

IDENTIFICANDO SIMILARIDADES:  
UMA APLICAÇÃO PARA A AVALIAÇÃO DE RISCO DE ÁGUA DE LASTRO

Alexandre de Carvalho Leal Neto

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM PLANEJAMENTO ENERGÉTICO.

Aprovada por:

---

Prof. Luiz Fernando Loureiro Legey, Ph.D.

---

Profa. Roseli Suzi Wedemann D.Sc.

---

Profa. Alessandra Magrini D.Sc.

---

Prof. Luiz Pereira Calôba Dr.Ing.

---

Prof. Rubens Mendes Lopes D.Sc.

---

Prof. Silvio Jablonski D.Sc.

RIO DE JANEIRO, RJ - BRASIL

ABRIL DE 2007

LEAL NETO, ALEXANDRE DE CARVALHO

Identificando similaridades: Uma aplicação para a avaliação de risco de água de lastro [Rio de Janeiro] 2007

VIII, 179 p. 29,7 cm (COPPE/UFRJ, D.Sc., Planejamento Energético, 2007)

Tese - Universidade Federal do Rio de Janeiro, COPPE

1. Espécies invasoras
2. Água de lastro
3. Mapa auto-organizável
4. Análise multivariada de dados

I. COPPE/UFRJ II. Título ( série )

Aos meus pais  
Alexandre (*in memoriam*) e Marly

minha homenagem e gratidão

À Flavia, Thiago e Guilherme  
fontes de afeição

dedico

## AGRADECIMENTOS

Ao Prof. Luiz Fernando L. Legey um agradecimento especial pelo incentivo e interesse demonstrado como orientador desta tese.

Ao projeto GLO/99/G31/A/AG/19-GEF/UNDP/IMO Programa Global de Gestão de Água de Lastro pela oportunidade de capacitação proporcionada e, em especial, a Steve Raaymakers e Robert Hilliard incentivadores da atividade de avaliação de risco.

A Companhia Docas do Rio de Janeiro, em particular, à Direção do Instituto de Pesquisas Hidroviárias.

A Profa. Roseli Wedemann pelo apoio durante a realização deste trabalho.

Aos Profs. Andréa Junqueira, Sergio Exel e Silvio Jablonski pela rica troca de experiências.

A Leonardo Rodrigues pelo apoio na utilização de software estatístico.

À Georg Pölzlbauer da Vienna University of Technology, Áustria, e Susan Worner da Lincoln University, Nova Zelândia pelo interesse com que acolheram minhas dúvidas e pelo fornecimento de valiosas informações.

Aos funcionários do Programa de Planejamento Energético pela gentileza com que sempre me atenderam.

E a todos que de alguma maneira contribuíram para a realização deste trabalho.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

IDENTIFICANDO SIMILARIDADES:  
UMA APLICAÇÃO PARA A AVALIAÇÃO DE RISCO DE ÁGUA DE LASTRO

Alexandre de Carvalho Leal Neto

Abril/2007

Orientadores: Luiz Fernando Loureiro Legey

Roseli Suzi Wedemann

Programa: Planejamento Energético

Este trabalho apresenta um estudo sobre similaridade ambiental, desenvolvido para um conjunto de 357 portos, cujas características ambientais e geográficas estão representadas por dados relativos a 34 parâmetros. Na mineração de dados efetuada foram testados métodos da análise multivariada de dados, dividida em duas categorias, agrupamento (métodos Ward e k-médias) e ordenação (Análise de Componentes Principais e Escalonamento Multidimensional Não-métrico). Além desses métodos foi utilizado, também, um algoritmo de redes neuronais, o Mapa Auto-Organizável, que pode ser usado ao mesmo tempo para reduzir a quantidade de dados por agrupamento, e para projetar os dados de forma não linear em um visor de menor dimensão. Os resultados são apresentados e analisados visando entender, o tanto quanto possível, o fator de risco de similaridade ambiental, e contribuir para o desenvolvimento de um sistema de apoio à decisão para prevenção da introdução de espécies invasoras.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

IDENTIFYING SIMILARITIES: AN APPLICATION FOR  
BALLAST WATER RISK ASSESSMENT

Alexandre de Carvalho Leal Neto

April/2007

Advisors: Luiz Fernando Loureiro Legey

Roseli Suzi Wedemann

Department: Energy Planning

This work presents a study on environmental similarity, developed for a set of 357 ports, whose environmental and geographical characteristics are represented by data related to 34 parameters. In doing the data mining, multivariate data analysis divided in two categories, clustering (Ward and k-means methods) and ordination (Principal Components Analysis and Non-metric Multidimensional Scaling) were tested. Besides these methods, a neural network algorithm, the Self-Organizing Map was also used at the same time both to reduce the amount of data by clustering, and for projecting the data nonlinearly onto a lower dimensional display. The results are presented and analyzed aiming to understand, as far as possible, the environmental similarity risk factor, and to contribute for developing a decision support system for invasive species introduction prevention.

## SUMÁRIO

1 Introdução.....	1
2 Água de Lastro.....	5
2.1 Definição .....	6
2.2 Transladação de espécies.....	7
2.3 Espécies invasoras.....	8
2.4 Organismos na água de lastro .....	13
2.5 Tipos de impactos .....	15
2.6 Iniciativas.....	16
2.6.1 Programa GloBallast.....	17
2.6.2 Convenção Internacional sobre Gestão de Água de Lastro .....	19
2.6.3 Iniciativas nacionais .....	23
3 Avaliação de risco de água de lastro.....	25
3.1 Componentes da avaliação de risco de água de lastro .....	28
3.1.1 Similaridade ambiental .....	29
3.2 Avaliação de risco do Programa GloBallast.....	31
3.2.1 Obtenção dos dados portuários .....	32
3.2.2 Identificação dos portos fonte .....	33
4 Reconhecimento de padrões .....	37
4.1 Medidas de similaridade.....	38
4.2 Redução da dimensionalidade .....	43
5 Análise Multivariada de Dados .....	44
5.1 Análise de Agrupamentos.....	51
5.1.1 Métodos hierárquicos .....	51
5.1.2 Métodos não-hierárquicos.....	56
5.1.3 Matriz de Concordância.....	65
5.2 Métodos de Ordenação .....	70
5.2.1 Análise de Componentes Principais .....	70
5.2.2 Escalonamento Multidimensional não-métrico.....	80

6 Redes neurais .....	86
6.1 Regras de aprendizado competitivo .....	87
6.1.1 Inicialização dos vetores de pesos.....	91
6.2 Mapa Auto-Organizável .....	94
6.2.1 Algoritmo de Kohonen .....	95
6.2.2 Função de vizinhança .....	97
7 Implementação do SOM.....	102
7.1 Aprendizado.....	103
7.1.1 Treinamento seqüencial .....	104
7.1.2 Treinamento em lote.....	105
7.2 Vetores de peso .....	106
7.3 Avaliação da qualidade do treinamento .....	107
7.4 Resultados.....	108
7.4.1 Agrupamento do SOM.....	125
7.4.2 Avaliação do comportamento dos parâmetros .....	128
8 Discussão.....	134
9 Conclusão .....	140
Referências Bibliográficas .....	144
Anexo 1 - Vetores comuns de bioinvasão marinha .....	153
Anexo 2 - Parâmetros da análise de similaridade ambiental do sistema GloBallast/URS ....	155
Anexo 3 - Base de Dados GloBallast.....	157
Anexo 4 - Coeficiente de Similaridade do sistema GloBallast/URS.....	170
Anexo 5 - Lei Nº 11.200, de 24 de novembro de 2005.....	178



## 1 Introdução

As espécies exóticas invasoras são na atualidade reconhecidas como uma das maiores ameaças à biodiversidade global. Elas podem provocar sérios impactos econômicos, ambientais e na saúde humana. Em ambientes marinhos e costeiros, as espécies invasoras foram identificadas como uma das quatro grandes ameaças aos oceanos do mundo, junto com: fontes terrestres de contaminação; sobreexploração dos recursos marinhos; e alteração física ou destruição dos habitats. Plantas, animais e organismos microscópicos são transportados de seus ambientes nativos para novos locais de diferentes maneiras, seja intencionalmente ou não.

A água de lastro é um dos muitos vetores de transferência de espécies ao redor do mundo. Navios são projetados para navegar com seus cascos submersos até certo nível na água. Se um navio não está com sua capacidade de carga completa, necessita encher seus tanques de lastro com água para manter um calado apropriado e estabilidade durante sua viagem. Quando o navio recebe carga num porto, precisa descarregar parte ou toda sua água de lastro para compensar o peso da carga. Estima-se que aproximadamente 3 a 4 bilhões de toneladas de água de lastro são transferidas globalmente cada ano, e que 7.000 espécies podem ser transportadas num dia na água de lastro de navios. Espécies contidas na água de lastro tomada a bordo em um determinado local, podem ser descarregadas nas águas de uma outra região. Embora muitas dessas introduções de espécies exóticas tenham sido e continuem sendo inócuas, algumas tiveram consequências econômicas e ambientais severas, com consideráveis alterações na fauna e flora marinha, e causando eventualmente o desaparecimento de alguma espécie nativa.

Esta tese discute o problema de transporte de espécies exóticas na água de lastro de navios, apresentando um estudo sobre similaridade ambiental, desenvolvido para um conjunto de 357 portos, cujas características ambientais e geográficas estão representadas por dados relativos a 34 parâmetros. O objetivo da pesquisa foi entender, tanto quanto possível, o fator

de risco de similaridade ambiental, e contribuir para o desenvolvimento de um sistema de apoio à decisão para prevenção da introdução de espécies invasoras.

A Convenção Internacional sobre Controle e Gestão da Água de Lastro e Sedimentos de Navios foi adotada em 2004 com o intuito de prevenir, minimizar e, por fim, eliminar os riscos ao meio ambiente, à saúde pública, às propriedades e recursos decorrentes da transferência de organismos aquáticos nocivos e agentes patogênicos através do controle e gestão da água de lastro dos navios e dos sedimentos nela contidos. Várias diretrizes estão sendo desenvolvidas com o propósito de assegurar sua aplicação uniforme. A minuta das diretrizes para avaliação de risco dá significativa flexibilidade na determinação do método de abordagem: por compatibilidade ambiental ou por espécie alvo, e objetiva aplicar a regulamentação de uma maneira cientificamente robusta.

A compatibilidade ambiental assume que a probabilidade de estabelecimento de uma espécie é proporcional ao grau de similaridade entre o porto doador e o porto receptor. Embora essa abordagem receba críticas, pois muitas espécies invasoras têm tolerância muito ampla, por exemplo, à salinidade ou temperatura, este é, dentre outros, um importante fator de risco. Uma vez que os parâmetros sejam definidos, similaridades podem ser calculadas por meio de diversos métodos. Várias análises diferentes podem ser feitas para diferentes propósitos.

Uma avaliação de risco foi conduzida para o porto de Sepetiba pelo Programa Global de Gestão de Água de Lastro (GloBallast), e os dados foram compilados ao longo do estudo para 357 portos ao redor do mundo. Este banco de dados é bastante heterogêneo tanto na sua montagem quanto em seu conteúdo. Diferentes equipes trabalharam simultaneamente no levantamento dos dados, e várias foram as fontes de obtenção das informações. Eventualmente alguns dados foram estimados com base na localização geográfica do porto ou em valores obtidos para portos comparáveis mais próximos, não correspondendo, portanto, a valores reais, sejam médios ou extremos.

Essa heterogeneidade e as incertezas associadas com os descritores ambientais motivaram esta tese, que buscou ir além da simples medida direta da distância Euclidiana entre os descritores dos portos. O aumento na quantidade e complexidade dos dados vem estimulando o desenvolvimento contínuo de novas técnicas. Além disso, a maior capacidade computacional, hoje disponível, facilita a implementação de algoritmos existentes, mas anteriormente pouco empregados. Apesar de as técnicas de análise multivariada terem muitas características em comum, várias diferenças importantes surgem em seus princípios e objetivos. Foram escolhidos dois métodos de análise de agrupamentos (Ward e k-médias) e dois métodos de ordenação (Análise em Componentes Principais e Escalonamento Multidimensional não-métrico).

Outra linha de investigação dos dados foi o uso do algoritmo de redes neurais artificiais da classe de aprendizado competitivo o Mapa Auto-Organizável (SOM – do inglês *Self-Organizing Map*), que tem a propriedade de poder ser usado ao mesmo tempo para reduzir a quantidade de dados pelo agrupamento, e para projetar os dados não linearmente num visor de menor dimensão. O interesse nesse algoritmo deve-se à sua habilidade de organizar dados, de modo não supervisionado, em grupos homogêneos. Essa tarefa é realizada ajustando as distâncias de uma estrutura de espaço multidimensional numa projeção de baixa dimensão que preserva a topologia do espaço de entrada.

O trabalho está organizado em 8 capítulos, além desta Introdução. No capítulo 2 são descritos alguns tópicos relativos à questão da introdução de espécies aquáticas invasoras e agentes patogênicos por meio da água de lastro. Esse capítulo apresenta ainda propostas de gestão do problema, buscando mostrar a relevância e a atualidade da temática a ser abordada.

O capítulo 3 contém a exposição dos aspectos envolvidos no assunto que se pretende pesquisar: a avaliação de risco de água de lastro e, em especial, o componente de análise de similaridade ambiental.

Nos capítulos subseqüentes são exploradas e avaliadas diferentes técnicas e abordagens para identificar a similaridade entre os portos que compõem a amostra. O capítulo 4 faz uma breve consideração sobre o tema de reconhecimento de padrões e seu objetivo de classificação. Discute ainda como as medidas de similaridade são utilizadas para agrupar objetos de mesmas características. O capítulo 5 apresenta as técnicas de análise multivariada de dados e a justificativa para a utilização desses métodos no contexto do problema da água de lastro. No capítulo 6 o Mapa Auto-Organizável (SOM) é detalhado.

A implementação do algoritmo SOM é feita no capítulo 7 onde são realizadas diversas simulações, com diferentes parâmetros de treinamento, visando gerar mapas que permitam avaliar as similaridades entre os portos analisados. A partir da escolha de um mapa representativo, os resultados são interpretados e é analisada a influência dos descritores no agrupamento dos portos.

Uma discussão é apresentada no capítulo 8 onde se procurou comparar os métodos da análise multivariada e das redes neurais, cotejando-os com alguns trabalhos semelhantes de outros autores.

No nono e último capítulo são feitas as considerações finais do trabalho, comentando-se a potencialidade e aplicabilidade dos métodos empregados. A tese traz, ainda, anexos contendo informações complementares e os dados utilizados neste trabalho.

Finalmente, cabe ressaltar que no decorrer da tese, foi aprovada a Lei nº 11.200, de 24 de novembro de 2005 (Anexo 5), que alterou o nome do porto de Sepetiba para porto de Itaguaí. Como todos os textos de referência fazem menção ao nome “Sepetiba” e uma parte substancial da tese e das análises já tinham sido desenvolvidas, optou-se por manter o nome original, atribuído desde a inauguração do porto em 1982.

## 2 Água de Lastro

Diariamente, um grande número de vetores relacionados às atividades humanas move milhares de organismos marinhos ao redor do mundo. Embora vários vetores tenham sido identificados como responsáveis pela transferência de organismos entre áreas marítimas geograficamente separadas (Anexo 1), a descarga de água de lastro de navios é considerada como o vetor mais importante de movimentos, transoceânicos e intra-oceânicos, de organismos costeiros de águas rasas. A introdução não intencional de organismos exóticos resultou no estabelecimento de muitas espécies, fora de seus limites naturais, com o potencial para ameaçar ambientes e economias. Fatores como poluição ambiental e destruição de habitats podem prover condições que favoreçam a invasão de espécies exóticas (IUCN, 2000).

Muitas espécies de bactérias, plantas e animais podem sobreviver em uma forma viável na água de lastro e no sedimento levado em navios, mesmo após viagens com várias semanas de duração. A descarga subsequente da água de lastro ou sedimento nas águas portuárias pode resultar no estabelecimento de colônias de espécies nocivas e patogênicas que podem perturbar seriamente o equilíbrio ecológico existente. Desde a introdução, no fim do século XIX, de navios com casco de aço, as descargas de água de lastro aumentaram consideravelmente, em todas as partes do mundo, a probabilidade do estabelecimento próspero de populações de espécie não-nativas auto-sustentáveis. Gollasch (1997) relata que “a primeira menção da introdução de uma espécie indesejada foi feita por Ostenfeld (1908) após uma ocorrência maciça da alga fitoplanctônica asiática *Odontella (Bidulpphia) sinensis* no Mar do Norte em 1903”.

O incremento no tamanho e na velocidade dos navios aumentou o risco de dispersão de novas espécies pelos mares, com maiores volumes de água de lastro bem como o tempo de viagem tornando-se menor, o que significa que mais organismos podem sobreviver à jornada.

O potencial da descarga de água de lastro causar dano foi reconhecido não só pela Organização Marítima Internacional (IMO), mas também pela Organização Mundial de Saúde (WHO), preocupada a respeito do papel da água de lastro como um meio de dispersão de bactérias de doenças epidêmicas (ICS & INTERTANKO, 2000). Apesar da preocupação mundial com o assunto, não existem estatísticas detalhadas do volume de água de lastro transportado anualmente. Estimativas iniciais variavam em torno de 10 a 12 bilhões de toneladas por ano e cálculos mais recentes indicam volumes da ordem de 3 a 4 bilhões (Endresen *et al.*, 2003). De acordo com Royal Haskoning (2001), os lucros potenciais para sistemas de tratamento de água de lastro a bordo são avaliados em mais de um bilhão de dólares por ano após a entrada em vigor da Convenção Internacional sobre Gestão de Água de Lastro (ver item 2.6.2).

## 2.1 Definição

Lastro é qualquer material usado para fazer peso ou equilibrar um objeto. Exemplos são os sacos de areia carregados em balões de ar quente que podem ser descartados para aliviar a carga do balão, permitindo que ele suba.

A água de lastro é colocada em um navio para aumentar o calado de modo a permitir a operação efetiva (em casos onde os navios estão vazios, ou pouco carregados, para submergir o hélice e o leme na água); controlar o *trim*<sup>1</sup>; aumentar a estabilidade; e manter cargas de tensão dentro de limites aceitáveis. Normalmente, a água de lastro é carregada, por bombeamento ou gravidade, em tanques de lastro segregado ou em porões de carga vazios. A água de áreas rasas frequentemente contém sedimentos que se depositam no fundo desses tanques ou porões. A água de lastro é, então, descarregada, toda ou em parte, nos portos onde o navio recebe novas cargas.

---

<sup>1</sup> *Trim* é a diferença de imersão entre a proa e a popa do navio.

A água se tornou comumente usada como lastro a partir dos anos 1880, por ser mais fácil de carregar e descarregar, sendo mais eficiente e econômica que o lastro sólido, tal como pedras, areia, solo, e outros materiais baratos e pesados. Este movimento de “lastro seco”, composto de pedras e cascalho das praias próximas, conduziu ao espalhamento de milhares de espécies de insetos e outros artrópodes, moluscos e plantas (Carlton, 2001).

A maioria dos navios, mesmo aqueles que levam carga, têm um pouco de água de lastro a bordo que pode ser de diversas origens. Pode ser água doce, salobra, ou salgada, e carregada a bordo em um ou vários pontos ao longo da rota do navio dependendo da carga, quantidade de combustível que é levado, e o estado do mar. A quantidade de água que um navio pode levar varia enormemente, desde algumas centenas de toneladas de água residual até várias milhares de toneladas de lastro para os navios realmente grandes (Carlton, 1993).

De acordo com a Convenção Internacional sobre Controle e Gestão da Água de Lastro e Sedimentos de Navios – Convenção BWM 2004 – em seu artigo 1 ‘Definições’, “Água de Lastro significa água com seu material em suspensão tomada a bordo do navio para controlar trim, adernamento, calado, estabilidade ou tensões de um navio” (IMO, 2004).

## **2.2 Transladação de espécies**

Uma das principais ameaças para diversidade biológica nativa é reconhecida pelos cientistas e governos como sendo as invasões biológicas causadas por espécies invasoras exóticas. Os impactos dessas espécies são imensos, insidiosos, e normalmente irreversíveis. Elas podem ser tão prejudiciais para as espécies nativas e os ecossistemas numa escala global como a perda e degradação de habitats. A União Internacional para a Conservação da Natureza e dos Recursos Naturais (IUCN), ou *World Conservation Union*, define transladação como o movimento de organismos vivos de uma área com livre soltura em outra, dividindo em três classes principais: introdução; re-introdução; e re-estocagem (IUCN, 1987).

Durante milênios, as barreiras naturais dos oceanos e acidentes geográficos (montanhas, rios, desertos etc) proveram o isolamento genético essencial para a evolução e manutenção da diversidade de espécies e ecossistemas que compõem a riqueza biológica do planeta. Em algumas centenas de anos essas barreiras tornaram-se ineficazes por forças globais que, combinadas, ajudaram espécies a viajar grandes distâncias para novos habitats, e tornarem-se espécies invasoras bem sucedidas, dominando crescentemente grandes áreas. A perturbação desse isolamento interferiu com a dinâmica natural dos sistemas, causando a substituição de diversas comunidades autóctones ou a extinção de espécies nativas.

A globalização e crescimento no volume de comércio e turismo, juntamente com a ênfase do comércio livre, proveram mais oportunidades para as espécies se espalharem acidentalmente ou deliberadamente. Práticas de alfândegas e de quarentena, desenvolvidas no passado para resguardar contra doenças e pestes, são freqüentemente proteções inadequadas contra espécies que ameaçam a biodiversidade nativa. Assim, o fim inadvertido de milhões de anos de isolamento biológico criou grandes problemas que afetam países desenvolvidos e em desenvolvimento (IUCN, 2000).

### **2.3 Espécies invasoras**

Espécies introduzidas também são conhecidas como espécies invasoras, alienígenas, exóticas, estrangeiras, não-nativas, naturalizadas, imigrantes e não-indígenas. Às vezes estas palavras são tratadas como sinônimos, em outros momentos, cada uma tem significados diferentes. Uma espécie exótica é aquela que ocorre em uma região onde ela previamente não existia, pelo menos em uma base de tempo histórica. Nove critérios podem ser usados para testar se uma espécie é nativa ou exótica, sendo os seis primeiros de acordo com situações em nível regional e os demais de caráter mais global (Chapman & Carlton, 1991, 1994; Cranfield *et al.*, 1998):



1. a espécie apareceu subitamente num local onde não era encontrada antes?
2. a espécie ampliou sua distribuição em seguida?
3. a distribuição da espécie está associada a mecanismos humanos de dispersão?
4. a espécie está associada com, ou é dependente de, outra espécie introduzida?
5. a espécie predomina em, ou está restrita a, ambientes anteriormente não ocupados e/ou ambientes artificiais?
6. a distribuição da espécie é restrita em comparação com a distribuição das espécies nativas?
7. a espécie possui uma distribuição mundial disjunta (não contínua)?
8. a espécie possui capacidade de dispersão passiva ou ativa insuficiente para explicar a distribuição observada?
9. a espécie se encontra isolada de espécies geneticamente e morfologicamente similares existentes em outras partes do mundo?

O termo “espécies invasoras” se refere a um grupo, definido em termos gerais, de espécies introduzidas que traz, ou poderia trazer, alguma medida de dano (ambiental, econômico ou para saúde humana). Uma espécie introduzida pode se tornar prejudicial em algumas áreas e não em outras, normalmente como resultado de idiosincrasias muito difíceis, se não impossíveis, de prever com confiança a partir das características do histórico do organismo em seus limites naturais.

Uma espécie introduzida pode falhar em estabelecer uma população viável a longo prazo devido à inibição de uma fase chave do ciclo de vida por fatores locais que limitem o crescimento, restrinjam a reprodução ou recrutamento, e/ou causem uma alta mortalidade. Em tais casos, só repetidas inoculações possibilitarão a contínua (ou esporádica) presença de uma população detectável.

Organismos que sobrevivem aos rigores de uma viagem no interior de um tanque de lastro podem ser considerados como espécies relativamente resistentes. A maioria das espécies introduzidas não estabelece uma população a longo prazo, ou então permanecem relativamente inócuas e restritas em sua distribuição. Isto acontece, presumivelmente, porque o ecossistema local e suas comunidades nativas exercem respectivamente suficiente stress físico e pressões defensivas (competição, predação etc), e/ou porque estímulos críticos e condições para reprodução ou mecanismos que facilitam a fase de dispersão estão faltando.

Cohen & Carlton (1995) definiram espécies estabelecidas como “organismos introduzidos que estão presentes e reproduzindo-se na natureza”, e cujos números, amadurecimento, distribuição e persistência com o passar do tempo sugerem que eles continuarão a estar presentes – exceto em decorrência de esforços de erradicação ou um grande evento catastrófico natural. 'Na natureza' é considerado potencialmente confuso, pois muitas espécies podem se estabelecer prosperamente em áreas artificiais ou altamente transtornadas. Então, uma espécie estabelecida é definida como uma espécie introduzida, tendo pelo menos uma população que (Hilliard *et al.*, 1997):

- aparece viável em tamanho e exibe uma estrutura de idade madura;
- persistiu, por meios naturais, em pelo menos uma área, além de um certo período (i.e. durante vários anos); e
- tenha se dispersado, ou ameace se espalhar mais, por meios naturais ou antrópicos (i.e. é capaz de superar qualquer fator local que possa impedir ou destruir um estágio crucial de dispersão).

As propriedades chave de uma espécie potencialmente prejudicial não só têm que prover a ela habilidade para se espalhar, a partir de seu ponto inicial, para novos locais ao longo da costa, mas também têm que incluir um conjunto de características suficientemente

competitivas, invasoras ou nocivas. Para obter a condição de espécie nociva, após seu estabelecimento, a espécie tem que conseguir realizar, pelo menos, uma das seguintes ações (Hilliard *et al.*, 1997):

- prontamente colonizar áreas não necessariamente ‘vagas’ deslocando espécies nativas via competição por espaço, luz ou alimento; e/ou
- infectar, parasitar, predação ou danificar espécies recreativas ou comerciais importantes; com alto significado ecológico ou valor de conservação;
- lançar toxinas que se acumulam na cadeia alimentar com riscos para a saúde pública e/ou envenenar outros organismos (ex. acumulação de toxinas de dinoflagelados em ostras comerciais, causando interrupção ou fechamento nas atividades de aquicultura para evitar o risco de epidemias por envenenamento paralisante (*paralytic shellfish poisoning- PSP*)) ou por infecção direta (ex. vírus ou bactérias como *Vibrio* ou *Clostridium*);
- causar perdas econômicas significativas para infra-estrutura litoral (ex. por volumosa incrustação em sistemas de tomadas d’água, circuitos de refrigeração, molhes, embarcadouros, marinas, bóias, cascos de embarcações e outras superfícies entremarés ou submersas).

De acordo com a Convenção BWM 2004 em seu artigo 1 ‘Definições’, “Organismos Aquáticos Nocivos ou Patogênicos significam organismos aquáticos ou patogênicos que, se introduzidos no mar, incluindo estuários, ou em cursos de água doce, podem criar riscos para o meio ambiente, saúde humana, propriedade ou recursos, deterioração da diversidade biológica ou interferir com outros usos legítimos de tais áreas” (IMO, 2004).

Informações sobre espécies invasoras são geralmente dispersas numa variedade de banco de dados e relatórios publicados e não publicados. O Ministério do Meio Ambiente, por meio do Projeto de Conservação e Utilização Sustentável da Diversidade Biológica Brasileira

(PROBIO), financiou cinco projetos, com cobertura nacional, focados na busca de informações sobre espécies invasoras. Cada projeto lidou com um aspecto do problema relacionado com organismos que afetam: a saúde; os sistemas produtivos (pecuária, agricultura); os ambientes terrestres; a água doce; e o ambiente marinho. Os relatórios foram coordenados pela Fundação Oswaldo Cruz; Empresa Brasileira de Pesquisa Agropecuária; Instituto Hórus; Universidade Federal de Viçosa; e o Instituto Oceanográfico da Universidade de São Paulo, respectivamente. Os resultados dessa iniciativa foram compilados como um Informe Nacional de Espécies Exóticas Invasoras (a ser publicado) e apresentados no I Simpósio Brasileiro sobre Espécies Exóticas Invasoras, realizado em 2005 (Brasil, 2006). No ambiente marinho foram inventariadas 65 espécies invasoras (Tabela 1).

<b>Grupo</b>	<b>Detectadas</b>	<b>Estabelecidas</b>	<b>Invasoras</b>	<b>Contidas</b>	<b>Total</b>
Bactérias Pelágicas	1	-	-	-	1
Fitoplâncton	-	1	2	-	3
Zooplâncton	3	1	2	4	10
Macroalgas	1	3	1	5	10
Zoobentos	20	9	5	3	37
Ictiofauna	3	1	-	-	4
<b>Total</b>	<b>28</b>	<b>15</b>	<b>10</b>	<b>12</b>	<b>65</b>

Tabela 1 - Situação populacional das espécies exóticas invasoras marinhas no Brasil, conforme o grupo biológico (adaptado de informações do I Simpósio Brasileiro sobre Espécies Exóticas Invasoras, 2005)

Embora esse Informe Nacional não tenha exaurido as informações sobre o assunto no Brasil, configura um passo importante na formação de uma visão anteriormente inexistente. Mais dados poderão ser coletados sempre, porém os dados coletados são inicialmente suficientes para gerar diversas ações práticas no âmbito da prevenção, erradicação e controle de espécies exóticas invasoras (Romais *et al.*, 2005).

## 2.4 Organismos na água de lastro

Navios podem levar a bordo qualquer espécie presente na água de portos, baías e estuários ao redor do mundo. O conjunto de organismos que podem ser transportados na água de lastro é muito grande. Mesmo espécies que quando adultas dificilmente seriam encontradas na água de lastro, seja porque são muito grandes ou porque vivem junto ao fundo, podem ser transferidas através do lastro durante a sua fase planctônica (Figura 1). Muitas das espécies estão em seu estágio larval ou de dispersão, com hábitos bentônicos na fase adulta (e.g. Figura 2). Dentre elas pode-se citar anêmonas, vermes, cracas, caranguejos, caracóis, moluscos, mexilhões, ostras, briozoários, ouriços do mar, ascídias, macroalgas e muitas outras. Outras espécies vivem permanentemente como adultas na água, incluindo diatomáceas, dinoflagelados, copépodes, medusas entre outras. Organismos causadores de doença como vírus, bactérias e parasitas também foram detectados em água de lastro (Ruiz *et al.*, 2000). Assim, organismos na água de lastro variam desde tamanhos microscópicos até peixes de 30cm ou mais.

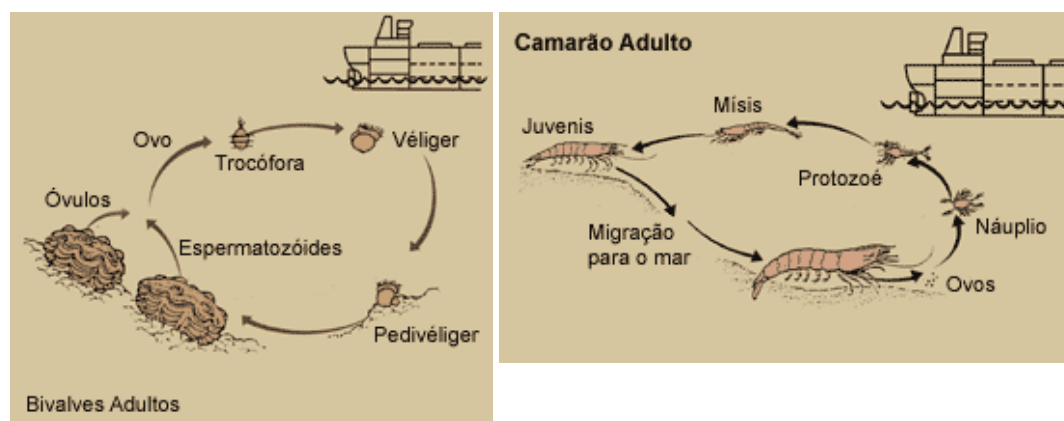


Figura 1 - Ilustração do ciclo de vida e fase planctônica dos bivalves e do camarão  
(<http://www.mma.gov.br/aguadelastro>)

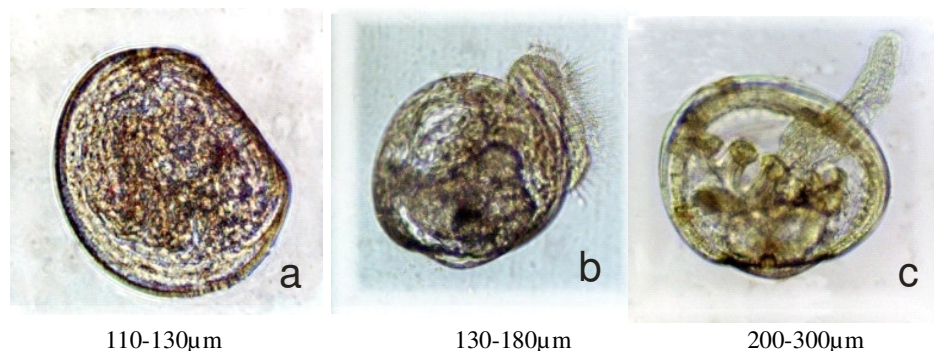


Figura 2 - Estágios larvais do mexilhão dourado *Limnoperna fortunei*: a: larva “D”; b: larva umbonada; c: larva pedivéliger (Cataldo & Boltovskoy, 2000 *apud* LATEC/COPEL, 2003).

Pelo menos 7.000 diferentes espécies de vida marinha são provavelmente transportadas, a cada dia, ao redor do mundo, e evidências apontam para uma diversidade ainda maior do que se suspeitava inicialmente entre os organismos do fitoplâncton e outros organismos da mesma faixa de tamanho (Carlton, 1999). De acordo com Carlton (2001) a água de lastro contém quatro tipos de comunidades vivas:

- plâncton: organismos que derivam passivamente, ou nadam moderadamente, na água;
- necton: espécies que nadam livremente na água;
- incrustantes: organismos fixos (incluindo filmes bacterianos) nas paredes verticais e estruturas horizontais dos compartimentos de lastro; e
- bentos: organismos de fundo ou bentônicos, tais como leito de lama com vermes marinhos e espécies associadas; estágios excistado ou de repouso de plâncton vegetal (fitoplâncton); e plâncton animal (zooplâncton).

Para algumas espécies de vida marinha, pode ser difícil determinar o modo de introdução. A água de lastro e os organismos contidos nela, por exemplo, são sempre descarregados quando um navio recebe carga. Entretanto, organismos incrustados nos cascos dos navios têm que se reproduzir, soltar-se ou nadar para fora do navio. Muitos organismos, incluindo cracas, mexilhões, hidróides, ascídias e macroalgas, podem ser transportados por meio de ambos os mecanismos, tornando-se difícil distinguir entre os dois vetores (Carlton, 2001).

A espécie aquática introduzida mais conhecida no Brasil, tendo a água de lastro como vetor, é o mexilhão dourado (*Limnoperna fortunei*), originário da China e sudeste da Ásia, que foi avistado pela primeira vez na América do Sul na desembocadura do rio da Prata, na costa argentina, em 1991. O mexilhão dourado teve seu primeiro registro no Brasil, no estado do Rio Grande do Sul, em 1999, e sua área de distribuição inicialmente atingiu grandes extensões do lago Guaíba, em Porto Alegre (Mansur, 2000). Na bacia do rio Paraná o mexilhão dourado foi detectado inicialmente em abril de 2001 na usina hidroelétrica de Itaipu (Fontes Júnior, 2002) e, posteriormente, em várias outras usinas como Porto Primavera em 2002; São Simão no rio Paranaíba em 2003; e Jupia e Ilha Solteira em 2004. O registro mais a montante no rio Uruguai ocorreu na represa de Salto Grande em setembro de 2001. A chegada do mexilhão dourado no Pantanal se deu em 2003 num ambiente ligado ao rio Paraguai, nas proximidades de Corumbá, e no ano de 2005 foram coletados exemplares em Cáceres que constitui o limite mais ao norte dessa espécie no Pantanal (Mansur & Darrigran, 2006). A presença do mexilhão dourado nessa latitude é preocupante pela proximidade com a bacia Amazônica.

## 2.5 Tipos de impactos

Uma espécie estabelecida, suficientemente expansiva, dispersiva e agressiva pode causar efeitos adversos. A introdução de organismos exóticos em uma determinada área pode causar impactos ambientais, econômicos e para a saúde humana, os quais podem variar de degradação de vias fluviais (e.g. pelo crescimento de mexilhões zebra nos Grandes Lagos), mortes de peixes até enfermidade humana ou mesmo mortes (e.g. devido a dinoflagelados tóxicos) (Royal Haskoning, 2001).

Organismos não precisam ter características de pestes para causar danos severos. Alguns invasores afetam a flora e fauna nativa competindo por alimento, espaço, iluminação, habitat e outros recursos. A redução da biodiversidade nativa pode se dar por alteração do

substrato, parasitismo ou infecção, sufocando e suplantando espécies nativas localmente importantes. O pior caso ecológico é a substituição de uma espécie nativa pelo invasor exótico. Isto pode afetar não apenas uma única espécie pela extinção, mas também outros organismos dependentes desta espécie como fonte de alimento ou habitat. Como resultado a estrutura da cadeia alimentar pode mudar significativamente depois da introdução de uma espécie.

## 2.6 Iniciativas

Em 1992, a IMO, a partir de solicitação da Conferência das Nações Unidas sobre Meio Ambiente e Desenvolvimento (UNCED), deu início a uma abordagem sistemática para a questão da água de lastro, buscando a definição de regras adequadas, para evitar a disseminação de organismos aquáticos não nativos das regiões nas quais as descargas são realizadas.

A IMO é uma agência especializada da Organização das Nações Unidas (ONU), contando com 167 Estados Membros<sup>2</sup>, e cujas principais responsabilidades referem-se, entre outras, à segurança da navegação; à facilitação do comércio marítimo; à prevenção da poluição do mar por navios; e aos regimes de compensação por danos. A resposta à demanda da UNCED concretizou-se pelas seguintes ações:

- formação de grupo de trabalho de água de lastro no âmbito de seu Comitê de Proteção ao Meio Ambiente Marinho (MEPC);
- adoção de Diretrizes para o controle e gerenciamento da água de lastro de navios para minimizar a transferência de organismos aquáticos nocivos e patogênicos;
- desenvolvimento de um novo instrumento legal de âmbito internacional (Convenção) sobre o gerenciamento de água de lastro a ser considerado para adoção por uma Conferência Diplomática; e

---

<sup>2</sup> Até janeiro de 2007.



- definição de uma iniciativa conjunta com o Fundo para o Meio Ambiente Mundial (GEF) e o Programa das Nações Unidas para o Desenvolvimento (UNDP) para identificar e avaliar barreiras ao efetivo trato da questão da água de lastro em algumas das regiões em desenvolvimento do mundo.

### **2.6.1 Programa GloBallast**

Essa última linha de ação, mencionada no item anterior, levou à aprovação, pelo Conselho do GEF, em 1999, do Programa Global de Gestão de Água de Lastro (GloBallast), tendo como base as recomendações das diversas missões preparatórias e acordos com os governos envolvidos. O projeto contou com recursos do GEF, repassados por intermédio do UNDP, sendo executado pela IMO, com o apoio dos Estados Membros e da indústria do transporte marítimo. O orçamento total para o projeto foi de US\$10.192.000 (incluindo os custos de administração da agência executora e a contribuição dos governos envolvidos de US\$2,8 milhões).

Esse projeto, denominado originalmente *Remoção de Barreiras para a Implementação Efetiva do Controle da Água de Lastro e Medidas de Gerenciamento em Países em Desenvolvimento*, visava reduzir a transferência de espécies aquáticas exóticas indesejáveis que tinham como vetor a água de lastro dos navios. Teve, ainda, como propósito ajudar os países em desenvolvimento a implementar as medidas de caráter voluntário previstas na Resolução de Assembleia da IMO A.868 (20) “Diretrizes para o Controle e Gerenciamento da Água de Lastro dos Navios para Minimizar a Transferência de Organismos Aquáticos Nocivos e Agentes Patogênicos” (Brasil, 1998).

Além disso, o GloBallast tencionava preparar os países, antecipadamente, para a implementação da Convenção BWM 2004 que estava sendo elaborada. Para que essa meta pudesse ser alcançada, foram providos, a seis países participantes, capacitação, assistência técnica, e reforço institucional, tendo como objetivo uma efetiva gestão da água de lastro.

Pretendia-se que seis locais de demonstração representassem as principais regiões em desenvolvimento do mundo, de acordo com as políticas de financiamento de projetos do GEF. Os estudos de caso desenvolvidos nesses países serviram, numa primeira etapa, como demonstração de dificuldades e experiências de sucesso de gestão do problema (Figura 3).

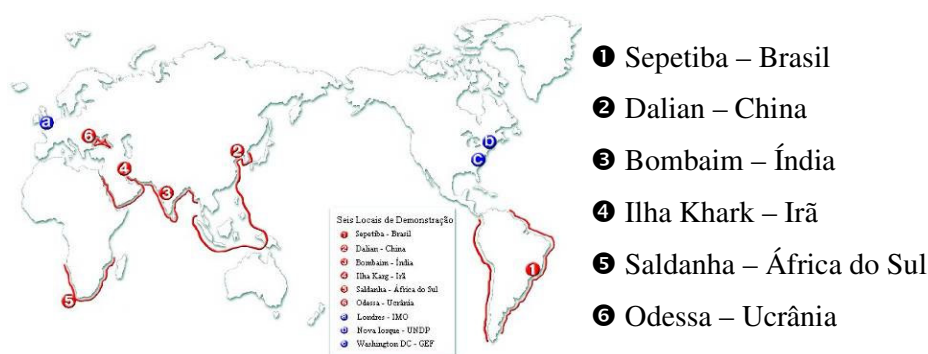


Figura 3 - Portos e países-piloto do Programa GloBallast  
(<http://www.mma.gov.br/aguadelastro>)

A implementação do projeto teve início no ano 2000, com uma duração prevista de três anos. Quando o GloBallast foi proposto, a comunidade internacional planejava adotar um regime regulador para a transferência de água de lastro em 2002. Entretanto, devido à complexidade do tema, as negociações entre os Estados Membros da IMO alongaram-se mais do que o esperado, e a Conferência Diplomática foi remarcada para 2004<sup>3</sup>. Assim, o intervalo entre o final do GloBallast, previsto para março de 2003, e a adoção da nova Convenção, determinaria uma perda do impulso já obtido pelo Programa. Considerando esses aspectos decidiu-se estender o Programa, mantido o orçamento total original, até dezembro de 2004.

O Ministério do Meio Ambiente (MMA) desempenhou o papel de Agência Líder para o Programa GloBallast, no Brasil, sendo a Secretaria de Qualidade Ambiental nos Assentamentos Humanos, do MMA, o Ponto Focal Nacional designado para o projeto. Diversas organizações nacionais participaram ativamente do projeto; juntamente com universidades e representações da área do transporte marítimo (Leal Neto & Jablonski, 2004).

<sup>3</sup> Decisão da 48ª reunião do *Marine Environment Protection Committee* (MEPC) da IMO em outubro de 2002.

Cada país elaborou seu Plano de Trabalho Nacional que esboçava as atividades a serem desenvolvidas no decorrer do Programa no local de demonstração; no caso brasileiro, o porto de Sepetiba, no Estado do Rio de Janeiro. Os principais componentes constantes deste plano foram: Comunicação, Educação e Mobilização; Avaliação de Risco da Água de Lastro; Levantamento da Biota do Porto; Medidas de Gestão de Água de Lastro; Treinamento; Legislação; Conformidade, Monitoramento e Efetivação; e Cooperação Regional.

Com o fim do Programa GloBallast, o GEF e o PNUD aprovaram um fundo para o desenvolvimento de uma proposta de projeto chamado GloBallast Partnerships. Esta proposta foi preparada pela IMO ao longo dos anos 2005 e 2006 para ser submetida ao Conselho do GEF. Caso a proposta seja aprovada, o projeto deverá ser iniciado em 2007 e ser desenvolvido em 5 anos, com um orçamento de US\$ 17 milhões, dos quais US\$ 7 milhões seriam fundos do GEF e o restante contribuições dos países participantes. Distintamente da primeira fase desenvolvida em seis países pilotos, nesta segunda fase o novo projeto pretende ter um enfoque regional, apoiando reformas legais e de políticas para minimizar os impactos adversos de espécies aquáticas invasoras transferidas por navios em países/regiões não cobertas na primeira fase como Caribe e ilhas do Pacífico (Fernandes & Leal Neto, 2006).

### **2.6.2 Convenção Internacional sobre Gestão de Água de Lastro**

Em 1993, foram publicadas pela IMO as primeiras diretrizes para prevenção e controle da introdução de espécies invasoras via água de lastro. Essas diretrizes visavam orientar os governos dos países membros no manejo de água de lastro para evitar a transferência de espécies invasoras e organismos patogênicos.

Posteriormente, em 1997, foram adotadas medidas mais abrangentes através da Resolução IMO A.868(20). As medidas propostas por esta Resolução são de caráter voluntário e incluem: limpeza regular dos tanques de lastro; evitar captação de lastro em áreas

em que tenham ocorrido irrupções ou infestações de organismos nocivos ou agentes patogênicos ou próximas à descarga de esgotos sanitários; evitar a descarga desnecessária de água de lastro na área portuária, realizar a troca de lastro em águas oceânicas; e descarregar a água de lastro em instalações adequadas.

Culminando esses esforços, a Convenção BWM 2004 foi adotada numa Conferência Diplomática em fevereiro de 2004 com a participação de 74 Estados Membros da IMO. Esta Convenção só entrará em vigor 12 meses após a data em que não menos do que 30 países, cujas frotas mercantes combinadas constituam não menos que 35% da arqueação bruta da frota mercante mundial, tenham assinado a mesma sem reservas no que tange a ratificação.

A abordagem adotada pela Convenção BWM 2004 para o combate ao problema foi função da inexistência, no momento, de tecnologia que viabilize um tratamento da água de lastro a bordo dos navios que atenda padrões considerados seguros. Assim, foram estabelecidas uma norma para troca de água de lastro (Regra D-1) e uma norma de desempenho de água de lastro (Regra D-2). A Tabela 2 resume os prazos para cumprimento dessas exigências em função da capacidade de lastro dos navios.

<b>Construção do navio</b>	<b>Capacidade de lastro (cal) (m<sup>3</sup>)</b>	<b>Exigência</b>
Antes de 2009	$1.500 \leq \text{cal} \leq 5.000$	D-1 após a entrada em vigor e D-2 após 2014
Antes de 2009	$\text{cal} < 1.500$ ou $> 5.000$	D-1 após a entrada em vigor e D-2 após 2016
A partir de 2009	$\text{cal} < 5.000$	D-2
Entre 2009 e 2012	$\text{cal} \geq 5.000$	D-1 após 2009 e D-2 após 2016.
A partir 2012	$\text{cal} \geq 5.000$	D-2

Tabela 2 - Prazos estabelecidos pela Convenção BWM 2004 para adaptação dos navios (Almeida & Leal Neto, 2005)

A Regra D-1 prevê que os navios que realizarem a troca da água de lastro em conformidade com esta regra deverão fazê-lo com uma eficiência de pelo menos 95% por cento de troca volumétrica. Já a Regra D-2, que para ser atingida depende ainda de avanços científicos, estabelece que os navios que realizarem gestão de água de lastro em conformidade com esta regra deverão descarregar menos de 10 organismos viáveis por metro cúbico com

dimensão mínima igual ou maior que 50 micrômetros e menos de 10 organismos viáveis por mililitro com dimensão mínima menor que 50 micrômetros e com dimensão mínima igual ou maior que 10 micrômetros. Os indicadores microbiológicos, como padrões de saúde pública, deverão incluir:

- *Vibrio cholerae* (O1 e O139): menos que uma unidade formadora de colônia (UFC) por 100ml ou por 1 grama de amostra de zooplâncton.
- *Escherichia coli*: menos que 250 UFC por 100ml.
- *Enterococos* fecais: menos que 100 UFC por 100ml.

A Convenção BWM 2004 foi aberta para assinatura na sede da IMO de 01/06/04 até 31/05/05 e oito países (Argentina, Austrália, Brasil<sup>4</sup>, Finlândia, Maldivas, Holanda, Espanha e Síria) assinaram o instrumento sujeito a ratificação. Desde 31/05/05 a Convenção está aberta a adesão de qualquer Estado, e Maldivas se tornou a primeira Parte Contratante depois de depositar o instrumento de ratificação pertinente em 22/06/05. Até 31 de março de 2007 apenas oito Estados tornaram-se Partes Contratantes: Espanha, Maldivas, Nigéria, Noruega, Quiribati, São Cristóvão e Neves, Síria e Tuvalu, representando somente 3,21% da arqueação bruta da frota mercante mundial.

Desde a adoção da Convenção, a IMO vem se dedicando ao desenvolvimento e análise de diretrizes que visem garantir a aplicação global e uniforme de seus requerimentos. A maior parte dessas diretrizes já foi adotada no âmbito do MEPC (ver Tabela 3). Alguns temas foram postergados sejam por questões de prioridade ou complexidade. A reunião do MEPC 56, em julho de 2007, deverá adotar as diretrizes G7 e G13, restando a G2 para ser concluída na reunião do BLG 12 (*Sub-Committee on Bulk, Liquid and Gases*), em 2008. A minuta de diretrizes de avaliação de risco, foi inicialmente preparada por representante da Noruega, passando depois para discussão num grupo de correspondência sob a coordenação da Austrália.

---

<sup>4</sup> O Brasil assinou a Convenção BWM 2004 em 25 de janeiro de 2005. Esta assinatura está sujeita a ratificação após a aprovação pelo Congresso Nacional.

<b>Diretrizes</b>		<b>Observações</b>
G1	Instalações de recepção de sedimentos	Adotada - Resolução MEPC 152(55)
G2	Amostragem de água de lastro	Minuta
G3	Conformidade equivalente para embarcações de passeio e de busca e salvamento com relação a gestão de água de lastro	Adotada - Resolução MEPC 123(53)
G4	Planos de gestão de água de lastro	Adotada - Resolução MEPC 127(53)
G5	Instalações de recepção de água de lastro	Adotada - Resolução MEPC 153(55)
G6	Troca de água de lastro	Adotada - Resolução MEPC 124(53)
G7	Avaliação de risco sob a regra de isenção do gerenciamento da água de lastro	Minuta
G8	Aprovação de sistemas de gestão de água de lastro	Adotada - Resolução MEPC 125(53)
G9	Procedimento para aprovação de sistemas de gestão de água de lastro que façam uso de substâncias ativas	Adotada - Resolução MEPC 126(53)
G10	Aprovação e supervisão de protótipos de tecnologias para tratamento de água de lastro	Adotada - Resolução MEPC 140(54)
G11	Padrões de projeto e de construção de navios para realização da troca da água de lastro	Adotada - Resolução MEPC 149(55)
G12	Padrões de projeto e de construção de navios para facilitar o controle de sedimentos nos navios	Adotada - Resolução MEPC 150(55)
G13	Medidas adicionais e situações de emergência	Minuta
G14	Designação de áreas para troca da água de lastro	Adotada - Resolução MEPC 151(55)
	Troca de água de lastro na área do Tratado Antártico	Diretrizes propostas mais recentemente ainda por serem desenvolvidas.
	Vistorias do sistema de gestão de água de lastro pelo Estado de Bandeira	
	Vistoria do controle do Estado do Porto	

Obs 1: O número entre parênteses em cada resolução indica a reunião do MEPC. Datas: MEPC 53 - julho 2005; MEPC 54 - março 2006; e MEPC 55 - outubro 2006.

Tabela 3 - Situação das Diretrizes desenvolvidas pela IMO após a 55ª reunião do MEPC (elaboração própria)

### 2.6.3 Iniciativas nacionais

Nos anos que antecederam a adoção da Convenção BWM 2004 o Brasil participou ativamente das discussões na IMO, inclusive com o desenvolvimento pela Petrobras de um método alternativo para troca de água de lastro em alto mar que passou a ser considerado, a partir da 44ª reunião do MEPC, em março de 2000, como mais um padrão de referência, além dos métodos sequencial e de fluxo contínuo constantes de Resolução IMO A.868(20). Este método, conhecido como método de diluição, consiste em carregar água de lastro pela parte superior do tanque enquanto, simultaneamente, descarrega a água pelo fundo do tanque, numa mesma vazão. O nível da água no tanque é conservado constante. Tal método permite aos navios manter sua condição normal de carregamento de lastro ao longo de toda a duração da viagem (Leal Neto, 2001).

A Agência Nacional de Vigilância Sanitária (ANVISA) teve uma contribuição relevante no que diz respeito à inclusão da temática relativa a agentes patogênicos, por estudos realizados em vários portos do Brasil e testes a bordo.

Num primeiro momento, as ações da IMO refletiram-se no Brasil, através da adoção pela Diretoria de Portos e Costas (DPC), do Comando da Marinha, da Norma da Autoridade Marítima NORMAM-08, de fevereiro de 2000; e pela ANVISA da Resolução de Diretoria Colegiada (RDC) 17, de 12 de janeiro de 2001. Essas duas normas internalizaram para a legislação nacional as orientações da Resolução IMO A.868(20) (Almeida & Leal Neto, 2005).

A ANVISA iniciou, em junho de 2000, uma coleta de formulários para informações sobre a água de lastro entregues de forma voluntária pelos navios. A partir de janeiro de 2001, com base na Resolução RDC 17, revista pela RDC 217 (Brasil, 2001), a submissão do formulário passou a ser obrigatória por todos os navios que solicitam a Livre Prática<sup>5</sup>.

---

<sup>5</sup> Livre Prática é a autorização emitida, pelo órgão de vigilância sanitária, a uma embarcação procedente ou não do exterior a entrar em um porto do território nacional e iniciar as operações de desembarque e embarque de cargas e viajantes.

A exigência da troca de água de lastro em alto mar passou a ter o caráter obrigatório com a entrada em vigor da NORMAM-20/DPC em 15 de outubro de 2005. Os requisitos da legislação nacional foram informados aos Estados Membros da IMO por meio da Carta Circular BWM.2/Circ.1 de 22 de setembro de 2005. As principais características da NORMAM-20/DPC (Brasil, 2005) são:

- Exige que os navios possuam Plano de Gestão de Água de Lastro e apresentem o Formulário de Água de Lastro;
- Dá diretrizes para troca e captação de água de lastro, bem como para descarga de sedimentos do tanque de lastro;
- Sujeita os navios que escalem em terminais brasileiros à inspeção para conformidade com esta norma;
- Estabelece a obrigatoriedade de troca de lastro para navios engajados em navegação de cabotagem entre bacias hidrográficas distintas.

Essa norma recebeu algumas modificações, principalmente relacionadas ao prazo para a aprovação do plano de gestão de água de lastro dos navios pelo Estado de Bandeira ou por Sociedade Classificadora (Portarias nºs 1/DPC e 66/DPC de 2006).



### 3 Avaliação de risco de água de lastro

Uma avaliação de risco ambiental envolve a busca pelo equilíbrio entre o benefício social e o risco ambiental, com a responsabilidade de proteger a sociedade e o meio ambiente dentro de um realismo econômico. É um processo de negociação em que várias combinações de riscos são comparadas e avaliadas contra ganhos sociais ou econômicos. Isso não implica, necessariamente, numa política de risco zero ou mesmo mínimo. Entretanto, os riscos devem ser compreendidos tanto quanto possível caso se deseje manejá-los eficazmente, considerando que isso geralmente tem que ser feito com recursos limitados e de forma custo efetiva (SCOPE, 1980).

A modelagem do problema é uma importante ferramenta de tomada de decisão mesmo onde dados científicos confiáveis são escassos e o modelo é um simples diagrama de fluxo exibindo causas, rotas de transmissão e seus efeitos. Os sistemas humanos e ecológicos podem ser expostos através de todos os compartimentos ambientais – ar, água, solo/sedimentos e biomassa (alimento). Os caminhos do risco podem ser ilustrados de uma forma simples, como na Figura 4, ou muito mais complexa, por exemplo, como na Figura 5.

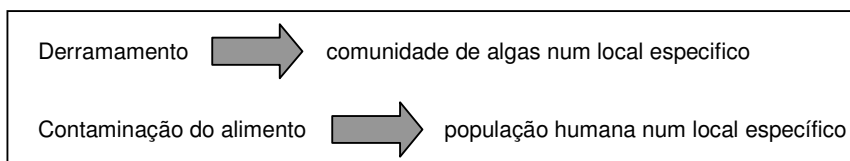


Figura 4 - Caminhos de risco simples (adaptado de Calow & Forbes (1997) *apud* MPP-EAS, 1999, p. 17)

O caminho destacado na Figura 5 ilustra o risco de introdução de espécies invasoras em novos ambientes tendo como vetor o comércio marítimo por meio da água de lastro e o casco dos navios.

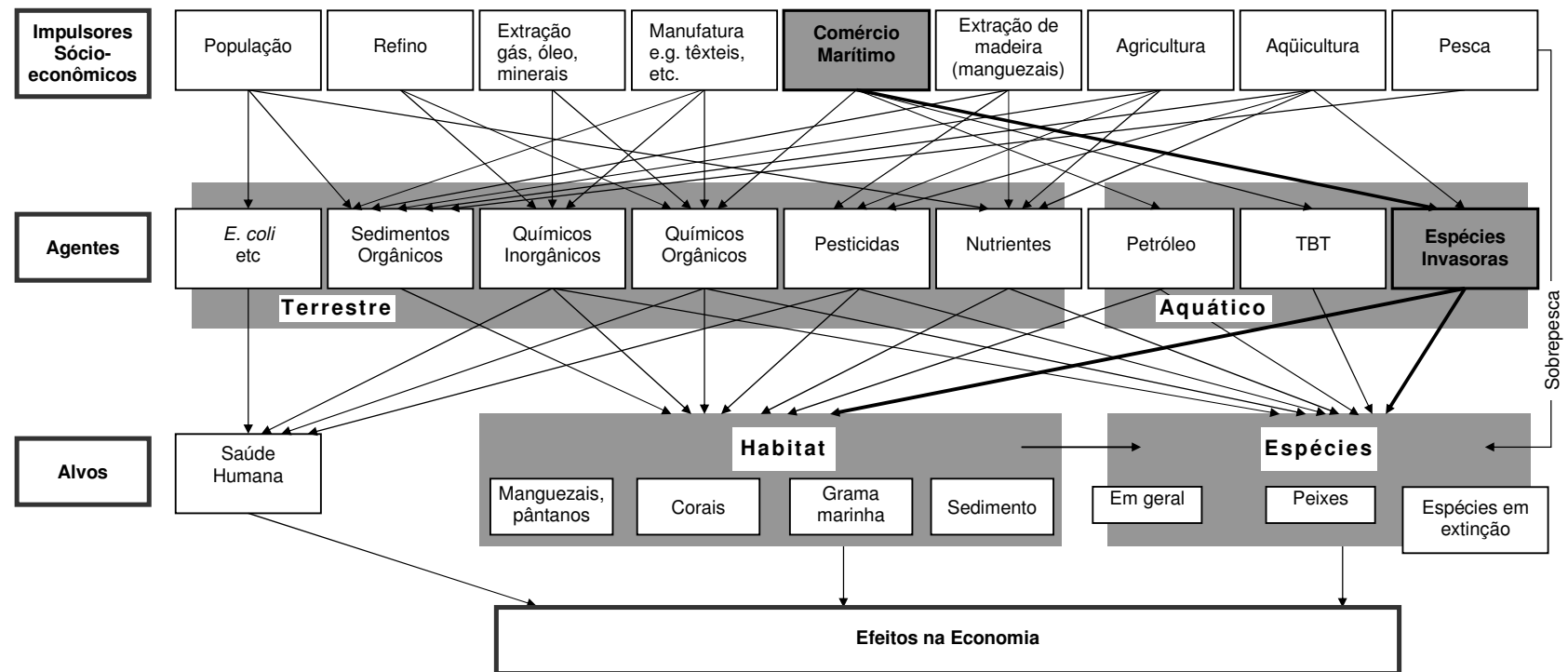


Figura 5 - Fluxos de riscos (adaptado de Calow & Forbes (1997) *apud* MPP-EAS, 1999, p. 18)

A Convenção BWM 2004 e as atuais diretrizes da IMO oferecem significativa flexibilidade aos Estados Membros na determinação da natureza e extensão de seus regimes nacionais de controle de água de lastro. As ações dos países se dão em diferentes níveis com responsabilidades como Estado de Bandeira (*Flag State*), garantindo o cumprimento dos procedimentos por parte dos navios do próprio país, assim como Estado do Porto (*Port State*), com um sistema de inspeção de navios estrangeiros, prevenindo-se tanto contra a importação quanto a exportação não intencional de organismos nocivos e patogênicos.

Um país pode aplicar um regime uniforme a todos os navios ou, buscando avaliar o risco relativo que os navios impõem aos recursos da região, aplicar procedimentos seletivos àqueles julgados como de maior risco. A opção de aplicação uniforme oferece as vantagens de um programa simplificado de administração no qual não há “julgamentos” a serem feitos (ou justificados) pelas autoridades do porto a respeito de quais navios devem ou não ser vistoriados. Adicionalmente, o sistema uniforme requer menos informações e oferece maior proteção em relação à introdução de espécies inesperadas e esta proteção não depende da eficiência do sistema de apoio à decisão, o qual pode não ser completo. Entretanto, algumas desvantagens advêm desta abordagem como custos adicionais de inspecionar navios que em princípio não precisariam ser fiscalizados. Como mais navios são envolvidos, a autoridade do Estado do Porto necessita estabelecer um sistema de controle de maior porte.

Alguns países estão experimentando sistemas que permitem uma seletividade baseada na avaliação de risco, pois esta abordagem reduz o número de navios sujeitos ao controle e monitoramento da água de lastro. A perspectiva de aplicar um programa de controle num número reduzido de navios é especialmente atrativa pelo ponto de vista de evitar organismos-alvo. Medidas mais rigorosas podem ser justificadas em navios julgados como sendo de “alto risco”, se menos restrições forem impostas a navios de “baixo risco”. Entretanto, a efetividade depende da qualidade da informação que sustenta o sistema. Essa abordagem pode deixar o

país/porto vulnerável a riscos desconhecidos oriundos de organismos não-alvo (Clarke *et al.*, 2004). Em contraste com uma abordagem uniforme, uma abordagem baseada na avaliação de risco pode fornecer uma ferramenta para melhor compreensão do problema da bioinvasão, para determinar melhores procedimentos de operação, para identificar rotas de risco e falhas no sistema (como falsos relatórios de água de lastro ou trocas inadequadas). Esta avaliação também pode auxiliar na estratégia de gerenciamento determinando que componentes do processo de invasão são mais susceptíveis ao controle (Hewitt & Hayes, 2002).

Para países/portos que elegerem uma abordagem seletiva será essencial estabelecer um modo organizado de avaliar o risco potencial imposto por cada navio que entre no porto, por meio de um Sistema de Apoio à Decisão. Somente dessa maneira torna-se possível tomar a decisão mais apropriada no que diz respeito a qualquer ação que seja necessária quanto à descarga de água de lastro de um navio. O Sistema de Apoio à Decisão é um instrumento de gestão que fornece mecanismos para avaliar todas as informações disponíveis relacionadas a um navio em particular, e suas medidas de gestão de água de lastro, de modo a, baseado na avaliação de risco, orientar a ação a ser tomada (Clarke *et al.*, 2004).

### **3.1 Componentes da avaliação de risco de água de lastro**

Ecologistas vêm buscando determinar um conjunto de fatores que permita prever o risco de invasão biológica. A probabilidade de uma espécie introduzida se estabelecer em novas regiões e causar problemas depende de vários fatores, principalmente relacionados às suas características biológicas e às condições ambientais nas quais a espécie for introduzida. Fatores adicionais são clima; número de espécies introduzidas (tamanho de população fundadora); competidores nativos; e disponibilidade de alimento. As espécies tendem a se estabelecer em ambientes que são semelhantes ao ambiente de origem delas (Gollasch, 1997).

Existem pelo menos 4 importantes hipóteses a respeito de esse risco ser proporcional:

(1) ao número de inoculações – A probabilidade de introdução aumenta com o número de inoculações num dado local pois à medida que o tempo passa as espécies se deparam com uma variedade de condições abióticas (temperatura, salinidade, habitats) e bióticas (densidade de competidores, abundância de predadores) e, portanto, mais possivelmente encontram uma situação apropriada.

(2) à quantidade de indivíduos transferidos – Cada espécie precisa de uma quantidade mínima de indivíduos para manter uma população inicial, embora muito mais indivíduos possam ser necessários para superar fortes variações ambientais ou catástrofes naturais. Organismos que se reproduzem vegetativamente estão hipoteticamente menos sujeitos às limitações impostas pela estocasticidade demográfica e genética para estabelecer-se num ecossistema.

(3) à similaridade ambiental – A probabilidade de estabelecimento aumenta com o grau de similaridade entre os ambientes da distribuição atual da espécie (nativa e/ou introduzida) e do local de potencial invasão.

(4) ao potencial das espécies de risco – Novas espécies são transferidas para os sistemas aquáticos na forma de cistos, gametas, larvas, juvenis ou adultos. Cada uma dessas formas tem um diferente potencial de inoculação em termos de probabilidade de sobrevivência e de habilidade para estabelecer uma população fundadora. As espécies chamadas “de risco” apresentam um maior potencial de invasão em virtude de poder ter nesses estágios uma maior tolerância a variações ambientais, menores taxas de mortalidade e/ou maior capacidade reprodutiva.

### **3.1.1 Similaridade ambiental**

Um dos componentes da avaliação de risco é a determinação do índice de similaridade ambiental entre o porto em estudo e cada um de seus portos fonte (origem da água de lastro).

Quanto mais um porto recebedor de água de lastro é ecologicamente similar a um porto doador de água de lastro, maior é a possibilidade dos organismos descarregados com a água de lastro “importada” poder tolerar o novo ambiente e manter um número de indivíduos suficientes para crescer, reproduzir e desenvolver uma população viável. A comparação de similaridades ambientais porto-a-porto fornece, conseqüentemente, uma medida relativa do risco da sobrevivência do organismo, bem como do estabelecimento e do potencial de propagação.

O grau de similaridade entre dois sítios terrestres é usualmente avaliado a partir de dados climáticos (temperatura, precipitação) disponíveis de um grande número de estações meteorológicas espalhadas pelo mundo. Já em ambientes aquáticos, a comparação é menos praticável porque: a) existem menos dados disponíveis para esses ambientes; e b) não há ainda uma definição consensual de quais variáveis ambientais são preditivas para o estabelecimento de espécies (Lodge *et al.*, 2006).

Até o momento, a bibliografia especializada contém poucos exemplos de abordagens. Análises simples têm aplicação limitada porque são baseadas puramente na comparação de salinidade (Tabela 4) ou de faixas de latitude (Tabela 5) entre a região de origem (doadora da água de lastro) e a região de destino (recebedora da água de lastro), onde as espécies podem ser introduzidas.

Uma análise multivariada foi proposta por Hilliard *et al.* (1997) num estudo conduzido em doze portos no nordeste da Austrália para o Ports Corporation of Queensland, e posteriormente revisada em Clarke *et al.* (2004) na metodologia desenvolvida para o Programa GloBallast pelo grupo consultores da URS Australia Pty que considerou uma lista de 34 parâmetros como descritores de cada porto (Anexo 2). Entre os 34 descritores existe uma predominância de parâmetros de temperatura (8) e de salinidade e relacionados com salinidade (também 8). Dados de temperaturas extremas do ar, precipitação e de maré foram incluídos devido a sua influência na sobrevivência de organismos intermarés e de habitats rasos.

