetary Dunes Systems.
URL <http://www.sciencedirect.com/science/article/pii/S0019103513003102>
- Ferri, F., Smith, P. H., Lemmon, M., & Rennó, N. O. (2003). Dust devils as observed by mars pathfinder. *Journal of Geophysical Research: Planets (1991–2012)*, 108(E12).
- Haberle, R. M., Pollack, J. B., Barnes, J. R., Zurek, R. W., Leovy, C. B., Murphy, J. R., Lee, H., & Schaeffer, J. (1993). Mars atmospheric dynamics as simulated by the nasa ames general circulation model: 1. the zonal-mean circulation. *Journal of Geophysical Research: Planets (1991–2012)*, 98(E2), 3093–3123.
- Hartmann, W. K., Malin, M., McEwen, A., Carr, M., Soderblom, L., Thomas, P., et al. (1999). Evidence for recent volcanism on mars from crater counts. *Nature*, 397(6720), 586–589.
- Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1st ed.
- Hayward, R., Fenton, L., & Titus, T. (2014). Mars global digital dune database: Global dune distribution and wind pattern observations. *Icarus*, 230, 38–46.
- Hobbs, S. W., Paull, D. J., & Bourke, M. C. (2010). Aeolian processes and dune morphology in gale crater. *Icarus*, 210(1), 102 – 115.
URL <http://www.sciencedirect.com/science/article/pii/S0019103510002289>
- Jain, A. K. (2010). Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8), 651 – 666. Award winning papers from the 19th International Conference on Pattern Recognition (ICPR) 19th International Conference in Pattern Recognition (ICPR).
URL <http://www.sciencedirect.com/science/article/pii/S0167865509002323>
- Jupp, M. . (2000). *Directional Statistics*. Wiley Series in probability and statistics.
- Kinch, K. M., Sohl-Dickstein, J., Bell, J. F., Johnson, J. R., Goetz, W., & Landis, G. A. (2007). Dust deposition on the mars exploration rover panoramic

- camera (pancam) calibration targets. *Journal of Geophysical Research: Planets (1991–2012)*, 112(E6).
- Marinova, M. M., Aharonson, O., & Asphaug, E. (2008). Mega-impact formation of the mars hemispheric dichotomy. *Nature*, 453(7199), 1216–1219.
- McKee (1979). <http://digital-desert.com/natural-formations/sand-dunes-06.html>.
URL <http://digital-desert.com/natural-formations/sand-dunes-06.html>
- Michaels, T., & Rafkin, S. (2008). Meteorological predictions for candidate 2007 phoenix mars lander sites using the mars regional atmospheric modeling system (mrams). *Journal of Geophysical Research: Planets (1991–2012)*, 113(E3).
- NASA (2012). 12-mile-high martian dust devil caught in act.
URL http://www.nasa.gov/mission_pages/MRO/news/mro20120404.html
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Rafkin, S. C., Haberle, R. M., & Michaels, T. I. (2001). The mars regional atmospheric modeling system: Model description and selected simulations. *Icarus*, 151(2), 228–256.
- Runyon, K., Bridges, N., Ayoub, F., & Mattson, S. (2014). Aeolian provinces and activity in herschel crater, mars. In *Lunar and Planetary Institute Science Conference Abstracts*, vol. 45, (p. 1495).
- Schaul, T., Bayer, J., Wierstra, D., Sun, Y., Felder, M., Sehnke, F., Rückstieß, T., & Schmidhuber, J. (2010). PyBrain. *Journal of Machine Learning Research*.
- Scott, T. K., D.H. (1986). Geologic map of the western equatorial region of mars, scale 1: 15,000,000, map, i-1802-a.
- Silvestro, S., Achille, G. D., & Ori, G. (2010a). Dune morphology, sand transport pathways and possible source areas in east thaumasia region (mars). *Geomorphology*, 121, 84 – 97. Planetary Dune Systems.
URL <http://www.sciencedirect.com/science/article/pii/S0169555X09003377>
- Silvestro, S., Fenton, L., Vaz, D., Bridges, N., & Ori, G. (2010b). Ripple migration and dune activity on mars: Evidence for dynamic wind processes. *Geophysical Research Letters*, 37(20).

- Silvestro, S., Vaz, D., Ewing, R., Rossi, A., Fenton, L., Michaels, T., Flahaut, J., & Geissler, P. (2013). Pervasive aeolian activity along rover curiosity's traverse in gale crater, mars. *Geology*, *41* (4), 483–486.
- Spagnuolo, M. G., Rossi, A. P., Hauber, E., & van Gasselt, S. (2011). Recent tectonics and subsidence on mars: Hints from aureum chaos. *Earth and Planetary Science Letters*, *312*, 13 – 21.
URL <http://www.sciencedirect.com/science/article/pii/S0012821X11005723>
- Tso, P. M., Brandt ; Mather (2001). *Classification Methods for Remotely Sensed Data*. Taylor and Francis.
- Vaz, D. A. (2010). Mars paleotectonic characterization using automatic image recognition techniques. *Tese de doutoramento em Geologia (Tectónica), apresentada à Faculdade de Ciências e Tecnologia da Universidade de Coimbra*.
- Vaz, D. A., & Silvestro, S. (2014). Mapping and characterization of small-scale aeolian structures on mars: An example from the msl landing site in gale crater. *Icarus*, *230*(0), 151 – 161. Third Planetary Dunes Systems.
URL <http://www.sciencedirect.com/science/article/pii/S001910351300345X>
- Ward, A., Doyle, K., Helm, P., Weisman, M., & Witbeck, N. (1985). Global map of eolian features on mars. *Journal of Geophysical Research: Solid Earth (1978–2012)*, *90*(B2), 2038–2056.
- Ward, W. R. (00/1992). Long-term orbital and spin dynamics of mars. *In: Mars (A93-27852 09-91)*, p. 298-320..
- Wenzel, M. J., Manga, M., & Jellinek, A. M. (2004). Tharsis as a consequence of mars' dichotomy and layered mantle. *Geophysical research letters*, *31*(4).