<b>Destino</b>	<b>Região de Origem</b>		
	Água doce	Água salobra	Água salgada
Água doce	alto	médio	baixo
Água salobra	médio	alto	alto
Água salgada	baixo	alto	alto

Tabela 4 - Probabilidade de colonização de espécies exóticas de acordo com a combinação de salinidade entre a região de origem e a região de destino (Carlton, 1985 *apud* Gollasch, 1997).

<b>Destino</b>	<b>Região de Origem</b>			
	Ártico & Antártico	Temperada baixa	Temperada média	Trópicos
Ártico & Antártico	alto	médio	baixo	baixo
Temperada baixa	médio	alto	médio	baixo
Temperada média	baixo	médio	alto	médio
Trópicos	baixo	baixo	médio	alto

Tabela 5 - Probabilidade de colonização de espécies exóticas de acordo com a combinação de clima entre a região de origem e a região de destino (Gollasch, 1997, p. 151).

### 3.2 Avaliação de risco do Programa GloBallast

Uma das atividades empreendidas pelo Programa GloBallast foi a experimentação de uma avaliação de risco da água de lastro com um método padronizado conduzido em cada um dos seis locais de demonstração. Essa avaliação visou determinar o nível de risco associado a espécies aquáticas invasoras que o porto enfrenta, avaliando os processos presentes na introdução e no estabelecimento bem sucedido dessas espécies. Um banco de dados foi estruturado reunindo os componentes necessários para a determinação dos coeficientes de risco primário integrantes do cálculo de um coeficiente de risco global:

- a frequência relativa do número de tanques de lastro provenientes de um determinado porto em relação ao total de tanques descarregados;
- a proporção do volume de água de lastro oriundo de um determinado porto em relação ao volume total descarregado no porto estudado;

- a similaridade ambiental entre o porto receptor e cada porto doador de água de lastro; e
- o grau de risco de cada porto doador, representado pelo número de espécies de risco (introduzidas, potencialmente nocivas e nocivas) presentes na biorregião desse porto.

O sistema foi desenvolvido utilizando-se um conjunto de softwares onde os resultados da avaliação de risco podem ser verificados no Microsoft Access ou na forma de mapas distintos elaborados no ESRI Arcview. A visão esquemática desse sistema, cuja metodologia é descrita em Junqueira & Leal Neto (2003) e Clarke *et al.* (2004), é apresentada na Figura 6.

### 3.2.1 Obtenção dos dados portuários

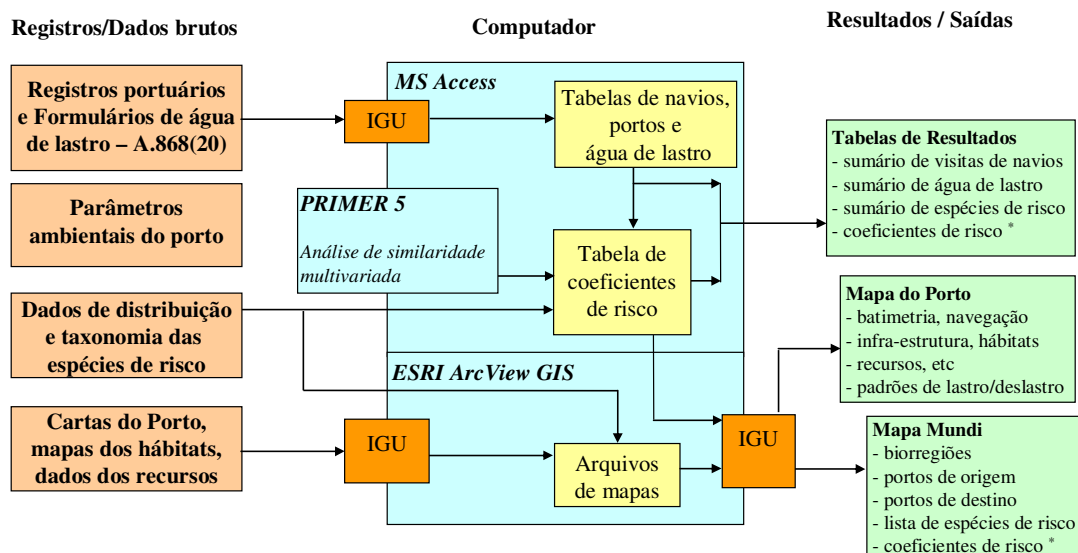
Antes de 1998 o porto de Sepetiba tinha a característica de ser, principalmente, exportador de água de lastro. Dessa forma, a análise de risco foi desenvolvida para o período compreendido a partir de 1998 até 2002. Foram utilizadas informações sobre a movimentação de navios do porto de Sepetiba, reunidas pela Companhia Docas do Rio de Janeiro (CDRJ) num banco de dados Access, para o que se chamou período “pré-formulário” de 1998 a 2000. Essas informações foram conferidas com os dados repassados pela Valesul e Ferteco (empresas que operam alumina e minério de ferro respectivamente) e pela Capitania dos Portos do Rio de Janeiro. Estimativas do volume de água de lastro descarregado pelos navios nesse período foram feitas com base numa planilha Excel fornecida pelos consultores do Programa GloBallast.

Este conjunto de dados foi consolidado juntamente com as informações dos formulários de água de lastro analisados entre 2001 e 2002 (até maio) em um único banco de dados contendo 919 registros de visitas (Tabela 6).

Ano	1998	1999	2000	2001	2002	Total
<b>Registros</b>	98	132	359	208	122	919

Tabela 6 - Número de registros por ano (Fonte: Junqueira & Leal Neto, 2003)





IGU = Interface Gráfica do Usuário

\* para os locais de demonstração relativos a cada porto fonte e destino

Figura 6 - Esquema do sistema de avaliação de risco (adaptado de Clarke *et al.*, 2004, p.6)

Verificou-se que uma notável proporção dos formulários de água de lastro foi preenchida incompleta e/ou incorretamente. Vários tipos de problemas foram identificados, sendo os mais comuns: diferentes tipos de formulários; diferentes unidades utilizadas (algumas vezes a falta de informação da unidade); falta de dados (e.g. data de chegada, nome e posto do oficial responsável); diferentes combinações de tanques na “fonte” e na “descarga” da água de lastro; cópias ilegíveis; escrita incompreensível; dados incoerentes entre as diferentes seções do formulário (número de tanques e/ou volumes); e confusão no campo “*Sea height (m)*” entre ‘profundidade onde aconteceu a troca de água de lastro’ e ‘altura da onda’.

### 3.2.2 Identificação dos portos fonte

No caso do porto de Sepetiba, a partir dos 919 registros de visitas e 1.540 tanques descarregados, o sistema de avaliação de risco identificou 148 portos como fontes de água de lastro. A Figura 7 mostra a saída do Sistema de Informação Geográfica com a representação e importância relativa desses portos fonte com respeito aos volumes descarregados.

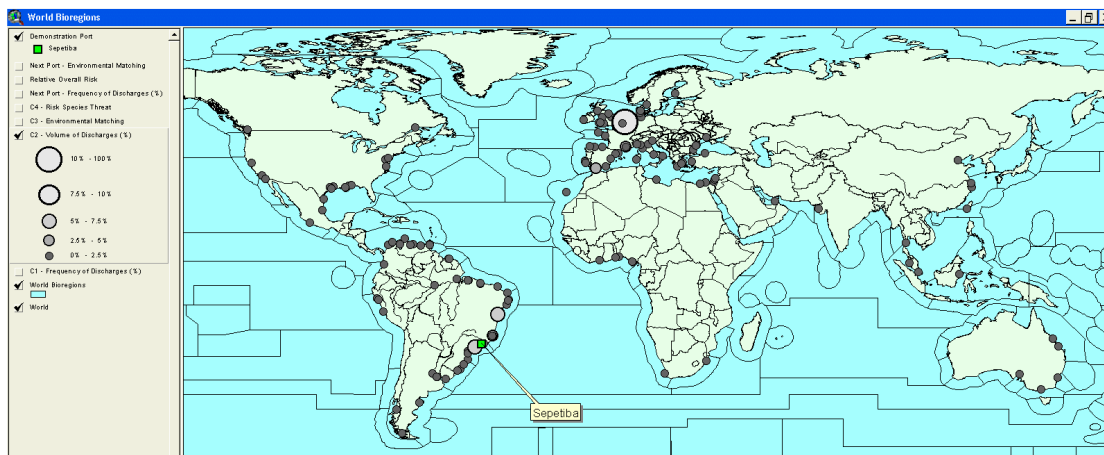


Figura 7 - Localização e importância relativa dos portos fonte com relação aos volumes descarregados no porto de Sepetiba (Clarke *et al.*, 2004, p.39)

Num esforço de reunir o maior número possível de informações sobre os portos fonte dos portos-piloto participantes do Programa GloBallast, foram levantados descritores para um total de 357 portos ao redor do mundo. As informações desses portos e os respectivos 34 parâmetros descritores são apresentadas numa lista que tem o porto de Sepetiba na primeira linha e os demais estão em ordem alfabética de acordo com seu código das Nações Unidas. Esse conjunto de dados, apresentado no Anexo 3, servirá como base para todas as análises que se seguem nesta tese, e será referido como “banco de dados GloBallast”.

As similaridades entre o porto de Sepetiba e seus portos fonte de água de lastro foram determinadas usando o software PRIMER (Clarke & Gorley, 2001). Das várias medidas da distância disponíveis, a distância Euclidiana normalizada (ver item 4.1) foi considerada a mais apropriada, pois a normalização dos vários parâmetros de entrada remove o problema de diferenças de escala, e esta medida pode lidar com uma mistura de valores escalares, inteiros e mesmo valores categorizados. Os valores produzidos pelo PRIMER dão uma medida direta, porém inversa, da similaridade. Uma distância de ‘0’ entre dois portos significa um casamento ambiental ‘perfeito’ e isso somente acontece caso todos os parâmetros tenham exatamente os mesmos valores. No sistema, as distâncias foram convertidas num coeficiente de similaridade

ambiental fazendo uma inversão e uma mudança de escala. Os valores na faixa de variação de 0 a 12,481 foram escalonados para a amplitude de 1,00 (similaridade máxima) a 0,05 (similaridade mínima). Um valor mínimo foi estabelecido (0,05), pois nenhum ambiente pode ser infinitamente distante de outro. Esses coeficientes estão listados no Anexo 4 para todos os 357 portos.

Esses cálculos indicaram que Sepetiba tem uma similaridade relativamente alta com um grande número de portos com os quais mantém comércio (i.e. faixa 0,60 a 0,80). Isso pode estar relacionado à sua localização limítrofe subtropical-tropical com significativa amplitude térmica, e mais o padrão de chuva que restringe qualquer desenvolvimento de salinidades extremas.

Analisando os coeficientes de similaridade apenas dos portos fonte do Anexo 4 nota-se que o porto mais ambientalmente similar com Sepetiba foi o Rio de Janeiro (0,86), junto com outros portos da costa brasileira com coeficientes na faixa de 0,70 a 0,80. Os portos estrangeiros mais similares foram Cingapura (300-SGSIN) (0,63); Trieste (168-ITTRS) na Itália (0,62); e Kembla (25-AUPKL) na costa Leste da Austrália (0,61). Os portos mais dissimilares ambientalmente foram vários fluviais, salobros e/ ou portos de águas frias na América do Norte, sul da Argentina e noroeste europeu (0,20 a 0,30).

É importante notar que foram obtidos dados ambientais suficientes para a análise de similaridade multivariada de apenas 58% dos portos fonte identificados para Sepetiba que respondem por 79,5% dos tanques descarregados registrados. Para os demais portos, nos quais não foi possível obter dados ou esses eram limitados, o coeficiente de similaridade foi estimado a partir do tipo de porto e do coeficiente dos portos mais próximos, com base no Princípio da Precaução. Esta estratégia levou à escolha de valores bastante conservativos em comparação aos coeficientes calculados.

O banco de dados GloBallast contém vários valores que a equipe que finalizou a compilação dos dados julgou não serem muito precisos. Esses valores em particular estão representados com fonte em itálico no Anexo 3. As incertezas associadas com esses descritores ambientais são de vários tipos. Pode ser que tenham sido estimados com base na localização geográfica do porto ou em valores obtidos para os portos comparáveis mais próximos, não significando valores médios/extremos reais. Pode denotar, também, ter como origem uma fonte de dados de pequena escala / baixa resolução, na qual apenas dados grosseiros estavam disponíveis.

Este banco de dados é bastante heterogêneo tanto na sua montagem quanto em seu conteúdo. Várias equipes, distribuídas nos seis países-piloto do Programa GloBallast e em algumas filiais da empresa URS, trabalharam simultaneamente no levantamento dos dados. Várias, também, foram as fontes de obtenção dos dados, seja diretamente ou derivados de informações selecionadas de uma ampla gama de publicações governamentais, portuárias e científicas; páginas da internet; cartas de temperatura da superfície do mar e de salinidade; bases de dados de clima; Atlas; tabelas de maré; cartas náuticas; mapas de sensibilidade de habitats; fotografias aéreas; e informações de especialistas locais.

Essa heterogeneidade e as imprecisões existentes nos dados motivaram o objetivo desta tese de avaliar, além da simples medida direta da distância Euclidiana entre os descritores dos portos, outras técnicas e abordagens para identificar similaridades, comparando métodos alternativos.

#### 4 Reconhecimento de padrões

Ross (1998) destaca o trabalho de Herbert Simon, psicólogo, ganhador do Prêmio Nobel de economia em 1978: “His central finding was that pattern recognition is critical. The more relevant patterns at your disposal, the better your decisions will be”<sup>6</sup> e enfatiza que “This is hopeful news to proponents of artificial intelligence, since computers can surely be taught to recognize patterns. Indeed, successful computer programs that help banks score credit applicants, help doctors diagnose disease and help pilots land airplanes depend in some way on pattern recognition”<sup>7</sup>.

A abordagem intuitiva na concepção de um “classificador” é o uso do conceito de similaridade: objetos que são similares devem ser designados à mesma classe. O conceito de similaridade é extensamente difundido na teoria psicológica. A percepção da similaridade tem sido assunto de muitos estudos e várias teorias são propostas, as quais fazem diferentes suposições a respeito de seus mecanismos (Santini & Jain, 1997). A similaridade desempenha um papel fundamental nas teorias do conhecimento e comportamento. Serve como princípio organizador pelo qual indivíduos classificam objetos, formam conceitos e fazem generalizações (Tversky, 1977).

Supõe-se que medidas de similaridade baseadas em distância não fornecem uma explicação adequada da similaridade percebida (Ashby & Perrin, 1998). Tem sido discutido que as representações dimensionais quantitativas são apropriadas para certos estímulos, mas não para outros. Conforme Tversky (1977) seria mais apropriado representar faces, países, ou personalidades em termos de várias características qualitativas do que em termos de grandezas quantitativas. A avaliação de similaridade entre tais estímulos, poderia ser mais

---

<sup>6</sup> Sua principal conclusão foi que o reconhecimento de padrões é crítico.

<sup>7</sup> Quanto mais padrões relevantes à disposição, melhores serão as decisões. Esta é uma notícia promissora para proponentes de inteligência artificial, pois computadores podem ser ensinados a reconhecer padrões. De fato, programas computacionais bem sucedidos, que ajudam bancos a classificar clientes que solicitam crédito; médicos a fazerem diagnósticos e pilotos a aterrissarem aviões, dependem de alguma forma de reconhecimento de padrões”.

bem descrita como uma comparação de características do que um cálculo de uma distância métrica entre pontos. No entanto, a análise teórica das relações de similaridade tem sido dominada por modelos geométricos. Esses modelos representam objetos como pontos em algum espaço de coordenadas tal que as dissimilaridades observadas entre objetos correspondem à distância métrica entre os respectivos pontos.

Para agrupar objetos de mesmas características é preciso antes descrevê-los de acordo com algum esquema. Medir uma similaridade significativa é uma dicotomia que se apóia em dois elementos: descobrir um grupo de descritores que codifiquem adequadamente as características do que se pretende medir; e dotar o espaço dos descritores com uma métrica apropriada (Santini & Jain, 1999). Após ter selecionado um conveniente grupo de descritores e ter caracterizado o objeto como um ponto num espaço vetorial propício, escolhe-se uma medida apropriada que quantifique o grau de associação entre os objetos ou descritores desses objetos.

#### **4.1 Medidas de similaridade**

As medidas de similaridade são grandezas numéricas que quantificam o grau de associação entre um par de objetos pelos seus descritores. Esses objetos podem ser descritos de três formas principais (Clarke & Warwick, 2001):

- a) números absolutos: dados quantitativos obtidos para cada descritor.
- b) números relativos: dados escalonados, tal padronização é essencial quando os números absolutos não são comparáveis seja por diferença de unidades ou pela base de medição.
- c) números binários - dados reduzidos à simples presença (1) ou ausência (0).

Várias medidas de similaridade podem ser encontradas na bibliografia especializada. Valentin (2000) cita como referências mais recentes: Sneath & Sokal (1973); Wolda (1981); Pielou (1984); Ludwig & Reynolds (1988); e Legendre & Legendre (1998). É necessário verificar a adequabilidade (vantagens e desvantagens) de cada medida e quais critérios e

condições satisfazem. Existem muitas maneiras de definir o conceito de similaridade entre pares de objetos, cada uma enfatizando um aspecto diferente do conjunto de dados representativos desses objetos. O que constitui a similaridade total, ou a dissimilaridade total, de dois objetos depende do coeficiente adotado.

Uma medida, particularmente comum em ecologia, é o coeficiente de Bray-Curtis que tem a propriedade de variar entre 1 (similaridade) e 0 (dissimilaridade) caso todos os valores de entrada sejam positivos. O cálculo da similaridade é baseado na razão entre as diferenças absolutas e as somas das entradas:

$$S_{A-B} = 1 - \frac{\sum_{i=1}^n |x_{Ai} - x_{Bi}|}{\sum_{i=1}^n (x_{Ai} + x_{Bi})} \quad (1)$$

Aqui  $x_{Ai}$  e  $x_{Bi}$  representam a entrada  $i$  do vetor de dados dos objetos “A” e “B”. Outras medidas têm formulação semelhante à de Bray-Curtis, tais como as medidas de Manhattan e de Cambera, e em certas aplicações os resultados são virtualmente indistinguíveis.

Os coeficientes de distância são preferencialmente adotados quando se pretende verificar a proximidade entre dois objetos, em função da composição específica ou de qualquer outro descritor desses objetos. Quanto mais próximos forem os objetos, i.e., menor a distância métrica entre os pontos representativos desses dois objetos, maior será a similaridade entre eles (Valentin, 2000).

Uma função de distância métrica,  $d$ , é uma escala que assinala a cada par de objetos um valor não negativo de acordo com certos axiomas métricos. A primeira exigência é que  $d(A,A) = d(B,B)$  para todos os objetos (constância de auto-similaridade). Um segundo axioma da função é a minimalidade  $d(A,B) \geq d(A,A)$ . O terceiro axioma, da simetria, implica que para todos os objetos  $A$  e  $B$ :  $d(A,B) = d(B,A)$ . O axioma final para a função é conhecido como a desigualdade do triângulo. Para três objetos  $A$ ,  $B$  e  $C$ , a desigualdade do triângulo expressa que:  $d(A,B) + d(B,C) \geq d(A,C)$  (Santini & Jain, 1996; Ashby & Perrin, 1998).

A distância Euclidiana, uma das mais conhecidas, e perfeitamente métrica, é calculada com base no teorema de Pitágoras, sobre a hipotenusa do triângulo retângulo (Figura 8).

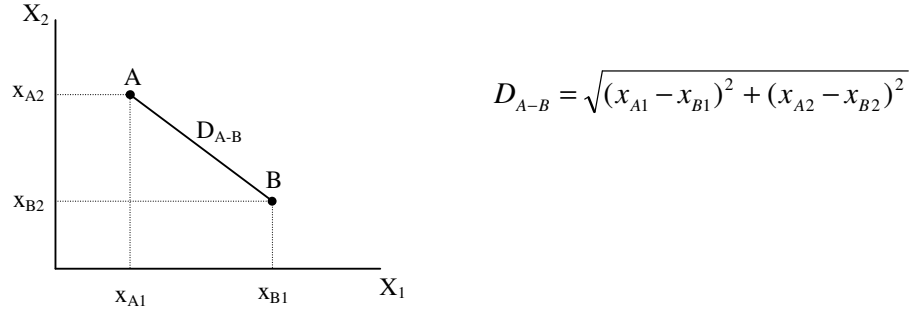


Figura 8 - Distância Euclidiana  $D_{A-B}$  entre dois objetos A e B (elaboração própria)

A distância Euclidiana não tem limite superior. Ela aumenta à medida que aumenta o número de descritores. Além disso, ela depende da escala de valores de cada descritor. Este inconveniente pode ser corrigido pela padronização dos dados. Generalizando para dois objetos contendo  $p$  descritores, a distância Euclidiana expressa-se por:

$$D_{A-B} = \sqrt{\sum_{i=1}^p (x_{Ai} - x_{Bi})^2} \quad (2)$$

Considerando que muitos algoritmos utilizam a métrica Euclidiana para medir distâncias entre objetos, é altamente desejável (mas não obrigatório) realizar um pré-processamento nos dados antes de analisá-los, seja por meio de simples transformações ou normalização. Quando se mede a distância entre objetos pode-se deparar com um problema de escala: descritores que tenham maiores valores irão quase que completamente dominar os resultados. De modo a obter uma comparação mais adequada, onde os descritores tenham igual importância, examinou-se a normalização e a normalização padronizada dos dados de entrada.

A normalização é feita dividindo cada componente do vetor pelo comprimento deste vetor. O comprimento é determinado pela raiz quadrada da soma dos quadrados de todos os componentes do vetor. Assim, se  $\mathbf{x}_k = x_{k1}, x_{k2}, \dots, x_{kp}$ :

$$\mathbf{x}'_k = \mathbf{x}_k / (x_{k1}^2 + x_{k2}^2 + \dots + x_{kp}^2)^{1/2} \quad (3)$$



onde  $\mathbf{x}_k'$  é um vetor unitário na mesma direção de  $\mathbf{x}_k$ , ou seja  $\mathbf{x}_k'$  é um vetor de comprimento unitário no espaço  $p$ -dimensional. A equação (3) generaliza o caso familiar de duas dimensões no qual o comprimento de um vetor é igual à hipotenusa do triângulo retângulo formado por seus componentes  $x$  e  $y$  (aplicação do teorema de Pitágoras (ver Figura 9)).

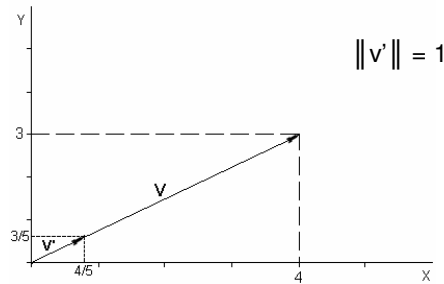


Figura 9 - Vetor Unitário (adaptado de Wasserman, 1989, p. 65)

A transformação conhecida como *norma-Z* faz com que os dados transformados tenham média zero e variância igual a um. A variável normal padronizada ou normal reduzida  $z$ , é obtida conforme a equação (4):

$$z_k = \frac{\mathbf{x}_k - \bar{\mathbf{x}}_k}{s_{xk}}, \quad (4)$$

onde:  $\bar{\mathbf{x}}_k$  e  $s_{xk}$  denotam, respectivamente, a média e o desvio padrão do componente  $\mathbf{x}_k$ .

Experimentalmente as duas maneiras de se normalizar conduziram a resultados divergentes. Para ilustrar isso se utilizou um conjunto de entradas com duas dimensões e escalas diferentes, formado por 1.000 dados divididos em quatro grupos representados por marcadores geométricos distintos (“□”, “◇”, “+” e “x”) como mostrado na Figura 10.

Um inconveniente ao se usar a normalização é que essa operação causa uma mudança na geometria dos dados de entrada. Os dados normalizados ficam mesclados (em formato de arco de círculo, com raio igual a 1) de forma que a separação existente entre os grupos se perde, enquanto que com a normalização padronizada, a disposição relativa entre os dados se mantém (Figura 11).

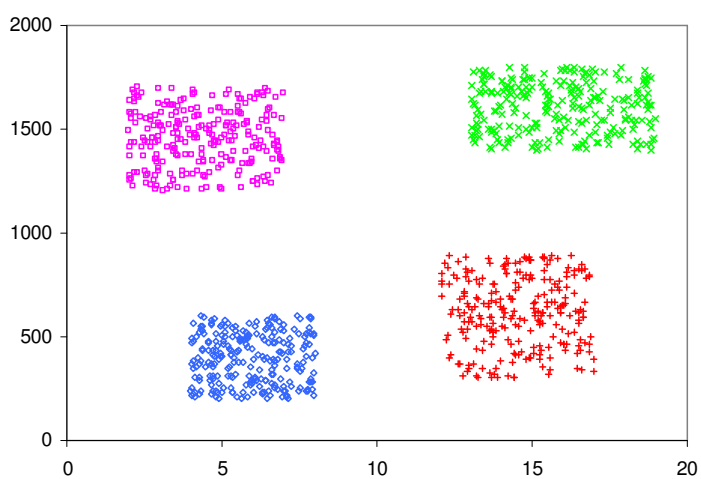


Figura 10 - Exemplo de dados de entrada (adaptado de Carvalho, 2001, p. 120).

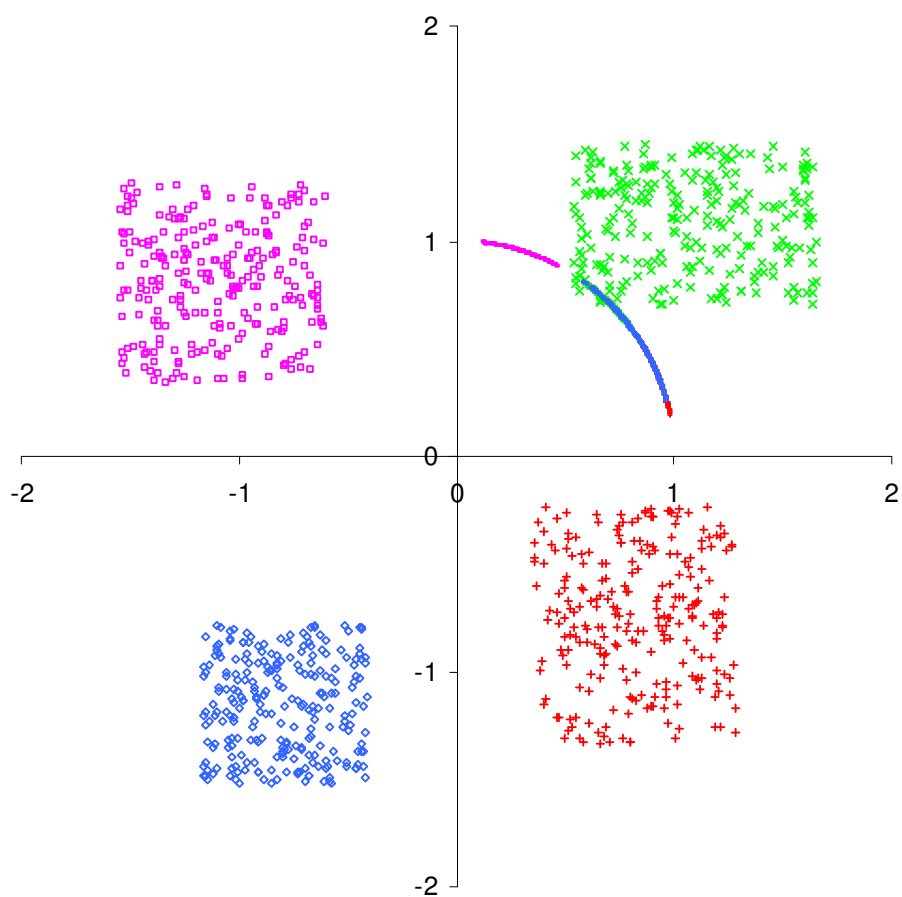


Figura 11 - Resultado da normalização – em formato de arco de círculo restrito ao primeiro quadrante – e da normalização padronizada (elaboração própria)

## 4.2 Redução da dimensionalidade

Existem duas razões principais para manter a dimensionalidade da representação (i.e. o número de descritores) tão pequena quanto possível: custos de medição e precisão da classificação. Um conjunto de descritores limitado, porém, relevante, simplifica ambos a representação e os classificadores que são construídos baseados na representação selecionada. Por outro lado, uma redução no número de descritores pode levar a uma perda no poder de discriminação e desse modo diminuir a precisão resultante do sistema de reconhecimento.

É importante fazer uma distinção entre a seleção e a extração de descritores. O termo seleção de descritores refere-se a algoritmos que selecionam o melhor subgrupo dentro do grupo de descritores de entrada. Métodos que criam novos descritores baseados em transformações ou combinações do conjunto original são chamados algoritmos de extração de descritores. Frequentemente a extração precede a seleção de descritores; primeiro, descritores são extraídos dos dados medidos (e.g. usando análise de componentes principais ou discriminante) e, então, alguns dos descritores extraídos com menor capacidade de discriminação são descartados.

A seleção de descritores leva à economia nos custos de medição (já que alguns são abandonados). Além disso, por reter sua interpretação física original, os descritores mantidos podem ser importantes para entender o processo que forma os padrões. Por outro lado, descritores transformados gerados pela extração podem fornecer uma melhor capacidade discriminativa que o melhor subgrupo dos descritores dados inicialmente, porém, esses novos descritores (obtidos por uma combinação linear ou não-linear) podem não ter um claro significado físico (Jain *et al.*, 2000).

## 5 Análise Multivariada de Dados

Conjuntos de dados ecológicos são oriundos de numerosos processos bióticos e abióticos que interagem entre si, de tal forma que seus diferentes efeitos não podem ser interpretados separadamente. A análise multivariada de dados refere-se a todos os métodos estatísticos que analisam simultaneamente múltiplas medidas feitas em amostras do objeto pesquisado (Hair *et al.*, 2005), ao contrário da estatística elementar que trata de variáveis independentemente. Os principais objetivos das técnicas de análise multivariada são detectar e descrever a estrutura de dependência entre descritores, quantificando seu grau de associação.

O sucesso da análise e interpretação é relacionado com a compatibilidade entre os dados e os métodos de processamento. Portanto, é importante levar em consideração os requisitos das técnicas diante dos dados amostrados. Essas técnicas podem ser divididas em dois grandes grupos (Legendre & Legendre, 1998): (1) agrupamento por algoritmos de aglomeração ou divisão; e (2) ordenação num espaço de dimensionalidade reduzida. Os itens a seguir as descrevem sucintamente. Informações mais detalhadas sobre os métodos estão nas publicações referidas.

O exame dos dados é uma parte essencial de qualquer análise multivariada. A Tabela 7 apresenta os principais parâmetros estatísticos (máximo, mínimo, média e desvio padrão) dos descritores dos portos contidos no banco de dados GloBallast. Para auxiliar esse exame foram elaborados gráfico de caixas<sup>8</sup> para os descritores escalares (Figura 12 a Figura 21). Em alguns gráficos certos rótulos foram suprimidos por questão de visualização.

---

<sup>8</sup> Conforme Hair *et al.* (2005), o gráfico de caixas é uma representação pictórica da distribuição dos dados. Os limites superior e inferior marcam os quartis superior e inferior, de modo que a caixa contém 50% dos valores centrais dos dados. A linha traçada no interior da caixa representa a mediana. Se a mediana se encontra próxima a um dos extremos da caixa, isso indica assimetria naquela direção. Quanto maior a caixa, maior a dispersão dos dados. As linhas que se estendem de cada caixa (chamadas de *whiskers*) representam a distância à menor e à maior observação que estão a menos de um quartil da caixa. As observações atípicas são as que variam entre 1,0 e 1,5 quartis de distância da caixa. Os valores extremos são as observações que estão a mais de 1,5 quartis do extremo da caixa.

Dos dezoito descritores, apenas 5 não tiveram valores fora da variação de um quartil, as temperaturas média e mínima da água e temperatura média do ar na estação fria, e a salinidade média e mínima do período chuvoso (MWNWT, LWNWT, WMART, MWSAL e LWSAL). Na Figura 14 pode-se verificar que o porto Bandar Imam Khomeyni (153-IRBKM) do Irã tem a maior temperatura máxima do ar do dia na estação mais quente de 52°C e na Figura 17 que o porto de Jubail (289-SAJUB) na Arábia Saudita tem a maior salinidade média e máxima durante o período seco de 50,0 e 52,0 g/l, respectivamente.

Nº	Descritor	Máximo	Mínimo	Média	Mediana	Desvio Padrão
1	PTYPE	6	1	3,17	3,00	1,41
2	MSUWT	34,9	8,0	24,64	25,00	5,04
3	USUWT	36,5	11,5	27,51	27,50	5,14
4	MWNWT	31,0	-1,0	16,25	16,50	7,51
5	LWNWT	30,0	-6,0	13,28	14,00	7,89
6	MSART	37,0	14,0	28,01	29,00	4,40
7	USART	52,0	21,0	35,91	36,00	5,04
8	MWART	28,0	-17,0	11,38	9,50	9,28
9	LWART	24,8	-30,0	4,17	3,00	10,76
10	MWSAL	49,0	0,0	26,49	30,00	11,09
11	LWSAL	48,0	0,0	22,73	24,00	12,47
12	MDSAL	50,0	0,0	29,35	32,20	10,08
13	UDSAL	52,0	0,0	31,39	34,00	9,36
14	MSPR	9,1	0,0	2,08	1,90	1,64
15	MNER	5,5	0,0	0,89	0,60	0,93
16	D6MRF	1.759,9	0	347,93	323,00	293,12
17	W6MRF	2.915,4	13,7	766,36	739,20	546,30
18	RNFL75	9	2	6,45	7,00	1,85
19	DISRVM	450	-250	22,86	5,00	71,82
20	SIZRVC	2.000.000	10	100.765,71	1.500,00	283.665,83
21	INASMW	5	3	4,96	5	0,22
22	INARKW	5	0	3,88	4	1,25
23	INAWP	5	0	3,51	4	1,33
24	INHTSM	4	0	1,52	2	1,29
25	INSNDB	5	0	3,03	3	1,13
26	INSTNB	5	0	2,35	3	1,47
27	INLTMF	5	0	2,75	3	1,47
28	INMANG	5	0	1,25	0	1,61
29	INRKSH	5	0	3,06	3	1,37
30	SUFSND	5	0	3,71	4	1,10
31	SUSFTM	5	0	4,30	5	1,10
32	SUSGRM	4	0	2,22	2	1,38
33	SURKRF	5	0	3,11	3	1,58
34	SUCORF	5	0	1,07	0	1,58

Tabela 7 - Estatística dos descritores dos portos contidos no banco de dados GloBallast

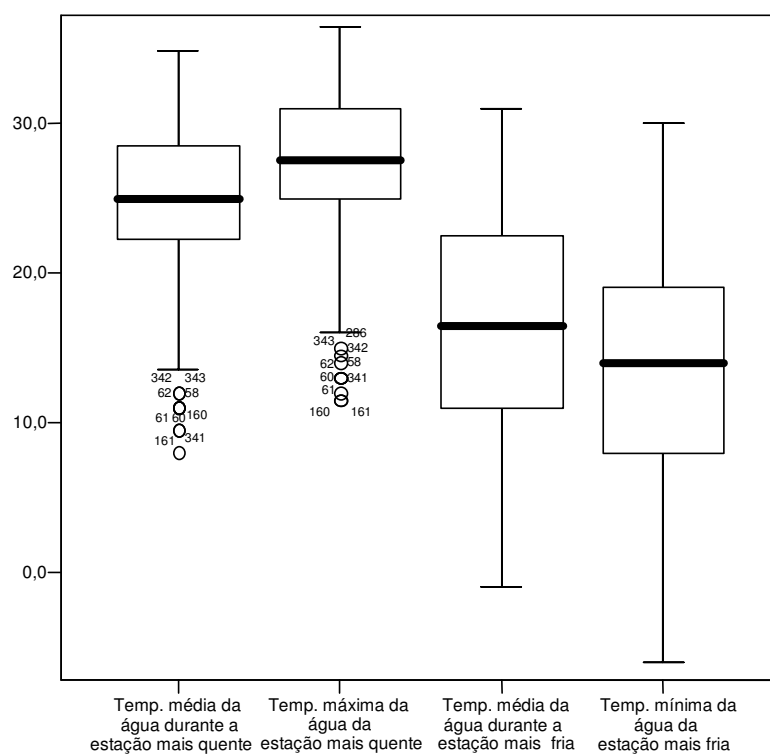


Figura 12 - Gráfico de caixas dos descritores MSUWT, USUWT, MWNWT e LWNWT (elaboração própria)

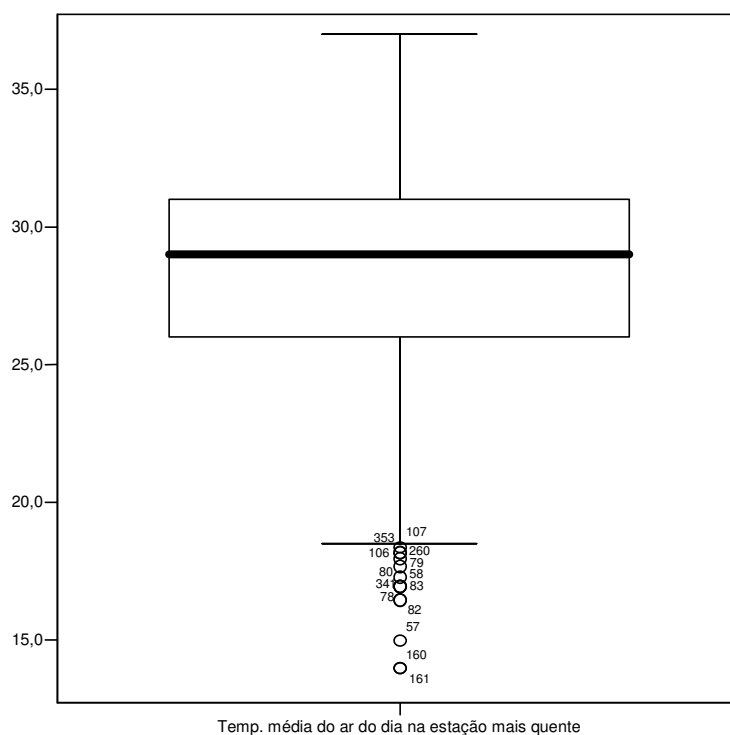


Figura 13 - Gráfico de caixas do descritor MSART (elaboração própria)

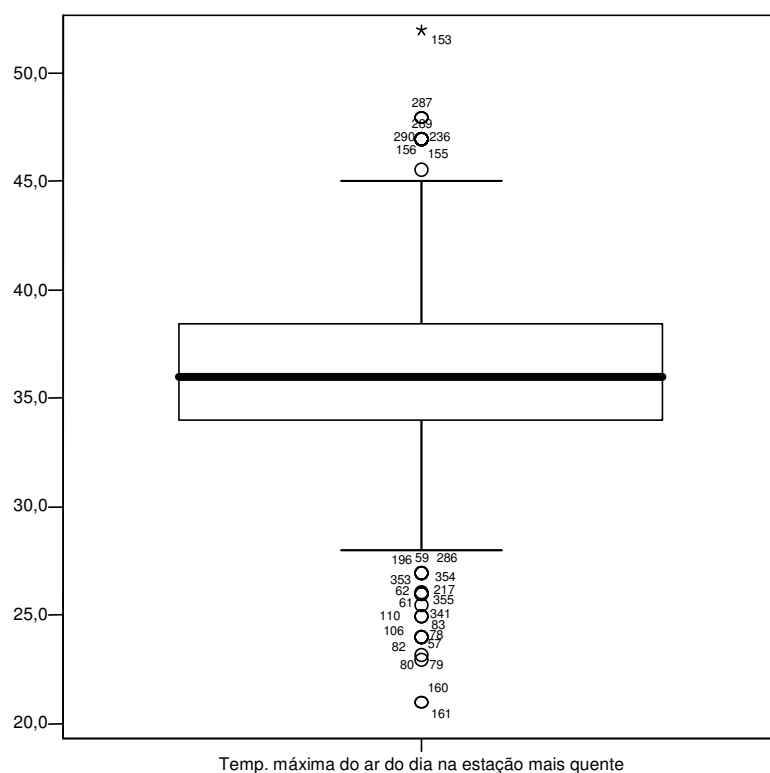


Figura 14 - Gráfico de caixas do descritor USART (elaboração própria)

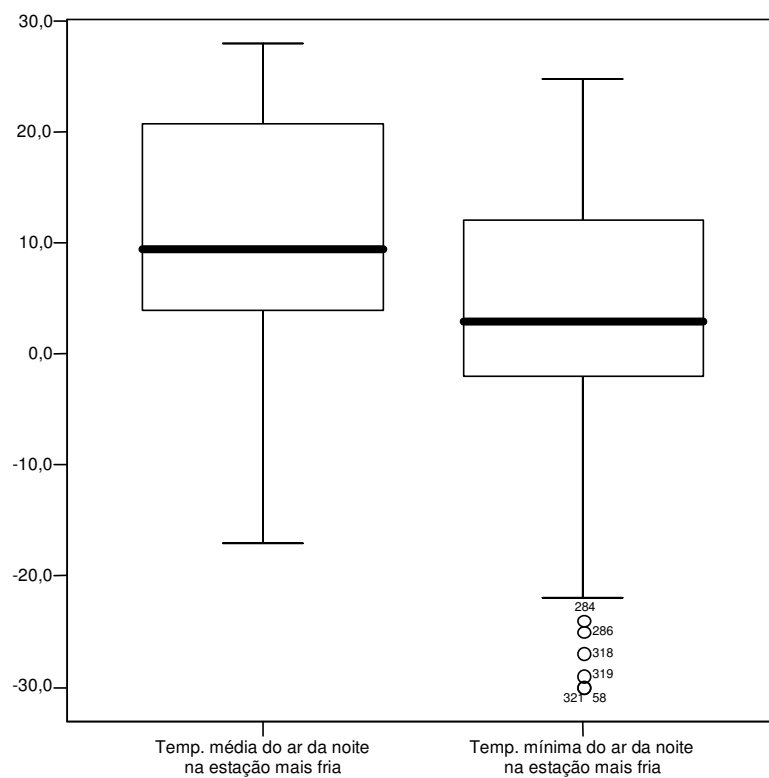


Figura 15 - Gráfico de caixas dos descritores MWART e LWART (elaboração própria)

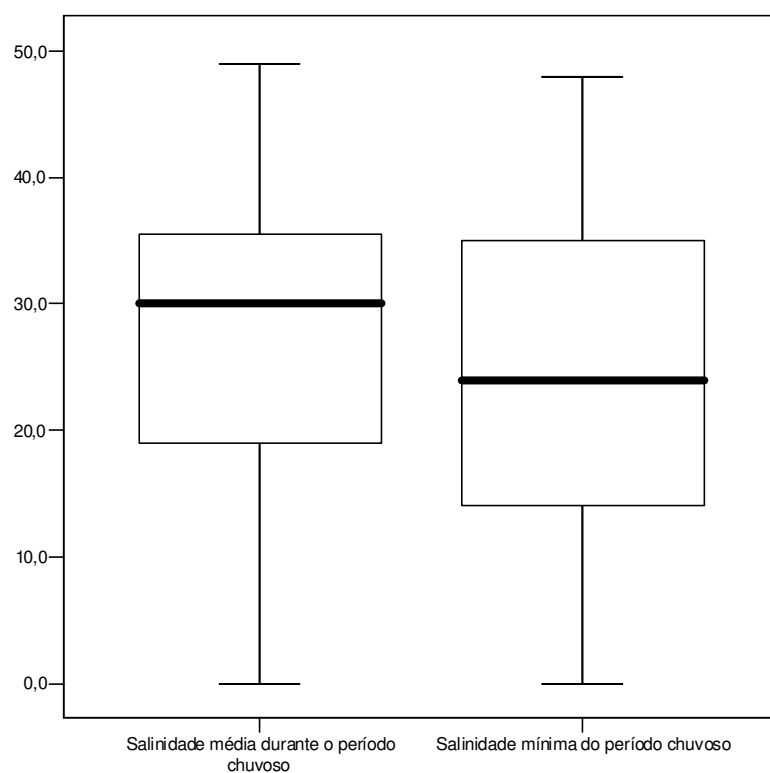


Figura 16 - Gráfico de caixas dos descritores MWSAL e LWSAL (elaboração própria)

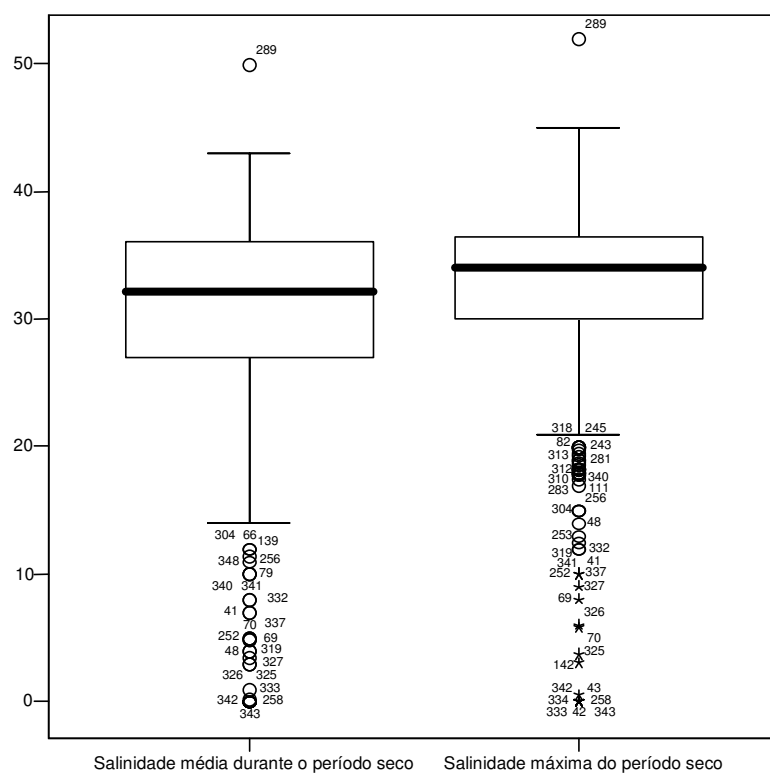


Figura 17 - Gráfico de caixas dos descritores MDSAL e UDSAL (elaboração própria)



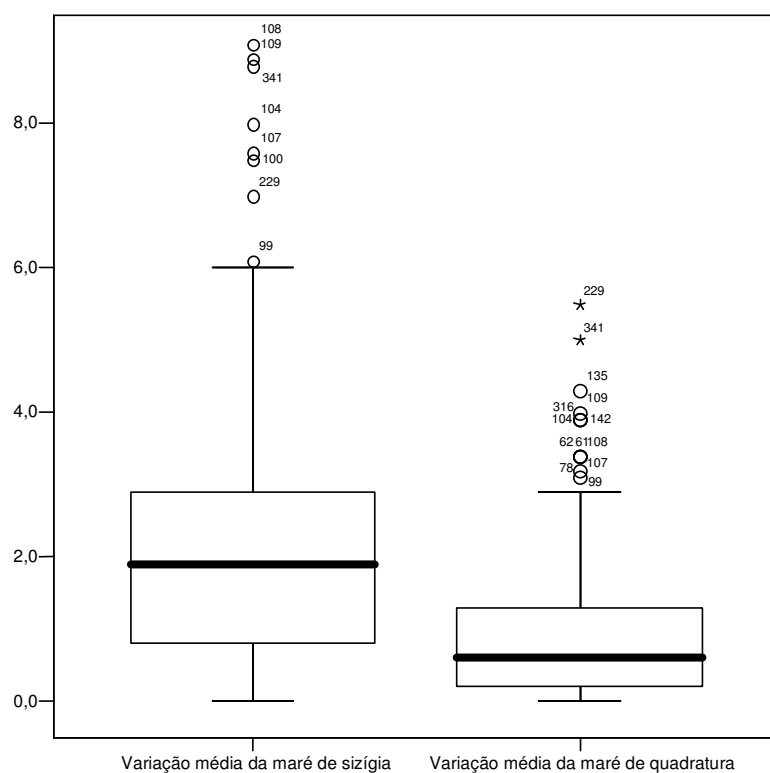


Figura 18 - Gráfico de caixas dos descritores MSPR e MNER (elaboração própria)

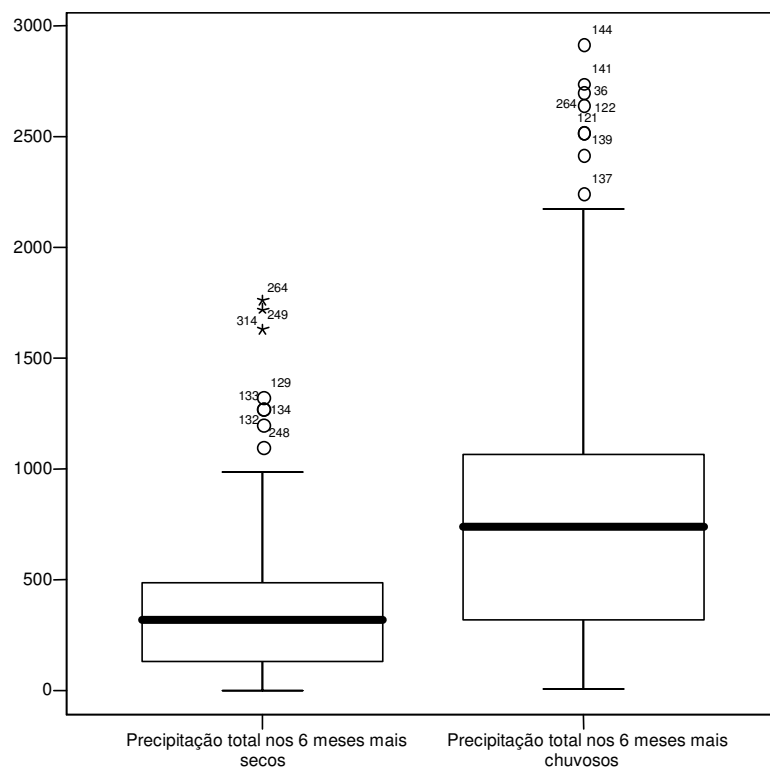


Figura 19 - Gráfico de caixas dos descritores D6MRF e W6MRF (elaboração própria)

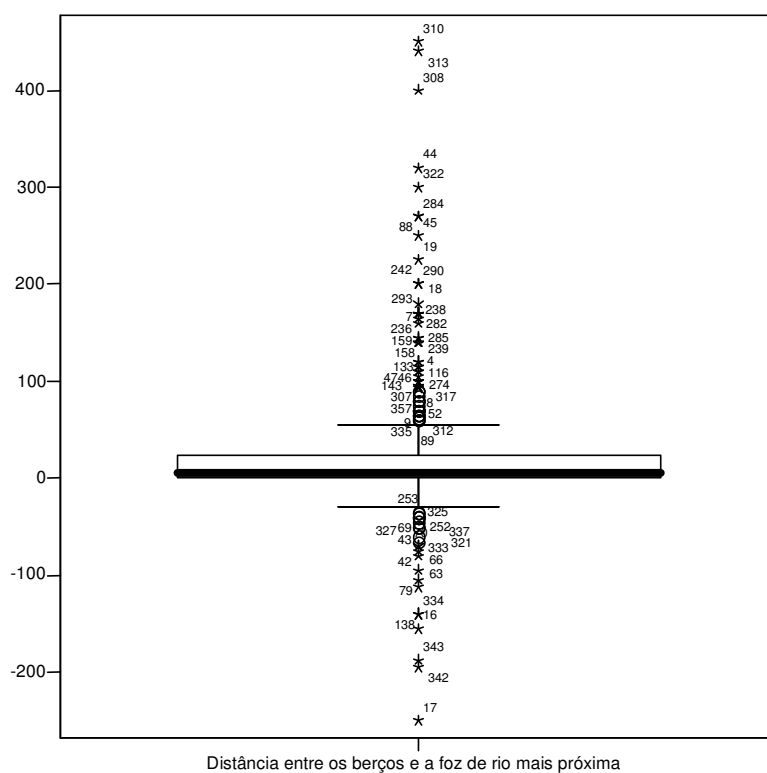


Figura 20 - Gráfico de caixas do descritor DISRVM (elaboração própria)

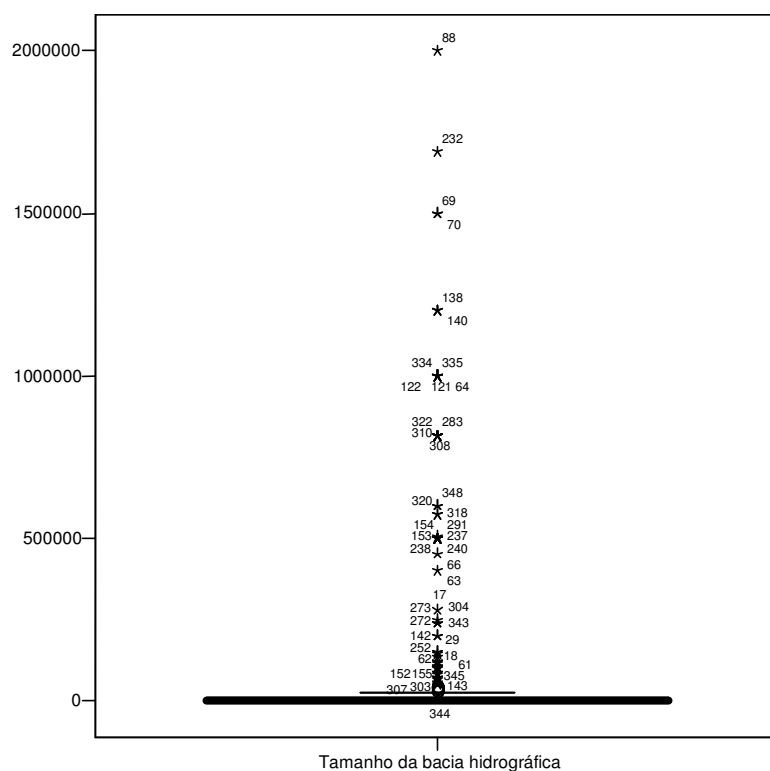


Figura 21 - Gráfico de caixas do descritor SIZRVC (elaboração própria)

## 5.1 Análise de Agrupamentos

O objetivo da análise de agrupamentos é classificar, com respeito a algum critério predeterminado, objetos em grupos mutuamente exclusivos. Para agrupar objetos é necessário reconhecer entre eles um grau de similaridade suficiente que permita reuni-los num mesmo grupo e, simultaneamente, perceber distinções ou separações entre os grupos. Desta forma, os grupos resultantes devem exibir homogeneidade interna (entre os objetos de um mesmo grupo) e heterogeneidade externa (entre objetos de grupos distintos). No caso do meio físico, a não ser que este seja fortemente descontínuo e que a amostragem tenha sido realizada de cada lado de um forte gradiente, tem-se geralmente dificuldade em definir nitidamente agrupamentos em virtude do conceito de *continuum* que caracteriza os ecossistemas (Valentin, 2000).

A análise de agrupamentos abrange diferentes algoritmos e métodos que podem ser divididos em sequenciais ou simultâneos; aglomerativos ou divisivos; monotéticos ou politéticos<sup>9</sup>; probabilísticos ou não-probabilísticos; e hierárquicos ou não-hierárquicos, também chamados métodos de partição (Legendre & Legendre, 1998). Na tese é utilizada essa última divisão.

### 5.1.1 Métodos hierárquicos

Nos métodos de agrupamento hierárquico as amostras pertencentes a um grupo são progressivamente reunidas num grupo maior, até formar um único grupo. O processo pode iniciar, também, com este único grupo, contendo todas as amostras, e prosseguir dividindo-se até chegar ao final com grupos individuais. Os métodos diferem de acordo com o critério usado para decidir quais dois pequenos grupos serão aglutinados ou qual grupo maior dividir. Para fazer os agrupamentos utiliza-se sempre o mesmo critério: reúnem-se amostras / grupos que tiverem maior similaridade entre si. O que caracteriza os diferentes algoritmos é

---

<sup>9</sup> O método politético utiliza todas as variáveis simultaneamente enquanto o monotético faz cada divisão de acordo com uma única variável.

exatamente o critério usado para definir a similaridade. Entre os principais métodos de agrupamento hierárquicos estão: ligação simples (ou vizinho mais próximo); ligação completa (ou vizinho mais distante); associação média; pesos proporcionais; e variância mínima.

O método escolhido para o agrupamento dos portos, realizado por meio do software SPSS, foi o método da variância mínima, também chamado de método Ward, pois este é considerado por vários autores (e.g. Valentin, 2000; Brandalize & Philips, 2002) como sendo bastante eficiente na formação de grupos. No princípio, têm-se um grupo para cada amostra componente da base de dados. Neste estágio inicial a variância em todos os grupos é nula, pois cada amostra é o próprio vetor médio do grupo. A cada etapa no procedimento de agrupamento, a pertinência à um grupo é avaliada calculando-se a variância intragrupo para todas as alternativas de aglomeração, escolhendo a que proporciona a menor. As possibilidades de aglutinação entre os grupos são verificadas, e é escolhido o agrupamento que causa o menor aumento no erro interno do grupo.

Na Figura 22 o resultado final obtido com o método Ward aplicado aos dados normalizados (norma Z) é representado graficamente numa estrutura chamada dendrograma<sup>10</sup> que mostra como os grupos estão relacionados entre si. Conforme a Tabela 8, o grupo 16 foi o maior, constituído de 54 portos, principalmente oriundos do Japão (45 portos). Não foram detectados dados extremos (*outliers*), normalmente identificados como ramos isolados da árvore e que acabam por encontrar os demais ramos somente quando próximos ao topo. O menor grupo foi o 21 com quatro portos (3 da Turquia e 1 da Ucrânia). O porto de Sepetiba situou-se no grupo 5, juntamente com outros 8 portos brasileiros (Salvador; Praia Mole; Tubarão; Vitória; Ponta Ubu; Rio de Janeiro; Santos; e Paranaguá) e 8 oito portos da Índia e um da Colômbia.

Grupo	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Portos	19	13	6	22	18	6	7	7	8	9	23	22	11	18	20	54	29	13	17	13	4	18

Tabela 8 - Número de portos em cada grupo formado pelo método de Ward (elaboração própria)

<sup>10</sup> Com origem na palavra dendrologia, ramo da botânica dedicado ao estudo das árvores.

Um problema com a análise de agrupamento é identificar um número ótimo de grupos. À medida que o processo de fusão prossegue, grupos cada vez mais dissimilares são aglutinados. A decisão do número de grupos para análise é muito subjetiva e é tomada, geralmente, a partir do exame do dendrograma, onde podem ser lidas as distâncias de ligação em que ocorrem nas junções dos pontos observados para formar grupos. Um grande salto nessas distâncias no dendrograma indica que a agregação reuniu dois grupos muito dissimilares e, em razão disso, deve-se definir o número de grupos anterior a este salto.

Ao cortar o dendrograma num nível desejado, um agrupamento dos dados em grupos disjuntos é obtido. A Tabela 9 apresenta várias soluções de agrupamentos de acordo com a distância de combinação de grupos. Aumentando a distância obtêm-se diferentes composições de grupos até a união total. No dendrograma percebe-se uma relativa homogeneidade até uma distância de combinação entre 4 e 6, com oito e cinco grupos formados respectivamente.

Distância de combinação de grupos								
Inicial	2	3	4	5	6	10	20	Final
1	W14-1	W10-1	W8-1	W6-1	W5-1	W4-1	W2-1	1
2	W14-2							
3								
4								
5	W14-3	W10-2	W8-2	W6-2	W5-2	W4-1		
6	W14-4							
7								
8								
9	W14-5	W10-3	W8-3	W6-3	W5-3	W4-2		
10								
11								
12	W14-6	W10-4	W8-4	W6-4	W5-4	W4-3	W2-2	
13	W14-7	W10-5	W8-5					
14	W14-8	W10-6	W8-6					
15								
16	W14-9	W10-7	W8-7	W6-5	W5-5	W4-4		
17	W14-10							
18	W14-11							
19	W14-12	W10-8	W8-8	W6-6				
20	W14-13	W10-9						
21	W14-14	W10-10						
22								

Tabela 9 - Agrupamentos de acordo com a distância de combinação de grupos (elaboração própria)

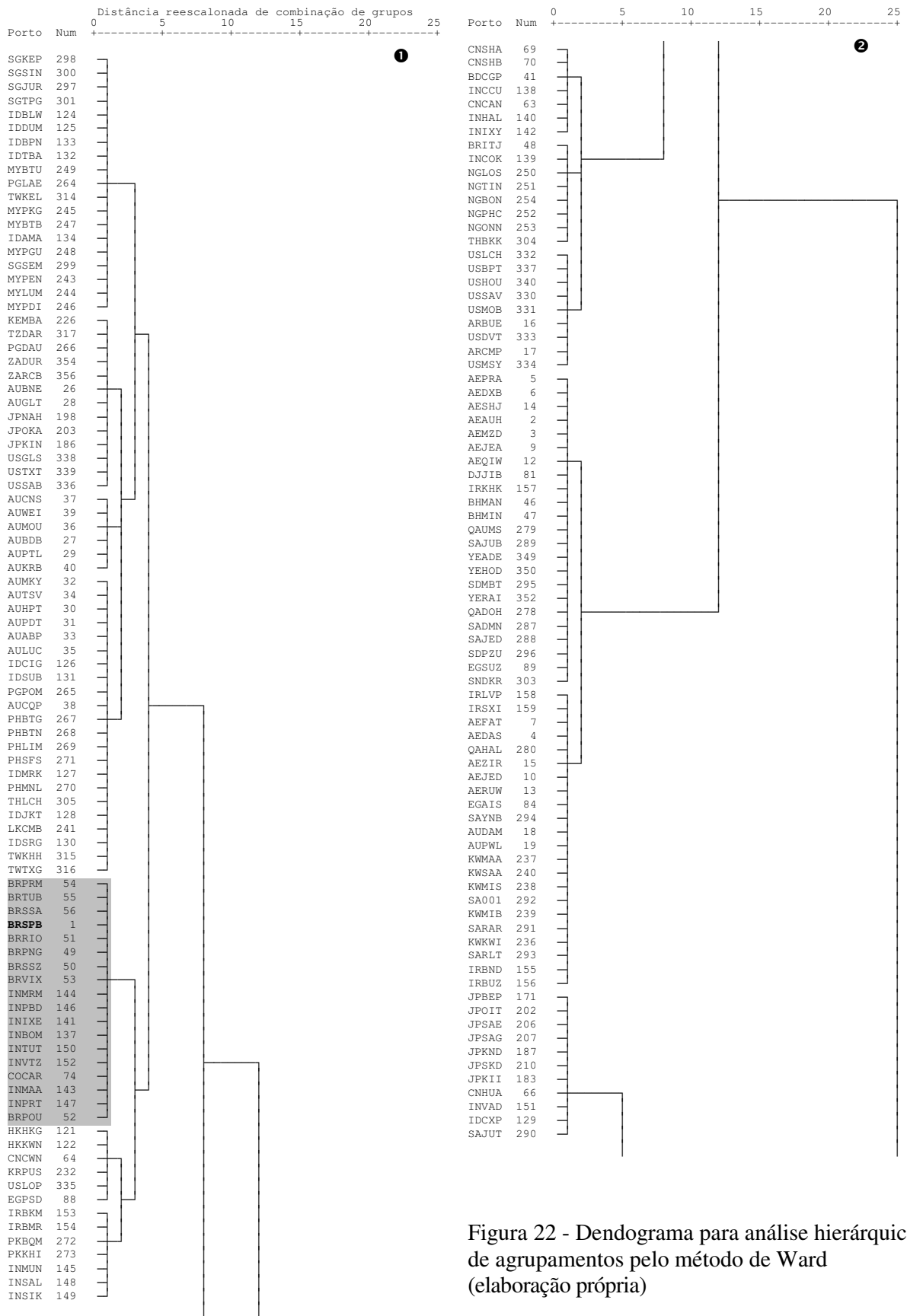


Figura 22 - Dendrograma para análise hierárquica de agrupamentos pelo método de Ward (elaboração própria)

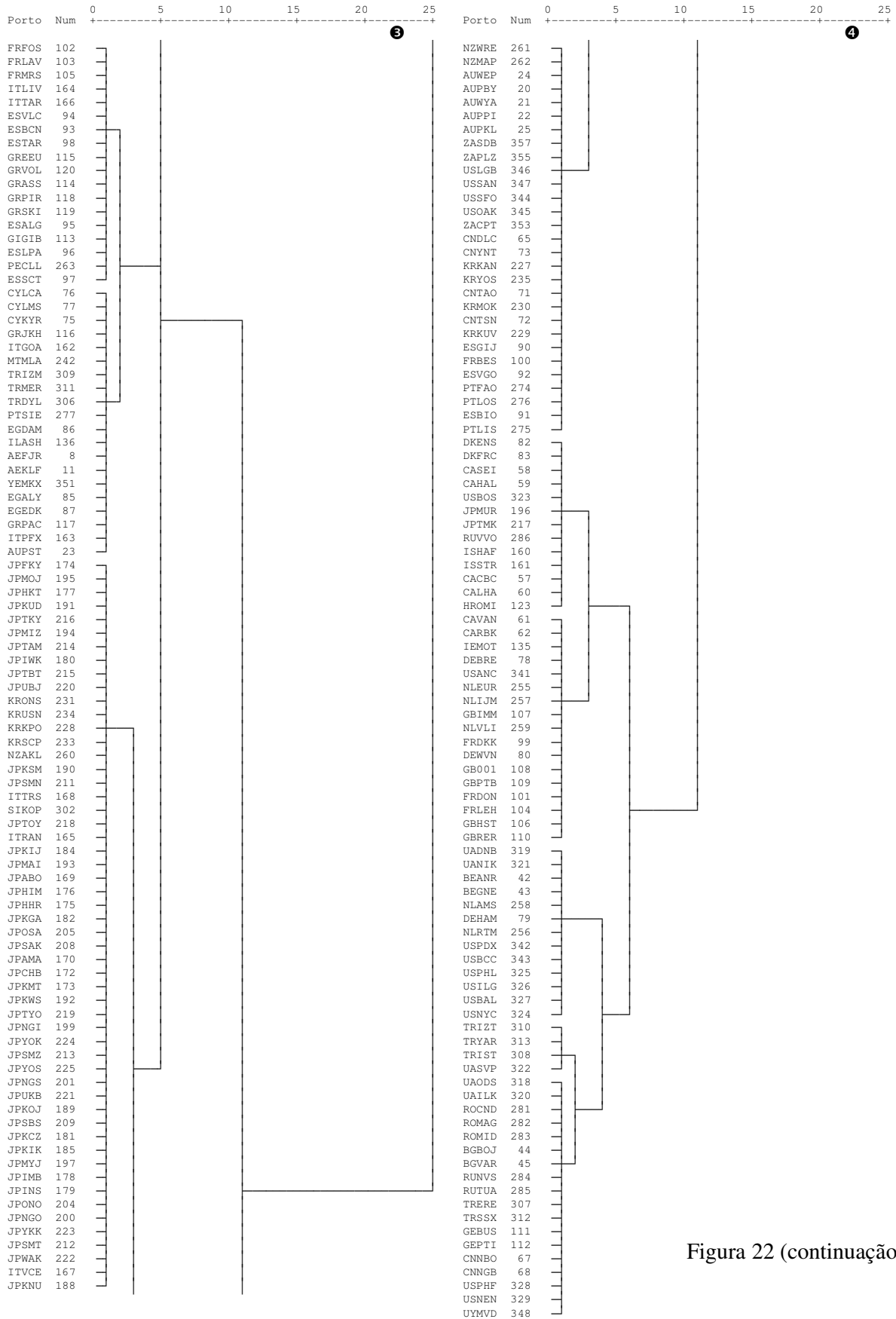


Figura 22 (continuação)

### 5.1.2 Métodos não-hierárquicos

Os métodos não-hierárquicos produzem uma única partição que otimiza a homogeneidade intragrupo em vez de uma série de partições hierárquicas que otimizam os caminhos hierárquicos entre as amostras e os principais grupos. Os métodos não-hierárquicos diferem entre si pela maneira que constituem a melhor partição, e o mais conhecido é frequentemente referido como método k-médias (MacQueen, 1967 *apud* Bottou & Bengio, 1995). Seu uso pressupõe a escolha a priori de um número  $k$  de partições desejadas. Ao contrário dos métodos hierárquicos, uma solução com  $k$  grupos não é apenas uma combinação de dois grupos da solução com  $k+1$  grupos, mas é baseada em encontrar a melhor solução com  $k$  grupos, onde os objetos são relocados se eles estiverem mais próximos de outro grupo do que daquele que originalmente lhes foi atribuído.

O algoritmo é iniciado escolhendo  $k$  centróides iniciais, seja de forma aleatória ou usando alguma heurística. Um centróide é o centro de massa de um grupo, cujas coordenadas num modelo Euclidiano são a média aritmética das coordenadas de todos os membros desse grupo. Em seguida, é calculada uma matriz de distância entre todos os pontos e os centróides. Constrói-se, então, uma partição pela associação de cada ponto com o centróide mais próximo. Os centróides são recalculados para os novos agrupamentos, e o algoritmo repete esses dois passos até a convergência, que é obtida quando os pontos não trocam mais de grupos (ou alternativamente, os centróides não se alteram mais). Finalmente, o algoritmo visa minimizar a função objetivo, neste caso uma função de erro quadrado:

$$J = \sum_{i=1}^k \sum_{j \in G_i} |x_j - c_i|^2 \quad , \quad (5)$$

onde:  $k$  é o número de grupos  $G_i$ ,  $i = 1, 2, \dots, k$  e

$c_i$  o centróide ou média de todos os pontos  $x_j \in G_i$



Em termos de desempenho do algoritmo não há garantia que um ótimo global seja atingido, ou seja, a função objetivo minimizada. A qualidade da solução final depende muito do conjunto inicial de grupos. Costuma-se usar o agrupamento hierárquico como sugestão para o estabelecimento do número de grupos a serem formados. A Tabela 10 reúne os resultados obtidos com o método k-médias, aplicado com o uso do software SPSS, para os mesmos números de grupos gerados pelo método Ward. Os destaques (negrito e realçado) na Tabela 10 indicam os grupos em que o porto de Sepetiba ficou inserido em cada partição. Nota-se que são grupos com grande número de portos, em geral o primeiro ou segundo mais numeroso.

Grupo	Número de grupos							
	22	14	10	8	6	5	4	2
<b>I</b>	29	58	8	52	10	<b>97</b>	30	<b>179</b>
<b>II</b>	1	<b>27</b>	2	11	30	131	<b>94</b>	178
<b>III</b>	12	19	20	32	<b>159</b>	43	77	
<b>IV</b>	7	2	20	80	80	11	156	
<b>V</b>	9	29	9	28	20	75		
<b>VI</b>	22	18	39	<b>90</b>	58			
<b>VII</b>	5	1	11	44				
<b>VIII</b>	56	11	110	20				
<b>IX</b>	7	18	<b>85</b>					
<b>X</b>	47	66	53					
<b>XI</b>	<b>21</b>	47						
<b>XII</b>	10	10						
<b>XIII</b>	15	2						
<b>XIV</b>	26	49						
<b>XV</b>	1							
<b>XVI</b>	14							
<b>XVII</b>	13							
<b>XVIII</b>	7							
<b>XIX</b>	1							
<b>XX</b>	45							
<b>XXI</b>	8							
<b>XXII</b>	1							

Tabela 10 - Número de portos em cada grupo formado pelo método k-médias (elaboração própria)

A solução em 6 grupos parece não ser um bom resultado, pois houve uma grande concentração de portos num único grupo, desproporcional quando comparada com as outras soluções. Diferentemente do método Ward, percebe-se que as partições em 10, 14 e 22 grupos, formadas pelo método k-médias, produziram grupos com apenas um ou dois portos:

Solução com 10 grupos: Vadinar Terminal (151-INVAD), Índia, e Al Juaymah Terminal (290-SAJUT), Arábia Saudita, no grupo K10-II.

Solução com 14 grupos: Port Said (88-EGPSD), Egito, e Pusan (232-KRPUS), Coréia do Sul, no grupo K14-IV; Al Juaymah Terminal (290-SAJUT) no grupo K14-VII; e Cilacap Java (129-IDCXP), Indonésia, e Kiire (183-JPKII), Kagoshima, Japão, no grupo K14-XIII.

Solução com 22 grupos: Cilacap Java (129-IDCXP) no grupo K22-II; Port Said (88-EGPSD) no grupo K22-XV; Al Juaymah Terminal (290-SAJUT) no grupo K22-XIX; e Bangkok (304-THBKK), Tailândia, no grupo K22-XXII.

Verificando-se os dados dos portos listados acima foi possível perceber alguns com descritores atípicos e extremos. Os portos de Said (88-EGPSD) e de Pusan (232-KRPUS) possuem as maiores áreas de drenagem (SIZRVC) no valor de 2,00 e 1,69 milhões de quilômetros quadrados, respectivamente (Figura 21). O porto de Al Juaymah Terminal (290-SAJUT) tem uma temperatura máxima do ar do dia durante a estação mais quente (USART) de 47°C (Figura 14). O porto de Cilacap Java (129-IDCXP) tem a precipitação total nos seis meses mais secos (D6MRF) de 1.325mm (Figura 19).

Como visto nas Figura 12 a Figura 21 vários outros portos também têm descritores com valores atípicos e extremos. No entanto, muitos desses portos foram agrupados de forma homogênea nas diferentes partições. Com o auxílio de gráficos de caixa foram analisadas essas partições e identificados os portos que se comportaram como atípicos ou extremos dentro de cada grupo. Esses portos estão listados na Tabela 11 e alguns notadamente possuem descritores atípicos ou extremos, a exemplo dos portos de altas latitudes como 107-GBIMM (53° 38,0N); 160-ISHA (61° 04,0N); 161-ISSTR (64° 03,0N); e 341-USANC (61° 13,8N). Entretanto, não foi possível perceber o motivo pelo qual determinados portos se situaram distantes dos centros de seus respectivos grupos.

Solução com 2 grupos (Figura 23) - No grupo K2-I: Port Said (88-EGPSD); Cilacap Java (129-IDCXP); Cochin (139-INCOK); Kandla (142-INIXY); Bandar Imam Khomeyni (153-IRBKM); Bandar Mushar (154-IRBMR); Port Harcourt (252-NGPHC); e Al Juaymah Terminal (290-SAJUT). No grupo K2-II: Calcutá (138-INCCU) e Anchorage (341-USANC).

Solução com 4 grupos (Figura 24) - No grupo K4-II: Cilacap Java (129-IDCXP) e Bandar Mushar (154-IRBMR). No grupo K4-III: Port Said (88-EGPSD); Bandar Imam Khomeyni (153-IRBKM); e Al Juaymah Terminal (290-SAJUT). No grupo K4-IV: Anchorage (341-USANC).

Solução com 5 grupos (Figura 25) - No grupo K5-I: Chittagong (41-BDCGP); Kandla (142-INIXY); Bandar Imam Khomeyni (153-IRBKM) e Bandar Mushar (154-IRBMR). No grupo K5-II: Come By Chance (57-CACBC); La Have (60-CALHA); Immingham (107-GBIMM); Hafnarfjordur (160-ISHAF); Straumsvik (161-ISSTR); e Anchorage (341-USANC). No grupo K5-III: Calcutá (138-INCCU). No grupo K5-V: Port Said (88-EGPSD).

Solução com 6 grupos (Figura 26) - No grupo K6-III: Karumba (40- AUKRB); Tanjung Bara Coal Terminal (132- IDTBA); Cochin (139-INCOK); Mundra (145-INMUN); Sikka (149- INSIK); Bandar Mushar (154-IRBMR); Lae (264- PGLAE); e Bangkok (304-THBKK). No grupo K6-IV: Bandar Imam Khomeyni (153-IRBKM). No grupo K6-V: Port Said (88-EGPSD) e Pusan (232-KRPUS). No grupo K6-VI: Anchorage (341-USANC).

Solução com 8 grupos (Figura 27) - No grupo K8-I: Fateh Oil Terminal (7-AEFAT); Dampier (18-AUDAM); Cape Flattery (38-AUCQP); e Bandar Imam Khomeyni (153-IRBKM), Bandar Mushar (154-IRBMR), Lavan Island (158-IRLVP); Sirri Island Oil Terminal (159-IRSXI). No grupo K8-III: Anchorage (341-USANC). No grupo K8-VII: Ponta Ubu (52-BRPOU); Damietta (86-EGDAM); e Ashdod (136-ILASH). No grupo K8-VIII: Port Said (88-EGPSD) e Pusan (232-KRPUS).

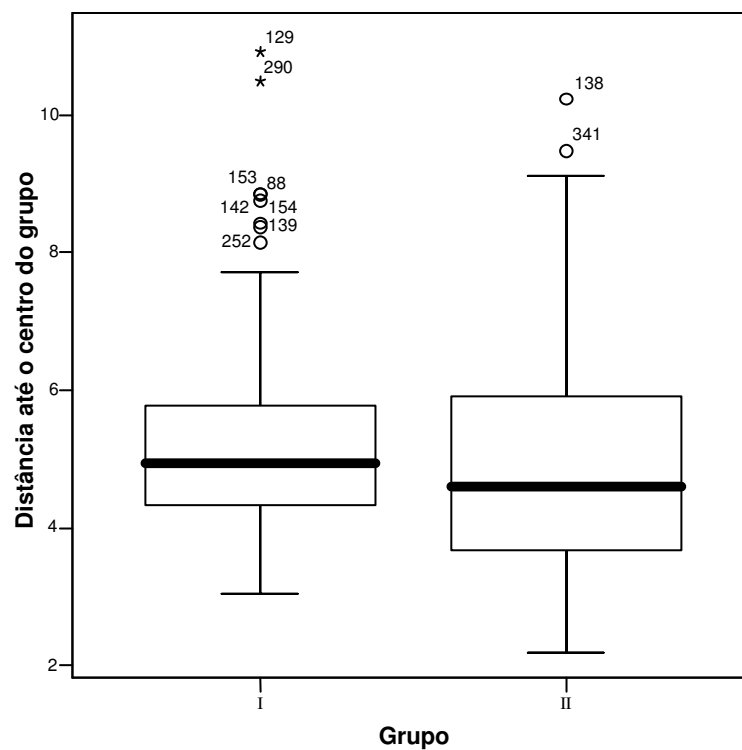


Figura 23 - Gráfico de caixas da partição em 2 grupos pelo método k-médias (elaboração própria)

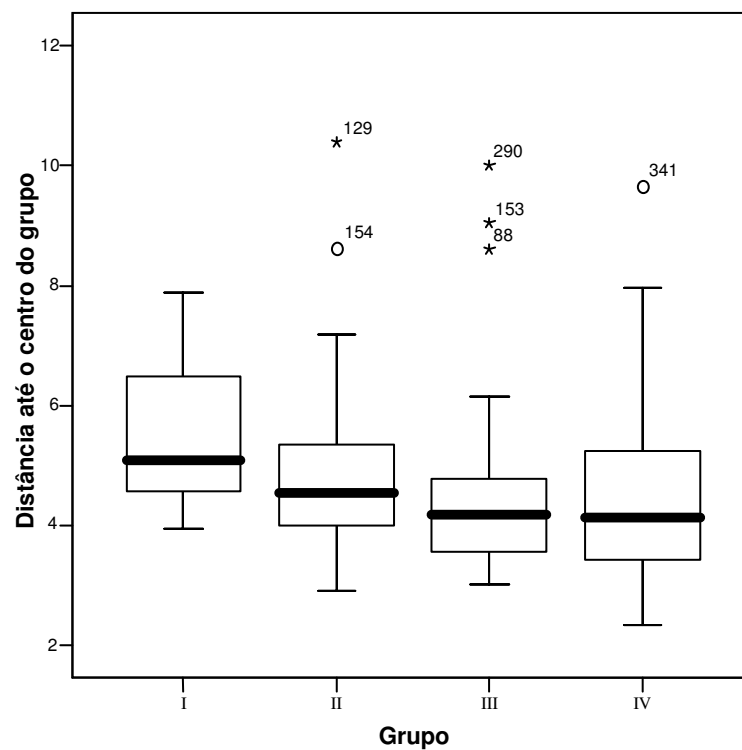


Figura 24 - Gráfico de caixas da partição em 4 grupos pelo método k-médias (elaboração própria)

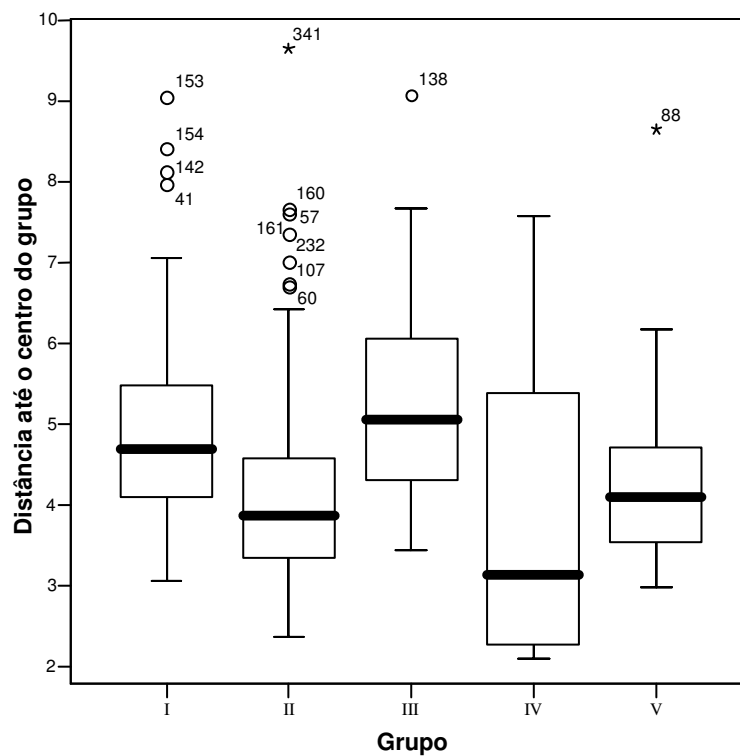


Figura 25 - Gráfico de caixas da partição em 5 grupos pelo método k-médias (elaboração própria)

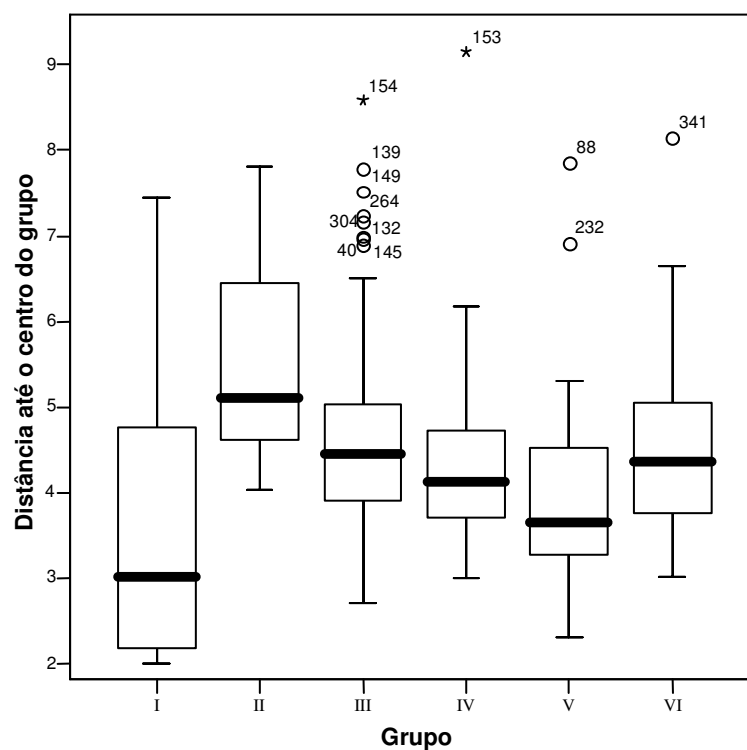


Figura 26 - Gráfico de caixas da partição em 6 grupos pelo método k-médias (elaboração própria)

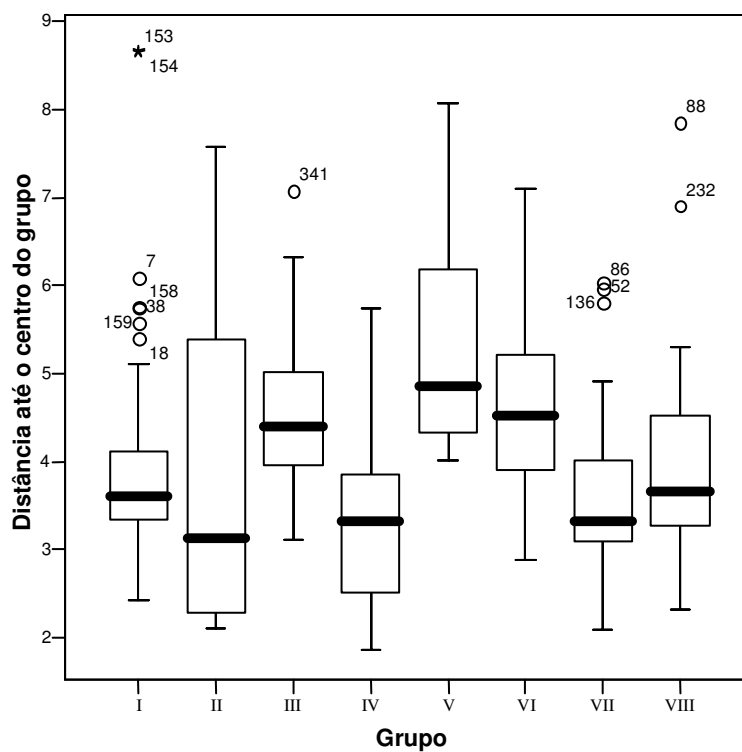


Figura 27 - Gráfico de caixas da partição em 8 grupos pelo método k-médias (elaboração própria)

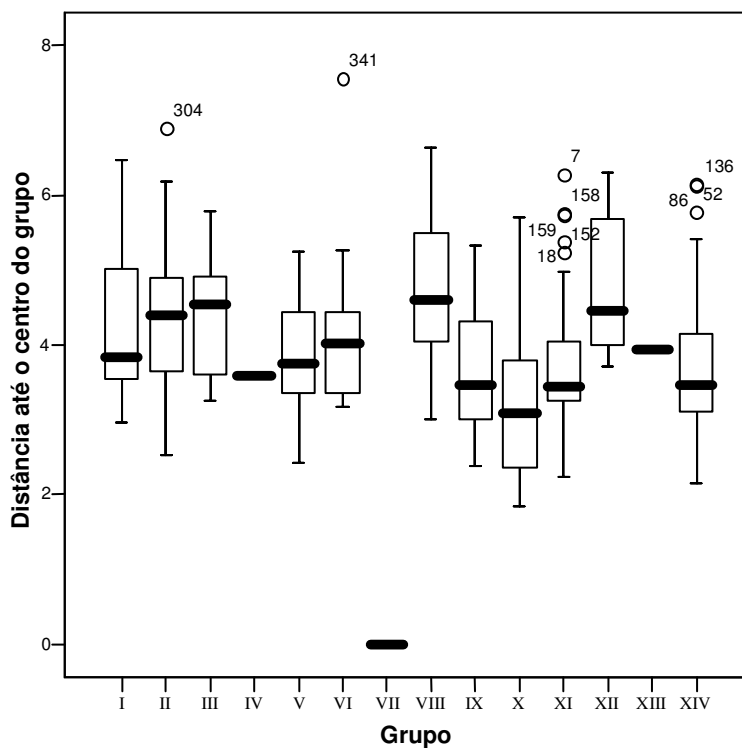


Figura 28 - Gráfico de caixas da partição em 10 grupos pelo método k-médias (elaboração própria)

Nº	Código	Porto	País	Latitude	Longitude
7	AEFAT	Fateh Oil Terminal	Emirados Árabes	25° 36.0N	54° 31.2E
18	AUDAM	Dampier	Austrália	20° 39.0S	116° 43.0E
38	AUCQP	Cape Flattery	Austrália	14° 59.0S	145° 21.0E
40	AUKRB	Karumba	Austrália	17° 29.0S	140° 50.0E
41	BDCGP	Chittagong	Bangladesh	22° 13.0N	91° 48.0E
52	BRPOU	Ponta Ubu	Brasil	20° 47.0S	40° 35.0W
57	CACBC	Come By Chance	Canadá	47° 28.8N	54° 00.6W
60	CALHA	La Have	Canadá	44° 17.0N	64° 21.0W
86	EGDAM	Damietta	Egito	31° 25.8N	31° 48.0E
88	EGPSD	Port Said	Egito	31° 15.6N	32° 18.6E
107	GBIMM	Immingham	Reino Unido	53° 38.0N	0° 11.0W
129	IDCXP	Cilacap Java	Indonésia	7° 44.0S	109° 00.0E
132	IDTBA	Tanjung Bara	Indonésia	0° 32.0S	117° 39.0E
136	ILASH	Ashdod	Israel	31° 50.0N	34° 38.0E
138	INCCU	Calcutá	Índia	22° 33.0N	88° 19.0E
139	INCOK	Cochin	Índia	9° 58.0N	78° 48.0E
142	INIXY	Kandla	Índia	22° 52.0N	70° 13.0E
145	INMUN	Mundra	Índia	22° 54.0N	69° 42.0E
149	INSIK	Sikka	Índia	22° 31.0N	69° 48.0E
151	INVAD	Vadinar Terminal	Índia	22° 30.0N	69° 42.0E
153	IRBKM	Bandar Imam Khomeyni	Irã	30° 25.0N	49° 04.0E
154	IRBMR	Bandar Mushar	Irã	30° 28.0N	49° 11.0E
158	IRLVP	Lavan Island	Irã	26° 47.0N	53° 20.0E
159	IRSXI	Sirri Island Oil Terminal	Irã	25° 57.0N	54° 32.0E
160	ISHAF	Hafnarfjordur	Islândia	61° 04.0N	21° 58.0W
161	ISSTR	Straumsvik	Islândia	64° 03.0N	22° 03.0W
183	JPKII	Kiire Kagoshima	Japão	31° 23.0N	130° 32.0E
232	KRPUS	Pusan	Coréia do Sul	35° 06.0N	129° 04.0E
252	NGPHC	Port Harcourt	Nigéria	4° 46.2N	7° 00.0E
264	PGLAE	Lae	Papua Nova Guiné	6° 44.0S	146° 58.0E
290	SAJUT	Al Juaymah Terminal	Arábia Saudita	26° 55.2N	50° 01.0E
304	THBKK	Bangkok	Tailândia	13° 42.0N	100° 34.0E
341	USANC	Anchorage Alaska	EUA	61° 13.8N	149° 52.8W

Tabela 11 - Portos atípicos ou extremos dentro dos grupos formados pelo método k-médias

A Tabela 12 mostra as distâncias Euclidianas entre os centros finais de cada grupo formado pelo método k-médias nas diferentes partições testadas. Esses centros refletem as características de cada grupo, de modo que distâncias maiores entre grupos correspondem a dissimilaridades maiores. Em cada solução os portos se deslocam de um grupo para outro em busca do centróide mais próximo e a composição de cada grupo varia de acordo com o número  $k$  de grupos predeterminado. Dessa forma, conjuntos de portos diferentes são reunidos em cada solução e a análise de proximidade entre grupos depende desse  $k$  escolhido.

Grupo	K4-I	K4-II	K4-III	K4-IV
K4-I	-	6,875	9,055	5,739
K4-II	6,875	-	4,461	5,094
K4-III	9,055	4,461	-	5,753
K4-IV	5,739	5,094	5,753	-

Grupo	K5-I	K5-II	K5-III	K5-IV	K5-V
K5-I	-	5,044	6,760	6,472	4,600
K5-II	5,044	-	4,821	6,329	5,754
K5-III	6,760	4,821	-	7,880	8,248
K5-IV	6,472	6,329	7,880	-	7,104
K5-V	4,600	5,754	8,248	7,104	-

Grupo	K6-I	K6-II	K6-III	K6-IV	K6-V	K6-VI
K6-I	-	8,656	6,081	7,021	8,050	7,324
K6-II	8,656	-	6,149	9,011	6,158	6,480
K6-III	6,081	6,149	-	4,112	5,901	4,984
K6-IV	7,021	9,011	4,112	-	7,106	7,023
K6-V	8,050	6,158	5,901	7,106	-	5,536
K6-VI	7,324	6,480	4,984	7,023	5,536	-

Grupo	K8-I	K8-II	K8-III	K8-IV	K8-V	K8-VI	K8-VII	K8-VIII
K8-I	-	7,571	9,453	6,179	9,672	4,747	4,649	7,937
K8-II	7,571	-	8,090	6,019	8,468	6,508	6,703	7,863
K8-III	9,453	8,090	-	4,409	6,858	7,801	6,450	6,296
K8-IV	6,179	6,019	4,409	-	5,823	4,504	3,538	4,885
K8-V	9,672	8,468	6,858	5,823	-	7,094	8,210	6,106
K8-VI	4,747	6,508	7,801	4,504	7,094	-	5,016	7,187
K8-VII	4,649	6,703	6,450	3,538	8,210	5,016	-	5,841
K8-VIII	7,937	7,863	6,296	4,885	6,106	7,187	5,841	-

Tabela 12 - Matrizes de distância final entre os grupos formados pelo método k-médias (elaboração própria)

No entanto, não existe nenhuma fórmula geral para encontrar o número ótimo de grupos para um determinado conjunto de dados. Uma forma de abordar esse problema é rodar o algoritmo para diferentes números  $k$  de grupos e contrapor os resultados com outros métodos de acordo com algum critério. No item a seguir os resultados dos dois métodos hierárquico e não-hierárquico são comparados.



### 5.1.3 Matriz de Concordância

Uma matriz de confusão exibe a distribuição dos dados nos termos de suas classes reais e classes previstas, e serve para avaliar a qualidade do classificador. A Tabela 13 mostra a matriz de confusão para um conjunto com duas classes, onde as entradas têm os seguintes significados: “a” verdadeiro negativo; “b” falso positivo; “c” falso negativo; e “d” verdadeiro positivo.

		Prevista	
		Negativo	Positivo
Real	Negativo	a	b
	Positivo	c	d

Tabela 13 - Exemplo de caso com duas classes “negativo” e “positivo” (adaptado de Kohavi & Provost, 1998)

O conceito dessa matriz é similar ao do teste de hipótese  $H_0$  onde existem dois tipos de erro conforme a Tabela 14.

		Decisão	
		$H_0$ aceito	$H_0$ rejeitado
Estado	$H_0$ verdadeiro	Decisão correta	Erro tipo I
	$H_0$ falso	Erro tipo II	Decisão correta

Tabela 14 - Resultados possíveis para o teste de hipótese (adaptado de Wonnacott & Wonnacott, 1990, p. 303)

A matriz de confusão é uma ferramenta de visualização tipicamente usada no campo do aprendizado supervisionado. No caso do aprendizado não-supervisionado é chamada de matriz de concordância. Várias são as medidas propostas para avaliar os resultados de uma matriz de confusão, por exemplo, o coeficiente de precisão, que é a proporção das previsões corretas (Kohavi & Provost, 1998). No caso da matriz de concordância, como as classes reais não são conhecidas, pode-se adaptar para um coeficiente de concordância “ $C_c$ ” que leva em consideração as classificações coincidentes de dois métodos distintos (a diagonal da matriz) em relação ao total de dados:

$$C_c = \frac{a + d}{a + b + c + d} \quad (6)$$

Alguns softwares que dispõem de ferramenta para construção da matriz de confusão foram identificados na internet (e.g. GeneXproTools<sup>11</sup>; GeneLinker<sup>12</sup>; DB2<sup>13</sup>). Contudo, como esses programas não estavam disponíveis, as matrizes foram elaboradas de forma manual.

Apesar de laboriosa, optou-se por realizar a análise comparativa dos métodos Ward e k-médias para as partições de 2, 4, 5, 6 e 8 grupos, conforme mostrado na Tabela 15. Como não existem classes “reais”, as matrizes foram montadas de modo a associar os grupos com maior número de coincidências. Na comparação das partições obtiveram-se os seguintes coeficientes de concordância: 85,7% (2 grupos); 69,5% (4 grupos); 71,4% (5 grupos); e 56,3% (6 grupos); e 86,8% (8 grupos). De modo geral as discordâncias pontuais observadas foram exatamente de portos identificados como atípicos ou extremos nos grupos formados pelo método k-médias (Tabela 11).

Um grupo constituído por portos banhados pelo Mar da Arábia, Mar Vermelho e Golfo Pérsico (Arábia Saudita, Emirados Árabes, Iêmen, Irã, Kuwait, Quatar e Sudão) e um porto do Senegal, Dakar (303-SNDKR), se mostrou muito coeso. Formado pelo método Ward, este grupo, com 45 portos, coincidiu integralmente com um único grupo formado pelo método k-médias nas partições de 4, 5, 6 e 8 grupos. No método hierárquico de Ward este grupo, formado inicialmente pelos grupos 11 e 12, foi o último a se fundir (Tabela 9).

---

<sup>11</sup> <http://www.gepsoft.com/gepsoft/APS3KB/Chapter09/Section2/SS03.htm>

<sup>12</sup> [http://www.improvedoutcomes.com/docs/WebSiteDocs/Plots/Classification\\_and\\_Prediction/Confusion\\_Matrix.htm](http://www.improvedoutcomes.com/docs/WebSiteDocs/Plots/Classification_and_Prediction/Confusion_Matrix.htm)

<sup>13</sup> <http://publib.boulder.ibm.com/infocenter/db2luw/v8/index.jsp?topic=/com.ibm.im.visual.doc/idmu0mst125.html>

Um outro grupo bastante consistente foi o 13 formado pelo método Ward. Onze portos compõem este grupo, sendo sete portos do Japão e mais os portos de Huangpu Guangdong (66-CNHUA), China; Cilacap Java (129-IDCXP); Vadinar Terminal (151-INVAD); e Al Juaymah Terminal (290-SAJUT). Curiosamente, estes três últimos portos, e mais Kiire (183-JPKII) do Japão, fazem parte dos portos listados como atípicos ou extremos nos grupos formados pelo método k-médias na Tabela 11 que, contudo, não tiveram este comportamento na partição em 8 grupos (Figura 27). Este grupo teve uma concordância de 100% nessa partição de 8 grupos (grupos W8-4 e K8-II são idênticos). Na partição em 5 grupos, o grupo K5-IV foi formado integralmente por esses onze portos, e na partição em 6 grupos o K6-I foi formado por dez dos onze portos do grupo 13, à exceção foi o porto Huangpu Guangdong (66-CNHUA), que se situou no grupo K6-II.

Pela análise comparativa dos dois métodos, a solução em 8 grupos figura como a melhor. Nesta partição quatro grupos formados pelo método k-médias se mostraram bem coincidentes com grupos obtidos com o método Ward, os grupos K8-II, K8-IV, K8-VI e K8-VIII, com uma concordância superior a 90%, sendo que no caso do grupo K8-II a concordância foi de 100%. O porto de Sepetiba ficou inserido no grupo K8-VI, formado pelo método k-médias, que teve 82 coincidências com o grupo W8-1; sete com o grupo W8-2; e uma com o grupo W8-6, formados pelo método Ward. Este conjunto de 82 portos revelou-se bastante homogêneo também na comparação das partições em 6 grupos (K6-III e W6-1).

A Tabela 16 relaciona os portos de cada grupo formado pelo método k-médias para a solução em 8 grupos. São realçados no grupo K8-VI os oito portos que não coincidiram com o grupo W8-1 do método Ward.

Partição com 2 grupos		Ward		
		W2-1	W2-2	
k-médias	<b>K2-I</b>	<b>144</b>	35	179
	K2-II	16	<b>162</b>	178
		160	197	357

Partição com 4 grupos		Ward				
		W4-1	W4-2	W4-3	W4-4	
k-médias	<b>K4-II</b>	<b>89</b>	-	5	-	94
	K4-III	8	<b>45</b>	24	-	77
	K4-IV	1	-	<b>102</b>	53	156
	K4-I	17	-	1	<b>12</b>	30
		115	45	132	65	357

Partição com 5 grupos		Ward					
		W5-1	W5-2	W5-3	W5-4	W5-5	
k-médias	<b>K5-I</b>	<b>83</b>	12	-	2	-	97
	K5-IV	-	-	-	11	-	11
	K5-V	7	-	<b>45</b>	23	-	75
	K5-II	1	-	-	<b>96</b>	34	131
	K5-III	-	12	-	-	<b>31</b>	43
		91	24	45	132	65	357

Partição com 6 grupos		Ward						
		W6-1	W6-2	W6-3	W6-4	W6-5	W6-6	
k-médias	<b>K6-III</b>	<b>82</b>	7	-	67	1	2	159
	K6-II	-	<b>17</b>	-	1	-	12	30
	K6-IV	7	-	<b>45</b>	28	-	-	80
	K6-I	-	-	-	<b>10</b>	-	-	10
	K6-VI	-	-	-	26	<b>29</b>	3	58
	K6-V	2	-	-	-	-	<b>18</b>	20
		91	24	45	132	30	35	357

Partição com 8 grupos		Ward								
		W8-1	W8-2	W8-3	W8-4	W8-5	W8-6	W8-7	W8-8	
k-médias	<b>K8-VI</b>	<b>82</b>	7	-	-	-	1	-	-	90
	K8-V	-	<b>16</b>	-	-	-	-	-	12	28
	K8-I	6	-	<b>45</b>	-	1	-	-	-	52
	K8-II	-	-	-	<b>11</b>	-	-	-	-	11
	K8-VII	1	-	-	-	<b>37</b>	6	-	-	44
	K8-IV	-	1	-	-	-	<b>73</b>	2	4	80
	K8-III	-	-	-	-	-	3	<b>28</b>	1	32
	K8-VIII	2	-	-	-	-	-	-	<b>18</b>	20
		91	24	45	11	38	83	30	35	357

Tabela 15 - Matrizes de concordância entre os métodos Ward e k-médias (elaboração própria)

K8-I	K8-II	K8-III	K8-IV	K8-V	K8-VI	K8-VII	K8-VIII
2 AEAUH	66 CNHUA	57 CACBC	22 AUPPI	16 ARBUE	1 BRSPB	8 AEFJR	44 BGBOJ
3 AEMZD	129 IDCXP	58 CASEI	24 AUWEP	17 ARCMP	26 AUBNE	20 AUPBY	45 BGVAR
4 AEDAS	151 INVAD	59 CAHAL	25 AUPKL	41 BDCGP	27 AUBDB	21 AUWYA	88 EGPSD
5 AEPPRA	171 JPBEP	60 CALHA	67 CNNBO	42 BEANR	28 AUGLT	23 AUPST	111 GEBUS
6 AEDXB	183 JPKII	61 CAVAN	68 CNNGB	43 BEGNE	29 AUPTL	52 BRPOU	112 GEPTI
7 AEFAT	187 JPKND	62 CARBK	71 CNTAO	48 BRITJ	30 AUHPT	75 CYKYR	232 KRPUS
9 AEJEA	202 JPOIT	65 CNDLC	72 CNTSN	63 CNCAN	31 AUPDT	76 CYLCA	281 ROCND
10 AEJED	206 JPSAE	78 DEBRE	73 CNYNT	69 CNSHA	32 AUMKY	77 CYLMS	282 ROMAG
11 AEKLF	207 JPSAG	80 DEWVN	90 ESGIJ	70 CNSHB	34 AUTSV	85 EGALY	283 ROMID
12 AEQIW	210 JPSKD	82 DKENS	91 ESBIO	79 DEHAM	36 AUMOU	86 EGDAM	284 RUNVS
13 AERUW	290 SAJUT	83 DKFRC	92 ESVGO	138 INCCU	37 AUCNS	87 EGEDK	285 RUTUA
14 AESHJ		99 FRDKK	123 HROMI	142 INIXY	39 AUWEI	93 ESBCN	307 TRERE
15 AEZIR		100 FRBES	165 ITRAN	252 NGPHC	40 AUKRB	94 ESVLC	308 TRIST
18 AUDAM		101 FRDON	167 ITVCE	256 NLRTM	49 BRPNG	95 ESALG	310 TRIZT
19 AUPWL		104 FRLEH	168 ITTRS	258 NLAMS	50 BRSSZ	96 ESLPA	312 TRSSX
33 AUABP		106 GBHST	169 JPABO	319 UADNB	51 BRRIO	97 ESSCT	313 TRYAR
35 AULUC		107 GBIMM	170 JPAMA	321 UANIK	53 BRVIX	98 ESTAR	318 UAODS
38 AUCQP		108 GB001	172 JPCHB	325 USPHL	54 BRPRM	102 FRFOS	320 UAILK
46 BHMAM		109 GBPTB	173 JPKMT	326 USILG	55 BRTUB	103 FRLAV	322 UASVP
47 BHMIM		110 GBRER	174 JPFKY	327 USBAL	56 BRSSA	105 FRMRS	348 UYMVD
81 DJJIB		135 IEMOT	175 JPHHR	331 USMOB	64 CNCWN	113 GIGIB	
84 EGAIS		160 ISHAF	176 JPHIM	332 USLCH	74 COCAR	114 GRASS	
89 EGSUZ		161 ISSTR	177 JPHKT	333 USDVT	121 HKHKG	115 GREEU	
150 INTUT		217 JPTMK	178 JPIMB	334 USMSY	122 HKKWN	116 GRJKH	
153 IRBKM		229 KRKUV	179 JPINS	337 USBPT	124 IDBLW	117 GRPAC	
154 IRBMR		255 NLEUR	180 JPIWK	340 USHOU	125 IDDUM	118 GRPIR	
155 IRBND		257 NLJUM	181 JPKCZ	342 USPDZ	126 IDCIG	119 GRSKI	
156 IRBUZ		259 NLVLI	182 JPKGA	343 USBCC	127 IDMRK	120 GRVOL	
157 IRKHK		286 RUVVO	184 JPKIJ		128 IDJKT	136 ILASH	
158 IRLVP		323 USBOS	185 JPKIK		130 IDSRG	162 ITGOA	
159 IRSXI		324 USNYC	188 JPKNU		131 IDSUB	163 ITPFX	
236 KWKWI		341 USANC	189 JPKOJ		132 IDTBA	164 ITLIV	
237 KWMAA			190 JPKSM		133 IDBPN	166 ITTAR	
238 KWMIS			191 JPKUD		134 IDAMA	242 MTMLA	
239 KWMIB			192 JPKWS		137 INBOM	263 PECLL	
240 KWSAA			193 JPMAM		139 INCOK	274 PTFAO	
278 QADOH			194 JPMIZ		140 INHAL	277 PTSIE	
279 QAUMS			195 JPMOJ		141 INIXE	306 TRDYL	
280 QAHAL			196 JPMUR		143 INMAA	309 TRIZM	
287 SADMN			197 JPMYJ		144 INMRM	311 TRMER	
288 SAJED			199 JPNGI		145 INMUN	346 USLGB	
289 SAJUB			200 JPNGO		146 INPBD	347 USSAN	
291 SARAR			201 JPNGS		147 INPRT	351 YEMKX	
292 SA001			204 JPONO		148 INSAL	357 ZASDB	
293 SARLT			205 JPOSA		149 INSIK		
294 SAYNB			208 JPSAK		152 INVTZ		
295 SDBMT			211 JPSMN		186 JPKIN		
296 SDPZU			212 JPSMT		198 JPNHA		
303 SNDKR			213 JPSMZ		203 JPOKA		
349 YEADE			214 JPTAM		209 JPSBS		
350 YEHOD			215 JPTBT		226 KEMBA		
352 YERAI			216 JPTKY		241 LKCMB		
			218 JPTOY		243 MYPEN		
			219 JPTYO		244 MYLUM		
			220 JPUBJ		245 MYPKG		
			221 JPUKB		246 MYPDI		
			222 JPWAK		247 MYBTB		
			223 JPYKK		248 MYPGU		
			224 JPYOK		249 MYBTU		
			225 JPYOS		250 NGLOS		
			227 KRKAN		251 NGTIN		
			228 KRKPO		253 NGONN		
			230 KRMOK		254 NGBON		
			231 KRONs		264 PGLAE		
			233 KRSCP		265 PGPOM		
			234 KRUSN		266 PGDAU		
			235 KRYOS		267 PHBTG		
			260 NZAKL		268 PHBTN		
			261 NZWRE		269 PHLIM		
			262 NZMAP		270 PHMNL		
			275 PTLIS		271 PHSFS		
			276 PTLIS		272 PKBQM		
			302 SIKOP		273 PKKHI		
			328 USPHF		297 SGJUR		
			329 USNEN		298 SGKEP		
			330 USSAV		299 SGSEM		
			344 USSFO		300 SGSIN		
			345 USOAK		301 SGTPG		
			353 ZACPT		304 THBKK		
			355 ZAPLZ		305 THLCH		
					314 TWKEL		
					315 TWKHI		
					316 TWTXG		
					317 TZDAR		
					335 USLOP		
					336 USSAB		
					338 USGLS		
					339 USTXT		
					354 ZADUR		
					356 ZARCB		

Tabela 16 - Grupos formados pelo método k-médias para a solução em  $k=8$  (elaboração própria)

## 5.2 Métodos de Ordenação

Verificar a dispersão de objetos com respeito a todos os possíveis pares de descritores é uma abordagem bastante laboriosa. A ordenação consiste em mapear objetos num gráfico, usualmente de dois ou três eixos, no qual as posições correspondem à relação ordenada entre esses objetos, representando a maior parte da variabilidade dos dados. Na ordenação os objetos são reunidos na tentativa de combinar as dissimilaridades correspondentes: pontos próximos têm características muito similares, os objetos que estão mais separados têm pouco em comum. O termo ‘tentativa’ é importante, pois não há uma maneira definida única em que isto possa ser conseguido. De fato, quando se tem um grande número de descritores, essencialmente uma situação multidimensional, pode ser impossível obter uma representação útil em duas ou três dimensões (Clarke & Warwick, 2001).

Os métodos de ordenação são freqüentemente chamados de *métodos de análise fatorial*, uma vez que são baseados na extração de autovetores ou *fatores* da matriz de associação (Legendre & Legendre, 1998). Desses métodos, optou-se pela utilização de dois, de certa forma por critérios opostos: a análise de componentes principais (ACP) e o escalonamento multidimensional não-métrico (*Non-metric Multidimensional Scaling-MDS*). A análise de componentes principais (Hotelling, 1933 *apud* Kaski, 1997) tem importância por ser um método de ordenação antigo e consagrado. Por outro lado, o MDS foi mais desenvolvido recentemente, pois seu complexo algoritmo requer uma capacidade de cálculo só obtida com o avanço da tecnologia computacional hoje disponível.

### 5.2.1 Análise de Componentes Principais

A análise de componentes principais estabelece, com base em uma matriz de correlações, variâncias-covariâncias, ou até mesmo similaridade, um conjunto de eixos

(componentes ou fatores) perpendiculares. Cada componente corresponde a um autovetor dessa matriz. Assim, para uma matriz de  $p$  variáveis, são calculados  $p$  autovetores (= eixos fatoriais) de comprimento decrescente em razão de sua contribuição à variância total dos dados. Desse modo, o primeiro eixo da ACP, sobre o qual serão ordenadas as amostras, representará a maior parte da variação dos dados. O resultado disso é um sistema reduzido de coordenadas, proporcionando informação sobre semelhanças das amostras.

Interpretar uma ACP consiste em tentar definir o que representa cada eixo em termos de fator responsável pela ordenação das amostras. Na medida em que a importância dos eixos, i.e., a sua participação à variância total vai diminuindo, a sua interpretação torna-se cada vez mais difícil. É natural citar os valores dessas variâncias (em relação ao seu total) como uma medida da ‘quantidade’ de informação contida em cada eixo (Valentin, 2000).

A ACP aplicada à base de dados GloBallast indica que o eixo do parâmetro SIZRVC (tamanho da bacia hidrográfica) contém uma parte substancial da variância total das amostras no espaço de descritores, fazendo com que a direção do eixo da componente principal 1 passe a ser ditada por este descritor apenas (Tabela 17). Isso é de se esperar considerando que o desvio padrão de SIZRVC é da ordem de  $10^5$ , muito superior aos demais.

Dados ambientais usualmente têm uma complexa mistura de escala de medidas (salinidade em ppm, profundidade em m, concentrações em  $\mu\text{g/g}$  etc). Numa visualização multidimensional da matriz de dados as amostras são pontos referidos a eixos. Se as unidades em cada eixo diferem, e não têm conexão natural um com o outro, então um ponto A pode se tornar mais próximo de um ponto B que um ponto C, ou mais próximo de um ponto C que de um ponto B simplesmente por uma mudança de escala em um dos eixos. Obviamente, seria errado para a ordenação ACP variar com tais mudanças de escala arbitrárias. Existe uma solução natural para isto que é fazer a ACP baseada em correlações, i.e. normalizar todos os eixos de modo que estes

tenham escalas comparáveis e adimensionais (Clarke & Warwick, 2001). A ACP obtida quando os dados não são normalizados é referida como baseada em covariância.

O grau que uma ACP bidimensional é bem sucedida em representar uma informação multidimensional é vista na percentagem da variância total explicada pelas duas primeiras componentes principais (CP). Um quadro mais equilibrado emergiu com o uso da ACP baseada em correlações, pois a normalização faz com que todos os descritores sejam de igual importância potencial na determinação das componentes principais (Tabela 18). No entanto, a análise dos resultados mostrou que a CP1 e a CP2 juntas não conseguiram explicar mais do que 40,7% da variância total, e mesmo as cinco primeiras componentes principais explicaram pouco mais do que 60%.

A Figura 29 mostra o gráfico com a plotagem dos portos em relação às duas primeiras componentes principais, onde nota-se a formação de duas aglomerações. Como visto no item 3.1.1, uma importante parcela dos descritores refere-se à caracterização climática dos portos, e para buscar identificar o significado dessas aglomerações, os dados foram separados de acordo com uma divisão dos portos em tropicais, localizados entre o Trópico de Câncer ( $23^{\circ} 27' N$ ) e o Trópico de Capricórnio ( $23^{\circ} 27' S$ ), e extra-tropicais, ao norte e ao sul desses trópicos respectivamente. Com este critério 96 portos ficaram na classe “tropical” e 261 na classe “extra-tropical”, sendo o porto de Straumsvik (161-ISSTR), localizado na Islândia a  $64^{\circ}03'N$ , o de mais alta latitude. Verifica-se na Figura 30 que o porto de Sepetiba ficou na fronteira entre as duas classes. A ACP conseguiu agrupar boa parte dos portos extra-tropicais, visualizados na parte superior esquerda do gráfico. No entanto, na parte inferior as duas classes estão misturadas. Embora uma boa visualização tenha sido obtida com essa pré-categorização, ela não foi suficientemente capaz de determinar grupos totalmente homogêneos.



CP	Autovalores	Varição (%)	Varição Acumulada(%)
1	$8,05 \times 10^{10}$	100	100
2	336.131,30	0	0
3	47.816,64	0	0
4	4.511,31	0	0
5	396,45	0	0

Descrição	Autovetores				
	CP1	CP2	CP3	CP4	CP5
PTYPE	0	0	0	0,009	-0,023
MSUWT	0	-0,001	0,005	-0,016	0,160
USUWT	0	0	0,005	-0,015	0,155
MWNWT	0	-0,004	0,008	-0,014	0,280
LWNWT	0	-0,004	0,008	-0,014	0,291
MSART	0	0	0,004	-0,01	0,119
USART	0	0,003	0,003	-0,017	0,102
MWART	0	-0,003	0,011	-0,024	0,332
LWART	0	-0,006	0,011	-0,011	0,397
MWSAL	0	0,006	0,010	-0,057	0,364
LWSAL	0	0,008	0,009	-0,061	0,386
MDSAL	0	0,004	0,011	-0,049	0,322
UDSAL	0	0,003	0,010	-0,044	0,301
MSPR	0	0	0	0,002	-0,002
MNER	0	0	0	0,001	-0,005
D6MRF	0	-0,359	-0,932	-0,025	0,026
W6MRF	0	-0,932	0,360	-0,029	-0,011
RNFL75	0	-0,001	-0,005	0,002	-0,031
DISRVM	0	0,035	0,01	-0,992	-0,103
SIZRVC	-1	0	0	0	0
INASMW	0	0	0	0	0
INARKW	0	0	0	0	-0,003
INAWP	0	0	0	0,004	-0,014
INHTSM	0	0	0,001	-0,001	0,014
INSNDB	0	0	0	-0,003	0,011
INSTNB	0	0	0	0	-0,025
INLTMF	0	0	0	0,005	-0,002
INMANG	0	-0,001	0,001	-0,001	0,040
INRKSH	0	0	0	-0,005	-0,004
SUFSND	0	0	0	-0,003	0,015
SUSFTM	0	0	-0,001	0,003	-0,009
SUSGRM	0	0	0	-0,004	0,032
SURKRF	0	0	0	-0,005	-0,012
SUCORF	0	0	0,001	-0,003	0,041

Tabela 17 - Parâmetros da ACP dos dados brutos (elaboração própria)

CP	Autovalores	Variação (%)	Variação Acumulada(%)
1	9,06	26,6	26,6
2	4,78	14,1	40,7
3	2,54	7,5	48,2
4	2,40	7,1	55,2
5	1,71	5,0	60,3

Descritor	Autovetores				
	CP1	CP2	CP3	CP4	CP5
PTYPE	-0,158	-0,217	-0,014	-0,039	-0,160
MSUWT	0,274	-0,162	-0,158	0,073	0,073
USUWT	0,269	-0,160	-0,189	0,062	0,076
MWNWT	0,271	-0,192	0,160	0,065	-0,038
LWNWT	0,263	-0,185	0,181	0,068	-0,048
MSART	0,249	-0,095	-0,132	0,070	0,108
USART	0,212	-0,007	-0,350	0,015	0,056
MWART	0,268	-0,183	0,070	0,037	-0,030
LWART	0,249	-0,183	0,212	0,049	-0,062
MWSAL	0,219	0,275	0,099	-0,130	-0,047
LWSAL	0,205	0,287	0,037	-0,097	-0,035
MDSAL	0,215	0,259	0,140	-0,156	-0,046
UDSAL	0,219	0,234	0,145	-0,150	-0,063
MSPR	-0,038	-0,052	0,261	-0,472	0,189
MNER	-0,059	-0,033	0,214	-0,495	0,207
D6MRF	-0,080	-0,186	0,300	0,227	-0,056
W6MRF	-0,006	-0,250	0,327	0,184	0,058
RNFL75	-0,191	-0,042	0,104	0,138	-0,170
DISRVM	0,089	0,179	-0,203	0,023	0,151
SIZRVC	-0,047	-0,111	-0,332	-0,020	0,182
INASMW	-0,012	0,007	-0,063	-0,063	-0,266
INARKW	-0,016	0,109	-0,113	0,025	-0,444
INAWP	-0,114	-0,152	0,032	-0,040	-0,280
INHTSM	0,109	-0,019	-0,144	-0,270	-0,134
INSNDB	0,096	0,132	-0,020	0,039	-0,448
INSTNB	-0,123	0,172	0,063	0,179	0,072
INLTMF	-0,030	-0,214	-0,085	-0,337	-0,247
INMANG	0,202	-0,211	0,060	-0,092	-0,138
INRKSH	0,014	0,237	0,193	0,141	-0,020
SUFSND	0,131	0,154	0,024	0,069	-0,225
SUSFTM	-0,102	-0,134	0,012	-0,184	-0,196
SUSGRM	0,186	0,084	0,053	0,056	-0,035
SURKRF	-0,024	0,203	0,254	0,169	0,089
SUCORF	0,219	-0,086	0,089	0,019	0,111

Tabela 18 - Parâmetros da ACP normalizada (elaboração própria)

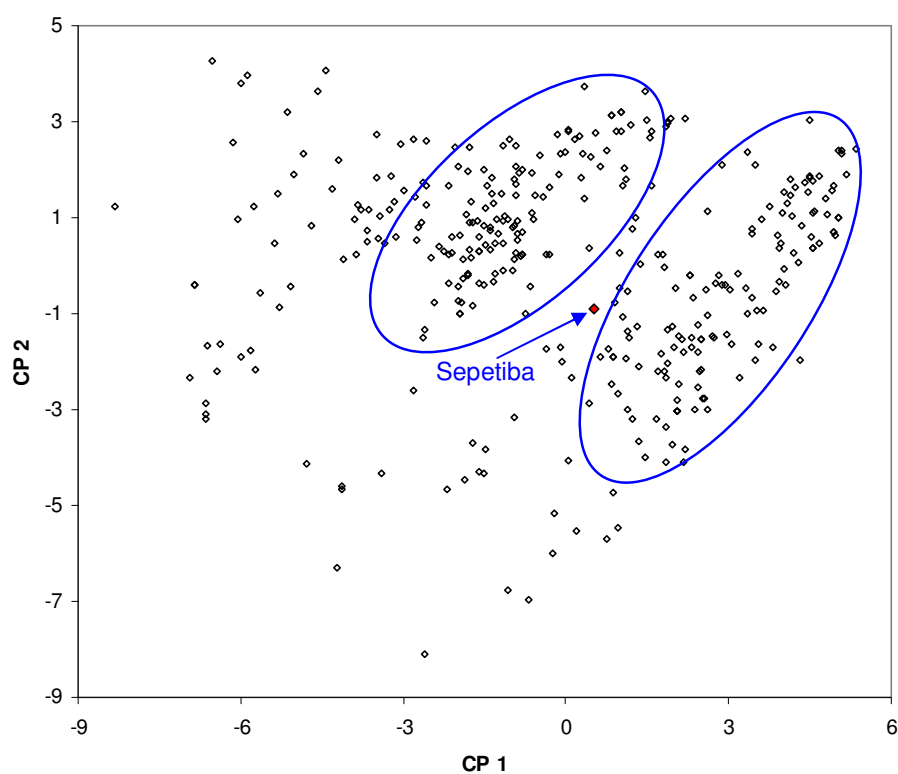


Figura 29 - Gráfico da ACP normalizada (elaboração própria)

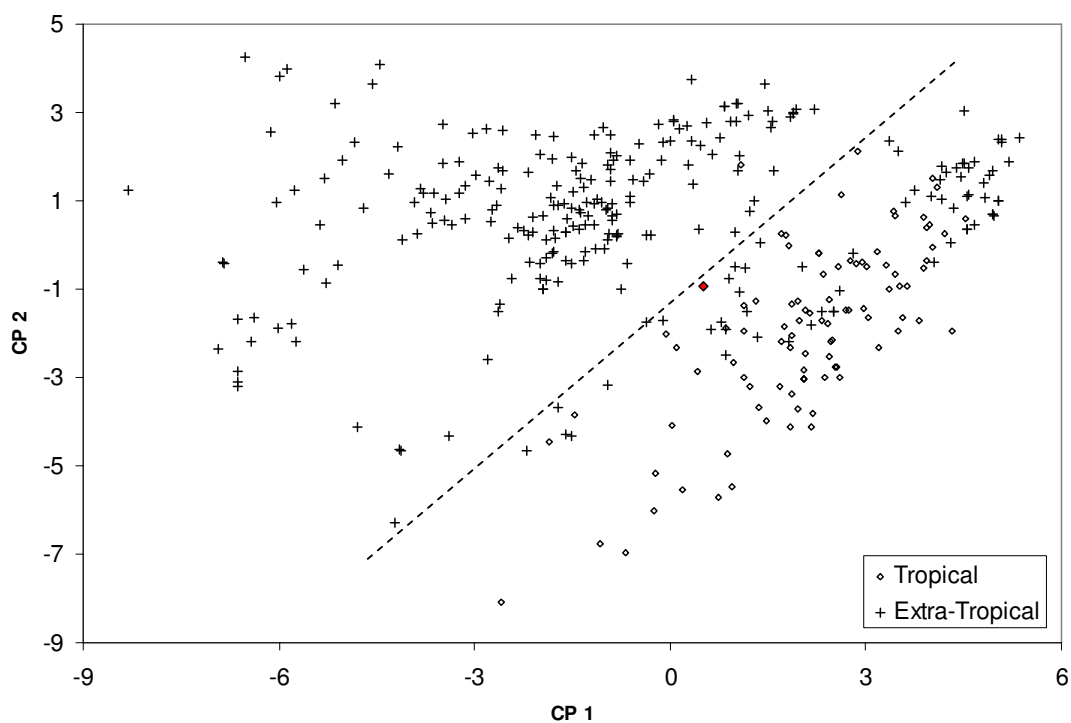


Figura 30 - Gráfico da ACP normalizada com duas classes: “Tropical” x “Extra-Tropical” (elaboração própria)

O mesmo gráfico foi construído tomando como base o agrupamento obtido com o método Ward para 5 grupos (Figura 31). Nota-se que, de modo geral, as posições dos grupos W5-4 e W5-5 correspondem à dos portos extra-tropicais, localizados na parte superior esquerda. Na parte inferior direita, os portos dos grupos W5-1 e W5-3 se mantiveram agregados em porções separadas do plano das componentes principais, enquanto que os portos do grupo W5-2 ficaram bem dispersos.

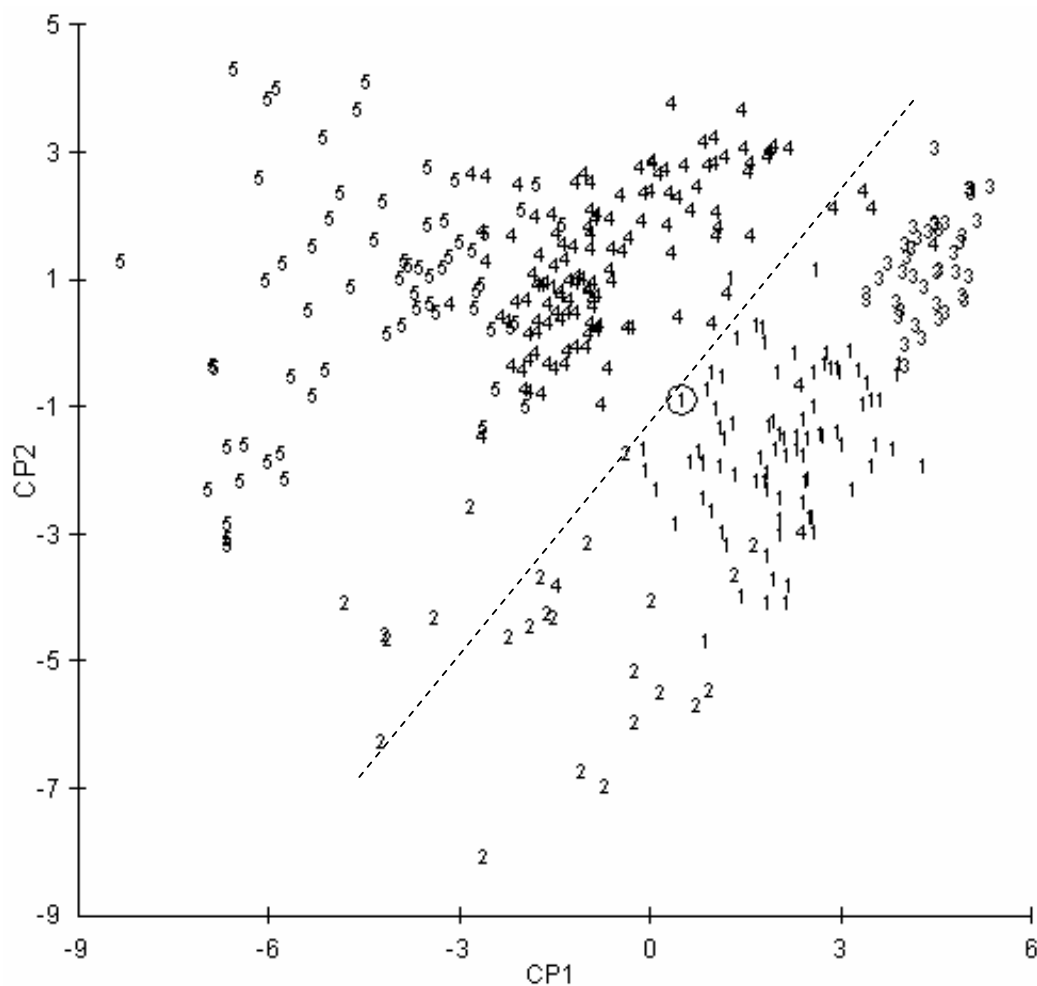


Figura 31 - Representação da ACP normalizada com 5 grupos obtidos pelo método Ward (elaboração própria)

Outra representação do gráfico com as duas primeiras componentes principais foi testada, desta vez com a partição em 8 grupos obtida com o método k-médias. De modo semelhante com a representação anterior, foi possível identificar alguns grupos na parte

superior esquerda (K8-III, K8-IV, K8-VII e K8-VIII) e outros na parte inferior direita (K8-I e K8-VI). Dois grupos não seguiram esta divisão, os grupos K8-II e K8-V. Alguns portos do grupo K8-II estão junto ao grupo K8-IV (de difícil visualização, pois estão superpostos) e outros dispersos no gráfico, exatamente os portos que se mostraram atípicos ou extremos nos grupos formados pelo método k-médias, destacados na Figura 32. O grupo K8-V ficou bem espalhado nessa representação. Na comparação entre os dois algoritmos de agrupamento o grupo K8-V teve uma concordância de 16 portos com o grupo W8-2 e 12 portos com o grupo W8-8, o que demonstra uma segmentação neste conjunto.

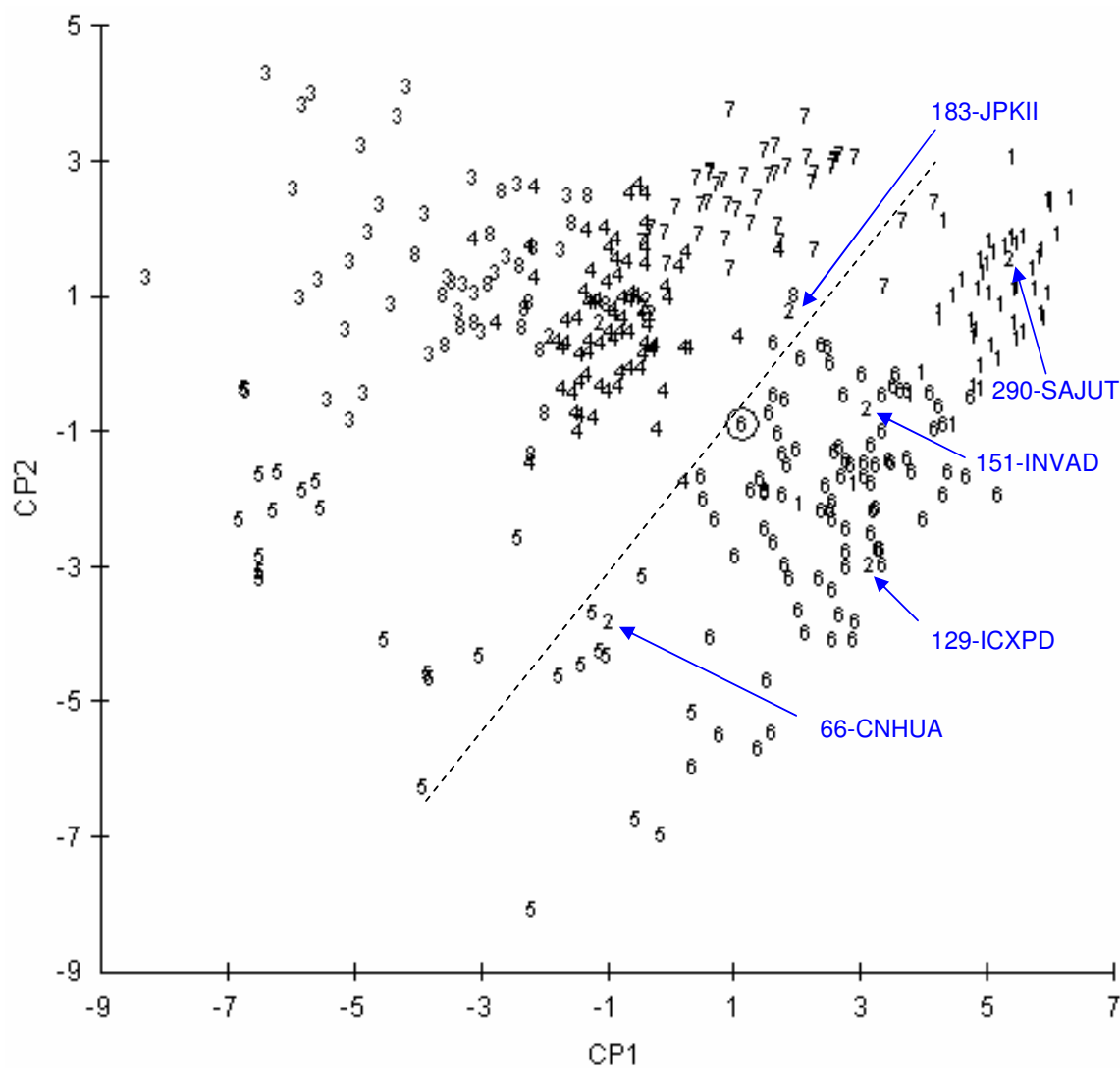


Figura 32 - Representação da ACP normalizada com 8 grupos obtidos pelo método k-médias com destaque para portos dispersos do grupo K8-II (elaboração própria)

A ordenação é essencialmente uma técnica para converter dissimilaridades entre amostras em distâncias (Euclidianas) entre essas amostras. A ACP ao fazer a projeção das distâncias num plano bidimensional de ordenação pode distorcer bastante algumas dessas distâncias (Clarke & Warwick, 2001). A Figura 33 localiza o número dos portos no gráfico das duas primeiras componentes principais. Devido ao grande número de portos houve muita superposição o que torna difícil a visualização individual de alguns índices.

Com a parte do gráfico correspondente ao porto de Sepetiba (1-BRSPB) destacada, constata-se que o porto do Rio de Janeiro (51-BRRIO) não está tão próximo como era de se esperar, já que Rio de Janeiro é o porto considerado mais similar ambientalmente com Sepetiba. Nas imediações de Sepetiba estão Brisbane (26-AUBNE); Port Alma (29-AUPTL); e Durban (354-ZADUR), enquanto que próximos ao Rio de Janeiro estão Santos (50-BRSSZ); Hong Kong (121-HKHKG); LOOP Terminal (335-USLOP); e Texas (339-USTXT), todos esses portos pertencentes ao mesmo grupo K8-VI.

A avaliação foi de que a ordenação ACP forneceu resultados coerentes com as análises de agrupamento vistas neste item. Porém, o exame desses resultados parece dependente de alguma pré-classificação dos dados. Para avaliação da relação de similaridade entre os portos, por sua pobre preservação de similaridades em distâncias projetadas no gráfico de ordenação, a ACP pode ser potencialmente enganosa.

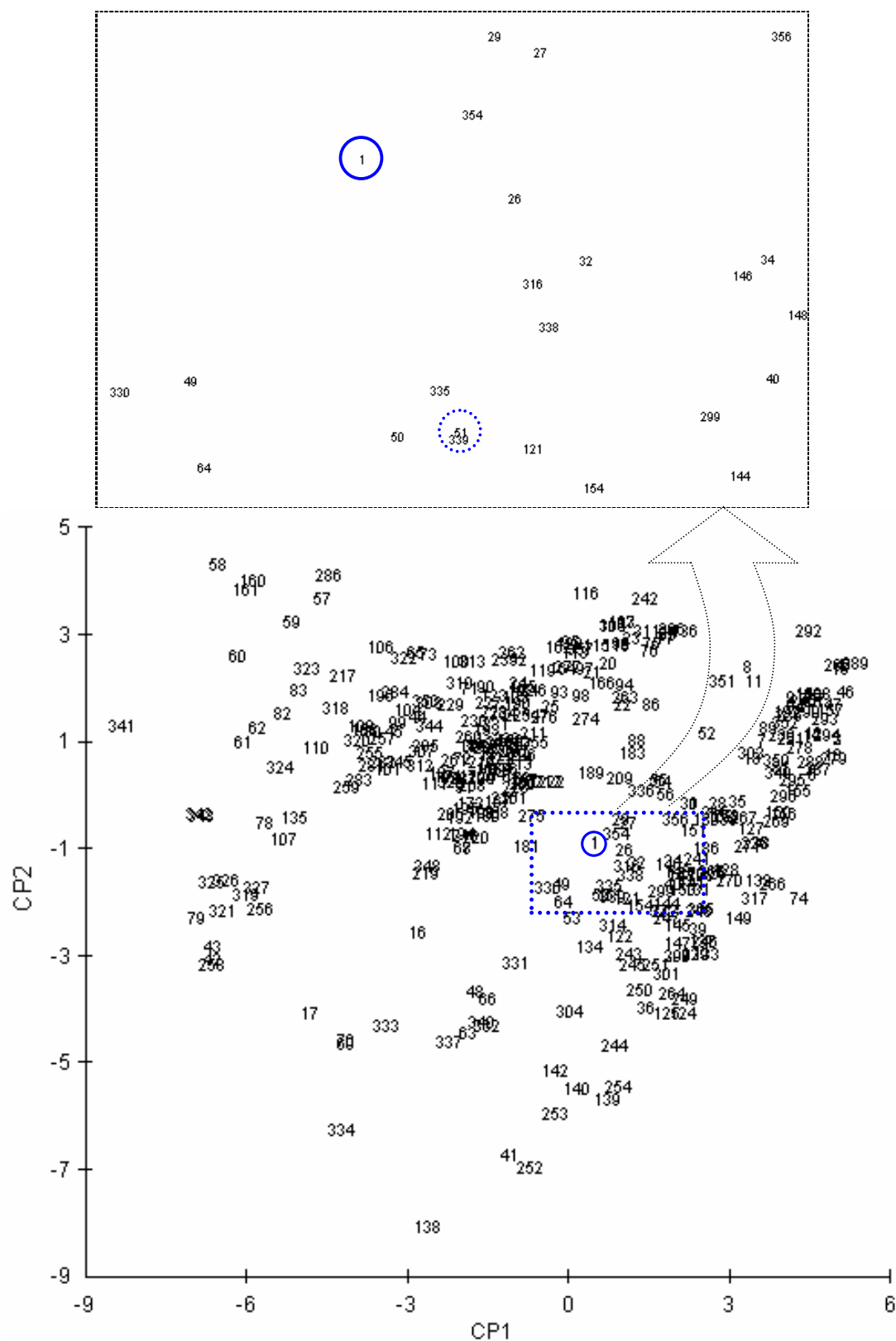


Figura 33 - Gráfico da ACP normalizada com destaque para o porto de Sepetiba (1) e o porto do Rio de Janeiro (51) (elaboração própria)

### 5.2.2 Escalonamento Multidimensional não-métrico

O escalonamento multidimensional (*Multidimensional Scaling - MDS*) é um conjunto de técnicas frequentemente usadas para explorar visualmente similaridades ou dissimilaridades em dados. O algoritmo do MDS observa os seguintes passos (Fasham, 1977; Legendre & Legendre, 1998; Clarke & Warwick, 2001):

1) Especificar o número de eixos de ordenação ( $m$ ), em geral escolhidos numa configuração de duas ou três dimensões.

2) Calcular uma matriz ( $p \times p$ ) de similaridade entre as amostras.

3) Construir uma configuração inicial para as  $p$  amostras que pode ser um conjunto aleatório de  $p$  pontos em  $m$  dimensões ou, por exemplo, o resultado da ACP<sup>14</sup>.

4) Desenhar um diagrama de dispersão com a distância ( $d_{jk}$ ) versus a dissimilaridade ( $\delta_{jk}$ ) para todos os  $p(p-1)/2$  pares de valores. Isso é chamado diagrama de Shepard.

5) Fazer a regressão das distâncias entre os pontos no gráfico de ordenação e as correspondentes dissimilaridades. Como definir esta relação é que caracteriza diferentes processos de ordenação. Existem duas abordagens principais:

5a) Ajustar uma regressão linear entre  $d$  e  $\delta$  de modo que a distância final seja proporcional à dissimilaridade original. Mais flexíveis podem ser ajustes de uma regressão curvilínea (quadrática, cúbica ou polinomial). Esses modelos paramétricos dão origem ao termo MDS métrico.

5b) Realizar um regressão não-paramétrica é a característica do MDS não-métrico<sup>15</sup>. A Figura 34 ilustra uma linha de regressão não-paramétrica. Essa linha tem a restrição de sempre ser crescente e, portanto, consiste numa série de degraus. O sucesso relativo do MDS

<sup>14</sup> A configuração inicial com resultados da ACP não está disponível no software Primer 5

<sup>15</sup> Conforme Clarke & Warwick (2001) “o método do escalonamento multidimensional não-métrico (*Non-metric Multidimensional Scaling*) foi introduzido por Shepard (1962) e Kruskal (1964) para aplicação a problemas da psicologia”.



não-métrico em preservar as relações entre as amostras nas distâncias no gráfico de ordenação vem da flexibilidade no formato da linha de regressão não-paramétrica.

6) Medir a qualidade do ajuste da regressão. O quanto os pontos estão desviados da linha de regressão mede a falha em ajustar a ordem de similaridades, e é chamado de stress<sup>16</sup>, medido pela fórmula de Krustal (1964):

$$stress = \sqrt{\sum_j \sum_k (d_{jk} - \delta_{jk})^2 / \sum_j \sum_k d_{jk}^2} . \quad (7)$$

Caso  $d_{jk} = \delta_{jk}$  para todos os  $p(p-1)/2$  pares de valores, o stress é zero.

7) Modificar a configuração corrente de modo a diminuir o stress. Esta mudança é baseada em técnicas de otimização numérica, e a idéia essencial é avaliar o stress para pequenas mudanças nas posições dos pontos no gráfico de ordenação.

8) Repetir os passos 5 a 7 até que o valor do stress alcance um valor predeterminado ou até que se chegue a um mínimo local e não se obtenha redução do stress.

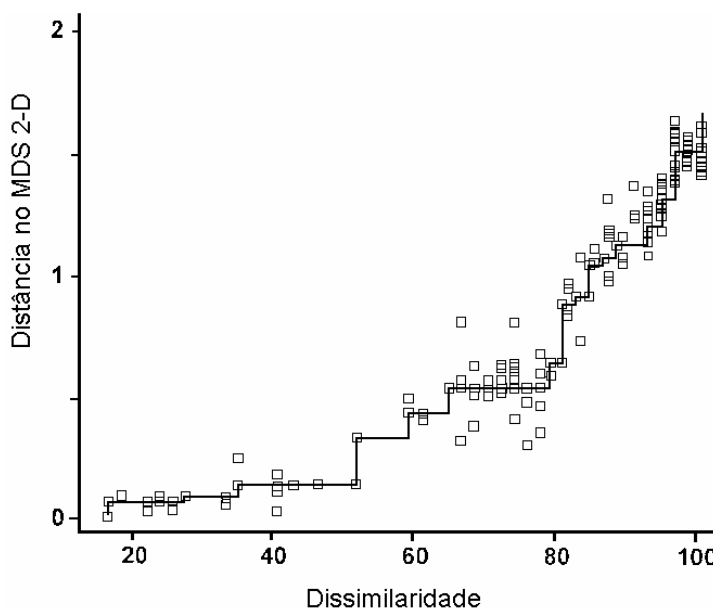


Figura 34 - Exemplo de diagrama de Shepard com ajuste por regressão não-paramétrica (adaptado de Clarke & Warwick, 2001)

<sup>16</sup> A edição em português de Hair *et al.* (2005) traduz stress como “medida de desajuste”.

O MDS tem a vantagem de reduzir a influência do pesquisador, uma vez que não requer a especificação de parâmetros iniciais. Por uma série de pequenos ajustes o método MDS gradualmente busca o mínimo da função de stress. Por vezes essa busca fica presa a um mínimo local e não existe uma forma de garantir que o mínimo global da função de stress tenha sido alcançado. Se um mínimo local ou global será encontrado depende muito do ponto de partida da busca, isto é, da configuração inicial dos pontos no gráfico de ordenação. A solução prática é repetir a análise MDS várias vezes começando com diferentes configurações iniciais, o que minimiza as chances de se prender num mínimo local nas iterações adiante.

Para realizar o MDS não-métrico foi utilizado o software PRIMER 5, cujo parâmetro “número de reinícios” permite testar configurações com diferentes posições iniciais aleatórias (passo 3 acima). Com a escolha de 20 reinícios para rodar o algoritmo, o stress mínimo de 0,18 ocorreu onze vezes. Testando-se uma quantidade maior de reinícios não foi obtido um stress menor, o que sugere que esse pode ser o mínimo global.

As posições multidimensionais podem ser rotadas em torno da origem, os eixos podem ser trocados (reflexão) ou a solução pode ser uniformemente expandida ou comprimida, tudo isso sem alterar as posições relativas dos objetos. Assim a Figura 35 mostra o resultado numa disposição na qual os pontos foram arranjados de forma a minimizar a superposição na área de interesse em torno do porto de Sepetiba.

Parte do grupo de portos identificados nas imediações de Sepetiba e do Rio de Janeiro na ACP também se localizou nesta circunvizinhança no MDS, como Brisbane (26-AUBNE); Texas (339-USTXT) e Durban (354-ZADUR). Porém, outros portos como Hong Kong (121-HKHKG) e LOOP Terminal (335-USLOP) já não mantiveram a mesma relação de proximidade, ao contrário dos portos de Mackay (32-AUMKY) e Galveston (338-USGLS), que na ACP não estavam tão próximos, e no MDS foram posicionados adjacentes ao porto de Sepetiba. A projeção do MDS não-métrico não mostra o porto do Rio de Janeiro como sendo o mais próximo.

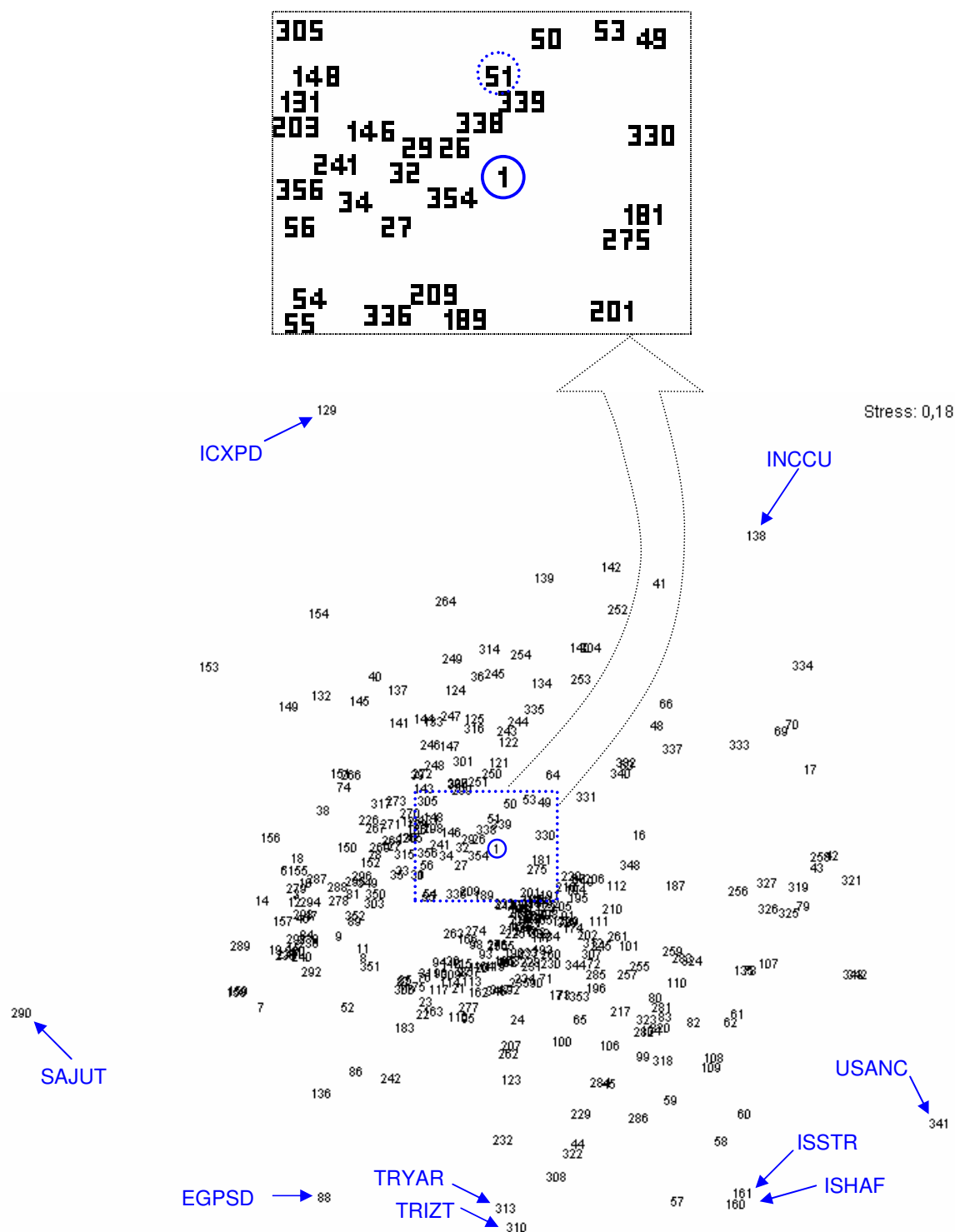


Figura 35 - Representação do MDS não-métrico com destaque para o porto de Sepetiba (1); o porto do Rio de Janeiro (51); e portos localizados nas extremidades da projeção – observação: alguns portos não estão representados no detalhe por questão de superposição (elaboração própria)

Os portos que se localizaram nas extremidades dessa projeção foram Port Said (88-EGPSD); Cilacap Java (129-IDCXP); Calcutá (138-INCCU); Hafnarfjordur (160-ISHAF); Straumsvik (161-ISSTR); Al Juaymah Terminal (290-SAJUT); Izmit (310-TRIZT); Yarimca (313-TRYAR); e Anchorage (341-USANC).

O posicionamento dos portos tropicais e extra-tropicais foi verificado no MDS por meio da Figura 36. A disposição com que os portos foram arranjados nessa projeção mostra que o MDS, tal como a ACP, conseguiu isolar parcialmente os portos extra-tropicais. Dois portos extra-tropicais ficaram mais distantes dos demais: Bandar Imam Khomeyni (153-IRBKM) e Bandar Mushar (154-IRBMR) no Irã.

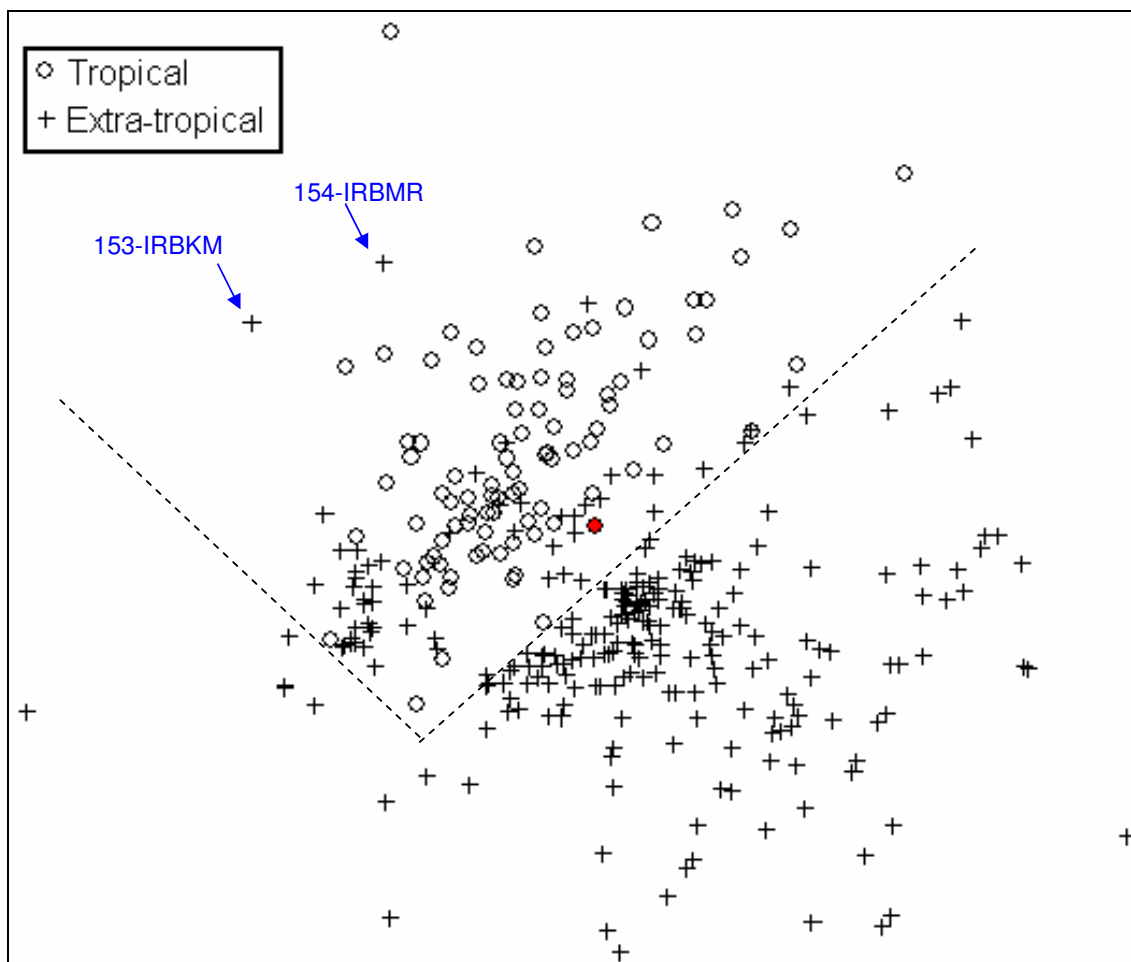


Figura 36 - Representação do MDS com duas classes: “Tropical” x “Extra-Tropical”  
(elaboração própria)

Repetindo, também, o procedimento adotado na ACP, a representação do MDS foi visualizada com os 8 grupos do método k-médias (Figura 37). Os grupos comportaram-se de modo semelhante no que diz respeito à divisão entre as duas classes “Tropical” x “Extra-Tropical”; e à aglomeração, sejam mais agrupados ou mais dispersos.

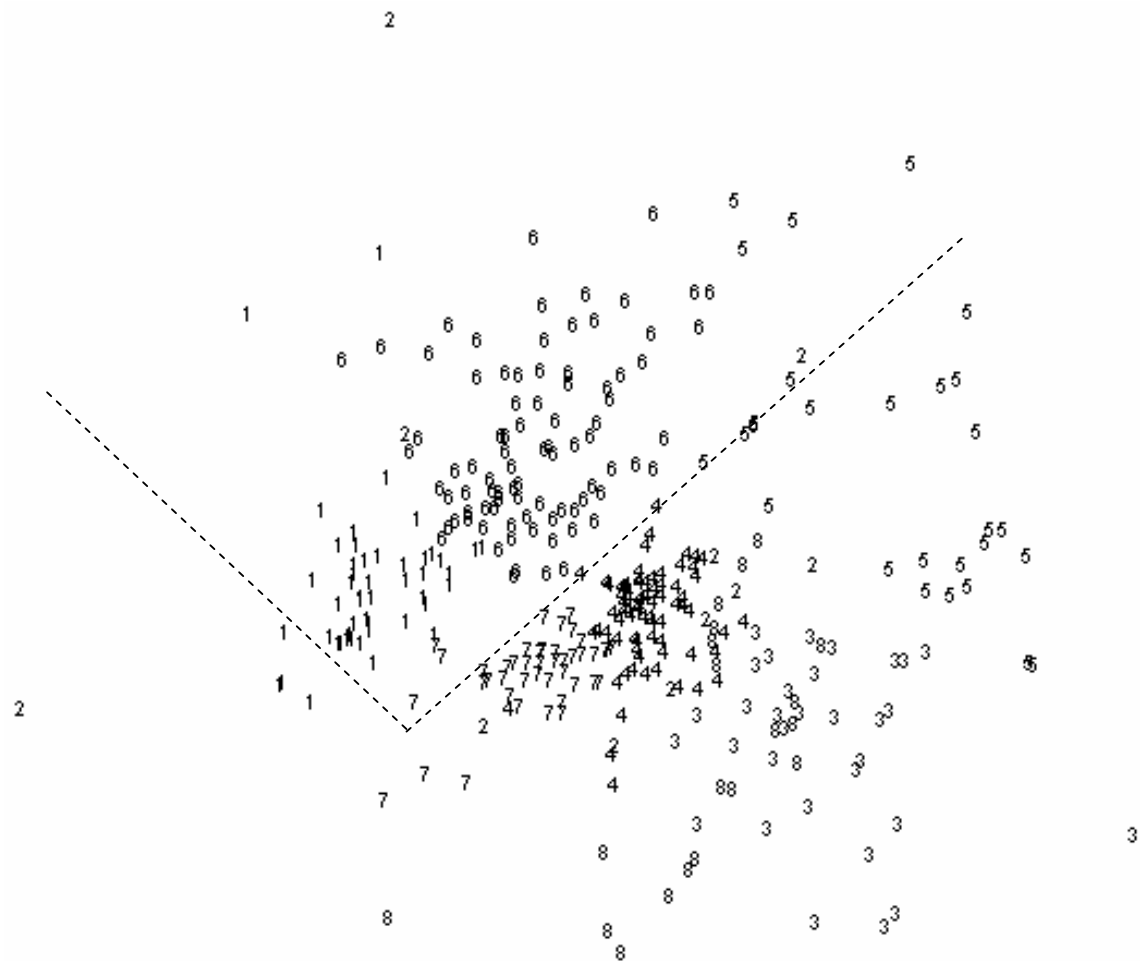


Figura 37 - Representação do MDS com 8 grupos obtidos pelo método k-médias (elaboração própria)

## 6 Redes neuronais

As redes neuronais artificiais consistem em um método de solucionar problemas de inteligência artificial, construindo um sistema que tenha circuitos que simulem o cérebro humano, inclusive seu comportamento, ou seja, aprendendo, errando e fazendo descobertas. São mais que isso, são técnicas computacionais inspiradas na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência. Uma grande rede neuronal artificial pode ter centenas ou milhares de unidades de processamento, enquanto que o cérebro de um mamífero pode ter muitos bilhões de neurônios (Figura 38).

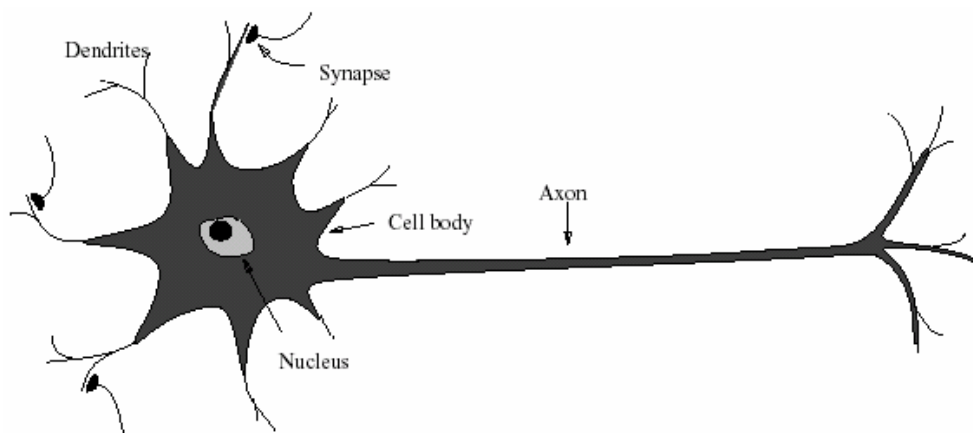


Figura 38 - Desenho esquemático de um neurônio biológico (Jain *et al.*, 1996, p. 6)

Entre os usos das redes neuronais está o mapeamento entrada-saída, “por exemplo, uma tarefa de classificações de padrões, na qual o objetivo seja atribuir um sinal de entrada representando um objeto físico ou evento a uma entre várias categorias (classes) preestabelecidas” (Haykin, 2001). Um modelo de redes neuronais comum para tarefas de classificação não supervisionada é o mapa auto-organizável, que é principalmente usado para agrupamento de dados e mapeamento de atributos. Esse algoritmo pertence à família de redes do tipo alimentada adiante, onde os neurônios são organizados em camadas com conexões estritamente unidirecionais de uma camada para outra.

A maioria dos modelos de redes neurais mais conhecidos são implicitamente equivalentes ou similares aos métodos estatísticos clássicos de reconhecimento de padrões. No entanto, as redes neurais diferem das técnicas multivariadas mais tradicionais exploradas anteriormente no sentido da formulação do modelo e, principalmente, nos tipos de relações mais complexas que podem ser acomodadas (Hair *et al.*, 2005). Apesar dessas semelhanças, as redes neurais oferecem várias vantagens tais como procedimentos flexíveis para encontrar soluções não-lineares (Jain *et al.*, 2000). Por exemplo, o método k-médias, pertencente à classe de técnicas de partição, faz as divisões do conjunto de dados por hiperplanos que não são capazes de levar em conta estruturas não-lineares (Ultsch, 1995).

### **6.1 Regras de aprendizado competitivo**

Ao contrário de outras formas de aprendizado (correção de erro, Boltzmann, Hebbian) onde múltiplas unidades de saída podem ser disparadas simultaneamente, no aprendizado competitivo as saídas competem entre si para serem ativadas. Como resultado dessa competição, apenas uma unidade de saída, ou apenas uma por grupo, é ativada num dado intervalo de tempo. Esse procedimento é freqüentemente chamado de *winner-take-all*.

O aprendizado competitivo foi encontrado em redes neurais biológicas. Experimentos neurobiológicos mostraram que o aprendizado competitivo tem um papel importante na formação de mapas topográficos no cérebro, e na auto-organização de células nervosas de orientação sensitiva no córtex estriado (face média do lobo occipital) (Jain *et al.*, 1996).

A Figura 39 mostra uma rede de aprendizado competitivo simples que consiste em conexões de alimentação adiante dos nós de entrada para uma única camada de unidades de saída (neurônios). Cada unidade de saída também se conecta lateralmente a todas as outras unidades de saída via pesos inibitórios ou excitatórios.

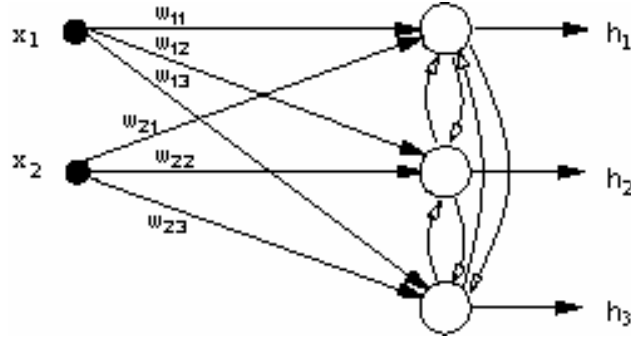


Figura 39 - Arquitetura de aprendizado competitivo simples (adaptado de Jain *et al.*, 1996, p. 25)

A saída do aprendizado competitivo é um agrupamento ou categorização dos dados de entrada. Padrões similares são agrupados pela rede e representados numa única unidade. Esse processo de agrupamento é feito pela rede a partir das correlações dos dados.

Cada unidade de saída na camada de imagem da rede se conecta a todas as unidades de entrada  $\mathbf{x}$  da camada de domínio via pesos  $w_{kj} > 0$ . O vetor representando o conjunto de sinais de entrada é denotado por:

$$\mathbf{x} = [x_1, \dots, x_k, \dots, x_p]$$

O vetor cujos componentes são os pesos sinápticos das arestas incidentes sobre o neurônio  $j$  é denotado por:

$$\mathbf{w}_j = [w_{1j}, w_{2j}, \dots, w_{pj}], \quad j = 1, \dots, N$$

onde  $N$  é o número de neurônios na camada de imagem. Como em outros tipos de redes, a função de ativação  $h_j$  de cada neurônio na rede é simplesmente um somatório de entradas ponderadas pelos pesos. Esta saída pode ser expressa como:

$$h_j = \sum_{i=1}^p x_i \cdot w_{ij} = \mathbf{x} \cdot \mathbf{w}_j, \quad j = 1, \dots, N \quad (8)$$

Assim, para achar o melhor ajuste do vetor de entrada  $\mathbf{x}$  com os vetores de pesos sinápticos  $\mathbf{w}_j$ , simplesmente compara-se os produtos  $\mathbf{x} \cdot \mathbf{w}_j$  e seleciona-se o maior.

$$\mathbf{x} \cdot \mathbf{w}_{j^*} \geq \mathbf{x} \cdot \mathbf{w}_j, \quad \forall j \quad (9)$$



O neurônio específico  $j^*$  que satisfaz esta condição é chamado o neurônio vencedor ou ‘*best matching*’ para o vetor  $\mathbf{x}$ . A saída desse neurônio ( $h_{j^*}$ ) é estabelecida como ‘1’ e as demais ‘zero’.

Ou seja:

$$h_j = \begin{cases} 1 & ; \quad j = j^* \\ 0 & ; \quad j \neq j^* \end{cases} \quad (10)$$

Na formulação do algoritmo o critério de melhor ajuste pode, também, ser expresso pela desigualdade (11) abaixo, onde  $\| \cdot \|$  denota a distância Euclidiana entre os vetores.

$$\| \mathbf{x} - \mathbf{w}_{j^*} \| \leq \| \mathbf{x} - \mathbf{w}_j \| \quad , \forall j \quad (11)$$

Uma regra simples de aprendizado competitivo pode ser obtida como segue, quando apenas os pesos da unidade vencedora  $j^*$  são atualizados. É também possível implementar um conjunto de unidades vencedoras com inibição lateral onde cada unidade inibe as outras conforme por exemplo a arquitetura de rede da Figura 39.

$$\Delta \mathbf{w}_j(t) = \begin{cases} \rho [\mathbf{x} - \mathbf{w}_j(t)] & ; \quad j = j^* \\ 0 & ; \quad j \neq j^* \end{cases} \quad (12)$$

O efeito desta regra de aprendizado é mover o peso das sinapses incidentes à unidade vencedora um pouco mais perto do padrão de entrada. A Figura 40 mostra esse processo geometricamente numa forma bidimensional, onde  $\mathbf{w}_j(t+1)$  é o novo valor de um peso conectando um componente de entrada  $\mathbf{x}$  ao neurônio vencedor;  $\mathbf{w}_j(t)$  é o valor prévio deste peso; e  $\rho$  é a taxa de aprendizado (modificação sináptica).

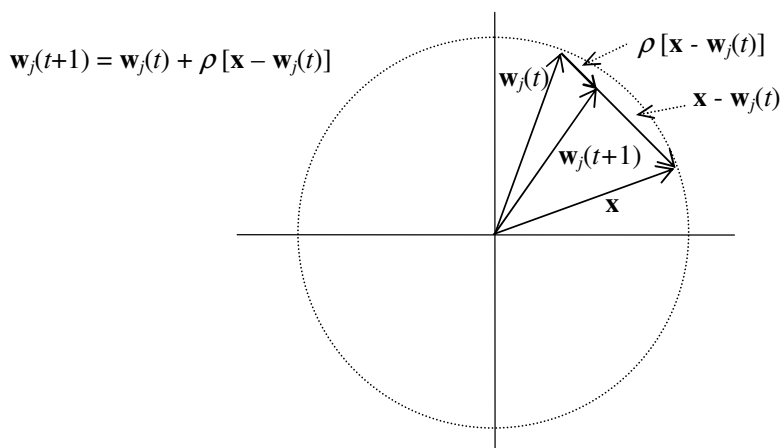


Figura 40 - Rotação de vetor de peso (adaptado de Wasserman, 1989, p. 68)

Cada peso associado com o neurônio vencedor é alterado por uma quantidade proporcional à diferença entre seu valor e o valor da entrada ao qual está conectado. A direção da mudança minimiza a diferença entre o peso e sua entrada. Como resultado, é mais provável que o neurônio vencedor vença a competição na próxima vez em que um vetor de entrada similar seja apresentado, e menos provável que vença no caso de apresentação de um vetor de entrada muito diferente.

A rede nunca irá parar de aprender, ou seja atualizar pesos. É possível que um padrão particular possa disparar diferentes unidades de saída (mudar as categorias) indefinidamente durante o aprendizado. Isso levanta a questão da estabilidade do sistema de aprendizado. Um sistema de aprendizado é dito estável, se nenhum padrão nos dados de treinamento muda sua categoria após um número finito de iterações de aprendizado. Uma forma de atingir a estabilidade é forçar a taxa de aprendizado a decrescer gradualmente à medida que o processo de aprendizado prossegue, eventualmente aproximando-se de zero. Entretanto, essa imobilização artificial do aprendizado causa um problema de plasticidade, que é definida como a habilidade de se adaptar a novos dados, conhecido como dilema estabilidade-plasticidade de Grossberg em aprendizado competitivo.

O compromisso entre plasticidade e estabilidade consiste justamente em como conseguir que o sistema aprenda um certo número de categorias estáveis, capazes de agrupar todo o conjunto de dados disponíveis, alocando cada elemento sempre à mesma classe quando este é apresentado, e por outro lado, mantendo suficiente plasticidade para acomodar novos dados, alterando as categorias quando necessário (Carvalho, 2001).

Existem algumas desvantagens genéricas da representação em unidades vencedoras (Caudill, 1989 apud Hertz *et al.*, 1991); a seguir listadas:

- a) precisa-se de um neurônio de saída (e as conexões associadas) para cada categoria envolvida -  $N$  unidades podem representar apenas  $N$  categorias;
- b) não é robusta à degradação ou falha; no caso de uma unidade de saída falhar, perde-se então toda a categoria; e
- c) não se pode ter representação hierárquica - dois padrões de entrada são agrupados juntos ou não, não existe modo de se ter subcategorias.

### **6.1.1 Inicialização dos vetores de pesos**

Devem ser designados valores iniciais a todos os vetores de pesos da rede antes de o treinamento começar, e é uma prática comum com redes neuronais atribuir valores aleatórios. Após o treinamento, esses vetores de pesos devem terminar próximos aos vetores de entrada. Os vetores de entrada usualmente não são uniformemente distribuídos e tendem a se agrupar numa porção relativamente pequena da superfície da hiperesfera. Selecionar os vetores de pesos de forma aleatória pode causar sérios problemas de treinamento, pois estes ficam uniformemente distribuídos em volta da hiperesfera e, por essa razão, a maioria fica tão distante de qualquer vetor de entrada que nunca se tornam a melhor opção. Estes neurônios terão sempre saída zero e serão desperdiçados. Além disso, os pesos restantes que se tornaram a melhor opção podem ser insuficientes para permitir a separação das categorias dos vetores

de entrada que estejam próximos na superfície da hiperesfera. Essas unidades distantes dos vetores de entrada que nunca vencem e, portanto, nunca aprendem, são chamadas unidades mortas (Hertz *et al.*, 1991).

Supondo que haja vários conjuntos de vetores de entrada, todos esses similares, mas que ainda tenham que ser separados em categorias diferentes, a rede deve ser treinada para ativar um neurônio diferente para cada categoria. Se a densidade inicial dos vetores de pesos é muito baixa na vizinhança dos vetores de treinamento, será impossível separar categorias similares, pois podem não existir vetores de treinamento suficientes nessa vizinhança para designar um para cada categoria de vetor de entrada.

Ao contrário, se vários vetores de entrada são pequenas variações de um mesmo padrão e devem ser aglomerados juntos, eles devem disparar um único neurônio. Se a densidade de vetores de pesos é muito alta próxima a um grupo de vetores de entrada ligeiramente diferentes, cada vetor de entrada pode ativar um neurônio diferente.

A solução mais desejável é distribuir os vetores de pesos de acordo com a densidade dos vetores de entrada que devem ser separados, colocando mais vetores de pesos na vizinhança do maior número de vetores de entrada. A seguir são apresentadas várias técnicas que buscam prevenir unidades mortas (Wasserman, 1989; Hertz *et al.*, 1991):

- a) Inicializar os pesos com amostras dos vetores de entrada, assegurando assim que estejam todos no domínio certo.
- b) Atualizar os pesos de todos os perdedores bem como dos vencedores, porém, com uma taxa de aprendizado  $\rho$  bem menor. Uma unidade que esteja sempre perdendo gradualmente move-se na direção da média das entradas até que eventualmente é bem sucedida em vencer uma competição. Essa técnica é chamada *leaky learning*.
- c) Nos estágios iniciais do processo de treinamento ajustar todos os pesos, não apenas aqueles associados com o neurônio vencedor. Isso move os vetores de pesos ao redor

da região dos vetores de entrada. À medida que o treinamento progride, os ajustes nos pesos são restringidos aos neurônios mais próximos do vencedor. Para isso, os neurônios são arranjados numa forma geométrica uni ou bidimensional. O raio de ajustamento é gradualmente reduzido até que são ajustados apenas os pesos associados com o neurônio vencedor. Esta é a essência do mapa auto-organizável de Kohonen que será visto adiante.

- d) No método de combinação convexa, todos os pesos são fixados com o mesmo valor  $1/\sqrt{p}$ , onde  $p$  é a dimensão do vetor de entrada e, portanto, o número de componentes dos vetores de pesos. Isto torna os vetores de pesos de comprimento unitários e coincidentes. Ademais, à cada componente  $x_i$  da entrada é dado o valor  $ax_i + \{[1/\sqrt{p}] (1 - a)\}$ . Inicialmente “ $a$ ” é atribuído um valor bem baixo, forçando todos os vetores de entrada a ter um comprimento próximo a  $1/\sqrt{p}$  e coincidente com os vetores de pesos. À medida que a rede treina, “ $a$ ” é gradualmente incrementado até o limite de 1. Isso permite que os vetores de entrada se separem e eventualmente assumam seus valores reais. Os vetores de pesos seguem um ou um pequeno grupo de vetores de entrada e o processo de treinamento termina produzindo o padrão desejado de saídas. O método de combinação convexa opera bem, mas retarda o processo de treinamento, visto que os vetores são ajustados a um alvo móvel.
- e) Adicionar ruído aos vetores de entrada faz com que eles se movam aleatoriamente, eventualmente capturando um vetor de pesos. Este método é ainda mais lento que a combinação convexa.
- f) Ainda um outro método proposto por DeSieno (1988) dá para cada neurônio uma “consciência”. Um ganhador freqüente sente-se “culpado” e, caso esteja ganhando mais do que uma parte justa do tempo (aproximadamente  $1/N$ , onde  $N$  é o número de

neurônios na camada de saída), cria temporariamente um limite que reduz suas chances de vencer, dando assim aos outros neurônios uma oportunidade de serem treinados.

Em muitas aplicações o problema da distribuição dos pesos pode afetar seriamente a precisão do resultado. Infelizmente, a efetividade das várias soluções não foi completamente avaliada e é certamente dependente do problema.

## 6.2 Mapa Auto-Organizável

O Mapa Auto-Organizável (SOM – do inglês *Self-Organizing Map*<sup>17</sup>) é um algoritmo neuronal não supervisionado, utilizado extensamente para visualizar e interpretar conjuntos de dados com dimensão elevada. Uma lista de referências com aplicações do SOM e suas variações em diferentes áreas, entre 1981 e 1997, é apresentada em Kaski *et al.* (1998), e entre 1998 e 2001 em Oja *et al.* (2002). Uma aplicação recente do SOM em estudos com espécies invasoras foi a determinação de padrões geográficos com modelagem da distribuição mundial de pestes de insetos (Gevrey *et al.*, 2006; Worner & Gevrey, 2006).

O SOM define uma rede elástica de pontos que são ajustados à distribuição dos dados de treinamento do espaço de entrada. Pode ser usado, assim, para dados multidimensionais, normalmente visualizados em uma treliça bidimensional (Laaksonen *et al.*, 2001). Tal rede pode aprender a detectar regularidades e correlações em suas entradas e adaptar futuras respostas a essas entradas. Os SOMs classificam vetores de entrada conforme estão agrupados no espaço de entrada. Eles diferem das redes competitivas simples, pois os neurônios vizinhos no SOM aprendem a reconhecer seções vizinhas do espaço de entrada. Por conseguinte, SOMs aprendem tanto a distribuição (como as camadas competitivas) quanto a topologia dos vetores de entrada no qual são treinados (MathWorks, 2005).

---

<sup>17</sup> Também chamado *Self-Organizing Feature Map* (SOFM)

O SOM tem a propriedade desejável de preservação topológica que captura um importante aspecto dos mapas de características no córtex dos cérebros dos animais mais desenvolvidos (Jain *et al.*, 1996). Por preservação topológica entende-se que padrões de entrada próximos devem ativar unidades de saída próximas no mapa. A arquitetura básica da rede de SOM é mostrada na Figura 41. Essa arquitetura consiste basicamente de um *array* bidimensional de unidades de saída onde cada uma é conectada a todos os nós  $p$  de entrada. De forma similar à rede de aprendizado competitivo, o vetor  $p$ -dimensional associado com a unidade no local  $(i,j)$  do *array* 2-D é denotado como  $\mathbf{w}_{ij}$ .

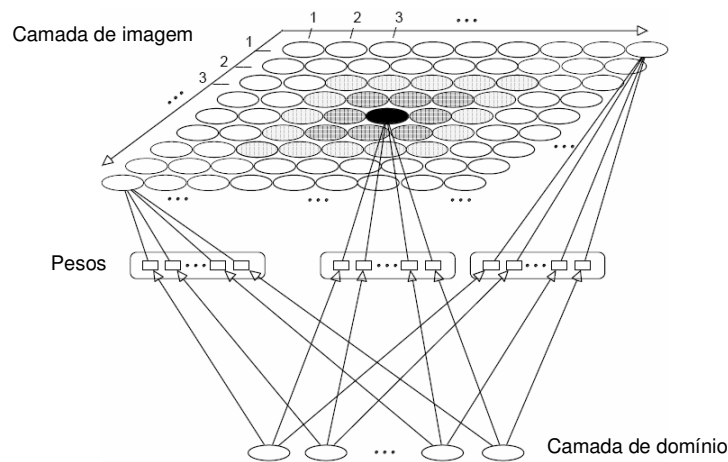


Figura 41 - Mapa auto-organizável de Kohonen (adaptado de Ultsch, 1995, p. 3)

### 6.2.1 Algoritmo de Kohonen

O algoritmo SOM, desenvolvido por Teuvo Kohonen, permite que um espaço de entrada contínuo seja mapeado em um conjunto discreto de neurônios (Figura 42). A motivação dessa idéia reside na redução de dimensionalidade ou compressão de dados. Vários padrões são apresentados à rede. Cada entrada apresentada causa a ativação na camada de saída de um grupo localizado de neurônios. Dependendo da aplicação de interesse, a resposta da rede pode ser tanto a posição na treliça do neurônio vencedor dada pelo índice  $j^*$  quanto o vetor de peso sináptico que está mais perto no sentido Euclidiano do vetor de entrada.

A ordem na qual os padrões de entrada são apresentados para a rede interfere no resultado final do mapeamento. Variações na sequência de entrada podem resultar em mapeamentos topologicamente incorretos, pois a ordem de apresentação afeta o processo competitivo na camada de saída influenciando nas direções de deslocamentos dos vetores de peso (Hsu & Wu, 1992 *apud* Miranda, 1998). Em geral para minimizar problemas com relação a viés, os padrões são escolhidos de forma aleatória dentro do espaço de entrada.

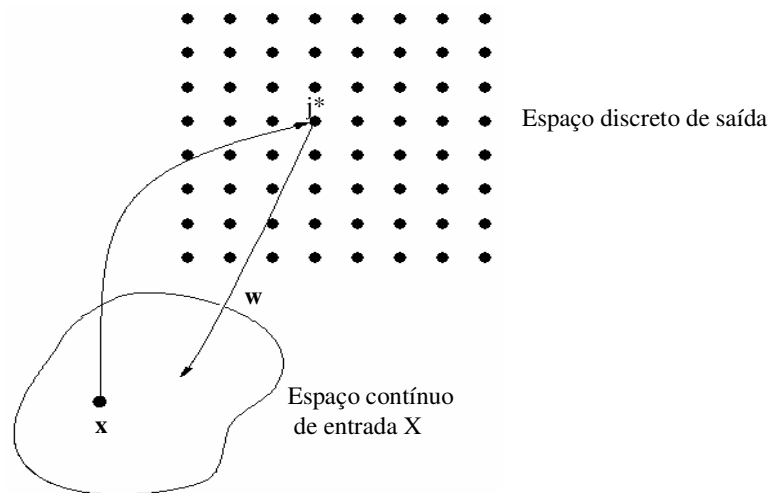


Figura 42 – Ilustração da relação entre os espaços de entrada e saída por meio do vetor de entrada  $x$  o vetor de peso  $w$  do neurônio vencedor  $j^*$  (adaptado de Haykin, 1994, p. 415)

Segundo Haykin (1994) os ingredientes essenciais da rede neuronal são:

- Uma treliça uni ou bi-dimensional de neurônios que calcula funções discriminantes simples de entradas recebidas com dimensão arbitrária.
- Um mecanismo que compara essas funções discriminantes e seleciona o neurônio com o maior valor.
- Uma rede interativa que ativa o neurônio selecionado e seus vizinhos simultaneamente.
- Um processo adaptativo que permite os neurônios ativados aumentar seus valores da função discriminante em relação aos sinais de entrada.



A procura pela unidade melhor ajustada domina o tempo de computação do algoritmo do SOM, e pode ser computacionalmente caro em entradas com alta dimensionalidade ou grandes redes SOM. O algoritmo básico usa procura completa, onde todas as unidades devem ser consideradas para achar a melhor ajustada.

A camada de Kohonen classifica os vetores de entrada em grupos que são similares. Isso é obtido pelo ajustamento dos pesos de forma que vetores de entrada similares ativem o mesmo neurônio, como uma rede de aprendizado competitivo. O treinamento de Kohonen é, portanto, um algoritmo auto-organizável que opera num modo não supervisionado. Por essa razão, é difícil (e desnecessário) prever qual neurônio específico será ativado para um dado vetor de entrada. É necessário apenas assegurar que o treinamento separa vetores de entrada dissimilares e agrupa os semelhantes (Wasserman, 1989).

### 6.2.2 Função de vizinhança

A rede do SOM identifica o neurônio vencedor  $j^*$  usando o mesmo procedimento empregado no aprendizado competitivo. Entretanto, ao invés de atualizar somente os pesos sinápticos do neurônio vencedor, todos os pesos dos neurônios situados dentro de uma certa vizinhança deste neurônio são atualizados. Assim, quando um vetor  $\mathbf{x}$  é apresentado, os pesos do neurônio vencedor e seus vizinhos próximos movem-se em direção a  $\mathbf{x}$ , conforme ilustrado na Figura 43. Conseqüentemente, após várias apresentações, neurônios vizinhos terão aprendido vetores similares um ao outro.

A Figura 44 ilustra o conceito de vizinhança em torno de um neurônio com raios diferentes, igual a um e dois. A vizinhança  $\phi_{j^*}(d)$  contém os índices para todos os neurônios situados dentro de um raio  $d$  do neurônio vencedor  $j^*$ .

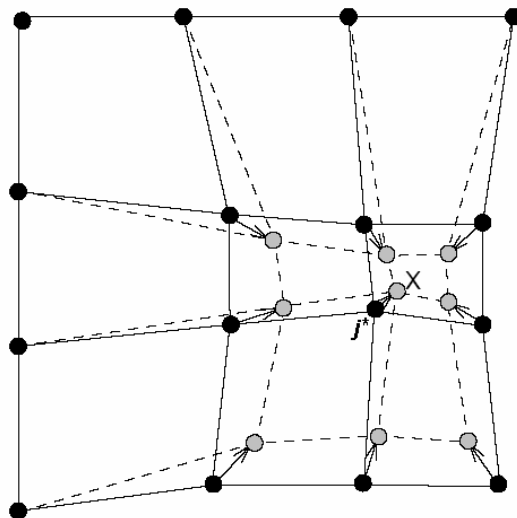


Figura 43 - Atualização dos pesos do neurônio vencedor  $j^*$  e dos seus vizinhos em direção à entrada apresentada  $x$ . Os círculos pretos e cinzas correspondem as situações antes e depois da atualização (adaptado de Vesanto, 2002, p. 13)

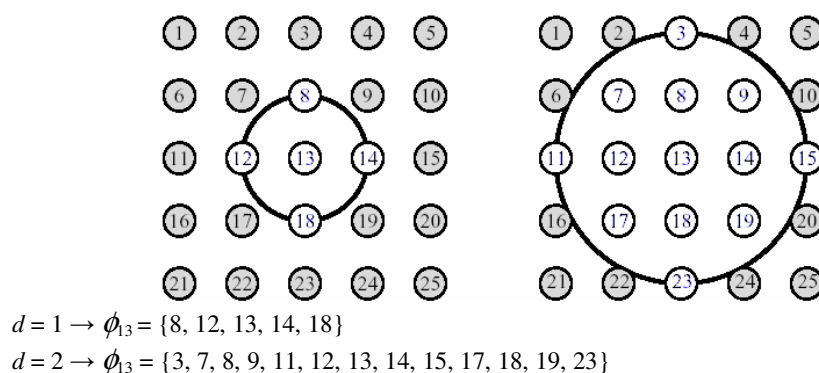


Figura 44 - Esquemas de vizinhança (MathWorks, 2005, p. 8-10)

Os neurônios de um SOM não necessitam estar dispostos num padrão bidimensional. Pode-se usar um arranjo unidimensional ou mapas de três ou mais dimensões que, contudo, não são tão comuns. No caso de um SOM unidimensional, um neurônio tem apenas dois vizinhos dentro de um raio igual a 1 (ou um único vizinho caso o neurônio esteja num extremo da linha). Os neurônios na camada de imagem de um SOM são dispostos originalmente numa determinada posição física de acordo com uma função de topologia (ex: grade, hexagonal, aleatória) como mostrado na Figura 45.

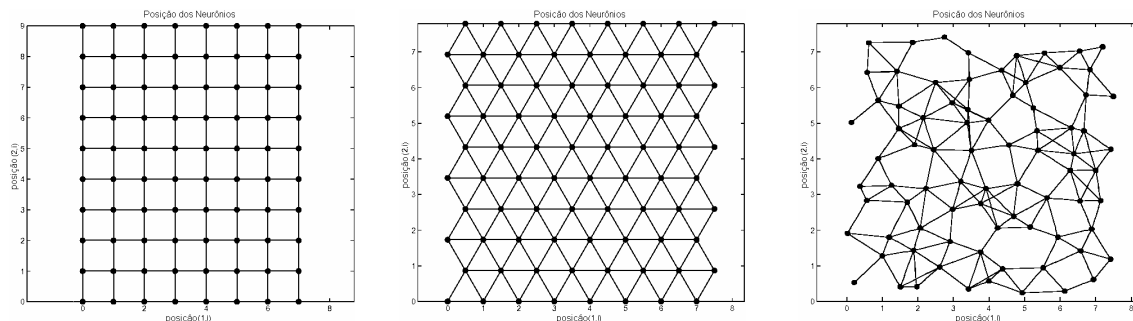


Figura 45 - Exemplos de topologias de mapas auto-organizáveis (adaptado de MathWorks, 2005, p.8-12 a 8-14)

É possível definir o formato da vizinhança local de diferentes maneiras usando, por exemplo, vizinhança quadrada, retangular, hexagonal ou circular (ver Figura 46). O desempenho da rede não é sensível à forma exata da vizinhança (MathWorks, 2005) que deve encolher com o tempo de acordo com alguma programação tal como uma função de decaimento exponencial.

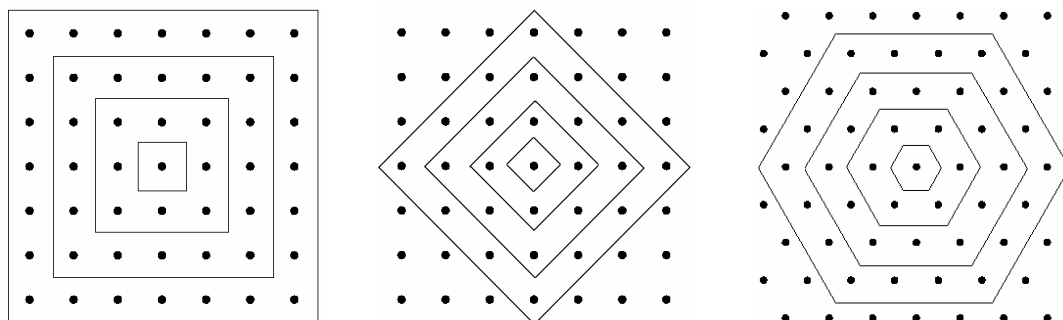


Figura 46 - Exemplos de vizinhanças discretas (elaboração própria)

Do ponto de vista neurobiológico, existe evidência de que um neurônio ao disparar tende a excitar mais os neurônios em sua vizinhança imediata do que aqueles distantes dele (Damásio, 1994). Seguindo essa motivação biológica, as interações laterais entre neurônios são usualmente descritas pela função do *Chapéu Mexicano*, cuja forma é ilustrada na Figura 47. As interações se dão numa estrutura *on-center/off-surround*. Nesta estrutura, podem-se distinguir duas áreas distintas, cada neurônio está em cooperação, por meio de sinapses excitatórias, com seus vizinhos imediatos, ao mesmo tempo em que está em competição, por meio de sinapses inibitórias, com seus vizinhos distantes (Carvalho *et al.*, 2003).

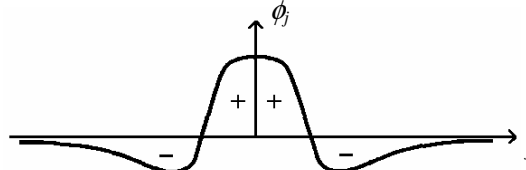


Figura 47 - Forma da função do *Chapéu Mexicano* (adaptado de Hertz *et al.*, 1991, p. 234)

Na forma tradicional do algoritmo SOM, a função de vizinhança  $\phi$  ao redor do neurônio vencedor  $j^*$  é assumida com uma amplitude constante. A implicação disso é que todos os neurônios localizados dentro desta vizinhança disparam com uma mesma taxa, e a interação é independente da distância lateral do neurônio vencedor  $j^*$ . Esta forma simplificada é ilustrada na Figura 48.

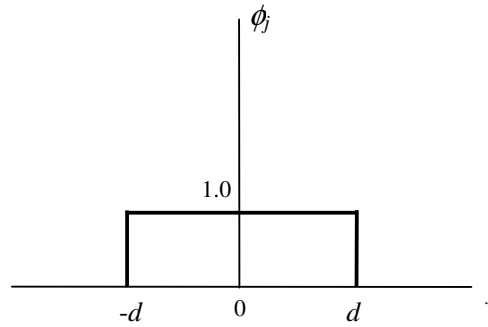


Figura 48 - Forma simplificada de vizinhança (adaptado de Haykin, 1994, p. 423)

Uma forma alternativa de levar em conta o efeito de decaimento com a distância lateral, é adotar uma função topológica gaussiana que estabelece o raio em torno do neurônio vencedor  $j^*$ , dentro do qual os pesos dos neurônios próximos são fortemente modificados (ocorre cooperação) enquanto os pesos dos neurônios distantes são pouco modificados ou permanecem inalterados (Figura 49 e Figura 50). Uma escolha típica para  $\phi_{j,j^*}$  é:

$$\phi_{j,j^*}(t) = \exp\left(-\frac{d_{j,j^*}^2}{2\sigma^2(t)}\right) \quad , \quad (13)$$

onde:  $\sigma(t)$  é a dispersão da função de vizinhança (mede a largura para  $\phi$  igual a  $e^{-1/8}$ )

e  $d_{j,j^*}$  é a distância lateral entre o neurônio  $j$  e o vencedor  $j^*$ .

A função  $\phi_{j,j^*}(t)$  tem seu máximo em  $d_{j,j^*} = 0$ , e a amplitude decai com o aumento da distância lateral, tendendo a zero para  $d_{j,j^*} \rightarrow \infty$ .

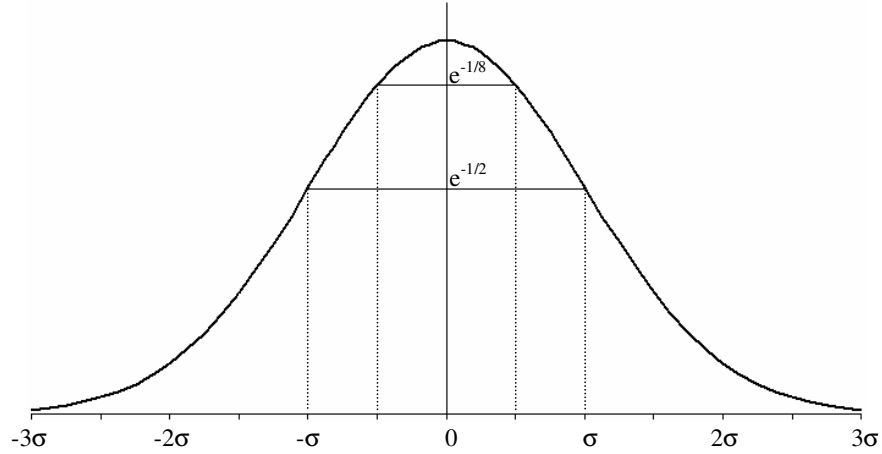


Figura 49 - Função de vizinhança gaussiana (elaboração própria)

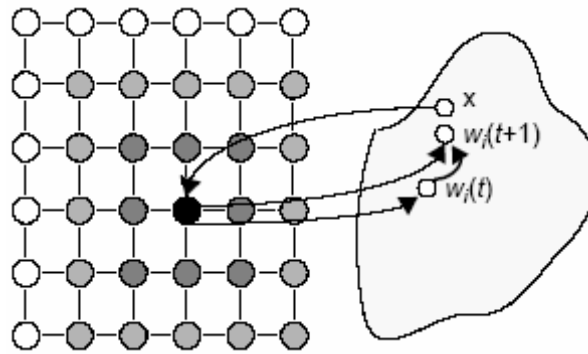


Figura 50 - Representação esquemática da atualização dos pesos sinápticos de um neurônio e seus vizinhos no SOM (Merkl, 1995, p.3)

O processo de modificação dos pesos sinápticos  $\Delta \mathbf{w}_j(t)$  para cada neurônio  $j$  é semelhante ao da equação (12), embora sejam levados em conta os neurônios vizinhos para cada passo de aprendizado  $t$  em que o estímulo  $\mathbf{x}$  é apresentado à rede neuronal (Figura 50). Assim, pode-se reescrever a equação (12) de modo a alterar todos os pesos dentro da vizinhança (enquanto neurônios distantes de  $j^*$  permanecem inalterados) da seguinte forma:

$$\Delta \mathbf{w}_j(t) = \rho(t) \cdot \phi_{j,j^*}(t) \cdot [\mathbf{x} - \mathbf{w}_j(t)] \quad (14)$$

## 7 Implementação do SOM

Os parâmetros de projeto da rede incluem a dimensionalidade do *array* de neurônios, número de neurônios em cada dimensão, forma da vizinhança, programação de encolhimento da vizinhança e taxa de aprendizado. A aplicação do algoritmo do SOM pode ser sumarizada em alguns passos básicos (Haykin, 1994; Jain *et al.*, 1996):

1. Inicialização: Escolha de valores para os vetores de peso iniciais  $\mathbf{w}_j(0)$ . A única restrição é que  $\mathbf{w}_j(0)$  sejam diferentes para  $j = 1, 2, \dots, N$ , onde  $N$  é o número de neurônios na camada imagem da rede. Definição da taxa de aprendizado inicial  $\rho(0)$  e da largura da vizinhança inicial  $\sigma(0)$ .

2. Para cada passo de aprendizagem ( $t$ ):

2.1. Apresentação: Os vetores de entrada  $\mathbf{x}$  são apresentados à rede numa ordem aleatória.

2.2. Similaridade: Para cada vetor de entrada apresentado encontra-se o neurônio vencedor  $j^*$  no tempo  $t$ , usando o critério de mínima distância Euclidiana:

$$\|\mathbf{x} - \mathbf{w}_{j^*}\| = \min \|\mathbf{x} - \mathbf{w}_j\|$$

2.3. Atualização: Ajuste de todos os vetores de pesos sinápticos dos neurônios com a seguinte regra de aprendizado:

$$\mathbf{w}_j(t+1) = \mathbf{w}_j(t) + \rho(t) \cdot \phi_{j,j^*}(t) \cdot [\mathbf{x} - \mathbf{w}_j(t)]$$

onde  $\rho(t)$  é o parâmetro da taxa de aprendizado; e  $\phi_{j,j^*}(t)$  é a função de vizinhança centrada ao redor do neurônio vencedor  $j^*$  na iteração  $t$ .

2.4. Revisão: Ambos  $\rho$  e  $\phi$  são variados dinamicamente durante o aprendizado: decrescer o valor de  $\rho(t)$  e encolher a largura  $\sigma(t)$ .

3. Prosseguimento: Voltar ao passo 2 até que a mudança nos valores dos pesos for menor que um limite pré-especificado, ou o número máximo de iterações tenha sido alcançado.

Para implementar o algoritmo do SOM decidiu-se inicialmente escrever um programa na linguagem ANSI C que incorporasse todos esses passos. Este exercício mostrou-se bastante proveitoso do ponto de vista da obtenção de um entendimento minucioso do funcionamento do algoritmo e de uma melhor compreensão das possibilidades de pré-processamento dos vetores

de entrada e inicialização dos vetores de pesos; e do comportamento das diferentes funções de decréscimo da taxa de aprendizado e encolhimento da vizinhança.

Um programa foi elaborado e checado inicialmente com bases de dados disponíveis para testes (e.g., banco de dados Iris<sup>18</sup>). Em seguida, foram realizadas rodadas com o banco de dados GloBallast. No entanto, a pouca experiência com programação em C, principalmente em desenvolver algoritmos para visualização e análise dos resultados, fez com que apesar de o programa estar operacional, se optasse por usar aplicativos como o SOM\_PAK (Kohonen *et al.*, 1995) e o SOM Toolbox para Matlab 5 (Vesanto *et al.*, 2000) para produzir os mapas.

Ambos os aplicativos são bem documentados, de domínio público e disponíveis na internet<sup>19,20</sup>, e foram desenvolvidos pelo *Laboratory of Computer and Information Science* da *Helsinki University of Technology*, sendo que o SOM\_PAK, escrito em C, teve a participação direta de Teuvo Kohonen. O SOM Toolbox funciona como uma biblioteca do software MATLAB composta por várias rotinas, e tem grande flexibilidade e variedade de funções para visualização. Um inconveniente nos dois pacotes é a falta de possibilidade de suporte.

## 7.1 Aprendizado

O processo de aprendizado subjacente à determinação de um SOM é de natureza estocástica, no qual a precisão do resultado depende do número de iterações do algoritmo. Além disso, a seleção dos principais parâmetros do algoritmo (taxa de aprendizado e função de vizinhança) é crítica para o sucesso da formação do mapa. Não existe uma base teórica para a escolha desses parâmetros, que são de modo geral determinados pelo processo de tentativa e erro.

---

<sup>18</sup> O conjunto de dados Iris é incluído como exemplo em muitas aplicações que lidam com estatística e análise de dados. Este conjunto contém 50 amostras de cada uma de 3 espécies da planta Íris (“Setosa”, “Virginica” e “Versicolor”) com valores de 4 características, “comprimento da sépala”, “largura da sépala”, “comprimento da pétala”, e “largura da pétala”.

<sup>19</sup> <http://www.cis.hut.fi/research/som-research/nnrc-programs.shtml>

<sup>20</sup> <http://www.cis.hut.fi/projects/somtoolbox/>

### 7.1.1 Treinamento seqüencial

As seguintes observações servem como guia para o treinamento seqüencial (Kohonen, 1990; Hertz *et al.*, 1991; Haykin, 1994):

- O parâmetro da taxa de aprendizado  $\rho(t)$  usada para atualizar o vetor de pesos sinápticos  $\mathbf{w}_j(t)$  deve variar no tempo. O valor de  $\rho(0)$  deve ser próximo da unidade; depois disso, durante as 1.000 primeiras iterações,  $\rho(t)$  deve decrescer gradualmente, mas ficando acima de 0,1. A forma exata da variação de  $\rho(t)$  com  $t$  não é crítica; pode ser linear, exponencial ou inversamente proporcional a  $t$ . É durante essa fase inicial do algoritmo que a ordem topológica dos vetores de pesos  $\mathbf{w}_j(t)$  é estabelecida. Esta fase do processo de aprendizado é chamada de *fase de ordenamento*. As iterações restantes do algoritmo (relativamente longas) são necessárias principalmente para o ajustamento fino dos vetores de peso; essa segunda fase do processo de aprendizado é chamada de *fase de convergência*. Para uma boa precisão estatística,  $\rho(t)$  deve ser mantido durante a fase de convergência num valor pequeno (da ordem de 0,01 ou menos) por um tempo razoavelmente longo, tipicamente de milhares de iterações.

- Para que o ordenamento topológico dos vetores de pesos  $\mathbf{w}_j$  aconteça, um cuidado especial deve ser dado à função  $\phi_{j^*}(t)$ . De modo geral, defini-se a função de vizinhança com  $\sigma(0)$  de forma a incluir todos os neurônios num raio de mais da metade do diâmetro da rede (Kohonen, 1990) reduzindo-a gradualmente com o tempo. Durante a fase inicial de 1.000 iterações, quando a ordem topológica dos vetores de pesos sinápticos está acontecendo, a largura de  $\sigma(t)$  é encolhida com o tempo  $t$  até um valor pequeno, de apenas uns poucos neurônios vizinhos. Durante a fase de convergência do algoritmo,  $\phi_{j^*}(t)$  deve conter apenas os vizinhos mais próximos do neurônio vencedor  $j^*$  ou, eventualmente, apenas o neurônio vencedor. Neste caso, o processo é reduzido a um aprendizado competitivo simples.



A variação no tempo de  $\sigma(t)$  e  $\rho(t)$  das equações (13) e (14) pode tomar várias formas. Em particular, podem ter decaimentos conforme a Figura 51. É interessante notar que, pelo apropriado uso de vizinhanças topológicas, o algoritmo SOM procura assegurar que os neurônios da rede não sejam subutilizados, um problema que aflige outras redes de aprendizado competitivo.

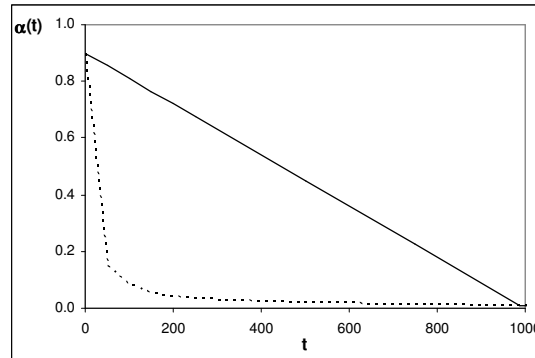


Figura 51 - Diferentes funções para decaimento: linear (linha sólida)  $\alpha(t) = \alpha(0) \times (1 - t/T)$  e inversa (linha tracejada)  $\alpha(t) = \alpha(0) / (1 + 100 \times t/T)$ , onde T é a duração do treinamento e  $\alpha(0)$  o valor inicial da taxa de aprendizado ou vizinhança (elaboração própria)

### 7.1.2 Treinamento em lote

O treinamento seqüencial descrito anteriormente pode ser substituído pelo treinamento em lote que é significativamente mais rápido e não requer a especificação da taxa de aprendizado. Seja  $V_j$  o conjunto de vetores de entrada  $\mathbf{x}$  capazes de atualizar o vetor de peso  $\mathbf{w}_j$ , em outras palavras, o neurônio vencedor  $j^*$  para cada  $\mathbf{x} \in V_j$  deve pertencer ao conjunto de vizinhos  $C_j$  do neurônio  $j$ . O domínio  $V_j$  no espaço de  $\mathbf{x}$  é também chamado região de influência do neurônio  $j$  ou conjunto de Voronoi. A união do conjunto  $V_j$  correspondente aos neurônios em  $C_j$  é denotado por  $U_j$ . A atualização dos vetores de peso se dá conforme a seguinte equação (Kohonen, 1998):

$$\mathbf{w}_{j^*} = \frac{\sum_{\mathbf{x} \in U_j} \mathbf{x}}{n(U_j)}, \quad (15)$$

onde  $n(U_j)$  significa o número de vetores de entrada  $\mathbf{x}$  que pertence ao conjunto  $U_j$ .

Como  $U_j$  depende de  $\mathbf{w}_{j*}$ , a equação (15) deve ser resolvida iterativamente. Kohonen (1997) descreve o procedimento com os seguintes passos:

1. Inicializar os vetores de peso iniciais  $\mathbf{w}_j(0)$ .
2. Apresentar todos os vetores de entrada  $\mathbf{x}$  à rede de uma vez.
3. Para cada neurônio  $j$  listar todos os vetores de entrada  $\mathbf{x}$  cujos neurônios vencedores estejam inseridos dentro da vizinhança do neurônio  $j$  na iteração  $t$ .
4. Atualizar os vetores de pesos de cada neurônio de acordo com a média de sua respectiva lista.
5. Repetir de 2 a 4 por algumas iterações até que a solução seja considerada estável.

O tamanho da vizinhança  $\phi$  e sua variação podem ser similares aos usados no treinamento sequencial. Nas duas últimas iterações,  $\phi$  deve conter apenas o neurônio  $j$ , e a última iteração do algoritmo é, então, equivalente ao método k-médias. Este algoritmo é particularmente efetivo caso os pesos iniciais sejam aproximadamente ordenados (ver item a seguir).

## 7.2 Vetores de peso

Cada vetor de peso tem dois componentes. O primeiro componente do vetor é a sua localização no mapa bidimensional. O segundo componente são os seus dados que têm a mesma dimensão dos vetores de entrada. Como comentado no item 6.1.1, a distribuição dos pesos no espaço  $p$ -dimensional pode afetar o resultado e existem múltiplas possibilidades de inicialização dos pesos  $\mathbf{w}_j(0)$  quando esses são escolhidos aleatoriamente de forma arbitrária.

Tanto o pacote SOM\_PAK como o SOM Toolbox possuem uma rotina (LININIT) na qual primeiramente são calculados os autovalores e autovetores dos dados de entrada. Em seguida, os vetores de peso são inicializados de uma forma ordenada ao longo dos 2 maiores autovetores da matriz de covariância dos dados de entrada. Esta opção leva a uma convergência mais rápida e a fase de ordenamento pode ser dispensada, sendo necessária apenas a fase de convergência (Kohonen *et al.*, 1995).

### 7.3 Avaliação da qualidade do treinamento

Como o SOM não possui uma função objetivo, é difícil decidir quando parar o treinamento ou avaliar a “qualidade” relativa de vários mapas obtidos com parâmetros ou inicializações diferentes. Costuma-se dividir o total das entradas em dois conjuntos: de treinamento e de teste. A rede é treinada com o primeiro conjunto e a classificação é, então, avaliada com base no desempenho com o conjunto de teste. O percentual de classificações erradas pode ser tomado como uma taxa de erro. No entanto, em algumas aplicações é difícil, ou mesmo impossível, dividir as entradas e, ainda, assinalar classes para o conjunto de treinamento.

Vários processos de aprendizado podem ser definidos começando o treinamento com diferentes valores iniciais para os parâmetros de aprendizado, e apresentando os vetores de entrada  $\mathbf{x}$  em diferentes seqüências. Além disso, existe uma multiplicidade de possíveis inicializações aleatórias de pesos  $\mathbf{w}_j(0)$  a serem testadas.

É óbvio que algum mapa ótimo para um dado conjunto de dados de entrada deva existir (Kohonen *et al.*, 1995). Portanto, é aconselhável selecionar o “melhor” SOM por meio de um índice de desempenho. Uma medida de avaliação é o erro de quantização médio, definido como a média de  $\|\mathbf{x} - \mathbf{w}_{j^*}\|$  avaliado para o mapa com os vetores de pesos treinados. Para cada vetor de entrada o neurônio vencedor é procurado no mapa e o respectivo erro de quantização é calculado. Espera-se que o melhor mapa tenha o menor erro de quantização médio, pois estaria mais bem ajustado aos vetores de entrada. Outra medida é o erro topográfico, calculado verificando-se para todas as entradas o neurônio mais bem ajustado e o segundo neurônio mais bem ajustado (*best and second-best matching*). O erro topográfico é a proporção das entradas para as quais esses neurônios não são adjacentes na treliça do mapa (Kiviluoto, 1996). O erro total é normalizado de 0 a 1, onde 0 significa a perfeita preservação topológica.

Essas duas medidas procuram avaliar a quantização vetorial e a projeção vetorial, chamada preservação topológica. Tecnicamente são objetivos competitivos, mas não opostos do SOM, pois existe um *tradeoff*, aumentar a qualidade de uma usualmente diminui a da outra (Vesanto *et al.*, 2003). No SOM acontece uma grande redução na dimensionalidade e, portanto, necessariamente há uma perda de informação no processo de projeção (e.g. no banco de dados GloBallast, a projeção de dados com 34 dimensões num mapa bidimensional).

Pözlzbauer (2004) fez um levantamento de outras medidas propostas para verificação da qualidade do mapa como a medida de distorção; *trustworthiness*; e produto topográfico. Este último pode ser usado para otimizar o tamanho do mapa para um certo conjunto de dados, propriedade que não é comum à maioria das medidas de qualidade. O erro de quantização médio, por exemplo, decresce monotonicamente com incremento no tamanho do mapa. Essas outras medidas não estão disponíveis nos pacotes SOM\_PAK e SOM Toolbox, e são de complexa implementação, fato pelo qual não serão utilizadas na tese.

## 7.4 Resultados

A escolha do tamanho do mapa pode ser feita de acordo com alguma regra ou por uma determinada conveniência. O SOM Toolbox usa uma fórmula heurística para o número de neurônios do mapa, caso este não seja explicitado, calculado como sendo igual a cinco vezes a raiz quadrada do tamanho da amostra. Considerando que o banco de dados possui 357 portos, por essa fórmula o número de neurônios sugerido é de  $5 \times (357)^{1/2} = 94$ . Em seguida, outra heurística é utilizada para determinar as dimensões do mapa. A razão entre os comprimentos dos dois lados do mapa é calculada como a raiz quadrada da razão dos dois maiores autovalores do conjunto de dados. Os comprimentos dos lados são, então, ajustados de modo que seu produto seja próximo do número desejado de neurônios do mapa. Esta rotina do SOM Toolbox determinou um tamanho de mapa de 13x8 e, portanto, com 104 neurônios.

Ainda no campo das rotinas automáticas do SOM Toolbox, existe o comando SOM\_MAKE que cria, inicializa, e treina um SOM usando parâmetros padrão, com o algoritmo de treinamento em lote com duas fases; a inicialização com a rotina LININIT; a duração do treinamento calculada em épocas (i.e. a apresentação de todo o conjunto de dados); a topologia dos neurônios em forma hexagonal e o formato da vizinhança dado pela função gaussiana; e o decréscimo linear da largura de vizinhança. O treinamento em lote não utiliza a taxa de aprendizado. Os demais parâmetros são determinados da seguinte forma<sup>21</sup>:

**Raio inicial:** Depende da inicialização, que caso seja aleatória, o raio inicial corresponde a 1/4 do maior lado do mapa, e caso se utilize a rotina LININIT, é adotado 1/8. O raio inicial pode ser função, também, da operação prévia de treinamento e, nessa circunstância, o raio inicial assume o valor do raio final do treinamento antecedente.

**Raio final:** O valor padrão é 1, mas caso seja uma fase de ordenamento, o raio final é estabelecido como 1/4 do raio inicial.

**Duração do treinamento:** Para a fase de ordenamento é de 10 vezes a razão entre número de unidades e quantidade de dados; e para a fase de convergência é de 40 vezes esta mesma razão.

Dessa forma, com o comando SOM\_MAKE, apenas com os argumentos padrão, um mapa (visto na Figura 52) é gerado com o seguinte treinamento e medidas de qualidade:

Fase 1: 3 épocas ( $10 \times 104 \div 357$ ) e redução do raio de vizinhança de 2 ( $13 \div 8$ ) para 1

Fase 2: 12 épocas ( $40 \times 104 \div 357$ ) e raio de vizinhança inicial e final igual a 1

- Erro de quantização médio: 3,4916

- Erro topográfico: 0,0168

---

<sup>21</sup> As fórmulas para obtenção dos valores usados como padrão em cada parâmetro foram verificadas diretamente no código fonte das rotinas, pois em alguns casos elas diferiam das fórmulas apresentadas no texto de ajuda do SOM Toolbox

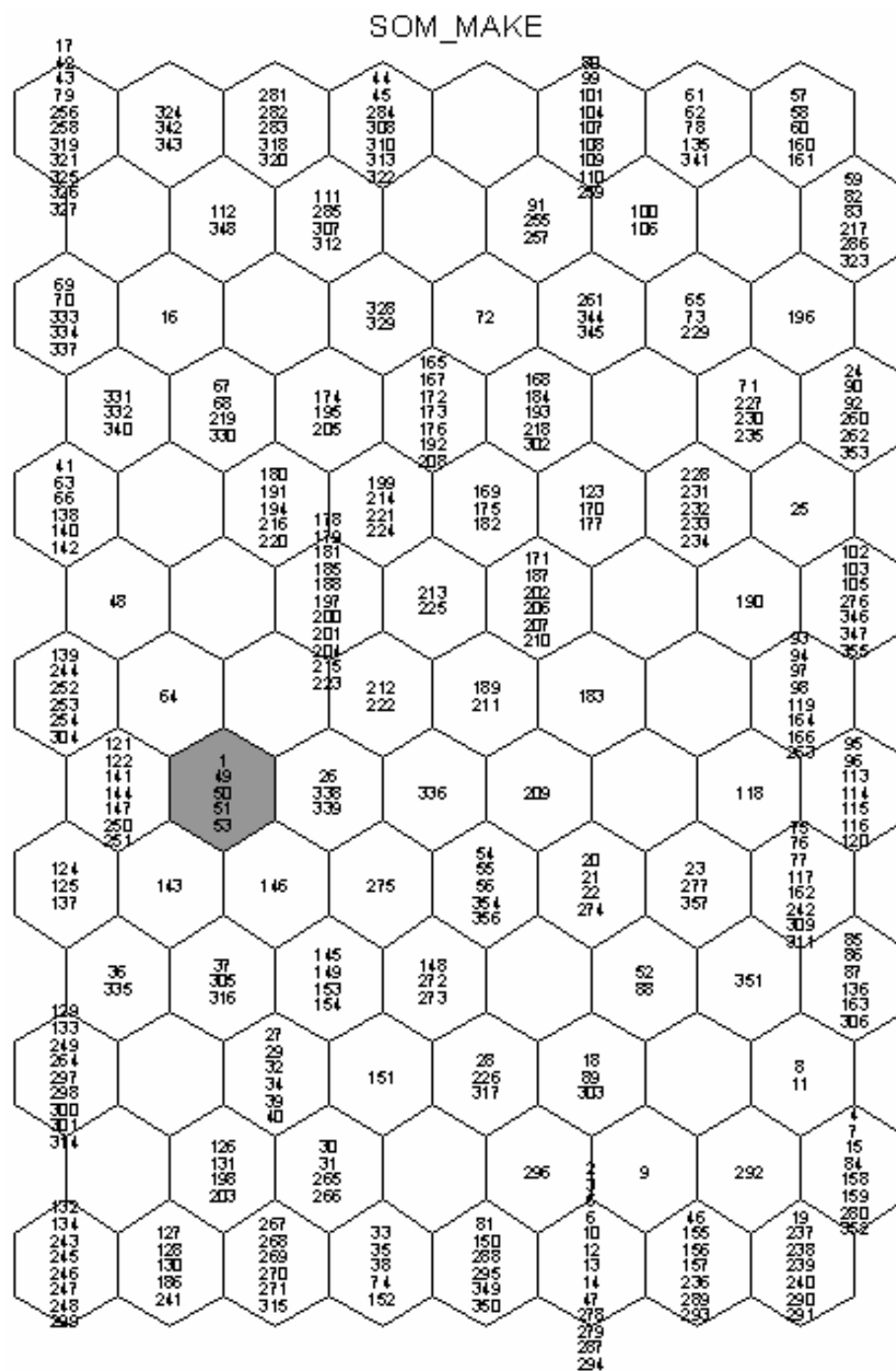


Figura 52 - Mapa gerado com o comando automático SOM\_MAKE (elaboração própria)

Como visto no item anterior, a medida de erro para comparar diferentes tamanhos de mapas não está disponível. De modo a obter uma visualização mais distribuída dos portos no mapa, optou-se por um número de neurônios maior, na ordem de 150, mantendo a razão

indicada pelo SOM Toolbox para a relação entre os lados do mapa, que ficaram definidos em  $15 \times 10$ . O mapa gerado em seguida (visto na Figura 53), também utilizou o comando SOM\_MAKE com a definição do número de neurônios em 150 com o seguinte treinamento e medidas de qualidade:

Fase 1: 5 épocas ( $10 \times 150 \div 357$ ) e redução do raio de vizinhança de 2 ( $15 \div 8$ ) para 1

Fase 2: 17 épocas ( $40 \times 150 \div 357$ ) e raio de vizinhança inicial e final igual a 1

- Erro de quantização médio: 3,2780

- Erro topográfico: 0,0140

Como esperado, por ser maior, o mapa  $15 \times 10$  apresentou menor erro de quantização médio, e nesta configuração o erro topográfico também foi menor. Curiosamente nesse mapa o porto de Sepetiba (1-BRSPB) ficou isolado e os portos de Paranaguá (49-BRPNG); Santos (50-BRSSZ); Rio de Janeiro (51-BRRIO) e Vitória (53-BRVIX), que estavam no mesmo neurônio que Sepetiba, passaram para um neurônio vizinho. Outras relações de vizinhança também mudaram. Os seguintes portos que estavam em neurônios adjacentes à Sepetiba (distância = 1) passaram para uma distância igual a 2: Mangalore (141-INIXE); Chennai (143-INMAA); Porbandar (146-INPBD); Paradeep (147-INPRT); Lagos (250-NGLOS); e Tin Can Island (251-NGTIN), enquanto o porto de Brisbane (26-AUBNE) passou para uma distância igual a 3. Os portos que mantiveram a vizinhança imediata nos dois mapas foram: Chiwan Guangdong (64-CNCWN); Hong Kong (121-HKHKG); Hong Kong Kowloon (122-HKKWN); Marmugao (144-INMRM); Galveston (338-USGLS); e Texas (339-USTXT).

Com o objetivo de encontrar o “melhor mapa”, uma exploração nos vários parâmetros de treinamento foi realizada. Alguns dos parâmetros foram fixados visando permitir a verificação da sensibilidade dos demais. A topologia do mapa foi fixada com neurônios posicionados em forma hexagonal com a dimensão em  $15 \times 10$ , conforme a numeração definida de acordo com a Figura 54. Considerou-se fixa, também, a forma de decréscimo da vizinhança com a função gaussiana.

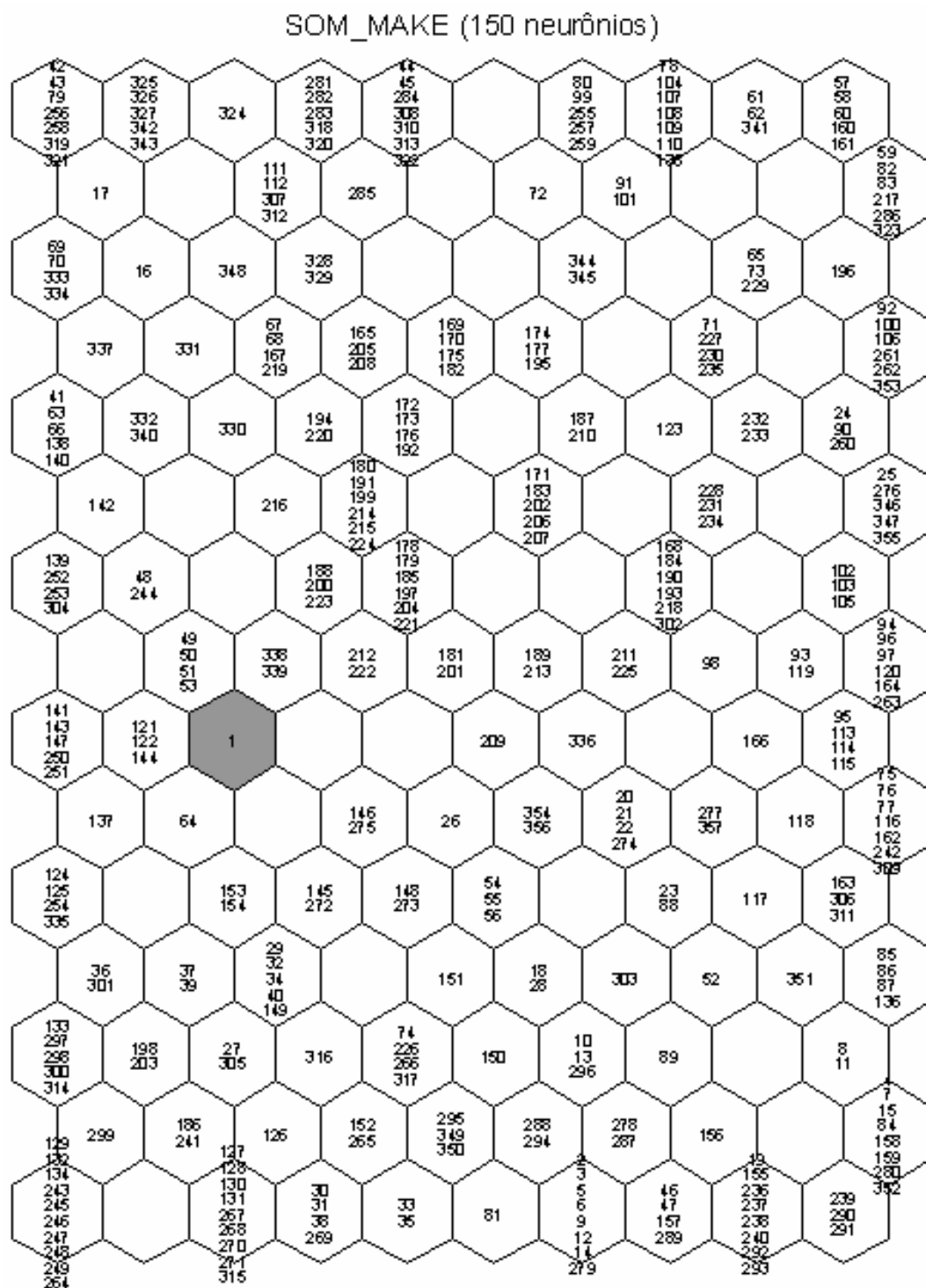


Figura 53 - Mapa gerado com o comando SOM\_MAKE para 150 neurônios (elaboração própria)



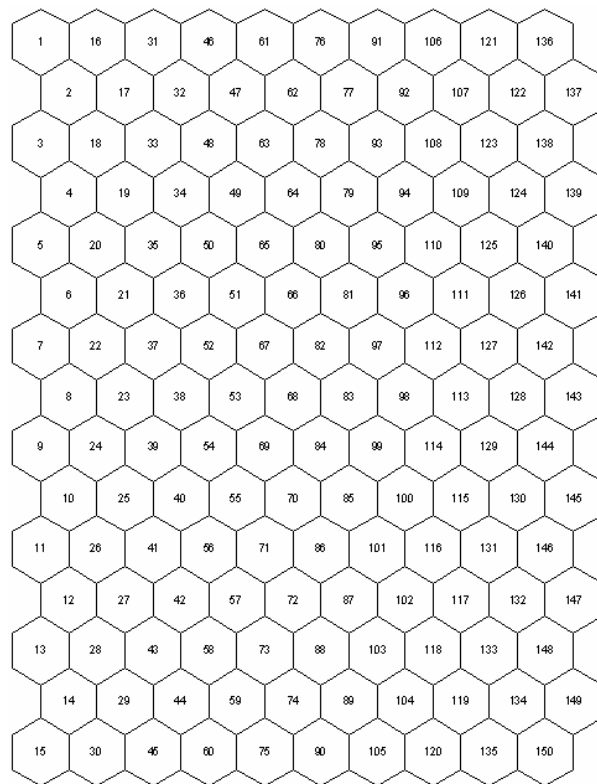


Figura 54 - Numeração dos neurônios no mapa 15×10 (elaboração própria)

No treinamento sequencial, Kohonen (1990) sugere uma “*rule of thumb*” onde o número total de iterações deveria ser pelo menos 500 vezes o número de neurônios (o que num mapa com 150 neurônios resulta em 75.000 iterações ou 210 épocas de 357 dados), porém, indica que 10.000 iterações ou menos podem ser suficientes em alguns aprendizados. De fato, na forma sequencial, o treinamento requer longa duração para minimizar o erro de quantização como mostra a Figura 55 onde a fase de convergência é testada com 25, 50 e 100 épocas (Tabela 19), sendo a taxa de aprendizado inicial de 0,1 (decaindo com a função inversa), e o raio inicial e final igual a 1. Nos três casos, a inicialização foi feita com a rotina LININIT e a fase de ordenamento foi realizada em 5 épocas, sendo a taxa de aprendizado inicial de 0,9 (decaindo com a função inversa); o raio inicial igual a 2; e o raio final igual a 1.

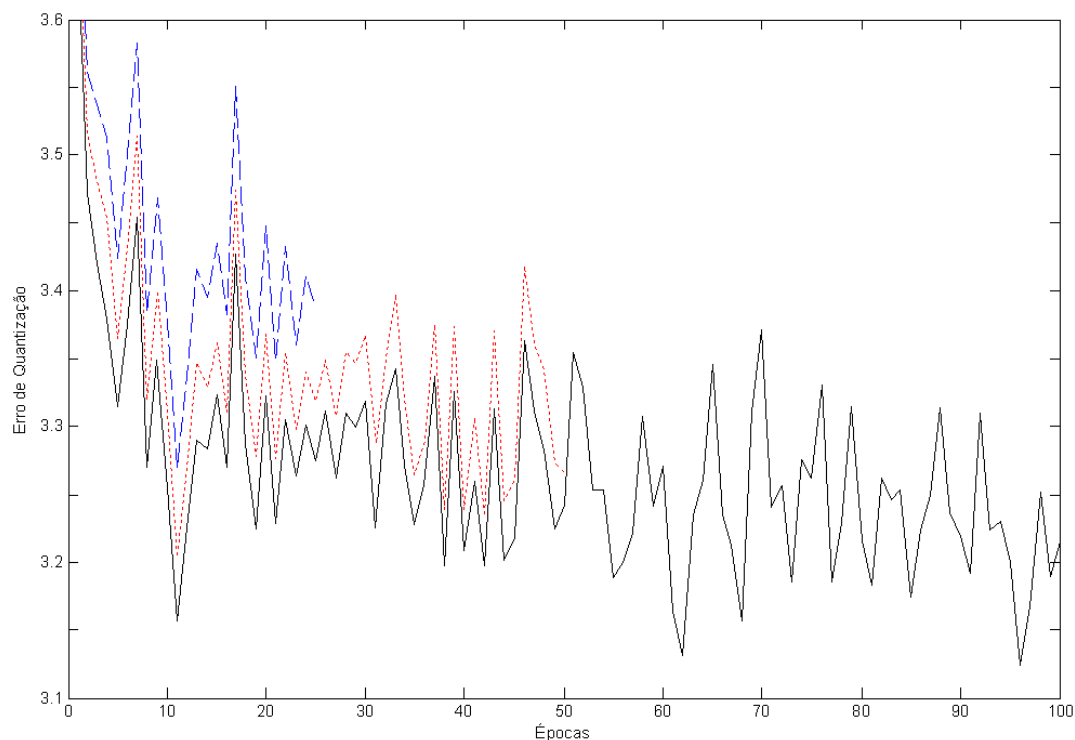


Figura 55 - Erro de quantização médio em cada época para 3 durações (25, 50 e 100 épocas) do treinamento sequencial (elaboração própria)

Duração (Épocas)	Duração (Iterações)	Erro de Quantização	Erro Topográfico
25	8.925	3,3816	0,0252
50	17.850	3,2977	0,0280
100	35.700	3,2532	0,0112

Tabela 19 - Medidas de qualidade de acordo com a duração do treinamento sequencial

Para verificar a influência da duração do treinamento em lote no erro de quantização, a configuração do SOM\_MAKE foi repetida mudando a inicialização para pesos aleatórios (rotina RANDINIT) com raio inicial da primeira fase aumentado para 4. Nos mapas gerados o erro de quantização médio na segunda fase estabilizou a partir da 6ª época na inicialização LININIT e a partir da 8ª época na inicialização RANDINIT (Figura 56), com as seguintes medidas de qualidade:

Inicialização	Erro de Quantização	Erro Topográfico
LININIT	3,2780	0,0140
RANDINIT	3,3005	0,0196

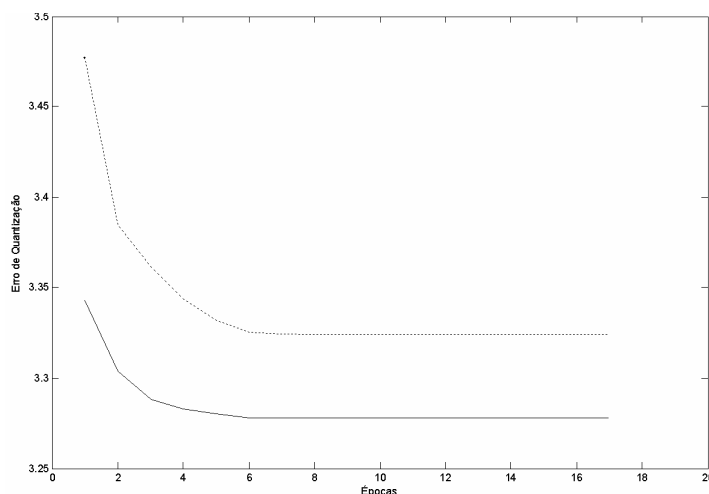


Figura 56 - Erro de quantização médio em cada época para o treinamento em lote com diferentes inicializações: LININIT e RANDINIT (tracejado) (elaboração própria)

Foi testada inicialmente a forma sequencial de treinamento com as duas possibilidades de inicialização (rotinas RANDINIT e LININIT). Na fase de ordenamento são necessárias aproximadamente 3 épocas para fazer as 1.000 iterações com 357 dados e adotou-se, também, a fórmula proposta no SOM Toolbox de 10 vezes a razão entre número de unidades e quantidade de dados, resultando em 5 épocas (1.785 iterações). Para a fase de ordenamento com inicialização aleatória foram estabelecidos raios de vizinhança iniciais de 7,5 e 5,0, respectivamente 1/2 e 1/3 do maior lado, e raio de vizinhança final igual a 2. A taxa de aprendizado inicial foi variada entre 0,6 e 0,9. Com a inicialização da rotina LININIT os parâmetros adotados foram raios de vizinhança iniciais de 4 e 3, e raio de vizinhança final igual a 2; taxa de aprendizado inicial foi variada entre 0,3 e 0,5. A fase de convergência foi fixada com a duração de 210 épocas, com taxa de aprendizado inicial igual a 0,1; o raio inicial definido em 2; e o raio final em 1. A Tabela 20 lista as várias configurações testadas inclusive com as formas linear e inversa para decréscimo da taxa de aprendizado.

De modo a permitir a reprodução dessas configurações, foram fixadas sementes para a função *rand* de geração de números aleatórios do MATLAB da seguinte forma: RANDINIT (*state* = 1); fase de ordenamento (*state* = 2); fase de convergência (*state* = 3).

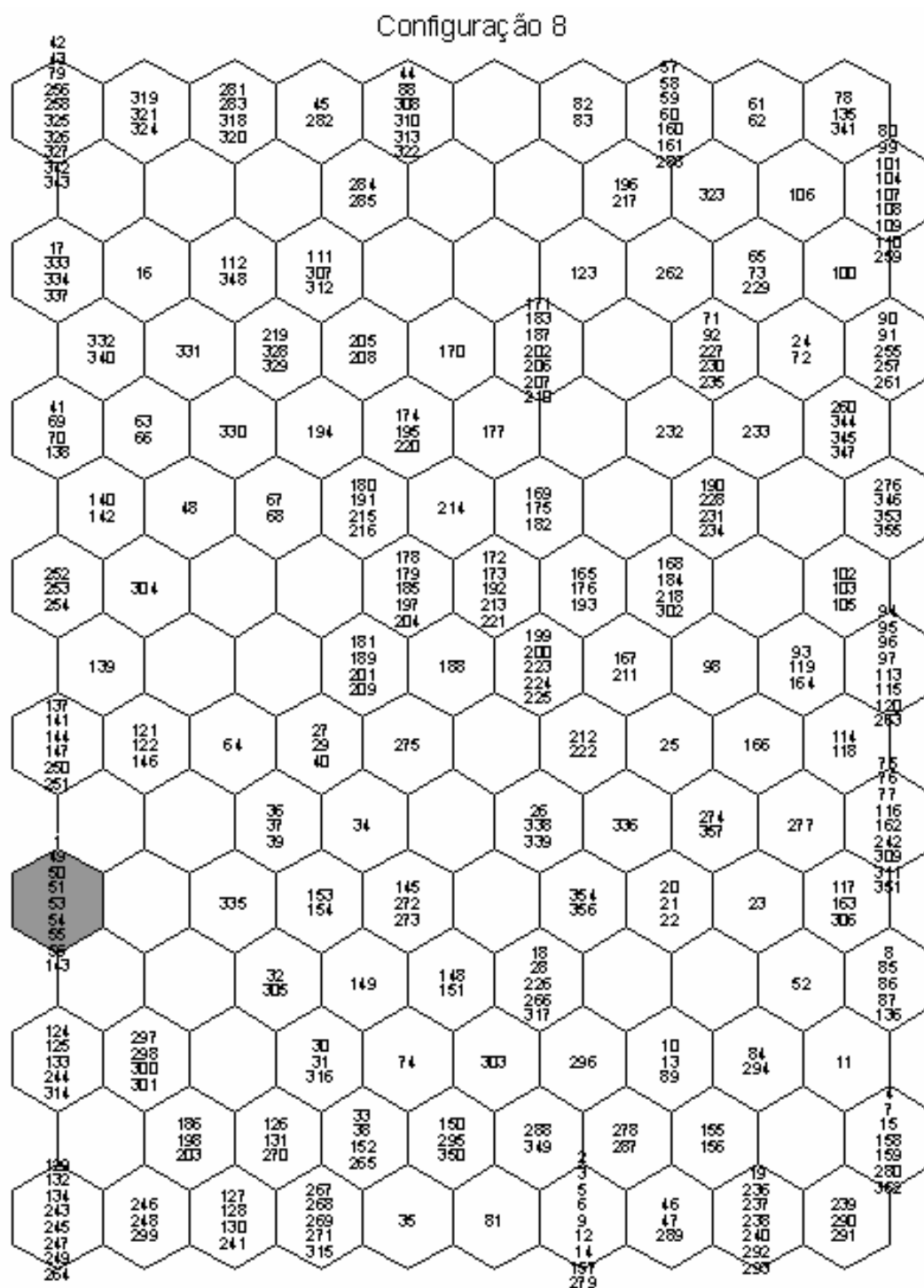
Configuração	$\rho(0)$	Decréscimo da taxa de aprendizado	$d_i$	$d_f$	Duração (Épocas)	Erro de Quantização	Erro Topográfico
Inicialização: RANDINIT							
1	0,9	inversa	7,5	2	3	3,4651	0,0056
2	0,8	inversa	7,5	2	3	3,4616	0,0084
3	0,7	inversa	7,5	2	3	3,4625	0,0028
4	0,6	inversa	7,5	2	3	3,4704	0,0112
5	0,9	linear	7,5	2	3	3,2882	0,0084
6	0,8	linear	7,5	2	3	3,2662	0,0028
7	0,7	linear	7,5	2	3	3,2758	0,0056
8	0,6	linear	7,5	2	3	3,2751	0
9	0,9	inversa	5	2	3	3,4457	0,0056
10	0,8	inversa	5	2	3	3,4615	0,0056
11	0,7	inversa	5	2	3	3,4348	0,0056
12	0,6	inversa	5	2	3	3,4453	0,0028
13	0,9	linear	5	2	3	3,2580	0,0028
14	0,8	linear	5	2	3	3,2917	0,0056
15	0,7	linear	5	2	3	3,2798	0
16	0,6	linear	5	2	3	3,2882	0,0084
Inicialização: LININIT							
17	0,5	inversa	4	2	3	3,4610	0,0056
18	0,4	inversa	4	2	3	3,4552	0,0140
19	0,3	inversa	4	2	3	3,4548	0,0084
20	0,5	linear	4	2	3	3,2729	0
21	0,4	linear	4	2	3	3,2882	0,0084
22	0,3	linear	4	2	3	3,2801	0
23	0,5	inversa	3	2	3	3,4666	0,0028
24	0,4	inversa	3	2	3	3,4793	0,0084
25	0,3	inversa	3	2	3	3,4606	0,0084
26	0,5	linear	3	2	3	3,2826	0,0028
27	0,4	linear	3	2	3	3,2882	0,0084
28	0,3	linear	3	2	3	3,2751	0
Inicialização: RANDINIT							
29	0,9	linear	7,5	2	5	3,2801	0
30	0,8	linear	7,5	2	5	3,2809	0
31	0,7	linear	7,5	2	5	3,2567	0,0028
32	0,6	linear	7,5	2	5	3,2567	0,0028
33	0,9	linear	5	2	5	3,2801	0
34	0,8	linear	5	2	5	3,2809	0
35	0,7	linear	5	2	5	3,2796	0
36	0,6	linear	5	2	5	3,2653	0,0028
Inicialização: LININIT							
37	0,5	linear	4	2	5	3,2801	0
38	0,4	linear	4	2	5	3,2653	0,0028
39	0,3	linear	4	2	5	3,2882	0,0084
40	0,5	linear	3	2	5	3,2801	0
41	0,4	linear	3	2	5	3,2801	0
42	0,3	linear	3	2	5	3,2653	0,0028

Tabela 20 - Lista das configurações testadas para a fase de convergência com o treinamento seqüencial

Com a forma linear a redução da taxa de aprendizado se dá de forma mais lenta e resultou em menores erros de quantização, fator pelo qual a maior duração para a fase de ordenamento só foi testada com essa função. A escolha de raios menores na fase de ordenamento – 5 (RANDINIT) e 3 (LININIT) – produziu, na maioria dos casos, mapas com menores erros.

Dependendo da configuração escolhida, a distribuição dos portos no mapa gerado pode ser bastante diferente. Foram escolhidas para análise algumas configurações cujos resultados tiveram erro topográfico igual a zero. O mapa gerado com a configuração 8 (idêntico ao da configuração 28) foi muito semelhante ao da configuração 37 (idêntico ao da configuração 40). No caso de Sepetiba apenas dois portos se localizaram em vizinhanças ligeiramente diferentes. O porto de Lagos (250-NGLOS) que na configuração 8 (Figura 57) estava numa distância igual a 2, na configuração 37 ficou numa distância igual a 3. O reverso aconteceu com o porto Laem Chabang (305-THLCH) que na configuração 8 estava numa distância igual a 3 e ficou numa distância igual a 2 na configuração 37 (Figura 58).

Chama atenção que nesses mapas os neurônios situados com a distância igual a 1 em torno de Sepetiba não contém nenhum porto. Em ambas as situações (configurações 8/28 e 37/40) uma indicação de os mapas serem idênticos foi terem os mesmo erros de quantização.

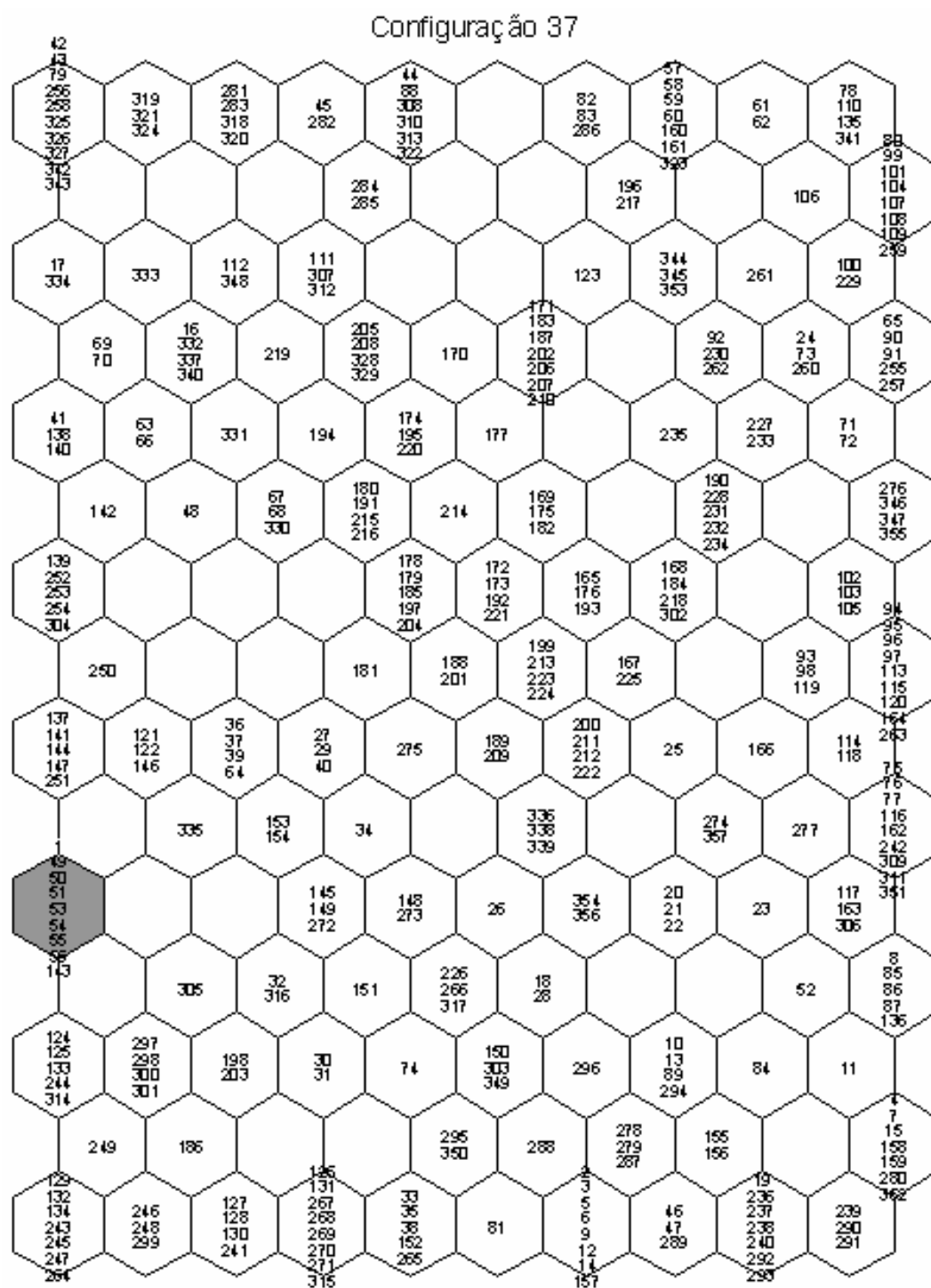


$d = 0$ : 49; 50; 51; 53; 54; 55; 56; 143

$d = 1$ : nenhum

$d = 2$ : 121; 122; 124; 125; 133; 137; 141; 144; 146; 147; 244; 250; 251; 297; 298; 300; 301; 314; 335

Figura 57 - Treinamento sequencial (Configuração 8) (elaboração própria)

 $d = 0: 49; 50; 51; 53; 54; 55; 56; 143$ 

$d = 1$ : nenhum

$d = 2$ : 121; 122; 124; 125; 133; 137; 141; 144; 146; 147; 244; 251; 297; 298; 300; 301; 305;  
314; 335

Figura 58 - Treinamento seqüencial (Configuração 37) (elaboração própria)

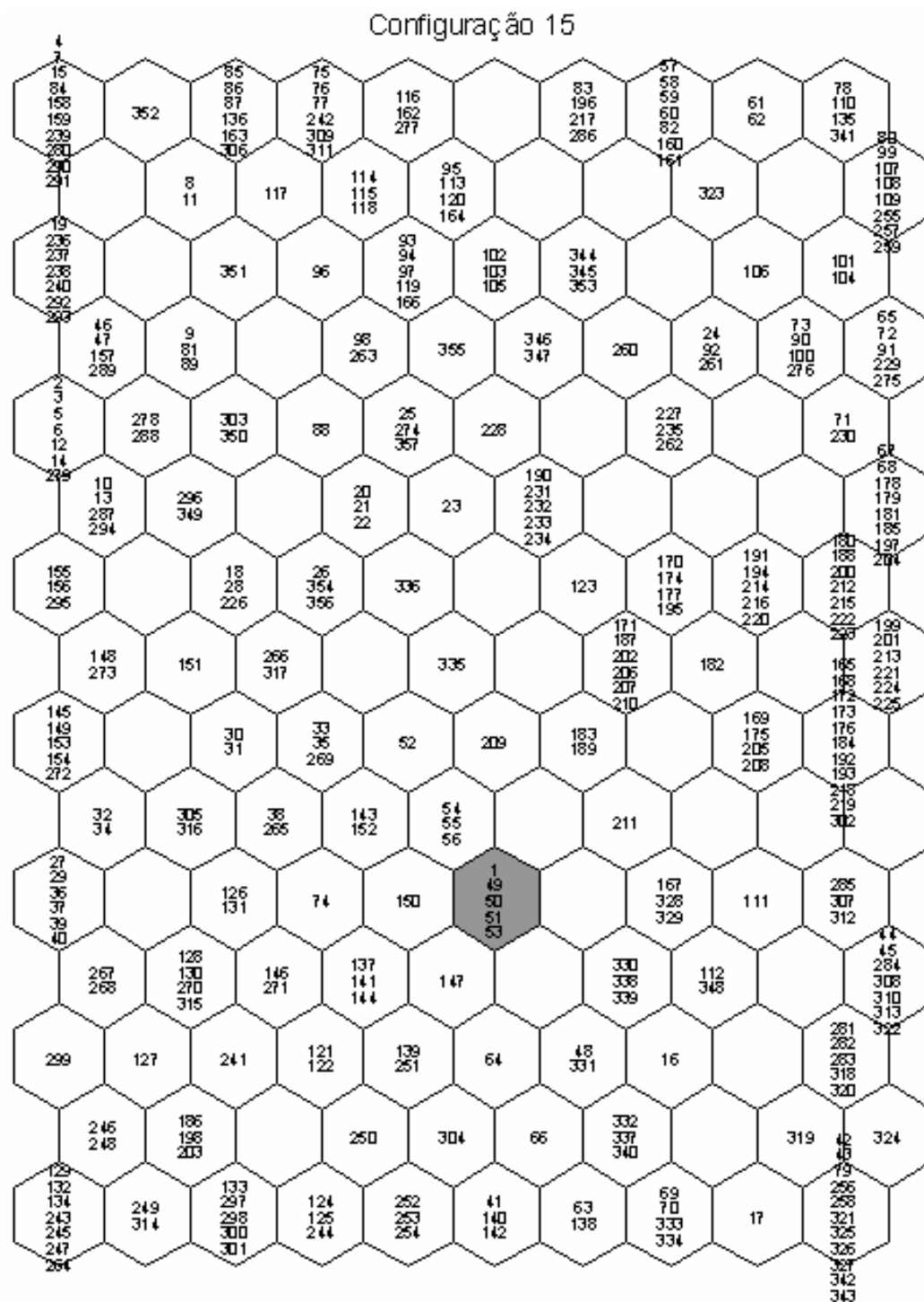
A forma como se dá a localização dos vetores de entrada no mapa varia muito em função dos parâmetros de treinamento. Um exemplo disso pode ser visto com o aumento da duração da fase de ordenamento de 3 para 5 épocas nas configurações 15 (Figura 59) e 35 (Figura 60). Houve um decréscimo mínimo do erro de quantização (de 3,2798 para 3,2796), no entanto, o posicionamento dos portos modificou consideravelmente. O porto de Sepetiba mudou do neurônio 86 para 127 e sua vizinhança na distância igual a 2 foi bastante alterada.

Uma segunda bateria de simulações foi realizada, dessa vez com a utilização do treinamento em lote, tomando como base a configuração dada pelo comando SOM\_MAKE para 150 neurônios. Buscou-se variar os raios iniciais e finais na fase de ordenamento e, eventualmente, sua duração. Diferentemente dos resultados alcançados com o treinamento seqüencial, não se obteve nenhuma configuração com erro topográfico zero (Tabela 21). Para verificar onde esses erros estavam localizados, após cada configuração testada, foram identificados os portos cujos primeiro e segundo neurônios mais bem ajustados não eram adjacentes no mapa. A partir da lista resultante (Tabela 22), observa-se que os portos Chiwan Guangdong (64-CNCWN), China, e Lumut (244-MYLUM), Malásia, contribuíram para o erro topográfico em várias configurações.

Configuração	Ordenamento			Convergência			Erro de Quantização	Erro Topográfico
	$d_i$	$d_f$	Duração (Épocas)	$d_i$	$d_f$	Duração (Épocas)		
SOM_MAKE	2	1	5	1	1	17	3,2780	0,0140
43	3	1	5	1	1	17	3,2741	0,0056
44	4	1	5	1	1	17	3,2895	0,0280
45	3	2	5	2	1	17	3,2724	0,0168
46	3	1,5	5	1,5	1	17	3,2718	0,0168
47	2,5	1	5	1	1	17	3,2834	0,0168
48	3	1	6	1	1	17	3,2714	0,0196
49	2,9	1	4	1	1	17	3,2572	0,0112
50	2,9	1	5	1	1	17	3,2698	0,0056
51	2,8	1	5	1	1	17	3,2703	0,0112

Tabela 21 - Lista das configurações testadas com o treinamento em lote



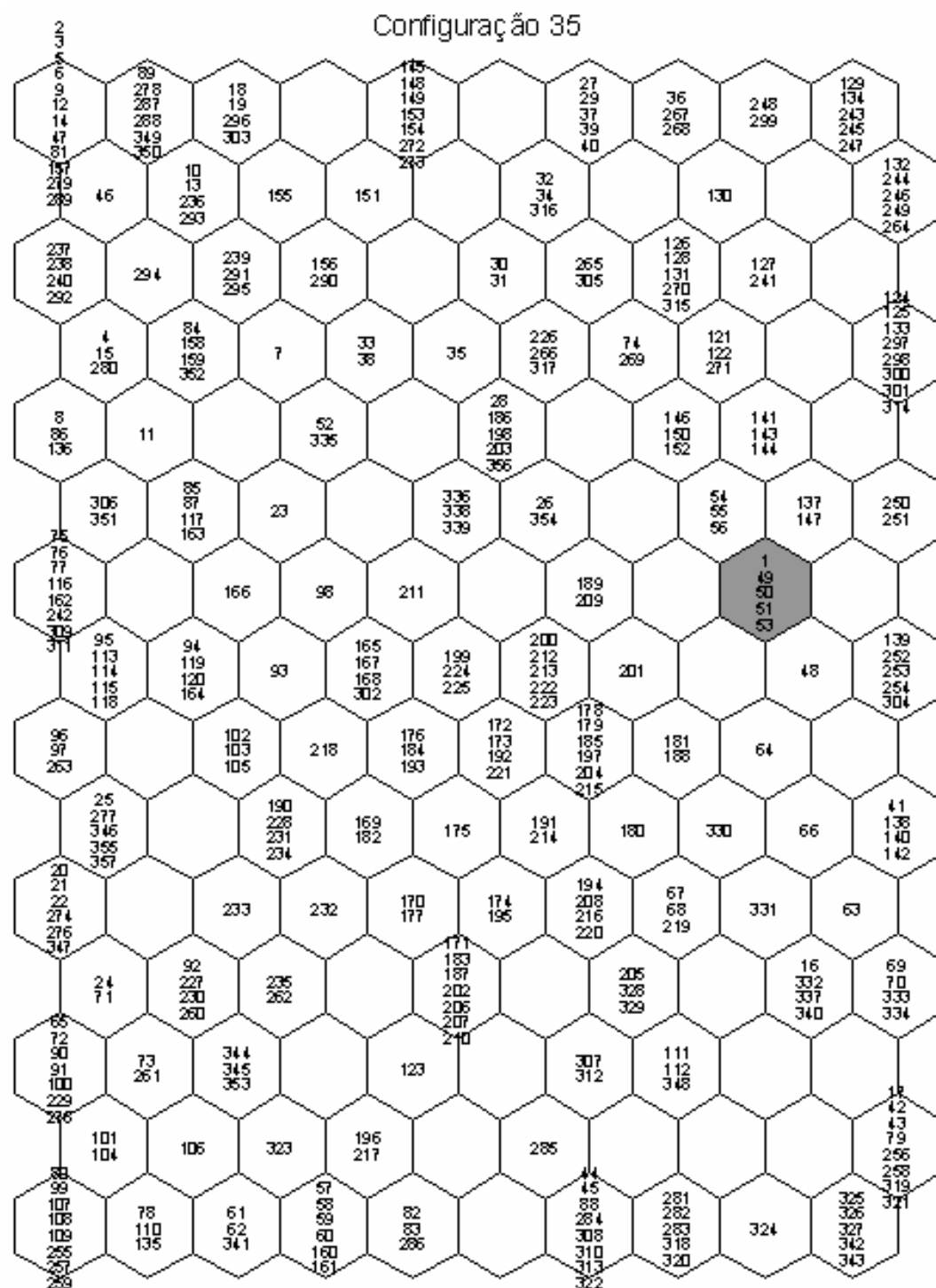


$d = 0$ : 49; 50; 51; 53

$d = 1$ : 54; 55; 56; 147; 150

$d = 2$ : 48; 52; 64; 74; 137; 139; 141; 143; 144; 152; 167; 183; 189; 209; 211; 251; 328; 329; 330; 331; 338; 339

Figura 59 - Treinamento sequencial (Configuração 15) (elaboração própria)



$d = 0$ : 49; 50; 51; 53

$d = 1$ : 48; 54; 55; 56; 137; 147

$d = 2$ : 64; 139; 141; 143; 144; 146; 150; 152; 181; 188; 189; 201; 209; 250; 251; 252; 253; 254; 304

Figura 60 - Treinamento sequencial (Configuração 35) (elaboração própria)

Configuração	Portos com erro topológico
SOM_MAKE	25, 165, 249, 294, 344
43	244, 354
44	58, 59, 60, 165, 167, 169, 263, 328, 336, 338
45	64, 138, 241, 351, 353, 354
46	84, 138, 241, 249, 329, 344
47	64, 138, 226, 254, 324
48	64, 129, 199, 244, 249, 294
49	244, 266, 275, 329
50	100, 244
51	64, 129, 226, 244

Tabela 22 - Portos com erro topológico nas configurações de treinamento em lote

Ao decidir pelo melhor mapa não se deve ficar restrito apenas a uma combinação arbitrária dos valores das medidas de qualidade (Vesanto *et al.*, 2003) e o mapa escolhido para a representação dos portos foi o da configuração 20 apesar de que outros mapas tiveram erro de quantização menor como, por exemplo, os gerados pelas configurações 31 e 32 (3,2567); 49 (3,2572); 13 (3,2580). Entre os mapas com erro topográfico zero, a configuração 20 obteve o menor erro de quantização (3,2729), porém, o critério de escolha desta configuração foi o fato de o porto de Sepetiba (1-BRSPB) estar localizado num neurônio central do mapa, onde as relações de vizinhança podem ser melhor observadas (Figura 61). Em geral, os vetores de pesos são comprimidos próximos da borda do mapa sofrendo um “efeito de fronteira” e causando uma diminuição na capacidade de representar a distribuição do espaço dos dados de entrada (Haykin, 1994). Para contornar esse efeito, existe o conceito do mapa sem fronteiras que é obtido com o uso de uma topologia em forma de toróide, raramente usada (Ultsch & Mörchén, 2005). Entretanto, essa topologia não foi aqui considerada.

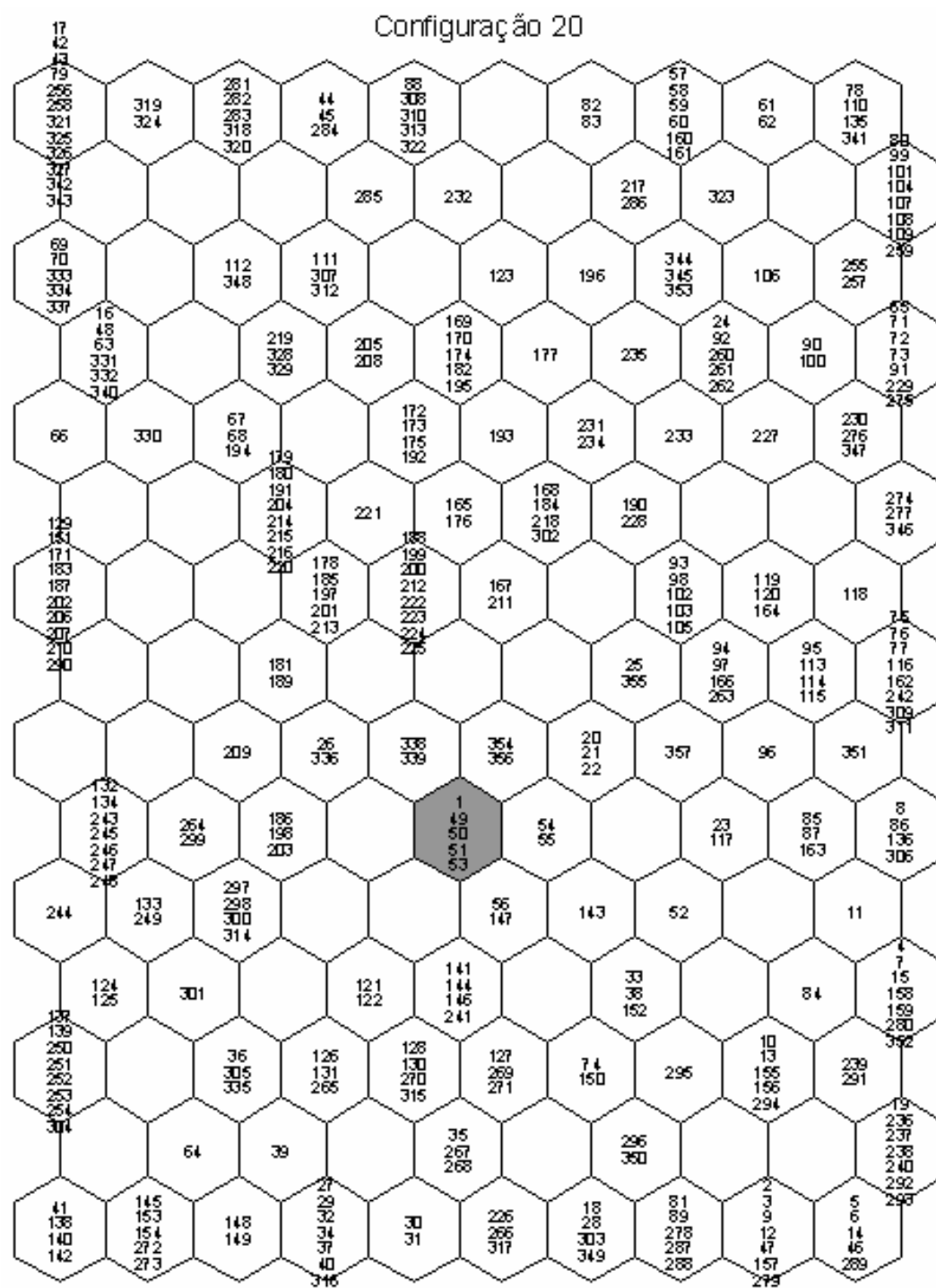


Figura 61 - Treinamento sequencial (Configuração 20) (elaboração própria)

#### 7.4.1 Agrupamento do SOM

Algumas combinações de metodologias podem ser especialmente úteis como a utilização de um método hierárquico de agrupamentos para estimar o número de grupos  $k$  para o método k-médias, como foi realizado no item 5.1.2. Num outro exemplo pode-se reduzir o número de dados por meio do SOM e, então, fazer uma análise de agrupamento dos pesos do mapa gerado. Assim, num primeiro estágio, os vetores de peso de cada neurônio são calculados e a saída do SOM torna-se entrada de um segundo estágio de aplicação do método k-médias. Dessa forma, empregando o k-médias, numa solução com 8 grupos, obteve-se a partição mostrada na Figura 62. Comparando o grupo no qual o porto de Sepetiba está inserido (que se denominará SOM-K8-1) com o grupo anteriormente formado com o método k-médias aplicado diretamente aos dados do banco de dados GloBallast (K8-VI) verifica-se que 81 portos foram coincidentes em ambos grupos. A Tabela 23 lista os nove portos do grupo K8-VI que não foram inseridos no grupo SOM-K8-1. Apesar de esses portos terem sido designados para outros três grupos, todos se localizaram em neurônios adjacentes ao grupo SOM-K8-1, sendo que seis deles ocuparam neurônios na vizinhança imediata do neurônio de Sepetiba, conforme detalhe do mapa visto na Figura 63. No grupo SOM-K8-1, dos sete portos que não pertencem ao grupo K8-VI (relacionados na Tabela 24) apenas o porto de Lucinda (35-AULUC) não coincidiu com os portos atípicos ou extremos dentro dos grupos formados pelo método k-médias listados na Tabela 11.

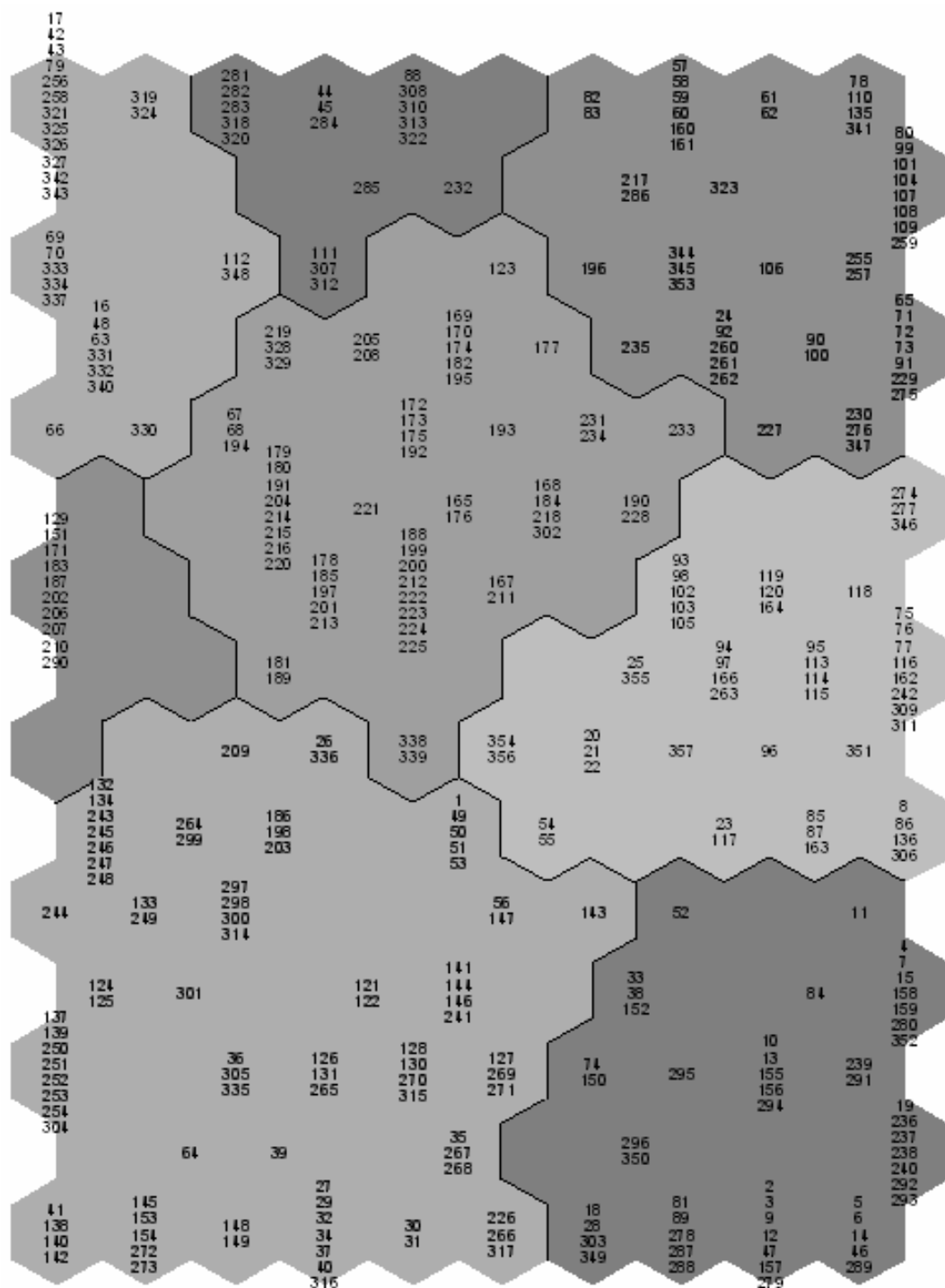


Figura 62 - Agrupamento dos neurônios do mapa da configuração 20 em oito grupos pelo método k-médias (elaboração própria)

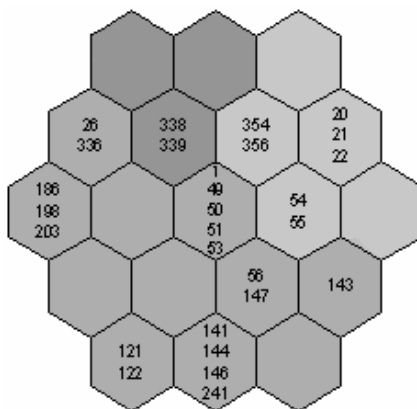


Figura 63 - Vizinhança em torno do neurônio no qual está localizado o porto de Sepetiba (elaboração própria)

Nº	Código	Porto	País	Latitude	Longitude
28	AUGLT	Gladstone	Austrália	23° 51.0S	151° 15.0E
54	BRPRM	Praia Mole	Brasil	20° 17.3S	40° 14.5W
55	BRTUB	Tubarão	Brasil	20° 17.3S	40° 14.6W
74	COCAR	Cartagena	Colômbia	10° 21.6N	75° 32.9W
152	INVTZ	Visakhapatnam	Índia	17° 14.0N	83° 18.0E
338	USGLS	Galveston Texas	EUA	29° 17.0N	94° 50.0W
339	USTXT	Texas City	EUA	29° 23.0N	94° 54.0W
354	ZADUR	Durban	África do Sul	29° 53.0S	31° 02.0E
356	ZARCB	Richards Bay	África do Sul	28° 48.0S	32° 03.0E

Tabela 23 - Portos do grupo K8-VI que não foram inseridos no grupo SOM-K8-1

Nº	Código	Porto	País	Latitude	Longitude
35	AULUC	Lucinda	Austrália	18° 31.0S	146° 19.0E
41	BDCGP	Chittagong	Bangladesh	22° 13.0N	91° 48.0E
138	INCCU	Calcutá	Índia	22° 33.0N	88° 19.0E
142	INIXY	Kandla	Índia	22° 52.0N	70° 13.0E
153	IRBKM	Bandar Imam Khomeyni	Irã	30° 25.0N	49° 04.0E
154	IRBMR	Bandar Mushar	Irã	30° 28.0N	49° 11.0E
252	NGPHC	Port Harcourt	Nigéria	4° 46.2N	7° 00.0E

Tabela 24 - Portos do grupo SOM-K8-1 não pertencentes ao grupo K8-VI

Dois grupos comentados na análise das matrizes de concordância tiveram comportamento semelhante nessa partição escolhida para o mapa. O grupo 13 formado pelo método Ward foi o único com concordância total na solução de 8 grupos dos métodos hierárquico e não-hierárquico (grupos W8-4 e K8-II, ver Tabela 15). Esse conjunto de onze portos ficou dividido na solução em 6 grupos do método k-médias, com dez portos integrando

o grupo K6-I e o porto Huangpu Guangdong (66-CNHUA) situando-se no grupo K6-II. Foi verificado no mapa que os mesmos dez portos se localizaram no neurônio 7, formando um único grupo afastado dos demais neurônios, e o porto Huangpu Guangdong (66-CNHUA) ficou próximo, isolado no neurônio 5, e num outro grupo.

Um outro grupo considerado bastante coeso, pois seus 45 portos coincidiram totalmente nas soluções com 4, 5, 6 e 8 grupos, nos dois métodos de agrupamento, localizaram-se integralmente no canto direito inferior do mapa (ver Figura 64). Apenas o porto de Gladstone (28-AUGLT), localizado no neurônio 105 junto com os portos Dampier (18-AUDAM); Dakar (303-SNDKR); e Aden (349-YEADE), não pertence ao grupo K8-1.

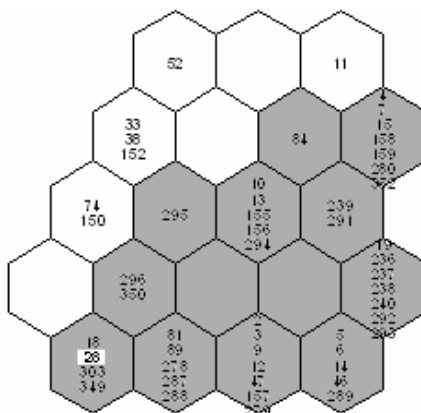


Figura 64 - Detalhe do canto inferior direito do mapa com o grupo K8-1 destacado (elaboração própria)

#### 7.4.2 Avaliação do comportamento dos parâmetros

Não importa o quão bom seja o método, pois os resultados irão essencialmente depender da qualidade e adequabilidade dos dados. As propriedades que emergem da aplicação dos métodos são determinadas pelos parâmetros escolhidos para representar os dados. A análise dos dados deve levar em consideração essa escolha para evitar conclusões enganosas na avaliação dos resultados.

Foi comentado no item 3.1.1 que não há ainda uma definição consensual de quais variáveis ambientais são preditivas para o estabelecimento de espécies. Não foi objetivo desta



tese questionar o conjunto de 34 parâmetros escolhidos como descritores dos ambientes portuários pelo grupo URS Australia Pty para realizar a avaliação de risco para o Programa GloBallast. No entanto, buscou-se verificar a influência desses descritores na formação do mapa. Com o uso de uma ferramenta do SOM Toolbox (SOM\_SHOW) é possível gerar mapas mostrando, por meio de um código de cores, a distribuição de determinado parâmetro, exibindo-se uma determinada dimensão do vetor de peso em cada neurônio.

Nos mapas dos parâmetros de temperatura (água e ar) e salinidade percebe-se a formação de gradientes claramente definidos. Principalmente pelas temperaturas de verão (MSUWT; USUWT; MSART; e USART ver Anexo2) identifica-se que os portos de mais baixa temperatura (e.g., Canadá, Dinamarca, Islândia) se concentraram nos neurônios do canto superior direito (Figura 65). A disposição dos parâmetros de salinidade se deu numa forma diagonal do canto superior esquerdo, com as salinidades mais baixas, para o canto inferior direito, com salinidades mais altas (Figura 66). A partir da visualização do aspecto dos gradientes é possível entender o posicionamento dos portos no mapa, por exemplo, a aglomeração do grupo K8-1 no canto inferior direito do mapa, como visto anteriormente. Por esta localização percebe-se que são portos de altas temperaturas e altas salinidades.

Alguns parâmetros foram decisivos individualmente pelo agrupamento de certos portos. O exemplo mais marcante disso foi o descritor de distância até “cais de paredes verticais” (INASMW), mostrado na Figura 67. Dos 357 portos do banco de dados, dois foram categorizados com o escalar “3” (5-10km), os portos de Cilacap Java (129-IDCXP) e Al Juaymah Terminal (290-SAJUT); nove com o escalar “4” (1-5km); e os demais com “5” (<1km). Esses onze portos com valores diferentes de “5” são exatamente os que compõem o grupo W8-4 ou K8-II observado nas análises anteriores como bastante consistente.

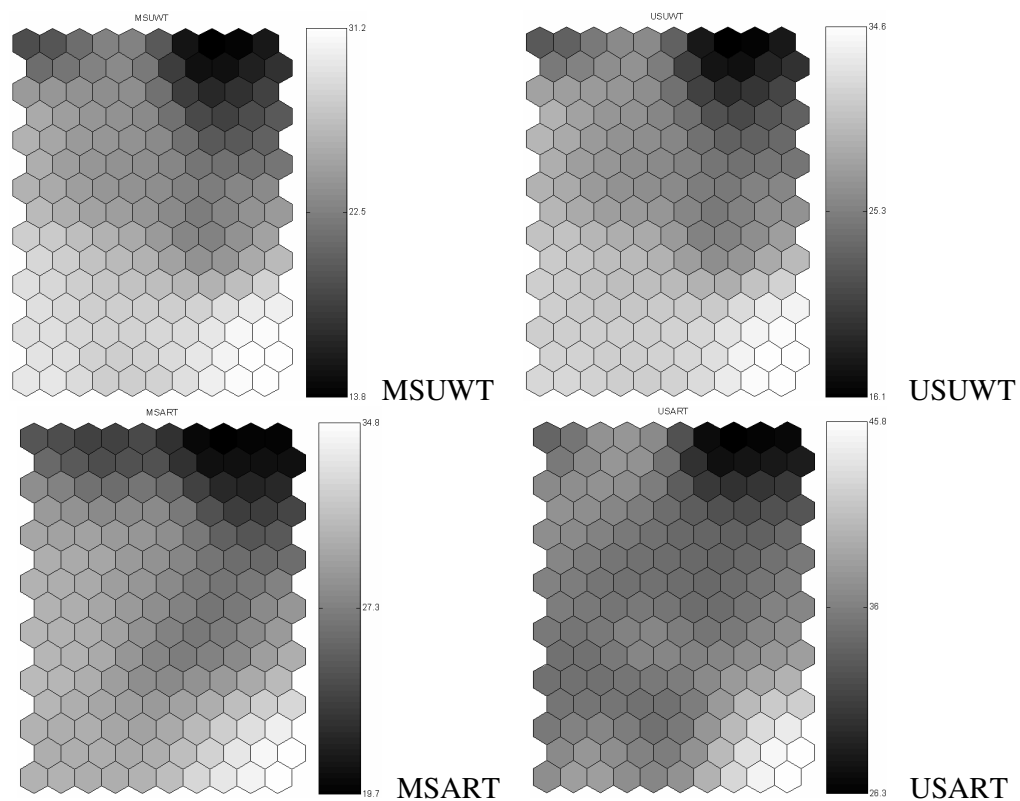


Figura 65 - Distribuição das dimensões de temperatura de verão do ar e da água

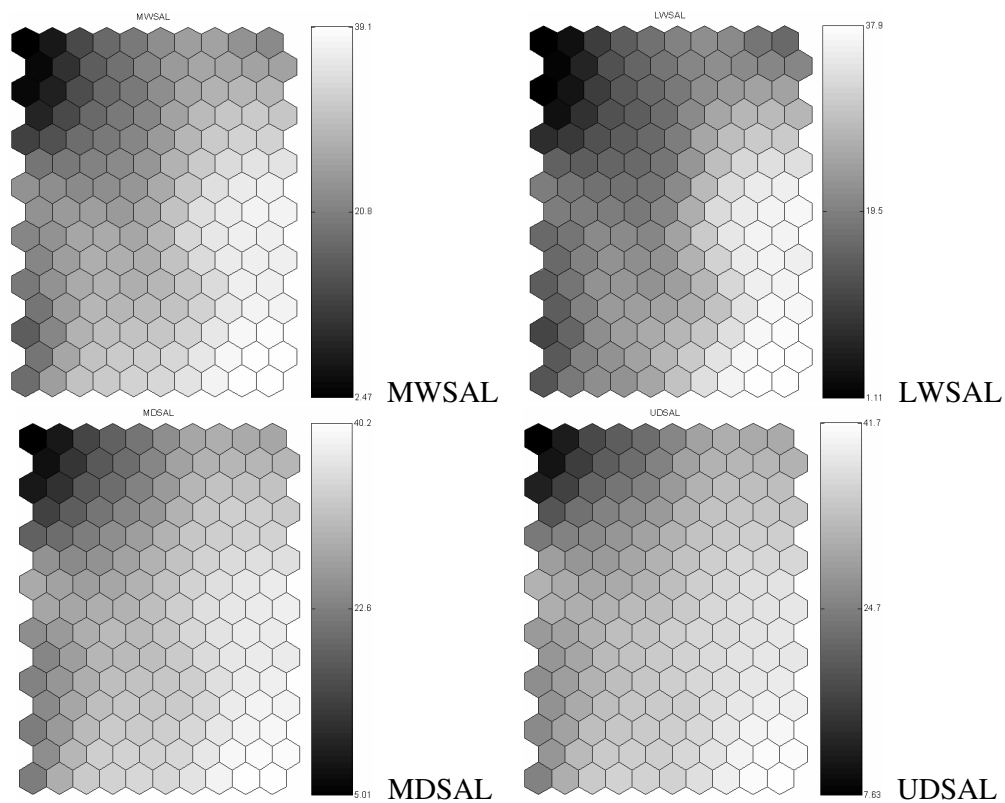


Figura 66 - Distribuição das dimensões de salinidade (elaboração própria)

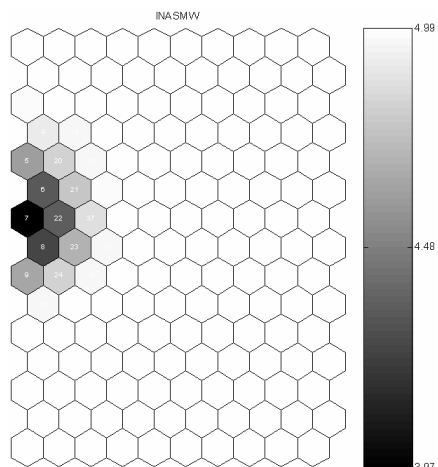


Figura 67 - Distribuição da dimensão 21 correspondente ao descritor INASMW (distância até “cais de paredes verticais”) (elaboração própria)

Outro descritor que contribuiu na aglomeração de portos em vários neurônios foi o “tamanho da bacia hidrográfica” (SIZRVC) (Figura 68). No neurônio 61 foram agrupados os portos Port Said (88-EGPSD) com 2.000.000km<sup>2</sup>; e Istanbul (308-TRIST); Izmit (310-TRIZT); Yarimca (313-TRYAR) e Sevastopol (322-UASVP) com 817.000 km<sup>2</sup>. No neurônio 3 foram reunidos os portos Shanghai (69-CNSHA) e Shanghai Baoshan (70-CNSHB) com 1.500.000km<sup>2</sup>; e Davant (333-USDVT) e New Orleans (334-USMSY) com 1.000.000km<sup>2</sup>. Outro neurônio com a reunião de portos com grandes bacias hidrográficas foi o 15 com Chittagong (41-BDCGP), Calcutá (138-INCCU) e Haldia (140-INHAL) de 1.200.000km<sup>2</sup>.

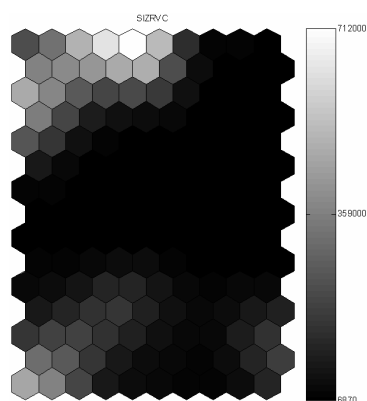


Figura 68 - Distribuição da dimensão 20 correspondente ao descritor SIZRVC (tamanho da bacia hidrográfica) (elaboração própria)

A Figura 69 mostra a distribuição dos outros descritores no mapa.

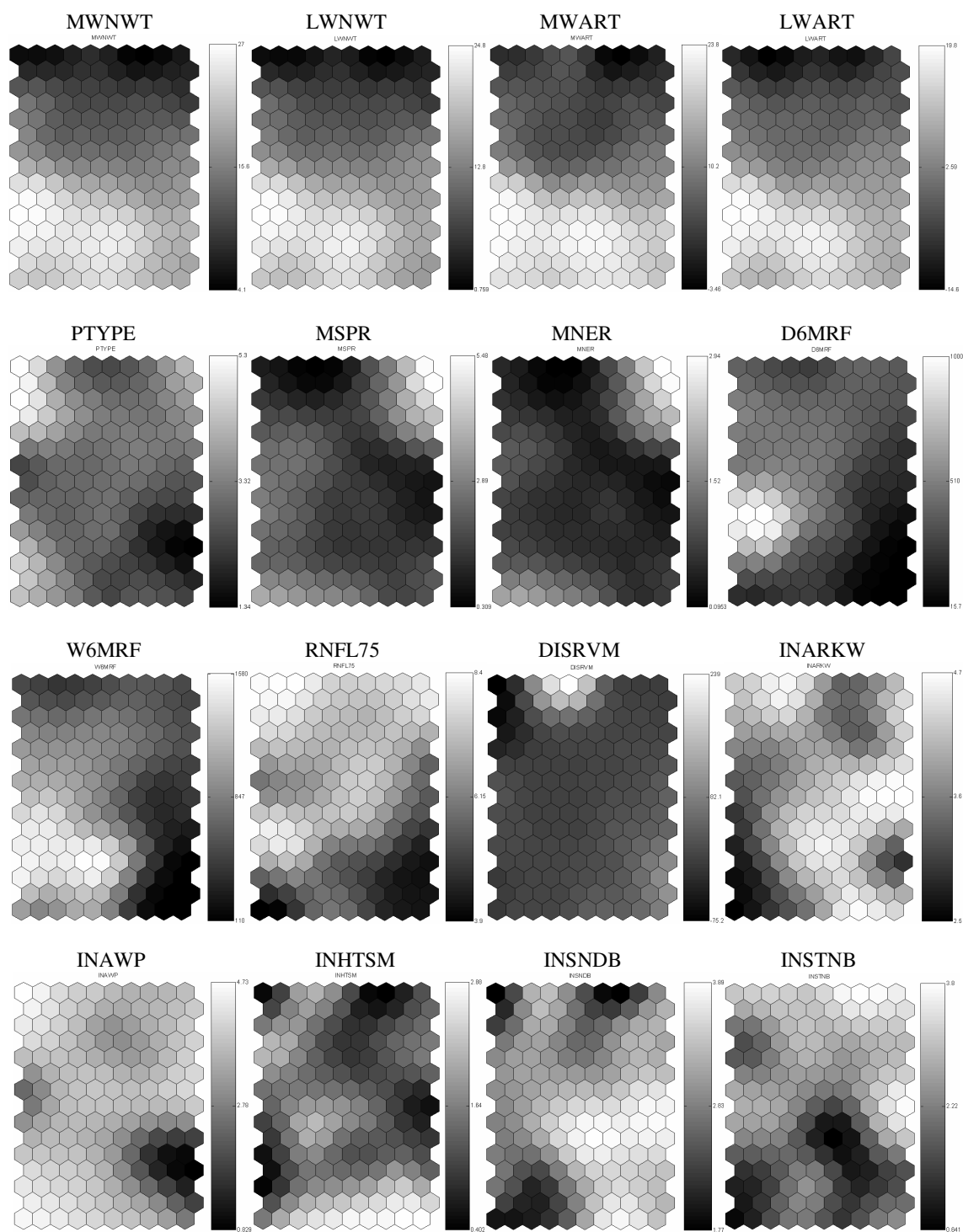


Figura 69 - Distribuição dos demais descritores no mapa (elaboração própria)

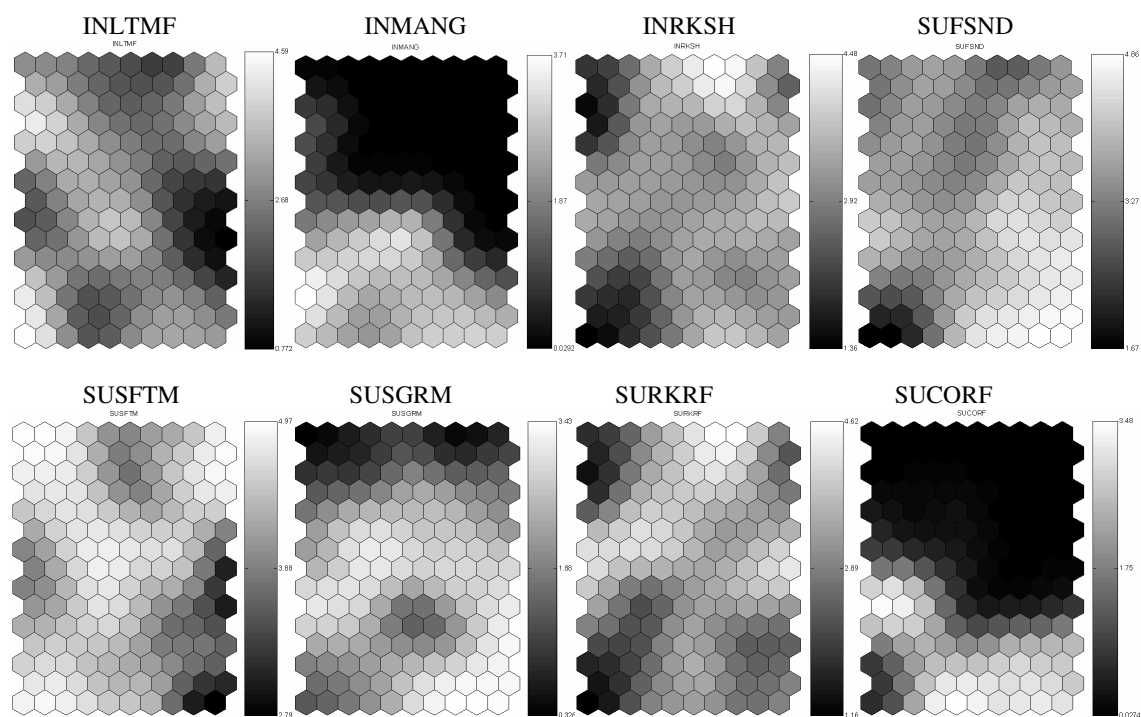


Figura 69 - Distribuição dos demais descritores no mapa (continuação)

## 8 Discussão

As relações entre objetos representados por dados multidimensionais podem ser investigadas por várias metodologias de mineração ou análise exploratória de dados. Comparar a capacidade de diferentes métodos, que se baseiam em diferentes princípios, de reconhecer padrões no espaço original de dados de entrada pode ser uma tarefa difícil, especialmente no caso em que os dados não estejam rotulados antecipadamente ou onde não haja, pelo menos, informação prévia sobre o número de classes existentes. Uma comparação, entretanto pode, eventualmente, ser feita com base na aplicação prática à qual a análise se destina. A seguir são feitas algumas dessas comparações, a partir do exame das propriedades do algoritmo do SOM e das metodologias empregadas na tese.

Embora o método do k-médias e o SOM sejam relacionados, certas diferenças podem ser ressaltadas, conforme as comparações efetuadas por Ultsch (1995) e Kaski (1997). No SOM, a distância entre cada vetor de entrada e todos os vetores de peso é levada em consideração, ponderada pela função de vizinhança, ao invés de apenas o vetor de peso mais próximo. No caso da adoção de vizinhança zero, o SOM funciona como um algoritmo convencional de agrupamento. Enquanto no k-médias o número de grupos  $k$  deve ser escolhido de acordo com o número de grupos existentes nos dados de entrada (ou que se supõe existir), no SOM um número bem maior de neurônios pode ser escolhido, independente do número de grupos esperados. Eventualmente a estrutura de grupos se tornará visível com o auxílio de técnicas especiais (e.g. método U-matrix (Ultsch, 1992 *apud* Ultsch, 1995)). Intrinsecamente o SOM, por meio do uso da função de vizinhança, é menos sensível às condições iniciais, e exibe uma convergência mais robusta do que o k-médias.

Os estudos efetuados por Chon *et al.* (1996) sobre a análise da ocupação espacial de peixes num lago, e por Brosse *et al.* (2001), sobre a classificação de comunidades de macroinvertebrados bentônicos, citam Melssen *et al.* (1993) ao discutir o aspecto da redução

de dimensão da ACP: “um grande número de descritores pode levar a muitas componentes principais significativas na ACP de modo que não se possa reter informação suficiente nos componentes principais usados para visualizar o espaço multidimensional dos dados”. Esse fato foi constatado no caso do banco de dados GloBallast, com 34 descritores, onde as duas primeiras componentes principais não conseguiram explicar, juntas, mais do que 40,7% da variância total.

Os resultados do SOM e do MDS não-métrico podem ser bem distintos, apesar de que a preservação de vizinhanças é o objetivo de ambos. A diferença essencial na projeção formada pelos dois métodos é que o SOM busca formar projeções corretas localmente, enquanto o MDS visa preservar todas as distâncias entre os pontos, mantendo a ordem dos postos (*rank*) dessas distâncias. Assim, no MDS, o cálculo do stress depende apenas da ordem das distâncias (similaridades) e não de suas magnitudes (Fasham, 1977).

No caso geral, não se consegue mapear precisamente um espaço de alta dimensão num de menor dimensão por nenhum método. Assim, como não se podem preservar todas as distâncias, é possível que o método do MDS preserve bem a maior parte dessas distâncias, e para algumas outras a preservação seja apenas aproximada. No SOM a ordem é determinada localmente pela função de vizinhança e a ordem global surge pelas interações entre neurônios, e porque o gradual encolhimento da vizinhança evita configurações desordenadas (Kaski, 1997).

Na análise exploratória dos dados pode-se interpretar que o método MDS forneça grosseiramente uma melhor ordem global dos dados projetados que tende a ser mais precisa, pois é feita num sistema de coordenadas livres enquanto que no SOM, a representação da camada de imagem é de uma treliça fixa de neurônios. Os mesmos pontos situados nas extremidades da projeção do MDS não-métrico, destacados na Figura 35, também se localizaram nas bordas do mapa obtido com o SOM (Figura 70). Porém, o distanciamento

para os demais portos não ficou tão evidente quanto na representação do MDS não-métrico que mostrou ser mais eficiente na indicação de dados atípicos ou extremos.

No SOM as estruturas globais são usualmente bem representadas, mas as locais são consideradas mais importantes e é onde a interpretação do mapa deve ser feita predominantemente, baseada nas relações próximas dos neurônios no mapa. Na comparação das vizinhanças geradas pelo MDS não-métrico e pelo SOM percebem-se mudanças nas relações de proximidade com Sepetiba entre os dois métodos. Dos portos listados na Tabela 25, o porto de Lisboa (275-PTLIS) foi o que se localizou mais distante no mapa gerado pelo SOM (ver Figura 71). Essa distância é coerente, pois os parâmetros de temperatura, principalmente de inverno, contribuem para que Lisboa fique próximo de outros portos de ar e água fria situados na parte superior direita do mapa. Nos demais portos não se identifica nenhum descritor, ou conjunto de descritores responsáveis por essa diferença. No entanto, pressupõe-se que a representação do SOM seja mais realista, pois este método busca acima de tudo garantir que dados projetados em neurônios próximos sejam similares, pelo qual a ordem e o agrupamento local são tão fidedignos quanto possível.

Nº	Código	Porto	País	Latitude	Distância em relação à Sepetiba
27	AUBDB	Bundaberg	Austrália	24° 46.0S	5
29	AUPTL	Port Alma	Austrália	23° 35.0S	5
32	AUMKY	Mackay	Austrália	21° 06.0S	5
181	JPKCZ	Kochi	Japão	33° 31.0N	3
189	JPKOJ	Kagoshima	Japão	31° 35.0N	3
201	JPNGS	Nagasaki	Japão	32° 45.0N	3
209	JPSBS	Shibushi Kagoshima	Japão	31° 28.0N	3
275	PTLIS	Lisboa	Portugal	38° 42.0N	8

Tabela 25 - Portos com distintas relações de proximidade com Sepetiba nos métodos SOM e MDS não-métrico



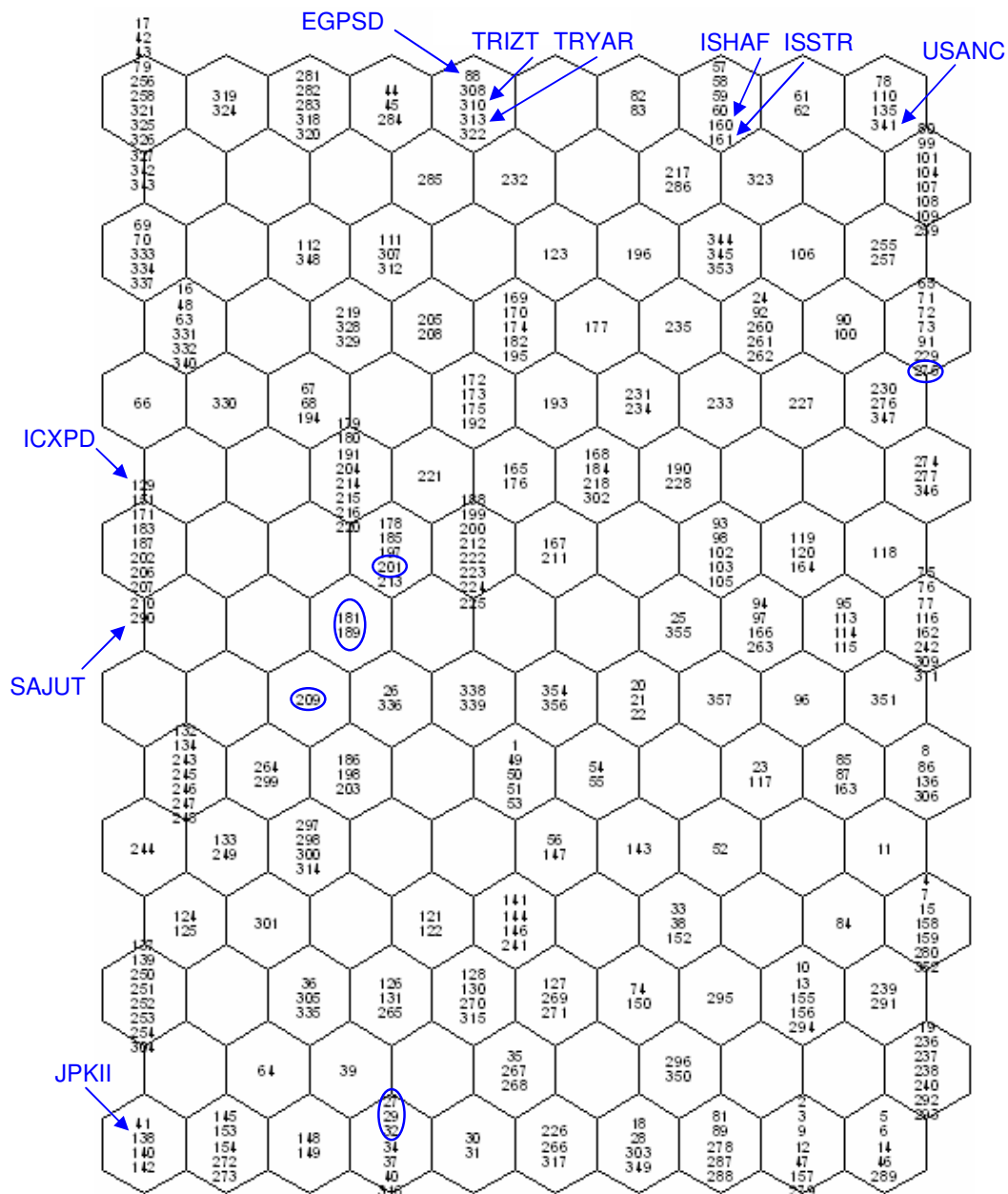


Figura 70 - Localização no mapa dos portos das extremidades da projeção do MDS não-métrico (setas) e portos realçados na Figura 71 (elipses) (elaboração própria)

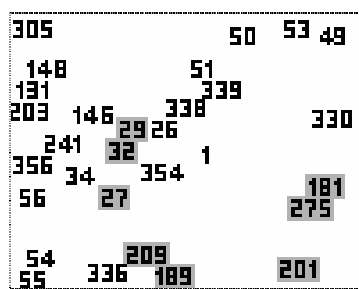


Figura 71 - Reprodução do detalhe da projeção do MDS não-métrico com destaque para portos com distintas relações de proximidade com Sepetiba em relação ao SOM

Com base na Figura 63, que mostra os portos dentro de um raio de dois neurônios em torno do neurônio onde se situou o porto de Sepetiba, os resultados do SOM foram comparados com a ordem obtida pelo coeficiente de similaridade determinado pela medida direta da distância Euclidiana utilizada no sistema de avaliação de risco do Programa GloBallast (dados do Anexo 4). Nessa comparação verificou-se, também, a pertinência desses portos aos grupos gerados com as técnicas de análise de agrupamento (Tabela 26).

No neurônio 70, junto com Sepetiba, localizaram-se portos muito próximos em termos da ordem do coeficiente de similaridade (2°, 5°, 6° e 8°) e no raio de um neurônio estão incluídos outros portos próximos (3°, 4° e 7°) e portos um pouco mais distantes (21°, 47°, 49°, 52° e 75°). Na vizinhança do raio de dois neurônios foram situados portos que, pelo coeficiente de similaridade, poderiam ser considerados todavia mais distantes ainda, com ordem até 202°. Chama a atenção que entre os 36 portos mais próximos de acordo com o coeficiente de similaridade (10% dos 357 portos), apenas 11 estão inseridos na vizinhança de até dois neurônios.

Por outro lado, entre os 28 portos circunscritos num raio de dois neurônios, apenas os portos Port Bonython (20-AUPBY); Whyalla (21-AUWYA); e Port Pirie (22-AUPPI) não pertencem ao grupo K8-VI obtido com o método k-médias. Pela Figura 62 observa-se que apesar de esses três portos estarem dentro dessa vizinhança do porto de Sepetiba, o agrupamento feito com os pesos do mapa indica que estão próximos, ambientalmente, de portos europeus, na Espanha, França, Grécia e Itália.

A apreciação desse resultado, corroborado por distintas metodologias – SOM e k-médias –, aponta que a simples medida direta da distância Euclidiana não se mostrou uma ferramenta eficiente para identificar similaridades. Por exemplo, nenhum porto com ordem do coeficiente de similaridade entre 9° e 20° foi posicionado na vizinhança mais imediata do neurônio 70, e de fato, é de se esperar que portos localizados, por exemplo no Japão não tenham condições ambientais tão similares com Sepetiba que justifiquem esta proximidade.

Distância no SOM	Porto	No.	Código ONU	Ordem de Similaridade GloBallast	Grupo Ward Inicial	Grupo k-médias K8-VI
0	Sepetiba	1	BRSPB	1º	✓	✓
	Paranaguá	49	BRPNG	8º	✓	✓
	Santos	50	BRSSZ	5º	✓	✓
	Rio de Janeiro	51	BRRIO	2º	✓	✓
	Vitória	53	BRVIX	6º	✓	✓
1	Praia Mole	54	BRPRM	3º	✓	✓
	Tubarão	55	BRTUB	4º	✓	✓
	Salvador	56	BRSSA	7º	✓	✓
	Paradeep	147	INPRT	47º	✓	✓
	Galveston Texas	338	USGLS	75º		✓
	Texas City Texas	339	USTXT	49º		✓
	Durban	354	ZADUR	21º		✓
	Richards Bay	356	ZARCB	52º		✓
2	Port Bonython	20	AUPBY	59º		
	Whyalla	21	AUWYA	95º		
	Port Pirie	22	AUPPI	88º		
	Brisbane	26	AUBNE	166º		✓
	Hong Kong Hong Kong	121	HKHKG	196º		✓
	Hong Kong Kowloon	122	HKKWN	202º		✓
	Mangalore (New Mangalore)	141	INIXE	170º	✓	✓
	Chennai (Ex Madras)	143	INMAA	74º	✓	✓
	Marmugao (Marmagoa)	144	INMRM	192º	✓	✓
	Porbandar	146	INPBD	69º	✓	✓
	Kinwan (Ishikawa) Okinawa	186	JPKIN	106º		✓
	Naha Okinawa	198	JPNAH	80º		✓
	Okinawa Okinawa	203	JPOKA	81º		✓
	Colombo	241	LKCMB	27º		✓
	Sabine	336	USSAB	36º		✓

Tabela 26 - Comparação do resultado obtido com o SOM com a distância Euclidiana e os métodos de agrupamento.

## 9 Conclusão

Diferentemente de algumas formas de poluição geradas por navios, o problema da transferência de espécies exóticas e/ou patogênicas decorre de uma atividade inerente à sua própria operação.

Empresas e pesquisadores estão buscando tecnologias que forneçam alternativas à troca de água de lastro em águas profundas, em mar aberto. O avanço de tais tecnologias, e seu eventual uso de forma a atender os requisitos da Convenção BWM 2004, enfrenta muitos desafios, principalmente o de conseguir desenvolver sistemas de tratamento em grande escala que possam ser acomodados nos limitados espaços das praças de máquinas dos navios. Como resultado, a troca em alto mar — único método aprovado para gestão da água de lastro a bordo — continuará a ser empregada por um longo período, apesar de todas as preocupações com sua efetividade. De fato, a Convenção BWM 2004 estabelece que a troca oceânica de lastro poderá ser utilizada até 2016 pelos navios construídos antes de 2009 e com capacidade de água de lastro menor que 1.500 ou maior que 5.000 metros cúbicos.

Em qualquer das formas de gestão a bordo, troca ou tratamento, o risco da introdução de organismos prejudiciais e espécies patogênicas aquáticas por meio da água de lastro dos navios não deixa de existir. Um sistema de avaliação de risco deve ser posto em prática por uma razão básica, qual seja, melhorar a gestão da água de lastro por parte das autoridades do controle do Estado do Porto, pois a vantagem de tal sistema é apontar para uma fiscalização mais profunda nos navios que representam maior risco. Uma vez que se decida por uma abordagem seletiva, será importante assegurar: um modo organizado de definir os alvos de interesse, os navios com risco potencial; e as ações necessárias quanto à descarga da água de lastro desses navios.

Toda avaliação de risco envolve incertezas. No caso da água de lastro, prever com confiança se uma espécie irá sobreviver à viagem e, posteriormente, em um novo ambiente ou

ainda, se esta espécie tornar-se-á invasora, é muito difícil, se não impossível. Dessa forma, quantificar o grau de similaridade entre os ambientes é um ponto de partida importante na avaliação de risco de introdução de espécies invasoras, juntamente com o conhecimento dos padrões de navegação (cabotagem/longo curso) e características da movimentação de carga (exportação/importação) existente no porto.

É importante mencionar que os resultados da avaliação de risco devem ser validados por meio de, pelo menos, duas ações: o levantamento das espécies exóticas presentes no porto em estudo, bem como a sua origem; e a amostragem da água de lastro dos navios que escalam naquele porto. Essa amostragem deve estar associada a métodos que permitam a identificação das espécies presentes na água que está sendo descarregada, com o uso, por exemplo, de ferramentas genéticas.

Não há ainda uma definição consensual de quais variáveis ambientais são preditivas para o estabelecimento de espécies aquáticas e, possivelmente, essas variáveis podem ser diferentes de acordo com o tipo de espécie. Além disso, muitos dados normalmente requeridos para descrever os ambientes aquáticos não estão disponíveis para certas regiões.

Na avaliação de risco conduzida pelo Programa GloBallast, foram eleitos 34 parâmetros como descritores dos ambientes portuários. O levantamento desses parâmetros envolveu muitas equipes e consulta a uma grande variedade de fontes de informações, alcançando reunir dados para 357 portos ao redor do mundo. Esse processo extensivo criou um ambiente de dados com diferentes graus de qualidade e confiabilidade.

Para um determinado conjunto de dados podem existir diversas representações que são úteis para propósitos distintos. Como diferentes métodos apresentam resultados que correspondem a diversas visões da realidade que se quer representar, a abordagem mais adequada seja, talvez, utilizar os vários métodos de forma complementar. Com esse objetivo

em mente, a tese procurou abordar o problema por intermédio de metodologias abrangendo desde a análise multivariada de dados até as redes neuronais.

Na análise de agrupamento o conjunto de dados foi explorado com o método hierárquico, onde se utilizou a técnica de ligação de Ward, e não-hierárquico, com o método k-médias. Quando comparados entre si, esses métodos geraram resultados similares, como foi observado por meio das matrizes de concordância. A flexibilidade de realocar os portos na busca da melhor solução com determinado número  $k$  de grupos — ao invés de simples combinações como no método hierárquico —, indica que método k-médias pode gerar grupos com melhor interpretação.

Como as categorias não eram previamente determinadas e nem todos os valores de  $k$  apresentavam, *a priori*, grupos plausíveis, aplicou-se o método k-médias para diferentes números de grupos, ou seja, valores de  $k$ . A construção das matrizes de concordância serviu para escolher, dentre os diversos resultados, a solução com 8 grupos como a melhor.

Os resultados da ACP não foram satisfatórios para a visualização da relação dos portos em termos de sua proximidade no espaço bidimensional. O resultado mais interessante da ACP está na interpretação das projeções sobre os novos eixos construídos pelo método (componentes principais). No entanto, as componentes obtidas não retiveram informação suficiente em relação à variância total, tornando difícil sua interpretação.

Por outro lado, a ACP mostrou ser um bom método de aproximação. Seus resultados são úteis como configuração inicial para o MDS não-métrico e para o SOM. No caso do MDS não-métrico, as configurações iniciais foram testadas somente com posições aleatórias, pois o software usado não dispunha da opção de utilizar o resultado da ACP para essas posições. No SOM, as componentes principais serviram como indicativo da razão entre os lados do mapa e para inicializar os pesos sinápticos de forma ordenada.

Como técnicas de visualização, o MDS não-métrico e o SOM são “competidores”. Na análise das representações gráficas produzidas pelos dois métodos foi possível verificar as diferenças nas propriedades dos resultados obtidos. O algoritmo do SOM concentra-se na preservação de relações de vizinhança dos dados, enquanto o MDS não-métrico na preservação de distâncias entre esses dados. Assim, no SOM, as estruturas locais são consideradas mais importantes, conduzindo a uma interpretação do mapa que seja feita, predominantemente, com base nas relações de proximidade dos neurônios no mapa.

Como o objetivo do trabalho foi identificar similaridades entre o porto em estudo e os demais portos levantados, o algoritmo do SOM provou sua utilidade, sendo uma alternativa não apenas viável, mas útil, principalmente se a análise for realizada aliada a outros métodos. No mapa obtido, é possível visualizar portos que, mesmo pertencendo a outros grupos, estão próximos, do ponto de vista de similaridade ambiental, à Sepetiba. Portanto, são portos aos quais os sistemas de inspeção de navios devem enfatizar suas ações de fiscalização.

Tendo em vista os resultados alcançados, observou-se a efetiva aplicabilidade das técnicas de análise de agrupamento e do MDS não-métrico, os quais se mostraram eficientes no tratamento de dados. Sem pretender comparar diferentes algoritmos, pode-se afirmar que, na aplicação prática estudada, foi possível mostrar que as habilidades do SOM produziram resultados iguais ou melhores aos obtidos pelos métodos de análise multivariada. Isso não significa afirmar que as análises feitas mostrem, em geral, uma superioridade do SOM sobre os outros métodos. Em outros conjuntos de dados o SOM pode ter um desempenho diferente. Essa é uma questão de pesquisa em aberto.

## Referências Bibliográficas

- Almeida, L.P., Leal Neto, A.C. *Convenção internacional para controle e gestão de água de lastro: Consequências para o porto de Suape*. 5º Seminário sobre Meio Ambiente Marinho. Sociedade Brasileira de Engenharia Naval. Rio de Janeiro, 2005.
- Ashby, F.G., Perrin, N.A. *Toward a unified theory of similarity and recognition*. Psychological Review, 95(1): 124-150. Janeiro, 1988.
- Brandalize, M.C.B., Philips, J. *Análise da distribuição de bens imobiliários no Brasil*. GEODÉSIA online 2. Revista da Comissão Brasileira de Geodésia. 2002. (Disponível na internet em <http://geodesia.ufsc.br/Geodesia-online/arquivo/2002/02/BP.htm>)
- Bottou, L., Bengio, Y. Convergence properties of the K-means algorithm. In *Advances in Neural Information Processing Systems*, volume 7. MIT Press. 1995.
- Brasil. *Resolução RDC nº 217, de 21 de novembro de 2001*. Aprova o Regulamento Técnico, anexo a esta Resolução, com vistas à promoção da vigilância sanitária nos Portos de Controle Sanitário instalados no território nacional, embarcações que operem transportes de cargas e ou viajantes nesses locais, e com vistas a promoção da vigilância epidemiológica e do controle de vetores dessas áreas e dos meios de transporte que nelas circulam. Agência Nacional de Vigilância Sanitária (ANVISA). Diário Oficial da União; Poder Executivo, de 21 de dezembro de 2001. 2001.
- \_\_\_\_\_. *Portaria Nº 52/DPC, de 14 de junho de 2005*. Norma da Autoridade Marítima para o gerenciamento da água de lastro de navios. Diretoria de Portos e Costas (DPC), Marinha do Brasil. Diário Oficial da União Nº 121, Poder Executivo, de 27 de junho de 2005. 2005.
- \_\_\_\_\_. *Diretrizes para o controle e gerenciamento da água de lastro dos navios para minimizar a transferência de organismos aquáticos nocivos e agentes patogênicos*. Resolução A.868(20)-IMO, Diretoria de Portos e Costas (DPC), Marinha do Brasil. 25p. 1998.
- \_\_\_\_\_. *Espécies exóticas invasoras: Situação Brasileira*. Ministério do Meio Ambiente, Secretaria de Biodiversidade e Florestas. 24 p. Brasília. 2006. (disponível em: <http://www.mma.gov.br/invasoras/capa/docs/invasoras.zip>)
- Brosse, S., Giraudel, J.L., Lek, S. *Utilisation of non-supervised neural networks and principal component analysis to study fish assemblages*. Ecological Modelling. 146:159-166. 2001.



- Calow, P., Forbes, V.E. *Malacca Straits: Initial risk assessment*. MPP-EAS/Info/97/117, 82p. GEF/UNDP/IMO Regional Programme for the Prevention and Management of Marine Pollution Prevention in the East Asian, Quezon City, Filipinas. 1997.
- Carlton, J.T. *Transoceanic and interoceanic dispersal of coastal marine organisms: The biology of ballast water*. *Oceanogr. Mar. Biol. Ann. Rev.*, 23, pp.313-371. 1985.
- \_\_\_\_\_. Marine species introductions by ships' ballast water: An overview. p. 23-25. In: M.R.Devoe. (ed.) *Introductions & transfers of marine species. Achieving a balance between economic development and resource protection*. Proceedings of the Conference and Workshop, October 30 - November 2, 1991, Hilton Head Island, South Carolina. S.C. Seagrass Consortium. Estados Unidos. 1993.
- \_\_\_\_\_. "The scale and ecological consequences of biological invasions in the world's ocean". In *Invasive Species and Biodiversity Management* (O.T. Sandlund, P.J. Schei, and Å. Viken, eds.) Kluwer Academic Publishers. pp. 195-212. Dordrecht, Holanda. 1999.
- \_\_\_\_\_. *Introduced species in U.S. coastal waters: Environmental impacts and management priorities*. Pew Oceans Commission, Arlington, Virginia, Estados Unidos. 2001.
- Cataldo, D. H., Boltovskoy, D. *Yearly reproductive activity of *Limnoperna fortunei* (Bivalvia) as inferred from the occurrence of its larvae in the plankton of the lower Parana River and the Rio de La Plata estuary (Argentina)*. *Aquatic Ecology* 34(3): 307-317. 2000.
- Carvalho, L.A.V. *Datamining: A mineração de dados no marketing, medicina, economia, engenharia e administração*. São Paulo: Ed. Érica. 2001.
- Carvalho, L.A.V., Mendes, D.Q., Wedemann, R.S. *Creativity and delusions: The dopaminergic control of cortical maps*. In: (Org.) *Lecture Notes in Computer Science*. Berlin: Springer Verlag, 2757: 511-520. 2003.
- Caudill, M. *Neural networks primer*. São Francisco: Miller Freeman. EUA, 1989.
- Chapman, J.W., Carlton J.T. *A test of criteria for introduced species: The global invasion by the isopod *Synoditea laevidorsalis* (Miers 1881)*. *Journal of Crustacean Biology* 11-3: 386-400. 1991.
- \_\_\_\_\_. *Predicted discoveries of the introduced isopod *Synidotea laevidorsalis* (Miers, 1881)*. *Journal of Crustacean Biology* 14-4: 700-714. 1994.

- Chon, T.S., Park, Y.S., Moon, K.H., Cha, E.Y. *Patternizing communities by using an artificial neural network*. Ecological Modelling 90: 69–78. 1996.
- Clarke, C., Hilliard, R., Junqueira, A. de O.R., Leal Neto, A.C., Polglaze, J., Raaymakers, S. *Ballast water risk assessment, Port of Sepetiba, Federative Republic of Brazil: Final Report*. GloBallast Monograph Series No. 14. IMO London. 2004.
- Clarke, K.R., Gorley, R.N. *PRIMER v5: User manual / Tutorial*. PRIMER-E: Plymouth, United Kingdom. 2001.
- Clarke, K.R., Warwick, R.M. *Change in marine communities: An approach to statistical analysis and interpretation*. 2<sup>nd</sup> Edition. PRIMER-E: Plymouth, United Kingdom. 2001.
- Cohen, A. N., Carlton, J. T. *Non-Indigenous aquatic species in a United States estuary: A case study of the biological invasions of the San Francisco bay and delta*. Report for the United States Fish and Wildlife Service, Washington D.C. and The National Sea Grant College Program, Connecticut Sea Grant. Estados Unidos. 1995. (Disponível na internet em: <http://www.sgnis.org/publicat/cc1.htm>)
- Cranfield, H.J., Gordon, D.P., Willan, R.C., Marshall, B.A., Battershill, C.N., Francis, M.P., Nelson, W.A., Glasby, C.J., Read, G.B. *Adventive marine species in New Zealand*. National Institute of Water and Air Research (NIWA). Technical Report 34. 48p. Auckland, Nova Zelândia. 1998.
- Damásio, A.R. *O erro de descartes: emoção, razão e cérebro humano*. 5a. Ed. Companhia das Letras. Rio de Janeiro. 1994.
- DeSieno, D. *Adding a conscience to competitive learning*. In: IEEE International Conference on Neural Networks (San Diego, 1988), vol. I, 117-124. Nova Iorque: IEEE. EUA, 1988.
- Endresen, Ø., Sjørgård, E., Behrens, H.L., Andersen, A.B. *How much ballast?*. In: Ballast Water News. Issue 14. pp. 6-7. Global Ballast Water Management Programme. 2003.
- Fasham, M.J.R. *A comparison of nonmetric multidimensional scaling, principal components and reciprocal averaging for the ordination of simulated coenoclines, and coenoplanes*. Ecology 58: 551-561. 1977.
- Fernandes, F.C., Leal Neto, A.C. *Agua de lastre como vía de introducción de especies a nivel global*. In: *Bio-invasión del mejillón dorado en el continente americano*. Eds. Darrigran, G. e Damborenea, M.C. Editorial de la Universidad de La Plata-EDULP, 220 pp., La Plata. Argentina, 2006.

- Fontes Júnior, H.M. *A presença do bivalve invasor Limnoperna fortunei na Hidroelétrica de Itaipu*. II Seminário Brasileiro sobre Água de Lastro. Resumos. 2002.
- Gevrey, M., Worner, S.P., Kasabov, N., Pitt, J., Giraudel, J-L. *Estimating risk of events using SOM models: A case study on invasive species establishment*. Ecological Modelling. 197: 361-372. 2006.
- Gollasch, S. *Removal of barriers to the effective implementation of ballast water control and management measures in developing countries*. GEF/IMO/UNDP Report. 197 pp. 1997.
- Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C. *Análise Multivariada de Dados*. 5ed. Porto Alegre: Bookman, 2005.
- Haykin, S. *Neural networks: A comprehensive foundation*. Macmillan, Englewood Cliffs, NJ, 1994.
- \_\_\_\_\_. *Redes neurais: Princípios e prática*. 2ª Edição. Bookman. 2001.
- Hertz, J.A., Palmer, G.R., Krogh, A.S. *Introduction to the theory of neural computation*. Santa Fe Institute studies in the science of complexity. Addison-Wesley, Redwood City, CA, 1991.
- Hewitt, C.L., Hayes, K.M. Risk assessment of marine biological invasions. In: Leppakoski, E., Gollasch, S. & Olenin, S. (eds). *Invasive aquatic species of Europe: Distribution, impacts and management*. Kluwer Academic Publishers, Dordrecht, Holanda. pp. 456-466. 2002.
- Hilliard, R.W., Hutchings, P.H., Raaymakers, S. *Ballast water risk assessment for 12 Queensland ports: Stage 4 Report - Review of candidate risk biota*. EcoPorts Monograph Series N° 13, published by Ports Corporation of Queensland, Brisbane. 76 pp. Austrália. 1997.
- Hotelling, H. *Analysis of a complex of statistical variables into principal components*. Journal of Educational Psychology. 24: 417-441, 498-520. 1933.
- Hsu, C-Y., Wu, H-E. *An improved algorithm for Kohonen's self-organizing feature maps*. Proceedings. IEEE International Symposium on Circuits and Systems, ISCAS '92. 1:328-331. San Diego, USA. 1992.

- ICS, INTERTANKO. *Model for a ballast water management plan*. International Chamber of Shipping & International Association of Independent Tanker Owners. 2<sup>nd</sup> edition. 68 pp. 2000.
- IMO. *Convenção Internacional sobre Controle e Gestão da Água de Lastro e Sedimentos de Navios*. International Maritime Organization. Londres. 2004.
- IUCN. *Position statement on translocation of living organisms: Introductions, reintroductions and re-stocking*. Approved by the 22<sup>nd</sup> Meeting of the International Union for the Conservation of Nature and Natural Resources Council, Gland, Suíça, Setembro, 1987.
- . *Guidelines for the prevention of biodiversity loss caused by alien invasive species*. Approved by the 51<sup>st</sup> Meeting of the International Union for the Conservation of Nature and Natural Resources Council, Gland, Suíça, Fevereiro 2000. (Disponível na internet: [http://www.iucn.org/places/medoffice/invasive\\_species/docs/iucn\\_guidline\\_prev\\_bio.pdf](http://www.iucn.org/places/medoffice/invasive_species/docs/iucn_guidline_prev_bio.pdf))
- Jain, A.K., Mao, J., Mohiuddin, K. *Artificial neural networks: A tutorial*. IEEE Computer Special Issue on Neural Computing. Março, 1996.
- Jain, A.K., Duin, R.P.W., Mao, J. *Statistical pattern recognition: A review*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no 1, pp. 4-37. Janeiro, 2000.
- Junqueira, A.O.R., Leal Neto, A.C. *Avaliação de risco de água de lastro*. 4<sup>o</sup> Seminário sobre Meio Ambiente Marinho. Sociedade Brasileira de Engenharia Naval. Rio de Janeiro, 2003.
- Kaski, S. *Data exploration using self-organizing maps*. Tese de D.Tech., Helsinki University of Technology. Finlândia, 1997.
- Kaski, S. Kangas, J., Teuvo Kohonen. *Bibliography of self-organizing map (SOM) papers: 1981–1997*. Neural Computing Surveys, 1: 102-350. 1998.
- Kiviluoto, J. *Topology preservation in self-organizing maps*. In Proceedings of the International Conference on Neural Networks (ICNN'96), 1: 294–299, Piscataway, New Jersey, USA, IEEE Neural Networks Council, June, 1996.
- Kohavi, R., Provost, F. *On applied research in machine learning. Editorial for the special issue on applications of machine learning and the knowledge discovery process*. Machine Learning. Kluwer Academic Publishers. Vol 30, N. 2/3, February/March. pp. 271-274. 1998.

- Kohonen, T. *The self-organizing map*. Proceedings of the IEEE, vol. 78, 9:1464-1480, Setembro, 1990.
- \_\_\_\_\_. *Self-organizing maps*. Springer series in information sciences. Volume 30. Springer, Berlin, Heidelberg, 2<sup>nd</sup> edition. 426p. 1997.
- \_\_\_\_\_. *The self-organizing map*. Neurocomputing 21: 1-6. 1998.
- Kohonen, T., Hynninen, J., Kangas, J., Laaksonen, J. *SOM\_PAK: the self-organizing map program package*. Version 3.1 Helsinki University of Technology, Laboratory of Computer and Information Science, Espoo, Finland, 1995.
- Kruskal, J.B. *Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis*. Psychometrika 29: 1-27. 1964.
- Laaksonen, J., Koskela, M., Laakso, S., Oja, E. *Self-organising maps as a relevance feedback technique in content-based image retrieval*. Springer-Verlag. Proceedings of Pattern Analysis & Applications 4:140-152. 2001.
- LATEC/COPEL. *Mexilhão dourado: Estado da arte das pesquisas relacionadas ao mexilhão dourado Limnoperna fortunei*. Projeto P&D ANEEL-CGER023. Instituto Tecnológico para o Desenvolvimento - LATEC / Companhia de Energia Elétrica do Paraná - COPEL. Curitiba, 2003.
- Leal Neto, A.C. *Ballast water management activities in Brazil*. 1<sup>st</sup> International Conference on Ballast Water Management: Best Practices & New Directions. Environmental Technology Institute. Cingapura, 2001.
- Leal Neto, A.C., Jablonski, S. O Programa GloBallast no Brasil. In: *Água de Lastro e Bioinvasão*. Eds. Silva, J.S.V. e Souza, R.C.C.L. Editora Interciência. Rio de Janeiro, 2004.
- Legendre, L., Legendre, P. *Numerical ecology*. 2<sup>nd</sup> English edition. Developments in environmental modeling, 20, Elsevier, New York, 853 pp. 1998.
- Lodge, D.M., Williams, S., MacIsaac, H.J., Hayes, K.R., Leung, B., Reichard, S., Mack, R.N., Moyle, P.B., Smith, M., Andow D.A., Carlton, J.T., McMichael, A. *Biological invasions: Recommendations for U.S. policy and management*. Ecological Society of America (ESA) Report. Ecological Applications, 16(6): 2035-2054. 2006.
- Ludwig, J.A., Reynolds, J.F. *Statistical ecology. A primer on methods and computing*. J. Wiley and Sons, Inc., New York, 1988.

- MacQueen, J.B. Some methods for classification and analysis of multivariate observations. In: *Proceedings of 5<sup>th</sup> Berkeley Symposium on Mathematics, Statistics and Probability*, Berkeley, University of California Press, 1:281-296. 1967.
- Mansur, M.C.D. *Moluscos bivalves asiáticos introduzidos no sul do Brasil: registros, densidades populacionais e conseqüências*. I Seminário Brasileiro sobre Água de Lastro. Resumos. 2000.
- Mansur, M.C.D., Darrigran, G. Distribución, abundancia y dispersión. In: *Bio-invasion del mejillón dorado en el continente americano*. Gustavo Darrigran y Cristina Damboronea Eds. 1ed. Editorial de la Universidad Nacional de La Plata - EDULP. La Plata, Argentina. 2006.
- MathWorks. *Neural network toolbox. User's guide - Version 4*. The MathWorks, Inc. 2005.
- McConnell, M. *GloBallast legislative review - Final report*. GloBallast Monograph Series N° 1. IMO. Londres. 2002.
- Melssen, W.J., Smits, J.R.M., Rolf, G.H., Kateman, G. *Two-dimensional mapping of IR spectra using a parallel implemented self-organising feature map*. Chemometrics and Intelligent Laboratory Systems 18: 195-204. 1993.
- Merkel, D. *Content-Based software classification by self-organization*. Proceedings of the IEEE International Conference on Neural Networks (ICNN'95). Nov 27 - Dec 1. pp 1086-1091. Perth. Australia. 1995.
- Miranda, J.M. *Redes neurais de Kohonen como modelos da topografia do sistema visual*. Dissertação de mestrado. COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 1998.
- MPP-EAS. *Environmental risk assessment manual: A practical guide for tropical ecosystems*. MPP-EAS Technical Report 21, 88p. GEF/UNDP/IMO Regional Programme for the Prevention and Management of Marine Pollution Prevention in the East Asian, Quezon City, Filipinas. 1999.
- Oja, M., Kaski, S., Kohonen, T. *Bibliography of self-organizing map (SOM) Papers: 1998-2001 Addendum*. Neural Computing Surveys 3, 1-156, 2002.
- Ostenfeld, C. H. *On the immigration of Biddulphia sinensis Grev. and its occurrence in the North Sea during 1903-1907*. Medd. Komm. Havunders, Ser. Plankton, 1(6): 1-46. 1908.

- Pielou, E.C. *The interpretation of ecological data. A primer on classification and ordination*. Wiley, New York, 263 pp. 1984.
- Pözlbauer, G. *Survey and comparison of quality measures for self-organizing maps*. Proceedings of the 5<sup>th</sup> Workshop on Data Analysis (WDA'04). pp. 67-82, 2004.
- Romais, D.K., Graf Neto, J., Zenni, R.D., Ribeiro, R.C., Ziller, S.R. *Resultados preliminares do informe nacional de espécies exóticas invasoras*. I Simpósio Brasileiro sobre Espécies Exóticas Invasoras. Brasília, Outubro, 2005.
- Ross, P.E. *As I see it: Flash of genius*. Forbes, pp. 98-104. November. 1998.
- Royal Haskoning. *Global market analysis of ballast water treatment technology*. Report 42810/001R/HSC/SKO. Holanda. 2001. (Disponível na internet em: <http://www.nemw.org/Haskoningreport.pdf>)
- Ruiz, G.M., Rawlings, T.K., Dobbs, F.C., Drake, L.A., Mullady, T., Huq, A., Colwell, R.R. *Global spread of microorganisms by ships*. Nature 408: 49-50. 2000.
- SCOPE. *Environmental risk assessment*. International White, A. & Burton, I. eds. International Council of Scientific Unions. Scientific Committee on Problems of the Environment-SCOPE. J. Wiley. 157 p. 1980.
- Santini, S., Jain, R. *The use of psychological similarity measure for queries in image databases*. Technical report, Visual Computing Laboratory, University of California San Diego, 1996.
- \_\_\_\_\_. *Similarity is a geometer*. Multimedia Tools and Applications, 5(3): 277-306. 1997.
- \_\_\_\_\_. *Similarity measures*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(9): 871-883. Setembro. 1999.
- Sneath, P.H.A., Sokal, R.R. *Numerical Taxonomy*. Freeman, San Francisco. 1973.
- Shepard, R.N. *The analysis of proximities: multidimensional scaling with an unknown distance function*. Psychometrika. 27: 125-140. 1962.
- Tversky, A. *Features of similarity*. Psychological Review, 84(4): 327-352. Julho, 1977.

Ultsch, A. *Self-organizing neural networks for visualization and classification*. In Proc. GfKI, Dortmund, Alemanha. 1992.

\_\_\_\_\_. *Self organizing neural networks perform different from statistical k-means clustering*. Gesellschaft für Klassifikation, Basel. 1995. (Disponível na internet em: <http://www.mathematik.uni-marburg.de/~databionics/en/downloads/papers/ultsch95kmeans.pdf>)

Ultsch, A., Mörchén, F. *ESOM-Maps: tools for clustering, visualization, and classification with Emergent SOM*. Technical Report No. 46, Dept. of Mathematics and Computer Science, University of Marburg, Germany. 2005. (Disponível na internet em: <http://www.mathematik.uni-marburg.de/~databionics/en/downloads/papers/ultsch05esom.pdf>)

Valentin, J.L. *Ecologia numérica: Uma introdução à análise multivariada de dados ecológicos*. Rio de Janeiro. Interciência. 2000.

Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J. *SOM Toolbox for Matlab 5*. Report A57. Finland. 2000.

Vesanto, J. Data exploration process based on the Self-Organizing Map. Dissertation for the degree of Doctor of Technology. Helsinki University of Technology. Espoo, Finland. 2002.

Vesanto, J., Sulkava, M., Hollmen, J. *On the decomposition of the self-organizing map distortion measure*. Proceedings of the Workshop on Self-Organizing Maps (WSOM'03), pp 11-16, Hibikino, Kitakyushu, Japão, Setembro, 2003.

Wasserman, P.D. *Neural computing: Theory and practice*. Van Nostrand Reinhold. Nova Iorque. 1989.

Wolda, H. *Similarities indices, sample size, and diversity*. Oecologia, 50: 296-302. 1981.

Wonnacott, T.H., Wonnacott, R.J. *Introductory statistics for business and economics*. John Wiley & Sons, Inc. Fourth Edition. 1990.

Worner, S.P., Gevrey, M. *Modelling global insect pest species assemblages to determine risk of invasion*. Blackwell Publishing. Journal of Applied Ecology. 43-5: 858-867. 2006.



**Anexo 1 - Vetores comuns de bioinvasão marinha**

	<b>Navios</b> <ul style="list-style-type: none"> <li>Organismos planctônicos e nectônicos na água de lastro</li> <li>Organismos livres ou incrustados no casco, quilha, leme, hélice, eixo da hélice e nos sistemas de circulação de água do mar, caixa de mar, água de lastro e em tanques de carga lastrados</li> <li>Organismos associados à âncora, amarras e caixa da âncora e amarras</li> <li>Organismos associados à carga, como troncos de madeira que flutuam durante o carregamento</li> </ul>		<b>Pesca, inclusive Aqüicultura Marinha (Maricultura)</b> <ul style="list-style-type: none"> <li>Transplante ou cultivo de produtos como ostras, mexilhões, vieiras, caranguejos, lagostas, peixes ou macroalgas em mar aberto, para crescimento ou rejuvenescimento; e outros organismos associados à estruturas de transporte e cultivo</li> <li>Lançamento intencional de espécies de moluscos, crustáceos, peixes e macroalgas, como resultado tanto de empreitada oficial e governamental como em iniciativa privada e ilegal</li> <li>Melhoramento de estoque, normalmente em andamento, além de organismos associados à espécie –alvo, transportados acidentalmente</li> <li>Movimento de organismos vivos para fins comerciais de revenda, mas descartados no ambiente</li> <li>Processamento de fruto do mar fresco ou congelado, seguido de descarte de restos no ambiente, o qual pode incluir organismos associados vivos ou encistados</li> <li>Movimento de iscas vivas com subsequente descarte no ambiente</li> <li>Descarte de material de embalagem - como macroalgas e organismos associados – usados em isca viva e frutos do mar</li> <li>Movimento, transporte e deriva de apetrechos de pesca como redes, flutuantes, armadilhas e dragas</li> <li>Lançamento de organismos como alimento para outras espécies</li> <li>Organismos transportados acidentalmente ou intencionalmente em pocetos ou tanques de coleta de água, embornais do navio ou outros recursos no convés</li> <li>Lançamento de estoques transgênicos – organismos geneticamente modificados (GM)</li> <li>Movimento de macroalga e organismos associados para servir de substrato quando da desova de peixes</li> </ul>
	<b>Plataformas</b> <ul style="list-style-type: none"> <li>Organismos incrustantes sésseis e vageis</li> <li>Organismos planctônicos e nectônicos na água de lastro</li> </ul>		
	<b>Diques Secos</b> <ul style="list-style-type: none"> <li>Organismos livres ou incrustantes</li> <li>Organismos planctônicos e nectônicos na água de lastro</li> </ul>		
	<b>Bóias de Navegação e Flutuantes</b> <ul style="list-style-type: none"> <li>Organismos livres ou incrustantes</li> </ul>		
	<b>Aviões-anfíbio e Hidroaviões</b> <ul style="list-style-type: none"> <li>Organismos livres ou incrustantes</li> <li>Organismos na água dos flutuadores</li> </ul>		
	<b>Canais</b> <ul style="list-style-type: none"> <li>Movimento de organismos através de canais de nível, de comportas e de irrigação</li> </ul>		
	<b>Aquários Públicos</b> <ul style="list-style-type: none"> <li>Descarte acidental ou intencional de organismos em exposição</li> <li>Descarte acidental ou intencional de organismos acidentalmente transportados com a espécie de interesse para exposição</li> </ul>		
	<b>Pesquisa</b> <ul style="list-style-type: none"> <li>Movimento e descarte de invertebrados, peixes, macroalgas, e gramas utilizadas em pesquisa (intencional ou acidental)</li> <li>Organismos associados ao equipamento de coleta e pesquisa, como material de mergulho (roupas ou apetrechos)</li> </ul>		
	<b>Detritos Marinhos Flutuantes</b> <ul style="list-style-type: none"> <li>Transporte de organismos em detritos gerados pelo Homem, como redes flutuantes e plástico</li> </ul>		<b>Aquários Domésticos</b> <ul style="list-style-type: none"> <li>Movimento e descarte de invertebrados, peixes, macroalgas e gramas (intencional ou acidental)</li> </ul>
	<b>Equipamento de Recreação</b> <ul style="list-style-type: none"> <li>Movimento de embarcações de recreação de pequeno porte, material de mergulho (roupas e apetrechos), jet skis e materiais similares</li> </ul>		<b>Restauração</b> <ul style="list-style-type: none"> <li>Movimento de vegetação de marismas e de manguezal, de dunas ou gramas marinhas e organismos associados</li> <li>Re-estabelecimento de populações nativas extintas ou dizimadas e transporte acidental de organismos associados</li> </ul>
			<b>Educação</b> <ul style="list-style-type: none"> <li>Descarte de espécies por escolas e faculdades após uso em aula prática</li> </ul>

**Anexo 2 - Parâmetros da análise de similaridade ambiental do sistema GloBallast/URS**

Nº	Descritor	Tipo	Código
1	Tipo do porto	Categorizada (1-6) <sup>1</sup>	PTYPE
2	Temperatura média da água durante a estação mais quente	Escalar (°C)	MSUWT
3	Temperatura máxima da água da estação mais quente	”	USUWT
4	Temperatura média da água durante a estação mais fria	”	MWNWT
5	Temperatura mínima da água da estação mais fria	”	LWNWT
6	Temperatura média do ar do dia na estação mais quente	”	MSART
7	Temperatura máxima do ar do dia na estação mais quente	”	USART
8	Temperatura média do ar da noite na estação mais fria	”	MWART
9	Temperatura mínima do ar da noite na estação mais fria	”	LWART
10	Salinidade média durante o período chuvoso <sup>2</sup>	Escalar (ppm)	MWSAL
11	Salinidade mínima do período chuvoso	”	LWSAL
12	Salinidade média durante o período seco <sup>2</sup>	”	MDSAL
13	Salinidade máxima do período seco	”	UDSAL
14	Variação média da maré de sizígia	Escalar (m)	MSPR
15	Variação média da maré de quadratura	Escalar (m)	MNER
16	Precipitação total nos 6 meses mais secos	Escalar (mm)	D6MRF
17	Precipitação total nos 6 meses mais chuvosos	Escalar (mm)	W6MRF
18	Número de meses com 75% do total da precipitação anual <sup>3</sup>	Inteiro	RNFL75
19	Distância entre os berços e a foz de rio mais próximo <sup>4</sup>	Escalar (km)	DISRVM
20	Tamanho da bacia hidrográfica	Escalar (km <sup>2</sup> )	SIZRVC
21	Cais de paredes verticais	Categorizada (0-5) <sup>5</sup>	INASMW
22	Quebra-mar / espigões	”	INARKW
23	Pier de pilares de madeira <sup>6</sup>	”	INAWP
24	Salina	”	INHTSM
25	Praia arenosa	”	INSNDB
26	Praia de seixos ou cascalho	”	INSTNB
27	Planície de maré (lamosa)	”	INLTMF
28	Manguezal	”	INMANG
29	Costão rochoso	”	INRKSH
30	Fundo arenoso com conchas	”	SUFSND
31	Fundo lamoso-siltoso	”	SUSFTM
32	Grama marinha	”	SUSGRM
33	Infralitoral rochoso	”	SURKRF
34	Recife de coral	”	SUCORF

Fonte: adaptado de Clarke *et al.* (2004) p.15

<sup>1</sup> - terminal offshore ou bóia / baía natural / porto abrigado dentro de molhes / canal de maré / estuário / rio

<sup>2</sup> - Para fins de uniformização, considerou-se para os portos do Brasil a estação chuvosa entre os meses de novembro a abril; e a estação seca de maio a outubro.

<sup>3</sup> - Duração do pico das vazões

<sup>4</sup> - O valor é negativo se o porto estiver localizado dentro do rio

<sup>5</sup> - Distâncias entre o porto e os habitats marinhos próximos. As distâncias medidas nas cartas náuticas da área do porto são categorizadas conforme a seguinte convenção:

<1 km	➔	5
1-5 km	➔	4
5-10 km	➔	3
10-50 km	➔	2
50-100 km	➔	1
>100 km	➔	0

<sup>6</sup> - Marinas antigas e pier de barcos de pesca

**Anexo 3 - Base de Datos GloBallast**

Nº	Nome do Porto	CÓDIGO		Lat			Long		PType	MSUWT	USUWT	MWNWT	LWNWT	MSART	USART	MWART	LWART	MWSAL	LWSAL	MDSAL	UDSAL	MSPR	MNER
1	Sepetiba	BRSPB	22	55.3	S	43	51.0	W	2	22.0	25.5	19.0	17.2	25.4	38.2	22.0	11.1	30.5	20.0	30.5	34.0	1.4	0.7
2	Abu Dhabi	AEAUH	24	32.0	N	54	23.0	E	3	30.5	35.5	21.0	16.0	36.5	45.0	19.0	8.0	38.0	38.0	39.5	42.0	1.6	0.4
3	Mina Zayed	AEMZD	24	32.0	N	54	23.0	E	3	30.5	35.5	21.0	16.0	36.5	45.0	19.0	8.0	38.0	38.0	39.5	42.0	1.6	0.4
4	Das Island	AEDAS	25	9.0	N	52	52.0	E	1	30.0	35.2	22.5	16.9	31.0	42.0	21.5	12.0	38.0	37.5	39.0	40.5	0.6	0.2
5	Port Rashid	AEPRR	25	16.0	N	55	16.0	E	3	30.5	34.6	23.0	16.0	33.0	47.0	19.0	8.0	38.5	37.0	40.0	42.0	2.1	0.3
6	Dubai	AEDXB	25	16.0	N	55	16.0	E	3	30.5	34.6	23.0	16.0	33.0	47.0	19.0	8.0	38.5	37.0	40.0	42.0	2.1	0.3
7	Fateh Oil Terminal	AEFAT	25	36.0	N	54	31.2	E	1	30.0	35.2	22.5	16.9	31.0	42.0	21.5	12.0	38.0	37.5	39.0	40.5	0.8	0.2
8	Fujairah	AEFJR	25	10.0	N	56	21.0	E	1	30.0	32.0	22.0	16.5	32.8	40.0	24.0	14.0	35.5	35.0	36.0	37.0	1.8	1.0
9	Jebel Ali	AEJEA	25	0.0	N	55	3.0	E	3	30.5	34.6	21.0	16.0	31.0	47.0	23.0	8.0	38.5	37.0	40.0	42.0	1.1	0.2
10	Jebel Dhanna	AEJED	24	12.0	N	52	40.0	E	1	30.0	35.2	22.5	16.9	31.0	42.0	21.5	12.0	39.0	38.0	39.5	41.0	0.8	0.2
11	Khor Al Fakkan (Tanker Anchorage)	AEKLF	25	24.0	N	56	22.0	E	1	30.0	32.0	21.5	16.5	32.8	42.0	24.0	14.0	35.5	35.0	36.0	37.0	1.8	1.0
12	Um Al Qiwain	AEQIW	25	36.0	N	55	37.0	E	4	30.5	34.6	21.0	16.0	31.0	47.0	23.0	8.0	38.5	37.0	40.5	42.0	1.1	0.6
13	Ruwais Oil Terminal	AERUW	24	7.8	N	52	43.8	E	1	30.0	35.2	22.5	16.9	31.0	42.0	21.5	12.0	39.0	38.0	39.5	41.0	0.8	0.2
14	Sharjah	AESHJ	25	22.0	N	55	23.0	E	3	30.5	34.6	23.0	16.0	33.0	47.0	19.0	8.0	38.5	37.0	40.0	42.0	2.1	0.2
15	Zirku Island	AEZIR	24	52.2	N	53	4.2	E	1	30.0	35.2	22.5	16.9	31.0	43.0	21.5	12.0	38.0	37.5	39.0	40.5	0.8	0.2
16	Buenos Aires	ARBUE	34	34.0	S	58	23.0	W	5	24.0	27.0	15.0	13.0	26.0	34.0	5.0	0.0	10.0	0.0	15.0	25.0	1.3	0.6
17	Campana	ARCMP	34	9.0	S	58	58.2	W	6	24.0	27.0	15.0	13.0	26.0	35.0	4.0	-3.0	0.0	0.0	0.0	0.0	0.0	0.0
18	Dampier	AUDAM	20	39.0	S	116	43.0	E	4	26.5	31.0	23.0	19.0	36.0	45.0	14.0	7.0	34.8	34.1	35.1	35.5	5.6	2.8
19	Port Walcott (Cape Lambert)	AUPWL	20	37.0	S	117	11.0	E	1	26.5	31.0	23.0	20.0	36.0	45.0	14.0	7.0	34.9	33.8	35.1	35.5	6.0	2.9
20	Port Bonython	AUPBY	33	1.0	S	137	46.0	E	2	20.0	22.0	14.5	12.0	27.0	35.0	7.5	1.0	36.1	36.0	36.2	36.3	2.6	1.6
21	Whyalla	AUWYA	33	1.8	S	137	24.0	E	3	19.5	21.5	14.5	12.0	28.0	36.0	5.8	0.5	35.4	35.1	35.5	35.8	2.3	1.6
22	Port Pirie	AUPPI	33	10.2	S	138	1.8	S	2	20.0	21.5	14.5	12.0	28.0	36.0	5.8	0.5	36.5	35.8	37.2	38.0	2.7	1.7
23	Port Stanvac	AUPST	35	6.0	S	138	28.0	E	1	20.5	22.0	14.5	13.0	26.0	39.0	14.0	3.0	36.3	35.8	36.4	36.5	2.1	1.4
24	Western Port (now Hastings; AUHAS)	AUWEP	38	18.0	S	145	13.2	E	2	18.0	20.0	13.5	11.0	25.2	34.0	6.5	-4.0	35.3	35.0	35.4	35.5	3.3	2.2
25	Port Kembla	AUPKL	34	28.2	S	150	54.0	E	3	18.5	21.0	17.5	15.0	25.0	36.0	8.8	0.0	34.5	32.0	35.2	35.6	2.0	1.1
26	Brisbane	AUBNE	27	22.8	S	153	9.6	E	5	24.0	27.0	20.0	17.5	29.2	39.0	9.0	2.0	26.0	4.0	33.0	35.0	2.3	1.3
27	Bundaberg	AUBDB	24	46.0	S	152	23.0	E	5	25.7	29.2	19.8	15.7	29.0	36.6	16.0	5.1	35.1	25.0	35.4	36.0	2.4	1.2
28	Gladstone	AUGLT	23	51.0	S	151	15.0	E	4	26.0	28.5	22.0	20.0	31.0	38.0	14.0	8.0	34.1	31.0	35.2	35.5	4.0	1.7
29	Port Alma	AUPTL	23	35.0	S	150	52.0	E	5	26.8	28.7	21.1	19.5	29.0	39.1	11.0	6.0	30.0	20.0	35.0	36.0	4.0	1.9
30	Hay Point (Dalrymple Bay)	AUHPT	21	16.0	S	149	19.0	E	1	27.0	30.0	21.5	19.0	29.0	36.6	14.0	6.7	34.9	33.0	35.3	36.5	4.8	2.2
31	Dalrymple Bay (Hay Point)	AUPDT	21	16.2	S	149	19.2	E	1	27.0	30.0	21.5	19.0	29.0	36.6	14.0	6.7	34.9	33.0	35.3	36.5	4.8	2.2
32	Mackay	AUMKY	21	6.0	S	149	20.0	E	3	26.2	31.0	19.5	18.0	30.5	36.6	14.8	6.7	35.0	31.0	35.3	36.5	4.5	2.1
33	Abbot Point	AUABP	19	53.0	S	147	5.0	E	1	28.4	32.4	23.4	19.3	31.0	37.4	18.0	9.4	34.7	24.3	35.3	36.5	2.0	0.8
34	Townsville	AUTSV	19	15.0	S	146	50.0	E	3	28.2	32.4	23.0	19.0	31.5	40.8	15.5	6.0	34.7	22.0	35.0	36.2	2.3	0.6
35	Lucinda	AULUC	18	31.0	S	146	19.0	E	1	28.0	31.0	24.0	21.0	31.0	37.0	18.0	14.5	34.4	20.0	35.0	37.0	2.2	0.6
36	Mourilyan	AUMOU	17	37.0	S	146	7.0	E	5	30.0	33.0	27.0	23.0	31.0	38.5	18.0	8.4	33.0	9.0	35.0	36.0	1.8	0.5
37	Cairns	AUCNS	16	55.0	S	145	47.0	E	4	27.0	31.0	23.6	18.5	32.0	38.3	17.5	9.2	34.4	5.0	35.0	35.3	1.8	0.4
38	Cape Flattery	AUCQP	14	59.0	S	145	21.0	E	2	28.0	30.0	24.0	21.0	31.0	38.4	22.0	11.3	33.6	32.0	35.0	35.5	1.7	0.4
39	Weipa	AUWEI	12	35.0	S	141	36.0	E	5	30.0	33.0	28.5	24.0	33.0	37.0	19.2	12.5	33.0	20.0	35.0	36.0	2.2	0.7
40	Karumba	AUKRB	17	29.0	S	140	50.0	E	6	30.0	32.0	28.0	24.0	32.0	38.0	18.5	14.2	24.0	0.0	35.0	39.0	3.3	2.5
41	Chittagong	BDCGP	22	13.0	N	91	48.0	E	5	29.0	32.4	25.0	19.0	30.0	33.0	19.0	13.0	2.0	0.0	7.0	12.0	4.9	2.1
42	Antwerpen	BEANR	51	14.0	N	4	28.0	E	6	17.0	20.0	5.0	0.0	21.5	30.0	2.2	-4.0	0.0	0.0	0.0	0.0	0.0	0.0
43	Ghent (Gent)	BEGNE	51	4.2	N	3	42.0	E	6	17.0	20.0	5.0	0.0	21.0	28.0	1.8	-5.0	0.0	0.0	0.0	0.0	0.0	0.0
44	Bourgas	BGBOJ	42	30.0	N	27	28.8	E	3	24.2	26.5	6.5	0.5	24.5	38.0	4.3	-16.0	17.0	16.4	18.1	19.5	0.1	0.0
45	Varna, Bulgaria	BGVAR	43	12.0	N	27	57.0	E	3	21.7	25.9	5.9	-0.5	24.0	38.0	3.5	-17.0	16.8	16.1	17.9	19.3	0.1	0.0
46	Sitra (Bahrain)	BHMAN	26	9.0	N	50	40.0	E	1	32.0	35.0	17.0	14.0	36.5	45.0	18.0	12.0	40.0	38.0	42.0	43.0	1.6	0.9
47	Mina Sulman (Al Manamah)	BHMIN	26	13.0	N	50	36.0	E	3	32.0	35.0	17.0	14.0	36.5	45.0	18.0	12.0	40.0	38.0	42.0	43.0	1.8	0.9
48	Itajai	BRITJ	26	54.0	S	48	39.7	W	6	26.2	28.0	19.1	17.5	23.0	30.6	17.6	1.3	3.0	0.0	4.9	15.0	1.2	0.3
49	Paranaguá	BRPNG	25	30.1	S	48	31.5	W	5	27.5	30.0	21.4	13.0	22.0	40.0	17.2	2.4	20.5	10.5	27.0	34.0	1.6	0.4
50	Santos	BRSSZ	23	55.0	S	46	20.0	W	4	26.5	30.5	21.5	17.0	23.7	39.0	19.0	6.4	28.0	24.0	28.5	33.5	1.3	0.4
51	Rio de Janeiro	BRRIO	22	53.8	S	43	11.0	W	2	27.5	31.0	23.4	19.0	25.4	38.2	22.0	11.1	23.4	17.2	27.6	30.9	1.2	0.9
52	Ponta do Ubu	BRPOU	20	47.0	S	40	35.0	W	1	24.0	27.0	23.0	21.0	25.8	39.0	22.7	14.2	26.0	35.0	37.0	39.0	1.4	0.6

Nº	Nome do Porto	CÓDIGO		Lat			Long		PType	MSUWT	USUWT	MWNWT	LWNWT	MSART	USART	MWART	LWART	MWSAL	LWSAL	MDSAL	UDSAL	MSPR	MNER
61	Vancouver (British Columbia)	CAVAN	49	16.8	N	123	7.2	W	5	11.0	13.0	3.0	-1.0	21.2	26.0	1.0	-4.0	10.0	2.0	24.0	28.0	4.1	3.4
62	Roberts Bank (British Columbia)	CARBK	49	2.0	N	123	9.0	W	5	11.0	13.0	3.0	-1.0	21.2	26.0	1.0	-4.0	10.0	2.0	24.0	28.0	4.1	3.4
63	Guangzhou Guangdong	CNCAN	23	6.0	N	113	14.0	E	5	23.0	28.0	19.7	16.0	28.3	38.7	13.2	2.0	3.9	0.1	19.0	21.0	3.5	1.0
64	Chiwan (Shenzhen) Guangdong	CNCWN	22	29.0	N	113	54.0	E	5	30.0	33.1	15.5	0.5	27.5	39.0	18.0	8.0	24.0	20.0	30.0	34.0	2.8	2.0
65	Dalian Liaoning	CNDLC	38	55.7	N	121	39.3	E	2	22.2	26.2	2.2	-1.9	24.1	34.4	-2.0	-15.4	28.9	26.9	30.3	32.0	3.9	2.6
66	Huangpu Guangdong	CNHUA	23	6.0	N	113	26.0	E	5	29.0	34.0	15.8	0.3	27.0	38.0	15.0	5.0	8.0	0.1	12.0	28.0	2.5	1.8
67	Beilun Zhejiang	CNNBO	29	56.0	N	121	53.0	E	5	22.5	26.1	11.0	8.0	29.4	39.5	8.5	-2.6	21.7	10.6	19.6	25.2	3.1	1.1
68	Ningbo (Beilun) Zhejiang	CNNGB	29	56.0	N	121	53.0	E	5	22.5	26.1	11.0	8.0	29.4	39.5	8.5	-2.6	21.7	10.6	19.6	25.2	3.1	1.1
69	Shanghai Shanghai	CNSHA	31	14.0	N	121	29.0	E	5	26.4	32.0	6.5	4.0	27.0	40.2	5.0	-10.0	0.8	0.1	4.9	9.0	4.2	1.2
70	Shanghai Baoshan	CNSHB	31	25.0	N	121	30.0	E	5	25.5	30.0	7.0	5.0	25.7	39.0	8.7	-5.0	0.5	0.1	5.0	5.8	5.5	2.8
71	Qingdao Shandong	CNTAO	36	5.0	N	120	18.0	E	3	24.3	27.0	4.2	2.1	25.2	35.4	-1.1	-15.5	31.8	31.6	32.4	32.6	3.4	1.8
72	Tianjin Tianjin	CNTSN	39	6.0	N	117	10.0	E	5	26.5	30.5	-0.1	-1.5	28.0	40.0	2.0	-18.3	31.4	26.5	31.9	35.7	3.8	2.0
73	Yantai Shandong	CNYNT	37	34.0	N	121	26.0	E	3	22.5	26.3	3.0	0.0	24.0	32.0	1.0	-10.0	31.0	29.5	32.0	33.0	2.8	1.8
74	Cartagena	COCAR	10	21.6	N	75	32.9	W	2	30.5	32.0	31.0	30.0	27.5	32.0	28.0	24.8	26.0	25.0	28.5	33.0	0.4	0.1
75	Kyrenia	CYKYR	35	20.0	S	33	19.0	W	3	25.6	28.5	18.0	16.0	30.5	37.0	10.0	6.0	38.6	38.0	39.2	39.3	0.5	0.1
76	Larnaca	CYLCA	34	55.0	N	33	39.0	E	3	25.6	28.6	18.2	16.6	31.0	37.0	9.0	5.0	38.6	38.0	39.2	39.4	0.6	0.1
77	Limassol	CYLMS	34	39.0	N	33	1.2	E	3	25.6	28.5	18.2	16.6	32.0	39.0	10.0	6.0	38.6	38.0	39.2	39.4	0.6	0.1
78	Bremen	DEBRE	53	0.0	N	8	46.8	E	5	14.0	16.5	6.4	3.0	17.0	24.0	1.2	-4.0	18.0	11.0	28.0	30.0	4.2	3.4
79	Hamburg	DEHAM	53	33.0	N	9	59.0	E	5	16.0	20.0	3.0	0.0	17.3	23.2	0.5	-5.5	4.0	0.0	11.0	18.0	3.0	1.0
80	Wilhelmshaven	DEWVN	53	32.0	N	8	8.0	E	2	17.0	21.0	4.0	2.0	17.0	24.0	1.2	-4.0	28.0	24.0	32.0	33.0	4.1	2.8
81	Djibouti (Djibouti)	DJIBI	11	36.0	N	43	8.0	E	3	29.5	32.0	23.5	20.5	32.2	40.0	26.3	16.0	35.8	35.3	36.9	37.3	1.0	0.5
82	Enstedvaerkets Havn	DKENS	55	1.0	N	9	26.0	E	2	17.0	20.5	3.5	1.5	16.5	24.0	0.0	-8.0	14.0	12.0	18.0	20.0	0.4	0.2
83	Fredericia	DKFRC	55	34.2	N	9	45.0	E	2	17.5	20.5	3.5	1.5	16.5	24.0	0.0	-8.0	19.0	18.0	21.0	24.0	0.4	0.2
84	Ain Sukhna	EGAIS	29	34.0	N	32	24.0	E	1	29.0	32.0	20.0	17.0	28.7	42.0	16.7	6.0	41.0	40.0	42.0	43.0	2.3	1.4
85	Alexandria (El Iskandariya)	EGALY	31	10.8	N	29	52.2	E	3	25.0	29.7	16.0	13.5	29.0	36.0	11.0	7.0	38.0	37.5	38.0	39.0	0.5	0.2
86	Damietta	EGDAM	31	25.8	N	31	48.0	E	3	25.0	29.7	16.0	13.0	29.0	36.0	11.0	7.0	25.0	20.0	33.0	36.0	0.5	0.2
87	El Dekheila	EGEDK	31	8.0	N	29	49.0	E	3	25.0	29.7	16.0	13.2	29.0	36.0	11.0	7.0	38.0	37.5	38.0	39.0	0.4	0.2
88	Port Said	EGPSD	31	15.6	N	32	18.6	E	3	25.0	29.7	16.0	13.2	26.0	33.0	15.0	8.0	37.0	34.0	38.5	39.5	0.6	0.2
89	Suez (El Suweis)	EGSUZ	29	58.0	N	32	33.0	E	3	29.0	31.4	20.0	17.6	34.0	44.0	18.0	6.0	40.5	39.3	42.0	42.5	1.6	0.9
90	Gijon	ESGIJ	43	34.0	N	5	41.0	W	3	18.0	20.0	13.0	11.0	25.0	35.0	7.0	4.0	35.2	34.5	35.4	35.6	4.6	2.2
91	Bilbao	ESBIO	43	21.6	N	3	4.2	W	5	20.5	22.0	13.0	11.0	25.1	34.5	6.8	4.0	33.0	25.0	35.0	35.5	4.8	2.1
92	Vigo	ESVGO	42	13.8	N	8	43.8	W	2	18.5	19.5	14.5	12.5	24.0	33.0	8.0	4.0	35.8	35.4	35.8	36.0	4.0	1.8
93	Barcelona	ESBCN	41	19.8	N	2	9.6	E	3	23.0	24.5	13.0	11.5	27.5	38.4	8.0	6.0	37.0	36.5	37.5	38.0	0.8	0.1
94	Valencia	ESVLC	39	27.0	N	0	18.0	W	3	25.0	27.0	13.5	12.0	28.0	39.0	10.0	7.5	37.2	37.0	37.6	38.0	0.2	0.0
95	Algeciras	ESALG	36	8.0	N	5	26.0	W	2	22.2	23.4	16.0	14.5	27.0	35.0	12.2	7.0	36.5	36.0	36.5	37.0	0.4	0.1
96	Las Palmas	ESLPA	28	9.0	N	15	25.0	W	3	22.3	24.0	20.0	17.5	27.2	35.0	14.1	11.0	36.6	36.4	36.6	36.8	2.6	1.0
97	Tenerife (Santa Cruz de Tenerife)	ESSCT	28	27.0	N	16	14.0	W	3	22.3	24.0	20.0	17.5	27.0	35.0	14.0	11.0	36.6	36.4	36.6	36.8	2.5	0.8
98	Tarragona	ESTAR	41	5.0	N	1	14.0	E	3	25.5	27.0	13.5	11.5	27.5	38.4	8.0	6.0	37.0	36.5	37.5	38.0	0.7	0.1
99	Dunkerque	FRDKK	51	3.0	N	2	22.0	E	3	17.5	21.0	7.0	3.0	21.0	30.0	3.0	-4.0	32.5	32.0	33.0	33.5	6.1	3.2
100	Brest	FRBES	48	24.0	N	4	30.0	E	3	17.0	19.5	11.0	9.0	22.0	33.0	4.5	-2.0	34.8	34.4	35.2	35.6	7.5	2.7
101	Donges	FRDON	47	18.0	N	2	4.0	E	5	19.5	21.0	11.0	9.0	21.0	29.0	4.0	-1.0	20.0	3.0	32.5	34.0	5.5	2.6
102	Fos sur Mer (Oil Terminal)	FRFOS	43	24.0	N	4	53.0	E	5	22.0	24.5	14.0	12.5	24.0	31.0	4.5	2.0	33.0	31.0	35.0	36.0	0.1	0.0
103	Lavera	FRLAV	43	24.0	N	5	0.0	E	5	22.0	24.5	14.0	12.5	24.0	31.0	4.5	2.0	33.0	31.0	35.0	36.0	0.1	0.0
104	Le Havre	FRLEH	49	29.0	N	0	6.0	E	5	18.0	20.0	9.0	7.0	21.0	29.0	3.0	-2.0	32.5	30.0	34.0	34.5	8.0	3.9
105	Marseilles	FRMRS	43	19.0	N	5	22.0	E	3	22.0	24.5	14.0	12.5	24.0	31.0	4.5	2.0	33.0	31.0	35.0	36.0	0.1	0.0
106	Hunterston	GBHST	55	45.0	N	4	53.0	W	2	14.5	16.5	7.0	4.5	18.2	25.0	0.9	-1.0	30.0	27.0	33.0	34.0	3.7	1.9
107	Immingham	GBIMM	53	38.0	N	0	11.0	W	5	16.0	18.0	6.5	2.5	18.4	26.0	1.1	-1.0	18.0	10.0	24.0	26.0	7.6	3.1
108	Burry Port (Llanelly)	GB001	51	40.0	N	4	15.0	W	5	17.0	19.0	8.5	7.0	21.0	27.0	3.0	-1.0	29.0	27.0	32.0	33.5	9.1	3.4
109	Port Talbot	GBPTB	51	34.0	N	3	48.0	W	3	17.0	19.0	8.5	6.0	18.5	26.0	2.0	-1.0	31.0	30.0	32.0	34.0	8.9	4.0
110	Redcar	GBRER	54	37.0	N	1	9.0	W	5	16.0	18.0	7.0	4.5	19.3	25.0	0.5	-2.5	20.0	5.0	31.0	33.1	5.2	1.9
111	Batumi, Georgia	GEBUS	41	39.0	N	41	37.8	E	3	23.8	28.6	10.3	6.8	21.6	40.0	7.9	-8.0	17.2	16.5	17.2	17.8	0.1	0.0
112	Poti, Georgia	GEPTI	42	7.8	N	41	39.0	E	3	24.0	28.5	9.6	6.1	21.7	41.0	6.3	-11.0	17.2	16.4	17.2	17.9	0.1	0.0
113	Gibraltar	GIGIB	36	8.0	N	5	21.0	W	3	22.2	23.4	16.0	14.5	27.0	36.0	12.5	7.5	36.5	36.0	36.5	37.0	0.5	0.

Nº	Nome do Porto	CÓDIGO		Lat		Long		PType	MSUWT	USUWT	MWNWT	LWNWT	MSART	USART	MWART	LWART	MWSAL	LWSAL	MDSAL	UDSAL	MSPR	MNER	
121	Hong Kong Hong Kong	HKHKG	22	17.0	N	114	10.0	E	3	26.0	28.5	19.9	18.1	27.2	34.0	20.8	15.0	18.6	10.0	32.5	34.0	0.8	0.6
122	Hong Kong Kowloon	HKKWN	22	17.0	N	114	10.0	E	2	26.0	28.5	19.9	18.1	27.2	34.0	20.8	15.0	18.6	10.0	32.5	34.0	0.8	0.6
123	Omislaj	HROMI	45	12.0	N	14	33.0	E	2	24.0	26.5	9.0	7.0	27.0	34.0	3.9	0.5	24.0	18.0	35.0	36.0	0.6	0.2
124	Belawan Sumatra	IDBLW	3	47.0	N	98	42.0	E	5	30.0	32.0	28.0	27.0	32.6	36.0	21.5	18.0	26.0	21.0	27.0	28.0	2.5	0.9
125	Dumai Sumatra	IDDUM	1	41.0	N	101	27.0	E	4	29.0	31.5	28.0	26.0	31.0	34.0	23.0	21.0	26.0	21.0	27.0	28.0	2.5	1.4
126	Cigading	IDCIG	6	1.0	S	105	57.0	E	2	28.4	30.5	27.0	26.0	30.5	36.0	24.0	23.0	32.0	30.0	33.0	34.5	0.7	0.2
127	Merak (inc. Anyer Terminal) Java	IDMRK	5	55.0	S	106	0.0	E	2	29.0	32.0	28.0	25.0	32.0	37.0	22.6	19.0	31.0	29.0	31.5	34.0	0.9	0.3
128	Jakarta Java	IDJKT	7	0.0	S	106	53.0	E	3	29.0	32.0	28.0	25.0	32.0	37.0	22.6	19.0	29.0	27.0	31.0	34.0	0.9	0.3
129	Cilacap Java	IDCXP	7	44.0	S	109	0.0	E	5	28.4	32.0	27.9	25.5	31.0	35.0	24.5	22.0	22.0	15.0	32.0	34.0	2.0	0.7
130	Semarang Java	IDSRG	6	57.0	S	110	25.0	E	2	28.5	30.5	27.5	25.0	30.0	36.0	24.2	23.0	30.0	25.0	33.0	34.0	1.0	0.7
131	Tanjung Perak (Surabaya) Java	IDSUB	7	12.0	S	112	44.0	E	2	28.6	30.4	28.0	25.0	30.5	36.0	23.1	23.4	28.0	23.0	30.0	32.0	1.5	0.2
132	Tanjung Bara Coal Terminal Kalimantan	IDTBA	0	32.0	S	117	39.0	E	1	29.0	32.0	28.0	26.0	31.0	35.0	23.5	20.0	29.0	28.0	30.0	31.0	2.5	0.8
133	Balikpapan Kalimantan	IDBPN	1	15.0	S	116	48.0	E	2	30.0	32.0	28.5	27.0	31.0	35.0	23.5	20.0	27.0	25.0	29.0	31.0	2.6	0.9
134	Amamapare Irian Jaya	IDAMA	4	49.0	S	136	58.0	E	5	28.5	30.0	27.5	25.0	30.5	36.0	22.0	19.0	12.5	0.0	15.8	28.0	2.4	0.7
135	Moneypoint	IE MOT	52	36.0	S	9	25.0	W	5	16.0	18.5	11.0	8.0	20.8	28.0	1.7	-3.0	10.0	0.0	22.0	27.0	5.9	4.3
136	Ashdod	ILASH	31	50.0	N	34	38.0	E	3	25.0	29.7	16.0	13.0	30.2	40.0	7.5	2.0	38.0	37.5	38.0	39.0	0.2	0.0
137	Mumbai (Ex Bombay)	INBOM	18	54.0	N	72	49.0	E	4	28.6	30.6	28.4	26.6	28.6	35.6	24.0	19.0	27.5	14.5	36.9	37.6	3.6	1.4
138	Calcutta	INCCU	22	33.0	N	88	19.0	E	6	29.0	32.4	25.0	19.0	30.0	33.0	19.0	13.0	0.0	0.0	0.0	0.0	4.2	2.1
139	Cochin	INCOK	9	58.0	N	78	48.0	E	5	30.0	31.9	28.0	25.0	29.0	31.3	23.5	19.0	5.0	1.1	11.4	22.0	0.6	0.2
140	Haldia	INHAI	22	2.0	N	88	5.0	E	5	29.0	32.0	25.0	22.0	30.0	34.0	20.0	17.0	15.0	0.0	18.0	23.0	4.9	1.6
141	Mangalore (New Mangalore)	INIXE	12	55.0	N	74	48.0	E	3	28.4	29.8	26.5	25.0	26.3	30.0	27.0	24.3	33.1	26.9	31.8	34.9	1.0	0.4
142	Kandla	INIXY	22	52.0	N	70	13.0	E	2	27.1	29.7	19.8	19.3	30.2	37.7	17.7	9.8	3.4	3.3	3.5	3.7	5.9	3.9
143	Chennai (Ex Madras)	INMAA	13	6.0	N	80	18.0	E	3	28.2	30.0	27.5	26.5	29.9	35.2	26.1	23.9	22.0	20.0	25.5	34.6	1.0	0.4
144	Marmugao (Marmagoa)	INMRM	15	25.0	N	73	47.0	E	5	27.8	30.4	27.8	26.0	29.8	31.7	22.7	20.5	28.4	22.4	32.2	33.3	1.4	0.8
145	Mundra	INMUN	22	54.0	N	69	42.0	E	2	27.9	30.5	22.0	20.0	29.1	40.0	22.8	7.0	26.0	21.0	32.0	33.0	5.2	2.6
146	Porbandar	INPBD	21	38.0	N	69	36.0	E	5	27.8	30.4	24.0	27.1	28.6	35.6	24.0	19.0	28.4	22.4	32.2	33.3	1.9	0.9
147	Paradeep	INPRD	20	15.0	N	85	40.0	E	3	30.8	31.4	27.2	26.6	29.6	36.2	18.4	12.9	14.9	9.0	29.3	33.2	1.9	0.7
148	Salaya	INSAL	22	18.0	N	69	34.8	E	2	28.0	31.0	23.0	20.0	29.0	40.0	23.0	11.0	27.0	22.0	32.0	33.0	4.1	2.8
149	Sikka	INSIK	22	31.0	N	69	48.0	E	2	27.9	30.5	24.0	27.3	29.1	40.0	22.8	7.0	36.0	35.5	35.0	35.0	4.8	2.6
150	Tuticorin (New Tuticorin)	INTUT	8	22.0	N	76	59.0	E	3	28.8	31.3	27.1	24.0	30.2	38.7	26.0	18.3	31.2	26.9	34.0	36.5	0.7	0.2
151	Vadinar Terminal	INVAD	22	30.0	N	69	42.0	E	1	28.0	31.0	23.0	20.0	29.0	40.0	23.0	11.0	27.0	22.0	32.0	33.0	4.2	2.8
152	Visakhapatnam	INVTZ	17	14.0	N	83	18.0	E	2	27.8	29.5	26.0	23.8	33.6	40.0	23.7	13.0	23.0	16.5	31.1	35.0	1.4	0.6
153	Bandar Imam Khomeyni	IRBKM	30	25.0	N	49	4.0	E	4	34.9	35.4	16.5	14.0	33.1	52.0	19.3	2.0	38.2	35.0	42.0	44.0	4.0	2.6
154	Bandar Mushar (Mushahr)	IRBMR	30	28.0	N	49	11.0	E	5	34.9	35.4	16.5	14.0	33.1	37.6	19.3	2.0	38.2	37.4	40.9	41.0	4.0	2.6
155	Bandar Abbas (Oil Jetty)	IRBND	27	11.0	N	56	17.0	E	1	34.5	34.9	20.0	19.0	34.2	45.6	18.5	7.2	36.6	35.0	37.0	37.5	3.1	1.4
156	Bushehr	IRBUZ	28	59.0	N	50	50.0	E	2	34.5	35.5	18.0	16.0	32.0	47.0	18.0	6.0	38.5	38.0	41.5	42.0	1.3	0.4
157	Khark Island	IRKHK	29	14.0	N	50	19.0	E	1	34.2	34.9	18.2	17.9	31.7	47.0	19.0	7.0	38.9	38.5	40.9	41.0	1.0	0.3
158	Lavan Island	IRLVP	26	47.0	N	53	20.0	E	1	33.0	34.0	21.0	19.0	33.0	37.0	23.8	10.0	36.0	35.0	36.0	37.0	1.1	0.2
159	Sirri Island Oil Terminal	IRSXI	25	57.0	N	54	32.0	E	1	33.0	34.0	21.0	19.0	33.0	37.0	23.8	10.0	36.0	35.0	36.0	37.0	1.2	0.4
160	Hafnarfjörður	ISHAF	61	4.0	N	21	58.0	W	2	9.5	11.5	3.0	-0.5	14.0	21.0	-5.0	-11.0	34.0	33.0	34.5	35.0	3.9	1.5
161	Straumsvík	ISSTR	64	3.0	N	22	3.0	W	3	9.5	11.5	3.0	-0.5	14.0	21.0	-5.0	-11.0	34.0	33.0	34.5	35.0	3.9	1.5
162	Genoa	ITGOA	44	24.0	N	8	55.2	E	3	23.5	24.0	14.0	12.5	26.0	34.0	6.0	2.0	37.0	36.0	37.5	38.0	0.2	0.0
163	Porto Foxi (Sarroch)	ITPFX	39	4.8	N	9	1.8	E	1	23.5	26.0	16.0	14.5	25.0	37.0	7.0	3.0	37.0	36.5	37.5	38.0	0.2	0.0
164	Livorno	ITLIV	43	33.0	N	10	18.6	E	3	24.0	26.0	15.0	13.0	27.5	37.0	5.0	0.5	37.5	37.0	38.0	38.5	0.3	0.0
165	Ravenna	ITRAN	44	28.8	N	12	16.8	E	3	25.5	27.0	12.0	10.0	29.0	37.0	4.0	0.5	18.0	10.0	36.0	37.0	0.5	0.2
166	Taranto	ITTAR	40	26.0	N	17	12.0	E	3	24.8	27.0	15.0	14.0	29.0	38.0	7.0	3.0	37.5	37.0	38.0	38.5	0.2	0.0
167	Venezia (=Fusina)	ITVCE	45	25.8	N	12	19.8	E	5	25.0	27.0	11.0	9.0	27.0	35.0	4.5	2.0	27.0	14.0	31.0	33.0	0.6	0.1
168	Trieste	ITTRS	45	39.0	N	13	45.0	E	3	24.0	26.0	10.0	8.0	27.0	34.0	3.9	0.5	27.0	22.0	35.0	36.0	0.9	0.2
169	Aboshi Hyogo	JPABO	34	45.0	N	134	34.0	E	3	25.5	27.0	11.0	9.0	29.0	36.0	5.0	1.0	25.0	20.0	28.0	30.0	1.6	0.3
170	Amagasaki Hyogo	JPAMA	34	41.0	N	135	23.0	E	5	24.0	26.0	11.5	9.0	29.0	34.0	6.0	1.0	18.0	16.0	20.0	25.0	1.5	0.5
171	Beppu Oita	JPBEP	33	20.0	N	131	31.0	E	2	24.0	27.5	16.0	12.0	29.0	34.0	3.0	-1.0	19.0	17.0	28.0	31.0	1.5	0.5
172	Chiba Chiba	JPCHB	35	35.0	N	140	6.0	E	2	23.0	26.0	12.0	8.0	27.0	35.0	7.0	-4.0	20.0	9.0	28.0	32.0	2.1	0.2
173	Kimitsu Chiba	JKMT	35	23.0	N	139	50.0	E	2	23.0	26.0	12.0	8.0	27.0	35.0	7.0	-4.0	20.0	9.0	28.0	32.0	2.0	0.2
174	Fukuyama Hiroshima	JPFKY	34	29.0	N	133	22.0	E	3	23.0	26.0	8.0	4.0	30.0	33.0	2.0	-2.0	17.4	16.3	18.0	22.0	2.9	1.4
175	Higashi-Harima Hyogo	JPHHR	34	43.0	N	134	50.0	E	3	25.0	27.0	9.5	8.0	29.0	35.0	4.0	-0.8	24.0	19.0	27.0	29.0	1.3	0.2
176	Himeji Hyogo	JPHIM	34	46.2	N	134	37.8	E	3	25.5	27.0	11.0	9.0	29.0	36.0	5.0	1.0	25.0	20.0	28.0	30.0	1.6	0.3
177	Hakata Fukuoka	JPHKT	33	35.0	N	130	23.0	E	2	23.6	25.0	10.8	9.5	31.0	31.8	4.0	0.0	18.4	17.0	28.0	30.0	2.4	0.8
178	Imabari Ehime	JPIMB	34	4.0	N	133	1.0	E	2	25.0	27.0	12.0	8.0	30.0	34.0	3.0	-1.0	23.0	20.0	28.0	32.0	3.3	1.2
179	Innoshima Hiroshima	JPINS	34	16.8	N	133	10.8	E	2	24.0	27.0	10.0	6.0	30.0	34.0	3.0	-1.0	23.0	20.0	28.0	32.0	3.8	1.3
180	Iwakuni Yamaguchi	JPWK	34	10.0	N	132	16.0	E	2	24.0	27.0	15.0	12.0	30.0	34.0	2.0	-1.0	18.0	14.0	21.0	24.0	2.9	1.1



Nº	Nome do Porto	CÓDIGO	Lat	Long	PTYPE	MSUWT	USUWT	MWNWT	LWNWT	MSART	USART	MWART	LWART	MWSAL	LWSAL	MDSAL	UDSAL	MSPR	MNER				
181	Kochi Kochi	JPKCZ	33	31.0	N	133	33.0	E	5	24.0	27.0	15.0	12.0	29.0	34.0	2.0	0.0	25.0	18.0	27.0	30.0	2.3	0.7
182	Kakogawa Hyogo	JPKGA	34	42.0	N	134	47.0	E	3	25.0	26.5	9.5	8.0	28.0	34.0	3.5	-0.8	24.0	18.0	26.0	29.0	1.3	0.5
183	Kiire Kagoshima	JPKII	31	23.0	N	130	32.0	E	2	25.0	28.0	19.0	17.0	30.5	35.0	6.0	3.0	33.0	29.0	33.0	34.5	3.0	0.5
184	Niigata Niigata	JPKIJ	37	54.0	N	139	4.0	E	5	23.0	26.0	10.0	8.0	28.0	33.0	0.5	-6.0	31.0	28.0	32.0	33.0	0.3	0.0
185	Kikuma Ehime	JPKIK	34	2.0	N	132	50.0	E	2	25.0	27.0	13.0	9.0	30.0	35.0	3.0	-1.0	23.0	20.0	28.0	32.0	3.3	1.1
186	Kinwan (Ishikawa) Okinawa	JPKIN	26	22.0	N	127	58.0	E	2	28.0	30.0	24.0	20.0	30.0	35.0	15.0	13.0	32.0	27.0	33.0	34.5	2.5	0.5
187	Kanda Fukuoka	JPKND	33	48.0	N	131	0.0	E	3	23.5	25.0	8.0	7.0	31.0	30.3	4.0	0.0	16.5	12.5	18.0	20.0	3.6	1.5
188	Kinuura Aichi	JPKNU	34	52.0	N	136	57.0	E	5	23.0	26.0	17.0	14.0	27.5	34.0	3.0	1.0	23.5	19.5	29.8	30.5	2.2	0.9
189	Kagoshima Kagoshima	JPKOJ	31	35.0	N	130	33.0	E	3	24.0	27.0	18.0	16.0	30.5	35.0	4.0	2.0	31.0	26.0	33.0	34.5	2.7	0.4
190	Kashima Ibaraki	JPKSM	35	55.0	N	140	42.0	E	3	23.0	25.0	12.5	8.0	28.0	33.0	2.5	-3.0	31.0	29.0	32.0	34.0	1.4	0.1
191	Kudamatsu Yamaguchi	JPKUD	34	0.0	N	131	51.0	E	3	23.0	26.0	11.5	10.0	30.0	34.0	7.0	5.0	18.0	14.0	21.0	24.0	3.0	1.0
192	Kawasaki Kanagawa	JPKWS	35	32.0	N	139	42.0	E	3	22.5	25.0	12.0	9.0	27.0	34.0	6.0	0.0	20.0	7.0	29.0	31.0	2.5	0.3
193	Maizuru Kyoto	JPMAI	34	28.0	N	135	21.0	E	2	24.5	26.5	13.0	10.0	29.0	34.0	0.3	-4.0	32.0	28.0	34.0	35.0	0.3	0.1
194	Mizushima Okayama	JPMIZ	34	30.0	N	133	45.0	E	5	26.1	28.0	11.0	9.0	30.7	34.0	2.0	-3.0	15.0	11.0	15.0	17.0	3.3	1.4
195	Moji (Kitakyushu) Fukuoka	JPMOJ	33	57.0	N	130	58.0	E	3	23.5	25.6	7.8	6.9	30.0	34.0	5.8	3.4	16.0	12.0	18.0	20.0	3.0	1.0
196	Muroran Hokkaido	JPMUR	42	20.0	N	140	58.0	E	3	15.0	18.0	8.0	2.0	22.0	27.0	3.0	-5.0	28.0	23.0	30.0	32.0	1.5	0.1
197	Matsuyama Ehime	JPMYJ	33	52.0	N	132	42.0	E	2	25.0	27.0	13.0	9.0	30.0	35.0	3.0	-1.0	23.0	20.0	28.0	32.0	3.4	1.1
198	Naha Okinawa	JPNAH	26	12.0	N	127	40.0	E	2	28.0	30.0	24.0	20.0	30.0	35.0	14.0	12.0	32.0	27.0	33.0	34.5	2.6	0.4
199	Negishi (Yokohama) Kanagawa	JPNGI	35	24.0	N	139	37.8	E	3	21.9	24.5	12.0	10.0	26.4	35.0	5.1	-1.0	22.0	8.0	31.0	33.5	2.0	0.5
200	Nagoya Aichi	JPNGO	35	4.0	N	136	51.0	E	3	22.3	26.0	17.0	13.0	27.1	34.0	6.0	0.0	23.5	19.5	29.8	30.2	1.9	0.6
201	Nagasaki Nagasaki	JPNGS	32	45.0	N	129	52.0	E	5	25.0	28.5	18.0	14.0	28.0	34.0	3.5	-0.5	28.0	21.0	33.0	34.5	2.9	1.0
202	Oita Oita	JPOIT	33	16.0	N	131	40.0	E	2	24.0	27.5	16.0	12.0	29.0	34.0	3.0	-1.0	19.0	17.0	28.0	31.0	1.6	0.6
203	Okinawa Okinawa	JPOKA	26	13.2	N	127	40.2	E	2	28.0	30.0	24.0	20.0	30.0	35.0	14.0	12.0	32.0	27.0	33.0	34.5	2.6	0.4
204	Onomichi Hiroshima	JPONO	34	22.0	N	133	11.0	E	3	24.0	27.0	10.0	6.0	30.0	34.0	3.0	-2.0	23.0	20.0	28.0	32.0	3.0	1.3
205	Osaka Osaka	JPOSA	34	38.0	N	135	25.0	E	5	24.0	26.0	11.0	8.0	30.0	36.0	6.0	2.0	18.0	14.0	20.0	25.0	1.4	0.2
206	Saiki Oita	JPSAE	32	58.8	N	131	55.8	E	2	25.0	28.0	17.0	13.0	29.0	34.0	3.0	-0.5	19.0	17.0	28.0	31.0	2.1	0.7
207	Saganoseki Oita	JPSAG	33	14.0	N	131	52.0	E	2	24.0	27.5	16.0	12.0	29.0	34.0	3.0	-1.0	19.0	17.0	28.0	31.0	2.1	0.7
208	Sakai Osaka	JPSAK	34	34.0	N	135	27.0	E	5	25.0	26.0	12.0	9.0	30.0	35.0	6.0	2.0	18.0	15.0	21.0	26.0	1.5	0.2
209	Shibushi Kagoshima	JPSBS	31	28.0	N	131	7.0	E	3	25.0	28.0	20.0	18.0	30.5	35.0	7.0	4.0	32.0	28.0	33.0	34.5	2.2	0.2
210	Sakaide Kagawa	JPSKD	34	21.0	N	133	50.0	E	2	24.0	25.7	11.0	9.5	28.0	33.0	5.0	-2.0	20.0	18.0	26.0	29.0	3.0	1.2
211	Sakaiminato Tottori	JPSMN	35	32.0	N	133	14.0	E	2	25.0	27.0	14.0	12.0	28.7	34.0	1.5	-3.0	32.0	28.0	34.0	35.0	0.9	0.6
212	Shimotsu Wakayama	JPSMT	34	7.0	N	135	8.0	E	2	23.0	26.0	17.5	14.5	30.0	35.0	2.0	-1.0	23.5	19.5	29.8	30.2	1.9	0.2
213	Shimizu Shizuoka	JPSMZ	35	1.0	N	138	30.0	E	2	23.0	26.0	17.0	15.0	28.0	34.0	5.0	0.0	26.0	21.0	31.5	33.0	2.6	0.2
214	Tamano (Uno) Okayama	JPTAM	34	28.8	N	133	57.0	E	3	26.1	28.0	11.0	9.0	30.7	36.0	2.0	-3.0	20.0	18.0	26.0	29.0	2.4	0.9
215	Tobata (Kitakyushu) Fukuoka	JPTBT	33	55.0	N	130	51.0	E	3	23.0	25.5	12.0	11.0	31.0	34.0	4.0	0.0	19.0	17.0	28.0	31.0	2.0	1.0
216	Tokuyama Yamaguchi	JPTKY	34	2.0	N	131	48.0	E	3	23.0	26.0	11.0	9.5	30.0	34.0	7.0	5.0	16.0	12.0	18.0	19.0	3.1	1.0
217	Tomakomai Hokkaido	JPTMK	42	37.0	N	141	37.0	E	3	15.0	17.0	7.0	2.0	21.2	25.5	-3.0	-18.0	28.0	23.0	30.0	32.0	1.7	0.2
218	Toyama Toyama	JPTOY	36	45.0	N	137	13.0	E	3	24.0	26.0	12.0	9.0	28.5	35.0	0.5	-5.0	31.0	27.0	32.0	33.0	0.3	0.1
219	Tokyo Tokyo	JPTYO	35	43.0	N	139	45.0	E	5	23.0	26.0	11.0	8.0	27.5	36.0	7.0	-3.0	15.0	5.0	25.0	28.0	2.1	0.1
220	Ube Yamaguchi	JPUBJ	33	56.0	N	131	14.0	E	3	25.0	28.2	10.0	8.0	30.0	34.0	7.0	4.0	14.4	9.0	16.0	18.0	3.0	1.6
221	Kobe Hyogo	JPUKB	34	41.0	N	135	12.0	E	3	25.5	27.5	10.0	5.0	30.0	35.0	4.7	-1.2	26.0	20.0	28.0	30.0	1.7	0.5
222	Wakayama Wakayama	JPWAK	34	13.0	N	135	9.0	E	3	22.3	25.0	17.0	14.5	30.0	35.0	10.5	6.5	23.5	19.5	29.8	30.2	2.1	0.3
223	Yokkaichi Mie	JPYKK	34	57.0	N	136	38.0	E	3	22.3	26.0	17.0	14.0	27.5	34.0	3.0	-1.0	23.5	19.5	29.8	31.0	2.0	0.8
224	Yokohama Kanagawa	JPYOK	35	27.0	N	139	39.0	E	3	21.9	24.5	12.5	9.5	26.4	35.0	5.1	-1.0	22.0	8.0	31.0	33.5	2.0	0.5
225	Yokosuka Kanagawa	JPYOS	35	17.0	N	135	39.0	E	2	22.0	24.0	12.0	10.5	26.0	34.0	7.5	-0.5	26.0	18.0	31.5	33.5	1.7	0.3
226	Mombasa	KEMBA	4	40.0	S	39	40.0	E	2	29.0	33.0	26.0	24.0	31.0	36.0	22.5	19.0	34.2	33.0	34.6	35.4	4.0	2.5
227	Kwangyang	KRKAN	34	54.0	N	127	42.0	E	3	18.5	24.5	11.0	5.0	28.0	34.0	2.0	-3.0	32.0	31.0	33.5	34.0	4.2	1.7
228	Pohang	KRKPO	36	2.0	N	129	26.0	E	3	19.0	22.5	12.0	6.0	27.0	34.0	-2.0	-7.0	33.4	31.0	33.8	34.5	0.2	0.1
229	Kunsan	KRKUV	35	58.0	N	126	37.0	E	2	18.5	24.0	10.5	3.0	28.0	34.0	-2.0	-7.0	28.0	24.0	32.0	33.0	7.0	5.5
230	Mokpo (Mogpo)	KRMOK	34	46.0	N	126	23.0	E	5	19.5	24.0	11.0	4.0	28.0	34.0	2.0	-3.0	31.0	28.0	32.0	33.0	4.5	2.1
231	Onsan	KRONs	35	28.0	N	129	24.0	E	5	19.0	23.0	13.0	5.0	28.5	35.0	-2.0	-7.0	33.0	29.0	33.0	34.0	0.5	0.4
232	Pusan	KRPUS	35	6.0	N	129	4.0	E	3	19.0	23.0	12.0	8.0	27.0	34.0	-0.5	-5.0	33.0	30.0	33.5	34.0	1.5	0.6
233	Samcheon Po	KRSCP	34	55.0	N	128	4.0	E	3	19.0	24.5	11.0	7.0	28.0	34.0	2.0	-3.0	32.0	31.0	33.5	34.0	3.0	2.2
234	Ulsan	KRUSN	35	29.0	N	129	24.0	E	5	19.0	23.0	14.0	7.0	28.5	35.0	-2.0	-7.0	33.3	30.0	33.9	34.5	0.6	0.3
235	Yosu (Yeosu)	KRYOS	34	44.0	N	127	45.0	E	2	18.5	24.5	11.0	6.0	27.5	36.0	0.0	-5.0	32.0	31.0	33.5	34.0	4.2	1.6
236	Kuwait (Shuwaikh; KWSWK)	KWKWI	29	21.0	N	47	55.0	E	3	32.0	36.0	17.0	14.0	36.8	47.0	13.0	5.0	38.5	37.0	39.0	41.0	3.5	1.4
237	Mina Al Ahmadi	KWMAA	29	4.0	N	48	9.0	E	1	33.0	35.8	17.0	15.0	36.8	48.0	14.5	4.0	38.9	38.0	39.0	40.0	3.0	0.7
238	Mina Saud	KWMIS	28	45.0	N	48	24.0	E	1	33.0	34.9	17.0	15.0	36.8	47.0	14.0	6.0	38.5	38.0	39.0	40.0	1.9	1.0
239	Mina Abdulla	KWMIB	29	2.0	N	48	11.0	E	1	32.0	34.0	17.0	15.0	36.5	47.0	14.0	6.0	38.5	38.0	39.0	40.0	3.0	0.8
240	Shuaiba	KWSAA	29	2.0	N	48	10.0	E	3	33.0	36.5	17.5	14.0	37.0	48.0	14.0	4.0	39.0	37.0	39.0	41.0	2.2	0.8

Nº	Nome do Porto	CÓDIGO		Lat		Long		PType	MSUWT	USUWT	MWNWT	LWNWT	MSART	USART	MWART	LWART	MWSAL	LWSAL	MDSAL	UDSAL	MSPR	MNER	
241	Colombo	LKCMB	6	57.0	N	79	51.0	E	3	29.0	32.0	27.0	24.0	30.0	35.0	26.0	19.0	31.0	26.0	33.0	35.5	0.8	0.2
242	Malta (Valletta)	MTMLA	35	54.0	N	14	31.2	E	3	24.0	26.0	16.5	15.0	31.0	40.0	10.0	6.0	37.5	37.0	38.0	38.5	0.3	0.0
243	Penang (Georgetown)	MYPEN	5	22.0	N	100	22.0	E	5	28.5	31.0	27.0	24.0	31.0	35.0	25.5	23.0	12.0	6.0	14.0	19.0	2.7	0.2
244	Lumut	MYLUM	4	16.2	N	100	39.0	E	5	29.0	31.0	28.0	26.0	31.0	36.0	26.0	22.0	12.0	6.0	14.0	20.0	3.0	0.9
245	Port Kelang	MYPKG	3	2.0	N	101	21.0	E	5	30.0	31.0	29.3	26.0	32.0	35.0	26.4	23.0	14.0	4.0	16.0	20.0	5.4	0.9
246	Port Dickson	MYPDI	2	31.0	N	101	47.0	E	1	29.0	31.0	28.0	26.0	31.0	35.0	28.0	22.0	18.0	14.0	20.0	25.0	3.0	1.0
247	Kapar Coal Terminal	MYBTB	3	5.0	N	101	18.0	E	2	29.0	31.0	28.0	27.0	30.0	35.5	26.0	19.0	17.0	13.0	19.0	24.0	4.1	1.3
248	Pasir Gudang Johor	MYPGU	1	26.0	N	103	55.0	E	3	28.5	31.0	27.5	25.0	31.0	34.3	25.7	21.0	26.0	22.0	27.0	29.0	3.0	0.3
249	Bintulu Sarawak	MYBTU	3	16.0	N	113	4.0	E	3	30.0	31.0	29.0	26.0	30.5	35.0	26.0	23.0	25.0	23.0	26.0	30.0	1.9	0.3
250	Lagos	NGLOS	6	25.0	N	3	25.0	E	5	28.5	30.0	24.0	22.5	31.0	36.0	23.0	19.0	18.0	10.0	30.0	33.0	1.0	0.6
251	Tin Can Island	NGTIN	6	25.0	N	3	18.0	E	5	28.4	29.1	24.5	23.0	31.0	35.0	23.0	20.0	20.0	15.0	31.0	34.0	1.0	0.6
252	Port Harcourt	NGPHC	4	46.2	N	7	0.0	E	5	29.0	31.0	26.0	24.0	31.0	35.0	26.0	24.0	0.0	0.0	4.0	10.0	2.6	1.4
253	Onne	NGONN	4	39.0	N	7	9.0	E	5	29.0	31.0	26.0	24.0	31.0	35.0	26.0	24.0	2.0	0.0	8.0	14.0	2.4	1.0
254	Bonny	NGBON	4	26.0	N	7	9.0	E	5	29.0	31.0	25.0	24.0	30.0	34.0	23.5	20.5	17.0	8.0	27.0	30.0	2.8	1.4
255	Europoort	NLEUR	51	58.0	N	4	8.0	E	5	17.5	19.0	6.5	3.0	21.5	28.0	2.5	-4.0	31.0	29.0	32.0	34.0	2.4	1.3
256	Rotterdam	NLRJT	51	54.0	N	4	29.0	E	5	18.0	20.0	6.5	2.0	21.0	28.0	1.8	-5.0	3.0	0.0	10.0	15.0	1.8	1.4
257	Ijmuiden	NLIJM	52	27.0	N	4	35.0	E	5	17.5	19.0	6.5	3.0	21.0	28.0	1.5	-4.0	31.0	29.0	32.0	33.0	2.6	1.3
258	Amsterdam	NLAMS	52	22.0	N	4	53.0	E	6	18.0	20.0	6.0	1.0	21.0	29.0	1.0	-5.5	0.0	0.0	0.0	0.0	0.0	0.0
259	Flushing (Vlissingen)	NLVLJ	51	27.0	N	3	36.0	E	5	17.5	19.0	6.5	3.0	21.5	28.0	2.0	-4.0	22.0	18.0	28.0	30.0	4.9	2.8
260	Auckland	NZAKL	36	51.0	S	174	48.0	E	3	18.8	22.4	13.0	10.5	17.7	32.2	12.9	0.2	33.5	28.0	35.0	36.0	2.6	1.9
261	Whangerei	NZWRE	35	46.0	S	174	21.0	E	5	19.0	23.0	13.0	11.0	19.5	33.0	13.0	1.0	29.8	22.0	32.0	34.0	3.1	1.5
262	Marsden Point	NZMAP	35	50.0	S	174	30.0	E	2	19.0	22.5	13.0	11.0	19.0	33.0	13.0	1.0	33.0	32.0	34.5	35.7	2.6	1.6
263	Callao (Lima)	PECLL	12	3.0	S	77	10.2	E	3	24.0	26.5	16.0	18.0	28.0	35.0	18.0	14.0	35.0	34.5	35.5	36.0	2.4	1.0
264	Lae	PGLAE	6	44.0	S	146	58.0	E	2	27.0	31.5	25.0	23.0	27.0	36.6	25.0	19.6	22.0	12.0	25.0	30.0	0.9	0.6
265	Port Moresby	PGPOM	9	26.0	S	147	6.0	E	3	28.0	32.0	26.0	24.0	31.0	36.0	24.0	20.0	33.0	31.0	33.5	34.5	2.0	0.9
266	Daru	PGDAU	9	4.0	S	143	12.0	E	1	28.0	31.0	26.0	24.0	32.0	36.0	26.0	21.0	30.0	24.0	32.0	33.5	3.7	1.3
267	Batangas (Luzon)	PHBTG	13	45.0	N	121	3.0	E	2	29.0	32.0	28.0	26.0	28.0	33.5	27.0	22.2	33.0	32.0	34.0	34.7	1.9	1.6
268	Bataan Mariveles	PHBTN	14	30.0	N	120	37.8	E	3	28.0	32.0	27.0	26.0	28.0	33.5	24.0	22.2	33.0	32.0	34.0	34.7	1.9	1.6
269	Limay	PHLIM	14	32.0	N	120	36.0	E	1	28.0	32.0	26.5	25.0	30.0	33.0	24.5	20.0	32.5	32.0	34.0	34.7	1.2	0.4
270	Manila	PHMNL	14	31.0	N	120	37.0	E	2	30.0	34.5	26.0	23.0	27.3	34.0	26.1	20.9	31.0	28.0	33.0	34.7	1.7	0.4
271	Subic Bay (Sana Clara)	PHSFS	14	35.0	N	120	58.0	E	3	29.0	33.0	28.0	26.0	27.4	30.4	24.0	22.4	33.0	32.0	34.0	34.7	1.3	0.9
272	Muhammad Bin Qasim	PKBQM	24	45.6	N	67	21.0	E	5	28.0	30.0	23.0	19.5	29.0	37.0	22.0	10.0	33.0	24.0	36.0	40.0	3.5	1.4
273	Karachi	PKKHI	24	48.0	N	66	59.0	E	4	27.5	30.0	23.0	21.0	29.0	37.0	22.3	10.0	35.0	20.0	37.0	40.0	2.8	1.1
274	Faro	PTFAO	37	0.0	N	7	55.2	W	4	22.0	24.0	17.5	15.5	24.7	37.0	11.0	6.0	35.5	35.0	35.5	36.0	2.5	0.8
275	Lisboa	PTLIS	38	42.0	N	9	6.0	W	5	25.0	34.0	15.0	13.0	24.5	36.0	9.0	5.0	32.0	15.0	35.0	36.0	4.6	2.4
276	Lagos (Portugal)	PTLOS	37	1.0	N	8	40.0	W	5	21.5	23.0	17.0	14.0	24.8	37.0	9.5	5.0	32.0	28.0	36.0	36.5	3.5	1.4
277	Sines	PTSIE	37	58.0	N	8	51.0	W	3	21.0	22.0	16.5	14.5	24.0	36.0	11.0	6.0	35.5	35.0	36.0	36.5	3.9	1.0
278	Doha	QADOH	25	16.8	N	51	33.0	E	3	31.0	35.0	17.0	13.0	35.0	44.0	19.0	12.0	38.5	37.0	40.0	42.0	1.5	0.2
279	Umm Said (Mesaieed)	QAUMS	24	54.0	N	51	34.0	E	3	31.0	35.0	17.0	13.0	35.0	43.0	20.9	12.6	39.0	39.0	40.0	41.0	2.4	0.4
280	Halul Island	QAHAL	25	39.0	N	52	26.0	E	1	30.0	35.2	22.5	16.9	31.0	42.0	21.0	11.0	38.0	37.5	39.0	40.5	0.8	0.1
281	Constanta	ROCND	44	10.0	N	28	39.0	E	3	23.3	24.0	4.9	0.5	22.2	38.0	2.4	-15.0	15.9	15.0	17.2	17.8	0.1	0.0
282	Mangalia	ROMAG	43	49.2	N	28	34.8	E	3	23.5	25.5	5.7	0.5	22.7	38.0	1.4	-15.0	17.0	15.4	17.6	18.0	0.1	0.0
283	Midia	ROMID	44	19.8	N	28	40.8	E	3	23.3	24.5	4.5	0.0	22.5	38.0	2.4	-16.0	15.0	12.0	17.0	17.5	0.1	0.0
284	Novorossiysk, Russia	RUNVS	44	43.2	N	37	46.8	E	3	22.1	26.2	7.9	0.5	21.8	41.0	4.3	-24.0	17.6	9.6	17.8	18.8	0.1	0.0
285	Tuapse, Russia	RUTUA	44	4.8	N	39	4.2	E	3	23.0	27.1	10.1	5.3	21.5	41.0	6.0	-19.0	16.6	12.7	17.3	18.6	0.1	0.0
286	Vladivostok	RUVVO	43	6.6	N	131	53.4	E	2	13.5	15.0	2.5	-1.0	22.0	27.0	-14.5	-25.0	32.0	30.0	33.0	34.0	1.4	0.3
287	Dammam	SADMN	26	30.0	N	50	12.0	E	3	32.0	35.0	17.0	14.0	36.0	48.0	19.0	10.0	41.0	39.0	43.0	45.0	2.3	0.6
288	Jeddah	SAJED	21	28.0	N	39	10.0	E	3	30.0	33.0	22.0	19.0	32.0	39.0	21.0	14.0	38.0	37.0	38.5	39.5	0.2	0.0
289	Jubail	SAJUB	27	3.0	N	49	40.0	E	3	32.0	36.2	16.0	11.6	36.0	47.0	15.0	9.0	49.0	48.0	50.0	52.0	1.2	0.7
290	Al Juaymah Terminal	SAJUT	26	55.2	N	50	1.0	E	1	31.0	34.0	16.8	13.0	36.0	47.0	15.0	9.0	40.0	38.0	42.0	44.0	2.4	1.5
291	Ras Al Khafji	SARAR	28	25.2	N	48	33.0	E	1	32.0	34.9	17.0	15.0	36.8	47.0	14.0	6.0	38.5	38.0	39.0	40.0	1.6	1.0
292	Ras Al Ghar	SAOOI	27	32.0	N	49	13.0	E	1	32.0	34.0	17.0	14.0	36.0	47.0	14.0	8.0	39.0	38.5	40.0	41.0	1.7	0.9
293	Ras Al Tannura	SARLT	26	39.0	N	50	10.0	E	1	31.0	33.8	16.8	13.0	36.0	47.0	21.0	9.0	40.0	38.5	40.5	42.0	2.4	1.5
294	Yanbu	SAYNB	24	5.0	N	38	3.0	E	3	30.0	33.0	22.0	18.0	32.0	39.0	19.0	10.0	39.0	37.5	39.5	41.0	0.8	0.3
295	Marsa Bashayer Oil Terminal	SDMBT	19	24.0	N	37	19.0	E	1	29.5	32.0	23.0	20.0	31.0	40.0	22.0	16.0	37.5	37.0	37.5	38.0	1.2	0.4
296	Port Sudan	SDPZU	19	36.0	N	37	13.0	E	3	31.0	34.0	22.0	19.0	32.0	42.0	21.0	16.0	38.0	37.0	38.5	38.5	1.2	0.4
297	Singapore Jurong	SGJUR	1	18.0	N	103	43.0	E	3	28.5	31.0	27.0	25.0	31.0	34.0	23.0	21.0	29.5	28.5	30.5	31.5	2.3	0.9
298	Singapore Keppel	SGKEP	1	16.2	N	103	52.3	E	3	28.5	31.0	27.0	25.0	31.0	34.0	23.0	21.0	29.5	28.5	30.5	31.5	2.3	0.9
299	Singapore Sembawang Port	SGSEM	1	16.0	N	103	50.0	E	4	29.0	31.5	28.0	25.0	31.0	34.0	23.0	21.0	26.0	21.0	27.0	28.0	2.3	0.9
300	Singapore Singapore	SGSIN	1	20.0	N	103	20.0	E	3	28.5	31.0	27.0	25.0	31.0	34.0	23.0	21.0	29.5	28.5	30.5	31.5	2.3	0.9

Nº	Nome do Porto	CÓDIGO		Lat			Long		PTYPE	MSUWT	USUWT	MWNWT	LWNWT	MSART	USART	MWART	LWART	MWSAL	LWSAL	MDSAL	UDSAL	MSPR	MNER
301	Singapore Pasir Panjan/Tanjung Pagar	SGTPG	1	15.6	N	103	51.0	E	3	28.5	31.0	27.0	25.0	31.0	34.0	23.0	21.0	29.5	28.5	30.5	31.5	2.3	0.9
302	Koper (Slovenia)	SIKOP	45	33.0	N	13	44.0	E	2	24.0	26.5	9.0	7.0	27.0	34.0	3.9	0.5	24.0	18.0	35.0	36.0	0.8	0.2
303	Dakar	SNDKR	14	40.2	N	18	38.4	W	3	26.0	27.5	23.0	20.0	33.5	39.0	21.0	16.0	34.5	34.0	35.0	35.5	2.5	1.2
304	Bangkok	THBKK	13	42.0	N	100	34.0	E	6	29.0	32.5	27.0	24.0	28.0	39.5	25.0	10.5	2.0	0.1	12.0	15.0	1.8	0.3
305	Laem Chabang	THLCH	13	4.0	N	100	50.0	E	3	27.5	30.0	26.0	24.0	28.0	36.0	25.0	14.0	32.0	30.0	33.0	34.0	1.9	1.3
306	Dortyol Oil Terminal	TRDYL	36	51.0	N	36	7.8	E	1	26.2	29.2	18.5	15.5	31.2	38.0	6.5	1.0	38.8	37.5	39.1	39.8	0.3	0.0
307	Eregli	TRERE	41	18.0	N	31	27.0	E	3	23.5	27.5	6.4	4.0	25.5	38.0	5.7	-5.0	17.5	16.9	17.5	18.2	0.2	0.0
308	Istanbul	TRIST	40	59.0	N	29	0.0	E	2	24.4	27.0	6.1	4.0	26.0	37.0	6.6	-11.0	17.5	16.3	17.5	18.3	0.3	0.0
309	Izmir (Smyrna)	TRIZM	38	25.2	N	27	4.2	E	2	24.4	25.5	14.0	11.5	32.0	38.0	6.0	0.0	38.0	37.0	38.7	39.2	0.4	0.0
310	Izmit (Tutuncifilik Oil Terminal)	TRIZT	40	45.0	N	29	55.0	E	2	24.0	27.0	7.0	5.0	26.0	38.0	8.0	-7.0	17.5	16.3	17.5	18.3	0.3	0.0
311	Mersin	TRMER	36	46.0	N	34	39.0	E	3	26.2	28.8	19.0	16.0	30.0	35.0	6.5	1.0	38.8	37.5	39.1	39.8	0.3	0.1
312	Samsun	TRSSX	41	21.0	N	36	34.2	E	3	24.6	28.0	7.6	6.0	25.0	38.0	7.9	-8.0	17.5	16.9	17.5	18.2	0.1	0.0
313	Yarimca	TRYAR	40	46.2	N	29	42.0	E	1	24.0	27.0	7.0	5.0	25.0	38.0	8.0	-7.0	17.5	16.3	17.5	18.3	0.3	0.0
314	Keelung (Chilung)	TWKEL	25	9.0	N	121	44.0	E	3	22.5	25.0	20.0	18.0	26.0	30.0	20.0	16.0	33.0	31.0	34.0	34.5	2.5	0.5
315	Kaohsiung	TWKHH	22	37.0	N	120	15.0	E	3	28.0	31.3	23.0	21.0	27.3	30.8	21.5	18.0	34.5	34.0	35.0	35.5	1.0	0.3
316	Taichung	TWTXG	23	17.0	N	120	30.0	E	3	27.0	30.0	18.0	16.0	26.6	30.1	18.6	15.6	26.0	17.0	33.0	34.5	4.8	3.9
317	Dar Es Salaam	TZDAR	7	10.0	S	39	17.0	E	2	29.0	32.0	26.0	24.0	31.0	37.0	22.0	17.0	30.0	20.0	34.0	35.5	3.8	1.5
318	Odessa	UAODS	46	30.0	N	30	52.8	E	3	18.4	24.5	2.6	-3.1	20.2	37.0	0.3	-27.0	13.5	5.4	16.1	20.0	0.1	0.0
319	Dnepro-Bugsky (Ochakov)	UADNB	46	45.0	N	31	55.0	E	5	21.4	26.0	1.3	-0.6	20.3	40.0	-0.6	-29.0	5.2	0.5	3.0	12.5	0.1	0.0
320	Ilyichevsk	UAILK	46	20.0	N	30	39.0	E	3	18.9	23.6	2.6	-0.7	22.6	38.0	0.5	-22.0	13.8	8.8	14.5	18.2	0.1	0.0
321	Nicolayev	UANIK	46	55.8	N	30	39.0	E	6	21.4	26.0	1.3	-0.6	20.4	39.0	-0.7	-30.0	0.1	0.1	0.2	0.5	0.0	0.0
322	Sevastopol	UASVP	44	37.2	N	33	31.8	E	2	21.5	25.9	7.8	1.7	20.5	38.0	4.3	-22.0	18.0	15.2	18.0	19.8	0.2	0.1
323	Boston Massachusetts	USBOS	42	21.0	N	71	4.8	W	2	14.0	17.0	0.5	-2.0	26.2	33.0	-5.0	-14.0	26.0	18.0	29.0	31.0	3.3	1.4
324	New York New York (New Jersey)	USNYC	40	42.0	N	74	1.0	W	5	14.0	20.0	2.0	-1.0	27.5	37.0	-3.0	-16.0	8.0	0.0	22.0	30.0	1.6	1.3
325	Philadelphia Pennsylvania (Port Richmond)	USPHL	39	57.0	N	75	10.0	W	5	18.0	22.0	2.0	-1.0	28.0	36.0	-3.0	-12.0	0.0	0.0	1.0	3.0	1.9	1.6
326	Wilmington Delaware	USILG	39	45.0	N	75	30.0	W	5	18.0	22.0	2.0	-1.0	28.0	36.0	-3.0	-12.0	0.0	0.0	3.0	6.0	1.8	1.6
327	Baltimore Maryland	USBAL	39	16.8	N	76	34.8	W	5	20.0	24.0	2.5	0.0	30.5	38.0	-1.0	-10.0	0.0	0.0	4.0	8.0	0.4	0.3
328	Hampton Roads	USPHF	36	58.0	N	76	20.0	W	5	23.0	27.0	12.0	7.0	29.1	36.0	0.5	-5.0	21.0	15.0	26.0	31.0	1.1	0.7
329	Norfolk-Newport News Virginia	USNEN	36	51.0	N	76	19.0	W	5	23.0	27.0	11.0	6.0	29.1	36.0	0.5	-5.0	21.0	15.0	26.0	31.0	1.2	0.8
330	Savannah Georgia	USSAV	32	5.0	N	81	5.0	W	5	27.0	30.0	19.0	16.0	31.8	37.0	4.8	-2.0	18.0	10.0	28.0	33.0	3.0	2.2
331	Mobile Alabama	USMOB	30	40.0	N	88	1.8	W	5	27.0	31.0	16.0	9.0	32.5	38.0	5.6	2.0	8.0	0.0	30.0	35.0	1.0	0.3
332	Lake Charles Louisiana	USLCH	30	13.2	N	93	13.2	W	5	27.0	29.0	20.0	15.0	32.0	39.0	6.0	1.0	0.0	0.0	7.0	13.0	0.5	0.0
333	Davant	USDVT	29	36.0	N	89	51.0	W	6	27.0	31.0	12.0	10.0	32.5	38.0	5.6	2.0	0.0	0.0	0.0	0.0	0.2	0.0
334	New Orleans	USMSY	29	57.0	N	90	4.0	W	6	27.0	31.0	17.5	15.0	32.0	39.0	7.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0
335	LOOP Terminal	USLOP	28	52.8	N	90	1.2	W	1	27.0	29.0	20.0	17.0	29.0	38.0	14.0	5.5	30.0	24.0	31.0	34.0	0.4	0.2
336	Sabine	USSAB	29	42.0	N	93	52.0	W	2	27.0	29.0	19.5	16.5	31.0	38.0	8.0	3.0	35.5	35.0	36.0	36.5	0.5	0.3
337	Beaumont	USBPT	30	5.0	N	94	5.0	W	6	28.5	32.0	16.0	13.0	33.0	41.0	5.8	1.5	0.0	0.0	5.0	10.0	0.5	0.1
338	Galveston Texas	USGLS	29	17.0	N	94	50.0	W	2	28.5	33.0	18.0	16.0	32.0	40.5	9.5	2.0	18.0	14.0	26.0	33.0	0.5	0.1
339	Texas City Texas	USTXT	29	23.0	N	94	54.0	W	2	28.5	34.0	18.0	16.0	32.0	40.5	9.5	2.0	18.0	14.0	26.0	32.0	0.4	0.1
340	Houston Texas	USHOU	29	45.0	N	95	19.8	W	5	28.5	32.0	16.0	14.0	33.0	41.0	5.8	1.5	2.0	0.0	10.0	18.0	0.4	0.1
341	Anchorage Alaska	USANC	61	13.8	N	149	52.8	W	5	8.0	12.0	1.0	-1.0	17.0	24.0	-12.0	-19.0	2.0	0.0	8.0	12.0	8.8	5.0
342	Portland Oregon	USPDX	45	35.0	N	122	44.0	W	6	12.0	14.0	1.0	-2.0	25.5	32.0	1.5	-3.0	0.0	0.0	0.0	0.0	0.0	0.0
343	Vancouver Washington	USBCC	45	36.0	N	122	40.0	W	6	12.0	14.5	1.0	-2.0	25.4	32.0	1.1	-3.0	0.0	0.0	0.0	0.0	0.0	0.0
344	San Francisco California	USSFO	37	48.0	N	122	25.2	W	5	15.0	20.0	12.0	11.0	22.5	31.0	6.0	2.0	28.0	10.0	30.0	32.0	2.0	0.6
345	Oakland California	USOAK	37	49.8	N	122	18.0	W	5	15.0	20.0	12.0	11.0	22.5	31.0	6.0	2.0	15.0	5.0	27.0	30.0	2.0	0.7
346	Long Beach California	USLGB	33	45.0	N	118	12.0	W	3	18.0	22.0	14.0	12.0	27.0	34.0	7.6	1.0	33.0	31.0	33.2	33.8	2.0	0.6
347	San Diego	USSAN	32	42.0	N	117	10.2	W	2	18.0	22.0	15.0	13.0	25.1	34.0	9.8	3.0	34.0	33.5	35.0	37.0	2.2	1.4
348	Montevideo	UYMVD	34	54.0	S	56	13.2	W	3	25.0	26.8	15.0	12.5	26.0	34.0	5.0	0.0	5.0	1.0	10.0	30.0	0.3	0.1
349	Aden (Yemen)	YEADE	12	48.0	N	44	54.0	E	2	29.0	31.0	23.0	21.0	33.0	39.0	26.3	16.0	36.0	35.5	36.1	36.5	1.5	0.5
350	Hodeidah (Yemen)	YEHOD	14	48.0	N	42	55.0	E	2	29.5	32.0	23.5	20.5	32.0	39.0	26.3	16.0	36.0	35.2	36.5	37.5	1.2	0.2
351	Al Mukallah (Yemen)	YEMKX	14	31.0	N	49	9.0	E	3	29.5	32.0	23.5	20.5	32.0	39.0	26.3	16.0	36.9	37.5	36.1	36.5	1.2	0.4
352	Ras Isa Marine Terminal (Yemen)	YERAI	15	7.8	N	42	36.0	E	1	29.5	32.0	23.5	20.5	32.0	39.0	26.3	16.0	36.0	35.5	36.5	37.0	1.1	0.3
353	Cape Town	ZACPT	33	54.0	S	18	26.0	E	3	14.0	16.0	13.5	11.5	18.2	26.0	11.0	9.0	20.0	10.0	34.3	34.8	1.5	0.6
354	Durban	ZADUR	29	53.0	S	31	2.0	E	5	24.5	25.5	21.0	19.0	26.3	26.1	15.4	14.5	28.0	18.0	35.5	35.5	1.8	0.5
355	Port Elizabeth	ZAPLZ	33	58.0	S	25	38.0	E	3	20.5	25.5	16.5	12.5	23.6	25.0	10.0	8.5	35.2	34.9	34.9	35.0	1.6	0.5
356	Richards Bay	ZARCB	28	48.0	S	32	3.0	E	4	25.5	26.5	22.0	20.0	28.0	29.5	15.5	12.0	39.2	37.0	3			

Nº	CÓDIGO	D6MRF	W6MRF	RNFL75	DISRVM	SIZRVC	INASMW	INARKW	INAWP	INHSTM	INSNDB	INSTNB	INLTMF	INMANG	INRKSH	SUFSND	SUSFTM	SUSGRM	SURKRF	SUCORF
1	BRSPB	750	750	7	5	2,500	5	3	3	1	4	0	4	4	4	3	5	0	4	0
2	AEAUH	3	92	4	8	30	5	5	4	3	4	2	4	4	4	5	2	4	4	4
3	AEMZD	3	92	4	8	30	5	5	4	3	4	2	4	4	4	5	2	4	4	4
4	AEDAS	5	76	4	110	120	5	4	0	1	4	0	1	1	4	5	3	4	4	4
5	AEPR	2	100	4	5	25	5	4	3	4	1	2	4	4	4	2	5	2	5	4
6	AEDXB	2	100	4	5	25	5	4	3	4	1	2	4	4	4	2	5	2	5	4
7	AEFAT	5	75	4	165	20	5	0	0	0	2	0	0	0	2	3	5	1	1	1
8	AEFJR	4	80	4	65	540	5	5	3	0	4	4	0	0	4	5	4	4	3	1
9	AEJEA	2	90	5	70	2,100	5	5	5	2	4	3	2	2	4	5	2	4	3	2
10	AEJED	5	75	4	20	120	5	4	1	4	4	0	4	4	4	4	5	4	2	2
11	AEKLF	4	80	4	42	540	5	4	2	0	3	3	0	0	4	5	4	4	3	1
12	AEQIW	2	90	5	22	1,000	5	4	4	4	5	4	4	4	4	5	2	4	3	3
13	AERUW	5	75	4	20	120	5	4	1	4	4	0	4	4	4	4	5	4	2	2
14	AESHJ	2	98	5	20	100	5	4	3	4	2	3	4	4	3	5	0	2	5	3
15	AEZIR	5	75	4	85	120	5	4	0	1	4	0	1	1	4	5	3	4	4	4
16	ARBUE	424	581	9	-155	600,000	5	4	4	0	4	2	4	0	1	4	5	3	1	0
17	ARCMF	424	581	9	-250	280,000	5	4	5	0	1	4	1	0	0	5	5	0	0	0
18	AUDAM	41	231	5	180	105,000	5	5	3	4	4	0	4	3	5	4	5	4	5	2
19	AUPWL	41	231	5	225	105,000	5	4	2	2	3	0	3	3	5	5	3	4	4	2
20	AUPBY	161	355	7	65	550	5	4	3	2	3	0	2	3	4	5	3	4	4	0
21	AUWYA	119	151	9	40	990	5	4	4	2	5	0	2	3	4	5	3	4	3	0
22	AUPPI	125	219	8	30	990	5	2	3	3	5	0	5	4	3	5	3	4	3	0
23	AUPST	280	440	6	30	1,500	5	5	0	1	4	0	3	0	4	5	3	4	3	0
24	AUWEP	245	302	9	25	900	5	1	4	3	4	1	4	0	4	5	5	3	4	0
25	AUPKL	457	913	8	12	400	5	5	4	1	3	0	2	1	4	5	5	4	5	0
26	AUBNE	369	779	7	-2	6,600	5	5	4	4	4	3	4	4	3	5	5	4	3	4
27	AUBDB	323	820	7	-5	3,300	5	3	4	4	2	3	1	1	3	3	5	2	3	5
28	AUGLT	244	704	7	4	9,000	5	5	3	4	4	0	5	4	5	5	5	4	5	4
29	AUPTL	245	558	7	-16	143,000	5	2	4	4	2	4	1	1	4	2	5	4	4	4
30	AUHPT	297	1312	6	6	500	5	4	3	2	2	2	2	2	2	5	4	2	2	4
31	AUPDT	297	1312	6	6	500	5	4	3	2	2	2	2	2	2	5	4	2	2	4
32	AUMKY	297	1312	6	5	2,500	5	4	5	2	1	2	2	3	2	3	5	2	2	2
33	AUABP	158	853	6	18	2,768	5	4	2	2	2	2	3	3	2	5	4	1	2	4
34	AUTSV	119	990	6	-1	200	5	4	5	3	2	2	2	2	2	3	5	1	2	3
35	AULUC	204	742	6	6	8,814	5	3	3	3	2	3	2	3	3	5	3	2	3	4
36	AUMOU	909	2643	7	-1	1,600	5	2	4	3	1	1	1	1	1	2	5	1	1	3
37	AUCNS	279	1726	6	-7	300	5	2	5	4	3	1	1	1	1	3	5	1	4	4
38	AUCQP	224	1586	6	14	114	5	4	0	1	1	1	3	3	1	4	1	3	1	4
39	AUWEI	58	1687	5	-5	4,107	5	3	4	4	1	2	2	1	2	2	5	2	2	3
40	AUKRB	36	884	5	-3	121,290	5	0	5	4	2	5	1	1	5	4	5	2	5	5
41	BDCGP	149	1484	4	-3	1,200,000	5	2	5	0	0	0	5	3	0	2	5	3	1	0
42	BEANR	334	460	8	-75	4,300	5	5	5	0	1	1	2	0	0	2	5	0	0	0
43	BEGNE	334	460	8	-51	2,150	5	5	5	0	2	2	3	0	0	2	5	1	0	0
44	BGBOJ	285	294	9	320	817,000	5	5	4	4	4	3	3	0	4	4	5	0	4	0
45	BGVAR	246	282	8	250	817,000	5	5	4	3	3	3	2	0	2	4	5	0	4	0
46	BHMAN	2	72	4	90	50	5	4	3	3	4	3	3	3	4	5	3	4	3	4
47	BHMIN	2	72	4	90	50	5	5	3	3	4	3	3	3	4	5	3	4	3	4
48	BRITJ	584	961	8	-3	15,500	5	4	5	4	4	0	5	5	4	4	2	0	4	0
49	BRPNG	648	1288	8	-15	797	5	5	3	0	4	0	4	5	4	4	4	0	5	0
50	BRSSZ	738	1343	7	-3	154	5	5	4	0	4	0	5	5	4	4	5	0	3	0
51	BRRIO	750	750	7	1	30	5	4	4	1	4	0	4	4	3	3	5	0	5	0
52	BRPOU	446	829	7	65	1,400	5	5	0	2	5	0	1	1	1	5	2	0	1	0
53	BRVIX	446	829	7	-6	1,400	5	5	5	0	3	0	5	4	4	3	5	0	5	0
54	BRPRM	446	829	7	4	1,400	5	5	2	1	4	0	4	3	5	4	5	0	4	0
55	BRTUB	446	829	7	4	1,400	5	5	2	0	4	0	4	3	5	4	5	0	4	0
56	BRSSA	727	1384	8	4	35,000	5	5	2	0	4	0	2	2	5	4	5	0	4	0
57	CACBC	530	660	8	5	200	5	0	0	0	0	4	0	0	5	3	3	0	5	0
58	CASEI	510	596	9	6	6,000	5	4	3	0	3	4	1	0	5	3	4	0	5	0
59	CAHAL	603	793	8	1	400	5	4	4	0	0	4	0	0	5	3	5	0	5	0
60	CALHA	603	793	8	-3	2,000	5	1	4	0	0	5	0	0	5	3	3	0	5	0

Nº	CÓDIGO	D6MRF	W6MRF	RNFL75	DISRVM	SIZRVC	INASMW	INARKW	INAWP	INHSTM	INSNDB	INSTNB	INLTMF	INMANG	INRKSH	SUFSND	SUSFTM	SUSGRM	SURKRF	SUCORF
61	CAVAN	300	807	6	0	85,000	5	4	4	0	2	4	3	0	4	4	5	0	5	0
62	CARBK	300	807	6	0	85,000	5	3	3	0	3	4	3	0	4	4	4	0	5	0
63	CNCAN	341	1362	6	-112	400,000	5	4	4	4	1	0	5	1	2	2	5	1	1	0
64	CNCWN	320	1604	6	4	1,000,000	5	3	5	0	3	2	1	3	3	2	5	2	3	2
65	CNDLC	183	457	7	35	1,500	5	5	5	2	3	4	4	0	5	4	5	2	4	0
66	CNHUA	326	1606	6	-95	452,600	4	3	5	4	4	2	5	1	2	3	5	1	2	1
67	CNNBO	464	947	8	-3	600	5	2	5	3	4	2	5	0	4	4	5	2	4	0
68	CNNGB	464	947	8	-3	600	5	2	5	3	4	2	5	0	4	4	5	2	4	0
69	CNSHA	480	840	8	-40	1,500,000	5	3	5	2	1	0	5	0	1	3	5	1	1	0
70	CNSHB	382	742	7	-45	1,500,000	5	4	4	2	1	0	5	0	1	3	5	1	1	0
71	CNTAO	192	577	6	24	8,800	5	5	4	2	2	1	3	0	4	3	5	3	4	0
72	CNTSN	278	603	8	0	71,600	5	5	5	2	3	1	3	0	2	4	5	2	1	0
73	CNYNT	190	500	7	30	1,200	5	5	4	0	3	4	4	0	5	4	5	2	4	0
74	COCAR	80	863	3	15	1,400	5	4	2	3	4	0	4	4	3	3	5	4	3	3
75	CYKYR	139	958	5	120	40	5	5	5	0	3	4	1	0	4	4	2	4	5	0
76	CYLCA	139	958	5	25	50	5	5	5	0	4	4	1	0	4	4	2	4	5	0
77	CYLSM	41	411	4	3	60	5	5	5	0	4	4	1	0	4	4	2	4	5	0
78	DEBRE	322	428	8	-15	6,500	5	2	5	0	1	2	5	0	2	2	5	1	2	0
79	DEHAM	429	325	9	-105	9,000	5	4	5	0	1	2	5	0	1	2	5	0	0	0
80	DEWVN	322	428	8	1	750	5	5	4	2	3	2	5	0	1	4	5	2	1	0
81	DJIB	14	33	6	5	900	5	5	4	2	4	3	2	2	4	5	3	3	0	4
82	DKENS	293	411	9	2	200	5	2	3	1	2	4	2	0	5	2	5	1	5	0
83	DKFRC	217	331	9	8	300	5	4	4	1	2	4	2	0	4	2	5	1	5	0
84	EGALS	3	29	4	3	200	5	3	0	2	4	0	2	2	4	4	5	4	2	0
85	EGALY	10	186	4	45	2,000	5	4	0	2	4	3	0	0	3	5	5	3	3	0
86	EGDAM	7	100	4	7	3,000	5	5	0	2	4	4	2	0	0	5	0	4	0	0
87	EGEDK	10	186	4	52	2,000	5	4	0	2	4	3	0	0	3	5	5	3	3	0
88	EGPSD	6	190	4	270	2,000,000	5	5	4	4	3	1	4	0	2	4	5	3	3	0
89	EGSUZ	5	100	4	63	200	5	5	4	2	4	0	3	3	4	5	3	1	2	1
90	ESGIJ	425	670	9	6	40	5	5	5	1	4	3	3	0	4	5	5	2	4	0
91	ESBIO	436	655	9	-2	4,500	5	4	4	2	3	3	5	0	3	3	5	2	2	0
92	ESVGO	503	1303	6	14	400	5	4	3	2	3	4	2	0	5	5	5	2	5	0
93	ESBCN	349	241	9	11	5,000	5	5	5	0	3	2	2	0	3	5	5	3	3	0
94	ESVLC	150	318	6	4	550	5	5	4	2	5	3	0	0	1	4	5	4	2	0
95	ESALG	690	146	5	35	1,600	5	5	5	0	5	5	0	0	4	3	5	3	5	0
96	ESLPA	17	159	4	1	60	5	5	3	0	4	3	0	0	4	5	5	3	5	0
97	ESSCT	46	396	9	2	20	5	5	3	0	2	3	0	0	4	5	5	3	5	0
98	ESTAR	349	241	9	75	6,000	5	5	4	1	3	3	4	0	3	5	5	3	3	0
99	FRDKK	264	347	9	27	1,800	5	5	3	2	4	4	4	0	0	4	5	0	0	0
100	FRBES	404	724	8	20	600	5	4	4	2	3	4	3	0	4	5	5	3	5	0
101	FRDON	336	475	8	-3	1,300	5	4	4	0	4	4	4	0	3	4	5	3	3	0
102	FRFOS	195	387	7	0	3,000	5	4	3	2	3	3	0	0	3	4	5	3	3	0
103	FRLAV	195	387	7	0	3,000	5	4	4	2	3	3	0	0	3	4	5	3	3	0
104	FRLEH	405	723	7	0	6,500	5	4	4	0	4	4	4	0	4	4	5	3	3	0
105	FRMRS	195	387	7	0	3,000	5	4	5	2	3	3	0	0	3	4	5	3	3	0
106	GBHST	443	662	8	30	2,400	5	4	3	1	5	3	4	0	4	5	5	0	4	0
107	GBIMM	271	330	9	-20	4,900	5	5	4	2	3	3	5	0	0	3	5	0	0	0
108	GB001	309	507	9	0	180	5	5	5	4	4	5	4	0	3	4	5	0	3	0
109	GBPTB	385	534	9	2	280	5	5	5	2	4	4	4	0	2	4	5	0	2	0
110	GBRER	285	330	9	-4	3,000	5	5	4	3	3	4	4	0	3	4	5	0	3	0
111	GEBUS	539	978	7	7	22,000	5	5	4	3	4	4	3	0	4	4	5	0	4	0
112	GEPTI	537	840	8	4	13,300	5	5	4	3	4	1	3	0	2	4	5	0	1	0
113	GIGIB	690	146	5	25	1,600	5	5	3	0	4	4	0	0	4	3	5	3	5	0
114	GRASS	180	278	6	4	5,600	5	5	4	0	5	4	2	0	4	5	4	2	3	0
115	GREEU	69	302	6	5	961	5	5	4	0	4	4	2	0	4	4	5	2	4	0
116	GRJKH	181	278	6	93	900	5	5	5	0	4	5	1	0	5	4	3	2	5	0
117	GRPAC	69	302	6	34	855	5	4	1	0	2	2	2	0	5	4	4	2	5	0
118	GRPIR	69	302	6	14	320	5	5	4	0	4	3	3	0	4	5	3	3	4	0
119	GRSKI	181	278	7	6	1,800	5	5	5	0	4	4	4	0	3	4	5	3	4	0
120	GRVOL	181	278	6	0	64	5	5	4	0	4	4	2	0	4	4	5	3	4	0

Nº	CÓDIGO	D6MRF	W6MRF	RNFL75	DISRVM	SIZRVC	INASMW	INARKW	INAWP	INHSTM	INSNDB	INSTNB	INLTMF	INMANG	INRKSH	SUFSND	SUSFTM	SUSGRM	SURKRF	SUCORF
121	HKHKG	206	2520	5	33	1,000,000	5	5	3	1	3	3	3	3	3	3	4	3	4	3
122	HKKWN	206	2520	5	33	1,000,000	5	5	5	1	2	2	4	3	3	3	5	3	3	3
123	HROMI	470	570	9	18	250	5	0	2	0	1	2	0	0	5	3	2	1	5	0
124	IDBLW	960	1150	9	12	550	5	0	4	0	3	0	5	5	3	4	5	3	3	2
125	IDDUM	985	1287	9	42	20,000	5	2	4	0	3	0	5	4	2	3	5	4	1	1
126	IDCIG	286	899	7	5	50	5	3	5	1	1	1	1	1	1	4	4	3	2	2
127	IDMRK	480	1335	7	20	700	5	4	4	0	3	1	2	2	4	5	3	3	5	4
128	IDJKT	268	1434	6	5	300	5	4	5	0	4	3	1	3	3	3	5	4	3	3
129	IDCXP	1325	2172	8	-1	900	3	2	3	1	3	2	1	1	2	5	3	4	3	3
130	IDSRG	312	1390	6	2	200	5	4	4	1	3	5	1	2	5	3	5	2	5	4
131	IDSUB	203	1277	6	1	200	5	3	4	0	1	2	1	2	2	4	4	2	2	2
132	IDTBA	1272	1494	9	80	100,000	5	0	0	0	3	0	2	3	4	5	4	4	4	4
133	IDBPN	1272	1494	9	100	100,000	5	4	5	0	3	0	4	4	4	4	5	3	3	3
134	IDAMA	1203	1330	9	0	500	5	4	5	1	1	5	1	1	5	5	3	3	4	4
135	IEMOT	544	878	8	-18	4,000	5	2	5	2	3	4	4	0	5	4	5	0	5	0
136	ILASH	7	100	4	3	20	5	5	0	0	4	4	0	0	0	5	0	4	0	0
137	INBOM	287	2246	2	10	9,800	5	1	2	2	0	2	0	5	4	4	4	0	4	0
138	INCCU	149	1484	4	-140	1,200,000	5	2	5	1	0	0	5	1	0	0	5	0	0	0
139	INCOK	498	2417	3	-3	6,170	5	5	5	0	5	0	5	5	1	2	2	0	1	0
140	INHAL	149	1484	4	-36	1,200,000	5	4	5	2	3	0	5	3	1	3	5	1	0	0
141	INIXE	268	2739	2	15	2,500	5	4	4	0	5	3	3	3	3	4	4	0	2	0
142	INIXY	3	338	2	-40	150,000	5	1	4	3	2	0	5	5	2	2	5	0	2	3
143	INMAA	341	863	7	110	50,000	5	5	2	0	5	0	1	1	1	4	5	1	1	0
144	INMRM	49	2915	4	0	2,500	5	5	4	3	4	3	4	3	5	4	5	0	4	0
145	INMUN	6	485	2	10	1,100	5	5	5	4	4	0	5	4	1	0	5	2	1	1
146	INPBD	350	1500	3	18	1,200	5	5	4	3	3	3	2	2	4	3	5	2	4	0
147	INPRT	198	1551	4	5	132,100	5	5	2	0	5	0	4	4	1	4	4	0	2	0
148	INSAL	150	900	3	2	800	5	4	4	3	3	2	5	4	3	3	5	2	4	0
149	INSIK	6	485	2	14	1,100	5	4	3	2	0	0	4	4	1	0	5	2	1	5
150	INTUT	158	506	3	15	14,400	5	5	3	0	5	0	4	4	5	4	4	2	5	3
151	INVAD	150	900	3	2	800	4	4	3	3	3	2	3	3	3	4	5	2	4	0
152	INVTZ	78	799	4	15	113,000	5	5	2	0	4	0	1	1	4	3	3	0	4	4
153	IRBKM	2	190	3	144	500,000	5	0	5	2	0	2	5	0	0	0	5	1	0	0
154	IRBMR	2	190	3	100	500,000	5	0	5	2	0	2	5	0	0	0	5	1	0	0
155	IRBND	11	172	4	30	42,000	5	5	0	3	3	1	5	2	1	5	5	2	2	2
156	IRBUZ	5	160	4	25	12,000	5	0	3	3	4	2	5	1	0	0	4	5	3	0
157	IRKHK	2	154	4	31	12,000	5	5	4	2	4	4	2	2	4	5	3	4	0	5
158	IRLVP	0	84	3	140	42,000	5	0	0	0	4	4	0	0	4	5	4	4	0	4
159	IRSSI	0	84	3	140	42,000	5	0	0	0	4	4	0	0	4	5	4	4	0	4
160	ISHAF	220	330	8	4	200	5	5	4	0	0	4	0	0	4	0	3	0	5	0
161	ISSTR	220	330	8	1	200	5	5	4	0	0	4	0	0	4	0	3	0	5	0
162	ITGOA	451	825	6	28	600	5	4	4	0	3	4	0	0	4	3	2	3	4	0
163	ITPEX	294	641	7	20	400	5	3	0	2	4	3	1	0	0	4	3	2	3	0
164	ITLIV	343	565	8	20	4,100	5	5	5	2	4	4	0	0	3	3	5	3	4	0
165	ITRAN	346	411	9	0	2,000	5	5	4	0	3	3	4	0	3	3	5	2	4	0
166	ITTAR	132	308	7	8	550	5	4	3	2	3	2	2	0	4	3	5	3	4	0
167	ITVCE	438	604	8	-2	1,200	5	5	4	2	3	1	3	0	3	3	5	3	2	0
168	ITTRS	469	573	9	20	3,000	5	4	4	0	3	3	3	0	3	3	5	2	3	0
169	JPABO	400	900	7	3	480	5	5	3	2	2	4	2	0	4	3	4	3	4	0
170	JPAMA	431	885	7	0	7,600	5	4	2	0	3	4	2	0	3	3	2	3	4	0
171	JPBEP	465	1176	6	10	1,500	4	3	2	1	3	4	2	0	4	3	4	2	5	0
172	JPCBH	580	1100	7	18	880	5	4	4	1	2	3	3	0	3	3	5	2	3	1
173	JPKMT	580	1100	7	11	880	5	4	4	1	2	3	3	0	3	3	5	2	3	1
174	JPFKY	342	834	7	5	649	5	3	2	1	3	2	2	0	2	3	3	2	5	0
175	JPHHR	390	950	6	10	1,656	5	4	3	2	3	4	3	0	4	3	4	3	4	0
176	JPHIM	400	900	7	10	480	5	5	4	2	2	3	2	0	3	3	5	3	4	0
177	JPHKT	509	1095	7	6	280	5	3	2	1	2	3	2	0	3	3	2	2	5	0
178	JPIMB	450	899	4	2	70	5	3	4	2	3	3	4	0	3	3	5	3	4	0
179	JPINS	340	850	6	15	60	5	3	4	2	3	3	4	0	3	3	5	3	4	0
180	JPIWK	499	1045	7	3	260	5	4	4	2	3	3	4	0	3	4	5	3	4	0

Nº	CÓDIGO	D6MRF	W6MRF	RNFL75	DISRVM	SIZRVC	INASMW	INARKW	INAWP	INHTSM	INSNDB	INSTNB	INLTMF	INMANG	INRKSH	SUFSND	SUSFTM	SUSGRM	SURKRF	SUCORF
181	JPCCZ	798	1841	4	-4	640	5	3	4	3	3	3	4	0	4	3	5	3	4	3
182	JPKGA	402	916	7	1	1,656	5	4	2	2	2	3	2	0	3	3	4	3	4	0
183	JPKII	632	1607	6	20	100	4	4	2	2	3	4	2	3	4	4	3	2	5	2
184	JPKIJ	724	1065	6	-2	1,800	5	4	4	0	3	2	3	0	5	3	5	2	4	0
185	JPKIK	450	899	4	1	10	5	3	4	2	3	3	4	0	3	3	5	3	4	0
186	JPKIN	818	1320	7	5	10	5	4	2	2	4	4	3	4	4	4	5	4	0	4
187	JPKND	554	1106	7	4	40	4	3	2	1	3	4	2	0	4	3	2	2	5	0
188	JPKNU	478	1057	7	-2	350	5	4	4	2	3	2	3	0	3	3	5	3	4	0
189	JPKOJ	632	1607	6	2	260	5	4	4	2	3	4	2	2	4	3	3	2	5	3
190	JPKSM	683	983	8	32	1,800	5	4	2	1	4	3	1	0	2	5	4	2	3	0
191	JPKUD	635	1210	7	1	30	5	4	3	1	3	1	3	0	5	3	5	3	5	0
192	JPKWS	570	1100	7	4	2,200	5	4	4	1	2	3	3	0	3	3	5	2	3	1
193	JPMAL	781	1020	8	12	1,600	5	4	4	0	2	4	3	0	5	2	5	2	4	0
194	JPMIZ	339	821	7	2	1,990	5	4	2	1	3	2	3	0	3	4	5	4	4	0
195	JPMOJ	554	1106	7	13	300	5	3	1	1	2	2	2	0	2	3	3	3	5	0
196	JPMUR	724	1065	6	15	740	5	4	4	0	3	3	3	0	5	3	5	2	4	0
197	JPMYJ	450	899	4	12	270	5	3	4	2	3	3	4	0	3	3	5	3	4	0
198	JPNAL	818	1320	7	2	15	5	4	4	3	4	2	3	3	3	4	5	4	0	4
199	JPNGI	548	1021	7	3	230	5	4	4	1	4	3	3	0	3	3	5	3	4	1
200	JPNGO	478	1057	7	2	900	5	4	4	2	4	1	3	0	3	3	5	4	5	0
201	JPNGS	562	1417	6	-5	480	5	3	5	0	3	4	3	0	4	3	5	3	5	1
202	JPOIT	465	1176	6	3	2,150	4	3	2	1	3	4	2	0	4	3	4	2	5	0
203	JPOKA	818	1320	7	2	15	5	4	4	3	4	2	3	3	3	4	5	4	0	4
204	JPONO	340	850	6	2	60	5	3	5	2	3	3	4	0	3	4	5	3	4	0
205	JPOSA	430	880	7	0	16,000	5	4	3	0	3	4	2	0	3	3	5	2	3	0
206	JPSAE	465	1176	6	18	2,150	4	3	2	1	3	4	2	0	4	3	4	2	5	0
207	JPSAG	465	1176	6	18	2,150	4	3	2	1	3	4	2	0	4	3	3	2	5	0
208	JPSAK	420	850	7	1	1,600	5	4	4	0	3	4	2	0	4	4	4	3	4	0
209	JPSBS	632	1607	6	2	140	5	4	4	2	3	4	2	3	4	4	3	2	5	2
210	JPSKD	358	775	7	4	120	4	4	3	0	2	2	2	0	2	4	4	2	5	0
211	JPSMN	848	1075	9	11	280	5	4	4	3	4	3	3	0	4	4	4	2	3	0
212	JPSMT	456	886	7	3	25	5	4	4	2	4	1	3	0	3	3	5	4	4	1
213	JPSMZ	738	1356	7	1	900	5	4	4	1	3	4	3	0	4	4	5	3	4	0
214	JPTAM	339	821	7	23	1,990	5	4	2	1	3	2	3	0	3	4	5	4	5	0
215	JPTBT	554	1106	7	5	10	5	3	4	1	3	2	4	0	2	4	4	4	5	0
216	JPTKY	631	1213	7	1	70	5	4	3	1	3	1	4	0	5	3	5	3	5	0
217	JPTMK	724	1065	6	0	10	5	4	4	0	3	3	3	0	5	3	5	2	4	0
218	JPTOY	327	739	7	3	4,800	5	4	4	0	3	2	3	0	3	3	5	2	3	0
219	JPTYO	585	1110	7	0	3,200	5	4	5	1	2	3	3	0	2	2	5	2	3	1
220	JPUBJ	631	1213	7	2	85	5	4	3	0	3	2	3	0	2	4	3	4	5	0
221	JPUBK	295	1021	6	2	50	5	3	4	0	3	4	4	0	4	3	5	4	4	0
222	JPWAK	420	850	7	2	3100	5	4	4	2	4	1	3	0	4	3	5	4	5	0
223	JPYKK	505	1082	7	13	900	5	4	4	2	3	1	3	0	3	3	5	4	5	0
224	JPYOK	548	1021	7	3	230	5	4	4	1	4	3	3	0	3	3	5	3	4	1
225	JPYOS	550	1000	7	20	230	5	4	4	1	4	3	2	0	4	4	5	4	5	1
226	KEMBA	355	789	8	100	9,000	5	4	5	4	4	2	4	4	4	5	5	3	3	3
227	KRKAN	344	958	7	7	2,100	5	3	4	1	3	3	2	0	4	4	4	3	2	0
228	KRKPO	284	774	7	7	2,300	5	4	2	2	3	2	3	0	2	3	5	3	1	0
229	KRKUV	344	931	4	0	500	5	4	3	1	2	3	2	0	4	3	5	3	4	0
230	KRMOK	344	931	4	7	1,000	5	4	3	2	2	2	2	0	5	3	5	2	4	0
231	KRONS	308	1175	6	3	900	5	4	3	2	2	2	2	0	1	3	3	4	4	0
232	KRPUS	356	1032	7	15	1,690,000	5	4	2	0	3	2	4	0	2	3	2	3	2	0
233	KRSCP	361	933	7	16	400	5	3	2	1	2	1	2	0	1	4	4	3	2	0
234	KRUSN	344	942	7	3	900	5	4	2	0	3	1	2	0	1	3	3	4	4	0
235	KRYOS	344	958	7	26	2,100	5	3	3	1	3	3	2	0	4	3	3	4	5	0
236	KWKWI	8	91	5	140	500,000	5	5	1	3	4	3	3	2	3	5	5	4	3	3
237	KWMAA	5	90	4	105	500,000	5	5	1	2	2	1	2	2	3	5	2	3	3	2
238	KWMIS	15	95	4	160	500,000	5	4	1	2	4	6	3	3	4	5	2	2	2	2
239	KWMIB	7	91	5	120	500,000	5	2	1	2	2	2	1	2	2	4	5	3	3	3
240	KWSAA	5	90	4	115	500,000	5	5	1	2	3	2	2	2	3	5	2	3	3	2

Nº	CÓDIGO	D6MRF	W6MRF	RNFL75	DISRVM	SIZRVC	INASMW	INARKW	INAWP	INHSTM	INSNDB	INSTNB	INLTMF	INMANG	INRKSH	SUFSND	SUSFTM	SUSGRM	SURKRF	SUCORF
241	LKCMB	644	1597	7	22	880	5	5	4	1	4	3	2	2	3	4	5	3	3	1
242	MTMLA	117	493	7	200	120	5	5	5	0	3	4	0	0	4	4	0	4	5	0
243	MYPEN	770	1351	8	0	480	5	4	4	0	4	3	1	1	3	5	4	4	4	4
244	MYLUM	790	1450	8	-2	400	5	3	5	0	3	2	5	3	3	4	5	3	4	3
245	MYPKG	885	1305	8	-1	650	5	3	3	0	2	4	1	1	4	5	4	3	5	5
246	MYPDI	713	1600	7	6	400	5	2	3	0	4	2	3	3	4	4	5	4	4	4
247	MYBTB	885	1305	8	1	140	5	2	3	0	2	3	1	1	3	5	3	3	5	5
248	MYPGU	1101	1433	8	10	2,800	5	2	4	3	4	3	1	2	3	4	4	3	4	4
249	MYBTU	1632	1993	9	9	3,700	5	4	4	0	2	1	3	3	1	5	3	3	2	2
250	NGLOS	405	1336	6	0	18,000	5	4	5	0	3	0	5	4	2	3	5	3	2	0
251	NGTIN	405	1336	6	16	18,000	5	4	4	0	4	0	5	4	2	3	5	3	2	0
252	NGPHC	561	1798	6	-66	120,000	5	0	5	0	1	2	5	4	2	4	5	0	1	0
253	NGONN	560	1800	6	-40	120,000	5	4	4	0	2	3	5	4	2	4	5	0	1	1
254	NGBON	605	1444	6	0	8,000	5	0	5	0	2	0	5	4	0	2	5	3	0	0
255	NLEUR	362	469	8	0	2,500	5	5	3	0	3	3	4	0	1	2	5	1	0	0
256	NLRMT	362	469	8	-10	2,500	5	5	4	0	2	2	4	0	0	2	5	1	0	0
257	NLIJM	475	365	9	-1	300	5	5	4	1	4	3	4	0	1	4	5	2	0	0
258	NLAMS	472	360	9	-18	5,000	5	5	5	0	2	2	5	0	0	2	5	1	0	0
259	NLVL	480	370	9	-3	600	5	5	4	2	4	3	4	0	1	4	5	1	0	0
260	NZAKL	497	687	8	3	200	5	4	3	2	2	2	2	0	3	3	5	2	2	0
261	NZWRE	487	673	9	0	600	5	0	4	1	5	3	5	0	5	4	5	2	4	0
262	NZMAP	487	673	9	20	600	5	0	2	1	4	3	2	0	5	5	3	2	5	0
263	PECLL	6	14	6	10	400	5	5	4	2	4	3	2	0	3	5	5	3	4	0
264	PGLAE	1760	2699	8	2	7,980	5	4	4	2	2	1	1	2	1	5	2	3	1	4
265	PGPOM	236	919	6	7	85	5	5	5	2	2	1	2	2	1	5	3	1	1	1
266	PGDAU	250	960	6	35	55,600	5	0	5	3	4	0	4	4	4	5	5	4	4	3
267	PHBTG	365	1372	6	2	500	5	3	4	4	2	4	1	1	4	5	4	3	4	4
268	PHBTN	216	1607	6	25	900	5	3	4	4	2	4	1	1	4	5	4	3	4	3
269	PHLIM	216	1607	6	4	900	5	4	2	3	3	2	3	3	3	5	4	4	4	3
270	PHMNL	216	1607	6	6	120	5	4	4	1	2	2	1	3	2	3	5	3	4	4
271	PHSFS	228	1797	6	4	1,800	5	5	3	1	3	3	3	4	3	5	3	3	4	3
272	PKBQM	50	185	3	-30	240,000	5	4	4	3	2	0	5	4	2	2	5	1	1	0
273	PKKHI	41	156	3	-4	240,000	5	4	5	4	4	0	5	4	3	2	5	2	2	0
274	PTFAO	87	434	5	80	8,250	5	4	4	3	3	2	4	0	4	3	4	3	3	0
275	PTLIS	164	538	6	0	11,000	5	4	5	3	2	3	5	0	4	3	5	2	3	0
276	PTLOS	87	434	5	0	40	5	4	5	2	4	4	4	0	4	5	5	3	4	0
277	PTSIE	164	538	4	30	700	5	4	4	2	3	3	2	0	4	4	1	0	4	0
278	QADOH	5	76	4	5	300	5	5	4	3	4	2	4	3	4	5	4	4	2	3
279	QAUMS	5	76	4	10	400	5	4	3	4	4	2	3	4	4	5	3	4	0	4
280	QAHAL	5	76	4	95	300	5	4	0	1	4	0	1	1	4	5	3	4	4	4
281	ROCND	196	224	9	115	817,000	5	5	4	3	4	4	3	0	2	4	5	0	1	0
282	ROMAG	191	222	9	145	817,000	5	5	4	3	4	4	3	0	2	4	5	0	1	0
283	ROMID	196	224	9	95	817,000	5	5	4	2	4	4	4	0	2	4	5	0	1	0
284	RUNVS	320	488	9	270	1,410	5	5	4	2	3	3	2	0	5	4	5	1	4	0
285	RUTUA	609	930	7	140	1,410	5	5	4	1	3	2	1	0	4	4	4	1	4	0
286	RUVVO	90	631	5	25	15,000	5	4	5	0	3	3	4	0	5	5	5	0	5	0
287	SADMN	15	80	4	3	80	5	5	4	2	4	1	4	4	3	4	5	4	3	3
288	SAJED	5	60	3	20	800	5	5	4	3	4	0	3	3	3	4	5	4	3	4
289	SAJUB	15	100	4	170	1,000	5	5	4	3	4	4	3	3	3	5	3	4	0	3
290	SAJUT	15	100	4	200	1,000	3	3	1	2	3	0	2	2	3	4	5	2	2	2
291	SARAR	15	95	4	170	500,000	5	2	1	2	2	1	2	2	2	4	5	3	3	2
292	SAO0I	15	95	4	170	1,000	5	5	1	2	4	3	2	2	4	5	3	4	4	1
293	SARLT	15	100	4	170	1,000	5	4	2	3	4	2	3	3	2	4	5	4	3	3
294	SAYNB	2	55	3	2	400	5	4	0	3	4	0	3	3	3	5	3	3	3	4
295	SDBMT	10	40	4	30	1,000	5	3	3	2	3	0	2	2	2	4	5	3	3	4
296	SDPZU	10	40	4	2	1,000	5	5	4	3	3	0	4	3	4	4	5	3	3	2
297	SGJUR	927	1103	9	5	200	5	4	4	3	2	2	4	4	2	3	5	2	2	2
298	SGKEP	927	1103	9	5	200	5	4	4	3	2	2	4	4	2	3	5	2	2	2
299	SGSEM	927	1103	9	2	2,800	5	4	4	3	4	4	1	1	4	3	5	3	4	4
300	SGSIN	927	1103	9	5	200	5	4	4	3	2	2	4	4	2	3	5	2	2	2



Nº	CÓDIGO	D6MRF	W6MRF	RNFL75	DISRVM	SIZRVC	INASMW	INARKW	INAWP	INHTSM	INSNDB	INSTNB	INLTMF	INMANG	INRKSH	SUFSND	SUSFTM	SUSGRM	SURKRF	SUCORF
301	SGTPG	927	1103	9	5	200	5	3	5	3	1	1	3	3	1	3	5	2	2	2
302	SIKOP	470	570	9	0	400	5	4	4	0	3	3	3	0	3	3	5	2	3	0
303	SNDKR	9	494	3	170	40,000	5	5	4	2	4	0	2	2	4	5	5	3	3	2
304	THBKK	190	1307	6	-25	250,000	5	2	5	0	5	5	1	4	5	1	5	4	1	1
305	THLCH	327	1000	6	8	330	5	3	4	0	2	1	2	4	1	2	5	3	3	4
306	TRDYL	102	652	5	5	400	5	2	2	2	3	4	1	0	4	4	2	3	5	0
307	TRERE	194	551	6	70	65,000	5	5	4	0	4	4	3	0	4	3	5	0	4	0
308	TRIST	190	523	7	400	817,000	5	5	5	1	2	4	2	0	5	3	5	1	5	0
309	TRIZM	95	602	5	45	5,000	5	4	4	2	3	4	1	0	4	4	3	2	5	0
310	TRIZT	190	523	8	450	817,000	5	3	5	1	3	4	1	0	5	4	2	0	4	0
311	TRMER	102	652	5	1	450	5	5	3	2	4	4	1	0	4	4	2	3	5	0
312	TRSSX	296	482	8	60	78,200	5	5	4	0	3	3	1	0	4	3	5	0	3	0
313	TRYAR	190	520	7	440	817,000	5	4	3	1	3	4	1	0	5	3	3	1	5	0
314	TWKEL	1721	2009	9	8	45	5	5	3	0	3	1	2	5	1	2	4	3	2	4
315	TWKHH	158	1593	5	10	100	5	5	3	1	3	4	1	4	3	3	4	3	3	4
316	TWTXG	370	935	7	10	1,600	5	5	4	0	4	1	1	4	1	2	4	3	5	5
317	TZDAR	248	810	6	95	100,000	5	5	5	3	3	0	5	4	2	5	5	3	3	4
318	UAODS	148	242	8	54	573,810	5	5	3	2	5	4	2	0	3	2	5	4	3	0
319	UADNB	227	237	9	0	505,810	5	2	5	0	2	2	3	0	1	2	5	0	1	0
320	UAILK	155	179	9	80	573,810	5	4	4	4	5	3	3	0	2	4	5	2	3	0
321	UANIK	156	191	9	-80	68,000	5	1	5	2	2	2	1	0	1	2	5	0	2	0
322	UASVP	162	238	9	300	817,000	5	5	4	1	2	3	1	0	5	4	3	2	4	0
323	USBOS	496	594	9	0	2,000	5	4	5	0	3	4	3	0	4	3	5	0	5	0
324	USNYC	556	633	9	-5	25,000	5	5	4	0	2	4	2	0	2	4	5	0	4	0
325	USPHL	485	559	9	-50	8,400	5	5	5	0	0	4	3	0	4	3	5	0	3	0
326	USILG	485	559	9	-30	8,400	5	5	5	0	1	4	3	0	4	3	5	0	1	0
327	USBAL	518	574	9	-70	25,000	5	5	5	0	0	4	4	0	4	3	5	0	4	0
328	USPHF	498	635	8	-10	25,000	5	5	3	2	3	2	4	0	4	5	5	0	4	0
329	USNEN	498	635	8	-10	25,000	5	4	3	2	3	2	4	0	4	5	5	0	4	0
330	USSAV	433	705	9	0	45,000	5	4	3	2	3	0	4	0	0	5	5	3	3	1
331	USMOB	758	915	9	-2	151,000	5	4	4	3	3	1	4	2	0	3	5	2	0	0
332	USLCH	630	786	7	-26	7,500	5	4	4	3	3	1	4	2	0	4	5	2	0	0
333	USDVT	758	915	9	-60	1,000,000	5	5	3	1	4	4	4	1	1	4	5	1	1	0
334	USMSY	718	946	8	-140	1,000,000	5	3	5	0	0	4	5	0	1	0	5	0	1	0
335	USLOP	729	944	8	60	1,000,000	5	2	1	1	1	0	1	1	0	2	5	2	0	1
336	USSAB	723	941	8	5	600	5	4	3	3	4	3	4	2	3	4	5	3	0	0
337	USBPT	540	681	8	-70	12,500	5	5	5	2	3	3	5	2	0	4	5	1	0	0
338	USGLS	459	609	9	0	8,000	5	5	3	3	4	2	5	4	2	4	5	4	1	0
339	USTXT	459	609	9	-12	8,000	5	5	3	3	3	2	5	4	2	3	5	3	1	0
340	USHOU	540	681	8	-35	8,000	5	4	5	2	3	1	5	2	1	2	5	2	1	0
341	USANC	117	280	7	0	15,000	5	5	0	0	2	4	3	0	4	3	5	0	5	0
342	USPDX	231	689	6	-195	200,000	5	4	5	0	3	5	1	0	5	5	5	0	5	0
343	USBCC	230	764	6	-188	200,000	5	4	5	0	3	5	1	0	5	5	5	0	5	0
344	USSFO	44	453	5	15	30,000	5	4	3	2	3	3	3	0	4	4	5	0	4	0
345	USOAK	44	453	5	14	30,000	5	4	4	2	3	3	3	0	4	4	5	0	4	0
346	USLGB	11	165	2	5	2,500	5	5	3	2	3	3	3	0	2	5	5	0	3	0
347	USSAN	22	233	5	14	550	5	4	4	2	4	3	4	0	2	5	5	0	0	0
348	UYMYD	540	591	9	0	600,000	5	4	5	3	4	3	4	0	4	5	5	0	3	0
349	YEADE	14	33	6	6	15,500	5	5	4	3	4	3	3	3	4	4	5	4	3	4
350	YEHOD	14	33	6	14	8,200	5	5	4	2	4	2	2	3	5	5	4	3	3	3
351	YEMKX	14	33	6	70	12,000	5	5	4	0	5	3	0	0	5	5	4	1	5	1
352	YERAI	14	33	6	55	8,200	5	2	2	1	3	2	1	1	4	5	4	4	3	3
353	ZACPT	197	630	6	3	217	5	4	4	2	4	4	2	0	4	4	5	2	4	0
354	ZADUR	332	758	7	3	180	5	4	4	3	4	0	3	4	4	5	4	3	2	0
355	ZAPLZ	363	783	8	1	84	5	5	5	2	4	0	2	0	2	5	5	2	3	0
356	ZARCB	462	757	8	2	183	5	4	4	4	4	0	5	4	2	5	5	4	3	0
357	ZASDB	69	274	5	70	7,900	5	5	4	2	4	0	3	0	4	5	4	3	4	0

**Anexo 4 - Coeficiente de Similaridade do sistema GloBallast/URS**

Posição	Porto	No.	Código ONU	Distância Euclidiana	Coefficiente de Similaridade
1	Sepetiba	1	BRSPB	0.000	1.000
2	Rio de Janeiro	51	BRRIO	2.360	0.861
3	Praia Mole	54	BRPRM	3.135	0.799
4	Tubarão	55	BRTUB	3.230	0.791
5	Santos	50	BRSSZ	3.419	0.776
6	Vitória	53	BRVIX	3.889	0.738
7	Salvador	56	BRSSA	4.059	0.725
8	Paranaguá	49	BRPNG	4.154	0.717
9	Wakayama Wakayama	222	JPWAK	4.890	0.658
10	Tin Can Island	251	NGTIN	5.016	0.648
11	Nagoya Aichi	200	JPNGO	5.089	0.642
12	Kudamatsu Yamaguchi	191	JPKUD	5.176	0.635
13	Yokohama Kanagawa	224	JPYOK	5.236	0.631
14	Negishi (Yokohama) Kanagawa	199	JPNGI	5.236	0.631
15	Singapore Jurong	297	SGJUR	5.242	0.630
16	Singapore Keppel	298	SGKEP	5.242	0.630
17	Singapore Singapore	300	SGSIN	5.242	0.630
18	Shimotsu Wakayama	212	JPSMT	5.279	0.627
19	Yokkaichi Mie	223	JPYKK	5.285	0.627
20	Lagos	250	NGLOS	5.317	0.624
21	Durban	354	ZADUR	5.327	0.623
22	Trieste	168	ITTRS	5.342	0.622
23	Toyama Toyama	218	JPTOY	5.344	0.622
24	Koper (Slovenia)	302	SIKOP	5.379	0.619
25	Kimitsu Chiba	173	JPKMT	5.382	0.619
26	Chiba Chiba	172	JPCHB	5.390	0.618
27	Colombo	241	LKCMB	5.390	0.618
28	Kinuura Aichi	188	JPKNU	5.407	0.617
29	Tokuyama Yamaguchi	216	JPTKY	5.426	0.615
30	Kawasaki Kanagawa	192	JPKWS	5.430	0.615
31	Shimizu Shizuoka	213	JPSMZ	5.435	0.615
32	Port Kembla	25	AUPKL	5.436	0.614
33	Tuticorin (New Tuticorin)	150	INTUT	5.437	0.614
34	Salaya	148	INSAL	5.483	0.611
35	Sakaiminato Tottori	211	JPSMN	5.485	0.610
36	Sabine	336	USSAB	5.489	0.610
37	Shibushi Kagoshima	209	JPSBS	5.511	0.608
38	Yokosuka Kanagawa	225	JPYOS	5.532	0.607
39	Taranto	166	ITTAR	5.577	0.603
40	Auckland	260	NZAKL	5.637	0.598
41	Norfolk-Newport News Virginia	329	USNEN	5.662	0.596
42	Ravenna	165	ITRAN	5.665	0.596
43	Laem Chabang	305	THLCH	5.689	0.594
44	Faro	274	PTFAO	5.698	0.593
45	Niigata Niigata	184	JPKIJ	5.705	0.593
46	Kikuma Ehime	185	JPKIK	5.706	0.593
47	Paradeep	147	INPRT	5.717	0.592
48	Matsuyama Ehime	197	JPMYJ	5.720	0.592
49	Texas City Texas	339	USTXT	5.724	0.591
50	Balikpapan Kalimantan	133	IDBPN	5.734	0.591
51	Venezia (=Fusina)	167	ITVCE	5.747	0.590

Posição	Porto	No.	Código ONU	Distância Euclidiana	Coefficiente de Similaridade
52	Richards Bay	356	ZARCB	5.749	0.589
53	Iwakuni Yamaguchi	180	JPIWK	5.770	0.588
54	Hampton Roads	328	USPHF	5.778	0.587
55	Karachi	273	PKKHI	5.782	0.587
56	Imabari Ehime	178	JPIMB	5.785	0.586
57	Tobata (Kitakyushu) Fukuoka	215	JPTBT	5.793	0.586
58	Dumai Sumatra	125	IDDUM	5.794	0.586
59	Port Bonython	20	AUPBY	5.799	0.585
60	Abbot Point	33	AUABP	5.812	0.584
61	Kagoshima Kagoshima	189	JKOJ	5.815	0.584
62	Innoshima Hiroshima	179	JPINS	5.828	0.583
63	Saldanha Bay	357	ZASDB	5.833	0.583
64	Beilun Zhejiang	67	CNNBO	5.868	0.580
65	Ningbo (Beilun) Zhejiang	68	CNNGB	5.868	0.580
66	Townsville	34	AUTSV	5.874	0.579
67	Singapore Pasir Panjan/Tanjung Pagar	301	SGTPG	5.879	0.579
68	Higashi-Harima Hyogo	175	JPHHR	5.884	0.579
69	Porbandar	146	INPBD	5.909	0.577
70	Himeji Hyogo	176	JPHIM	5.911	0.576
71	Maizuru Kyoto	193	JPMIA	5.917	0.576
72	Belawan Sumatra	124	IDBLW	5.919	0.576
73	Port Elizabeth	355	ZAPLZ	5.927	0.575
74	Chennai (Ex Madras)	143	INMAA	5.962	0.572
75	Galveston Texas	338	USGLS	5.978	0.571
76	Pachi	117	GRPAC	5.979	0.571
77	Kakogawa Hyogo	182	JKPGA	5.981	0.571
78	Elefsis (Eleusis)	115	GREEU	5.989	0.570
79	Onomichi Hiroshima	204	JPONO	5.998	0.569
80	Naha Okinawa	198	JPNAH	6.008	0.569
81	Okinawa Okinawa	203	JPOKA	6.008	0.569
82	Tamano (Uno) Okayama	214	JPTAM	6.015	0.568
83	Merak (inc. Anyer Terminal) Java	127	IDMRK	6.021	0.568
84	Port Dickson	246	MYPDI	6.027	0.567
85	Tarragona	98	ESTAR	6.030	0.567
86	Kashima Ibaraki	190	JKPSM	6.030	0.567
87	Whangerei	261	NZWRE	6.032	0.567
88	Port Pirie	22	AUPPI	6.051	0.565
89	Barcelona	93	ESBCN	6.064	0.564
90	Visakhapatnam	152	INVTZ	6.075	0.563
91	Bilbao	91	ESBIO	6.089	0.562
92	San Francisco California	344	USSFO	6.096	0.562
93	Callao (Lima)	263	PECLL	6.104	0.561
94	Gibraltar	113	GIGIB	6.107	0.561
95	Whyalla	21	AUWYA	6.118	0.560
96	Mackay	32	AUMKY	6.124	0.559
97	Muhammad Bin Qasim	272	PKBQM	6.134	0.558
98	Osaka Osaka	205	JPOSA	6.146	0.558
99	Aboshi Hyogo	169	JPABO	6.158	0.557
100	Tanjung Perak (Surabaya) Java	131	IDSUB	6.159	0.557
101	Port Sudan	296	SDPZU	6.159	0.557
102	Nagasaki Nagasaki	201	JPNGS	6.176	0.555
103	Lucinda	35	AULUC	6.177	0.555

Posição	Porto	No.	Código ONU	Distância Euclidiana	Coefficiente de Similaridade
104	Vigo	92	ESVGO	6.191	0.554
105	Mumbai (Ex Bombay)	137	INBOM	6.214	0.552
106	Kinwan (Ishikawa) Okinawa	186	JPKIN	6.215	0.552
107	Kobe Hyogo	221	JPUKB	6.219	0.552
108	Gijon	90	ESGIJ	6.221	0.552
109	Volos	120	GRVOL	6.243	0.550
110	Thessaloniki	119	GRSKI	6.246	0.550
111	Batumi, Georgia	111	GEBUS	6.247	0.549
112	Aspropyrgos	114	GRASS	6.250	0.549
113	Port Stanvac	23	AUPST	6.257	0.549
114	San Diego	347	USSAN	6.271	0.548
115	Western Port (now Hastings; AUHAS)	24	AUWEP	6.274	0.547
116	Limay	269	PHLIM	6.281	0.547
117	Piraeus	118	GRPIR	6.292	0.546
118	Manila	270	PHMNL	6.292	0.546
119	Suez (El Suweis)	89	EGSUZ	6.295	0.546
120	Muroran Hokkaido	196	JPMUR	6.299	0.545
121	Jakarta Java	128	IDJKT	6.301	0.545
122	Hakata Fukuoka	177	JPHKT	6.305	0.545
123	Marsa Bashayer Oil Terminal	295	SDMBT	6.310	0.544
124	Marsden Point	262	NZMAP	6.315	0.544
125	Daru	266	PGDAU	6.323	0.543
126	Tokyo Tokyo	219	JPTYO	6.327	0.543
127	Poti, Geórgia	112	GEPTI	6.341	0.542
128	Semarang Java	130	IDSRG	6.351	0.541
129	Pohang	228	KRKPO	6.359	0.541
130	Dar Es Salaam	317	TZDAR	6.366	0.540
131	Kwangyang	227	KRKAN	6.370	0.540
132	Port Moresby	265	PGPOM	6.373	0.539
133	Pasir Gudang Johor	248	MYPGU	6.377	0.539
134	Oakland California	345	USOAK	6.378	0.539
135	Gladstone	28	AUGLT	6.404	0.537
136	Samsun	312	TRSSX	6.405	0.537
137	Cartagena	74	COCAR	6.408	0.537
138	Sakai Osaka	208	JPSAK	6.415	0.536
139	Fukuyama Hiroshima	174	JPFKY	6.416	0.536
140	Lumut	244	MYLUM	6.425	0.535
141	Eregli	307	TRERE	6.440	0.534
142	Ponta do Ubu	52	BRPOU	6.440	0.534
143	Dakar	303	SNDKR	6.442	0.534
144	Mombasa	226	KEMBA	6.462	0.532
145	Lagos (Portugal)	276	PTLOS	6.483	0.531
146	Hay Point (Dalrymple Bay)	30	AUHPT	6.505	0.529
147	Dallrymple Bay (Hay Point)	31	AUPDT	6.505	0.529
148	Samcheon Po	233	KRSCP	6.516	0.528
149	Kochi Kochi	181	JPKCZ	6.526	0.527
150	Marseilles	105	FRMRS	6.529	0.527
151	Subic Bay (Sana Clara)	271	PHSFS	6.538	0.526
152	Porto Foxi (Sarroch)	163	ITPFX	6.556	0.525
153	Long Beach California	346	USLGB	6.564	0.524
154	Hodeidah (Yemen)	350	YEHOD	6.568	0.524
155	Cape Town	353	ZACPT	6.580	0.523

Posição	Porto	No.	Código ONU	Distância Euclidiana	Coefficiente de Similaridade
156	Cigading	126	IDCIG	6.586	0.522
157	Sines	277	PTSIE	6.597	0.521
158	Moji (Kitakyushu) Fukuoka	195	JPMOJ	6.604	0.521
159	Singapore Sembawang Port	299	SGSEM	6.613	0.520
160	Kaohsiung	315	TWKHH	6.618	0.520
161	Livorno	164	ITLIV	6.635	0.518
162	Yosu (Yeosu)	235	KRYOS	6.635	0.518
163	Cairns	37	AUCNS	6.636	0.518
164	Mokpo (Mogpo)	230	KRMOK	6.644	0.518
165	Aden (Yemen)	349	YEADE	6.644	0.518
166	Brisbane	26	AUBNE	6.658	0.517
167	Fos sur Mer (Oil Terminal)	102	FRFOS	6.665	0.516
168	Bundaberg	27	AUBDB	6.671	0.516
169	Las Palmas	96	ESLPA	6.673	0.515
170	Mangalore (New Mangalore)	141	INIXE	6.674	0.515
171	Lisboa	275	PTLIS	6.674	0.515
172	Algeciras	95	ESALG	6.678	0.515
173	Tenerife (Santa Cruz de Tenerife)	97	ESSCT	6.691	0.514
174	Jebel Dhanna	10	AEJED	6.698	0.513
175	Ruwais Oil Terminal	13	AERUW	6.698	0.513
176	Lavera	103	FRLAV	6.707	0.513
177	Ube Yamaguchi	220	JPUBJ	6.714	0.512
178	Montevideo	348	UYMVD	6.718	0.512
179	Izmir (Smyrna)	309	TRIZM	6.731	0.511
180	Tuapse, Russia	285	RUTUA	6.785	0.506
181	Mobile Alabama	331	USMOB	6.803	0.505
182	Savannah Georgia	330	USSAV	6.847	0.501
183	Ras Isa Marine Terminal (Yemen)	352	YERAI	6.847	0.501
184	Jeddah	288	SAJED	6.851	0.501
185	Genoa	162	ITGOA	6.853	0.501
186	Hunterston	106	GBHST	6.860	0.500
187	Amagasaki Hyogo	170	JPAMA	6.895	0.498
188	Qingdao Shandong	71	CNTAO	6.901	0.497
189	Al Mukullah (Yemen)	351	YEMKX	6.916	0.496
190	Keelung (Chilung)	314	TWKEL	6.946	0.493
191	Mizushima Okayama	194	JPMIZ	6.961	0.492
192	Marmugao (Marmagoa)	144	INMRM	6.962	0.492
193	Ulsan	234	KRUSN	6.971	0.491
194	Yantai Shandong	73	CNYNT	6.981	0.491
195	Chios	116	GRJKH	7.041	0.486
196	Hong Kong Hong Kong	121	HKHKG	7.084	0.482
197	Chiwan (Shenzhen) Guangdong	64	CNCWN	7.092	0.482
198	Itajai	48	BRITJ	7.098	0.481
199	Tomakomai Hokkaido	217	JPTMK	7.102	0.481
200	Bonny	254	NGBON	7.103	0.481
201	Omishaj	123	HROMI	7.146	0.477
202	Hong Kong Kowloon	122	HKKWN	7.146	0.477
203	Valencia	94	ESVLC	7.150	0.477
204	Tanjung Bara Coal Terminal Kalimantan	132	IDTBA	7.161	0.476
205	Ain Sukhna	84	EGAIS	7.163	0.476
206	Houston Texas	340	USHOU	7.175	0.475
207	Alexandria (El Iskandariya)	85	EGALY	7.189	0.474

Posição	Porto	No.	Código ONU	Distância Euclidiana	Coefficiente de Similaridade
208	Dortyol Oil Terminal	306	TRDYL	7.195	0.474
209	Mersin	311	TRMER	7.198	0.473
210	El Dekheila	87	EGEDK	7.202	0.473
211	Bintulu Sarawak	249	MYBTU	7.233	0.470
212	Yanbu	294	SAYNB	7.239	0.470
213	Kiire Kagoshima	183	JPKII	7.248	0.469
214	Oita Oita	202	JPOIT	7.252	0.469
215	Beppu Oita	171	JPBEP	7.254	0.469
216	Dammam	287	SADMN	7.255	0.469
217	Saiki Oita	206	JPSAE	7.262	0.468
218	Vadinar Terminal	151	INVAD	7.278	0.467
219	Guangzhou Guangdong	63	CNCAN	7.284	0.466
220	Bataan Mariveles	268	PHBTN	7.289	0.466
221	Taichung	316	TWTXG	7.309	0.464
222	Onne	253	NGONN	7.321	0.463
223	Onsan	231	KRONS	7.322	0.463
224	Batangas (Luzon)	267	PHBTG	7.324	0.463
225	LOOP Terminal	335	USLOP	7.326	0.463
226	Djibouti (Djibouti)	81	DJJIB	7.346	0.461
227	Doha	278	QADOH	7.354	0.461
228	Europoort	255	NLEUR	7.400	0.457
229	Donges	101	FRDON	7.406	0.457
230	Sakaide Kagawa	210	JPSKD	7.418	0.456
231	Khor Al Fakkan (Tanker Anchorage)	11	AEKLF	7.424	0.455
232	Boston Massachusetts	323	USBOS	7.433	0.454
233	Saganoseki Oita	207	JPSAG	7.436	0.454
234	Weipa	39	AUWEI	7.446	0.453
235	Buenos Aires	16	ARBUE	7.448	0.453
236	Kapar Coal Terminal	247	MYBTB	7.456	0.453
237	Larnaca	76	CYLCA	7.469	0.452
238	Mundra	145	INMUN	7.509	0.448
239	Ijmuiden	257	NLIJM	7.524	0.447
240	Port Alma	29	AUPTL	7.529	0.447
241	Bandar Abbas (Oil Jetty)	155	IRBND	7.576	0.443
242	Penang (Georgetown)	243	MYPEN	7.581	0.443
243	Lake Charles Louisiana	332	USLCH	7.637	0.438
244	Mourilyan	36	AUMOU	7.642	0.438
245	Kyrenia	75	CYKYR	7.653	0.437
246	Ras Al Tannura	293	SARLT	7.659	0.436
247	Dalian Liaoning	65	CNDLC	7.659	0.436
248	Brest	100	FRBES	7.670	0.435
249	Fujairah	8	AEFJR	7.680	0.435
250	Limassol	77	CYLMS	7.689	0.434
251	Cape Flattery	38	AUCQP	7.713	0.432
252	Jebel Ali	9	AEJEA	7.735	0.430
253	Zirku Island	15	AEZIR	7.751	0.429
254	Halul Island	280	QAHAL	7.758	0.428
255	Dampier	18	AUDAM	7.767	0.428
256	Tianjin Tianjin	72	CNTSN	7.774	0.427
257	Das Island	4	AEDAS	7.793	0.426
258	Fredericia	83	DKFRC	7.822	0.423
259	Ras Al Khafji	291	SARAR	7.838	0.422

Posição	Porto	No.	Código ONU	Distância Euclidiana	Coefficiente de Similaridade
260	Mina Abdulla	239	KWMIB	7.938	0.414
261	Wilhelmshaven	80	DEWVN	7.947	0.413
262	Abu Dhabi	2	AEAUH	7.960	0.412
263	Mina Zayed	3	AEMZD	7.960	0.412
264	Enstedvaerkets Havn	82	DKENS	7.970	0.411
265	Fateh Oil Terminal	7	AEFAT	7.972	0.411
266	Sitra (Bahrain)	46	BHMAN	7.974	0.411
267	Mina Saud	238	KWMIS	7.990	0.410
268	Flushing (Vlissingen)	259	NLVLI	8.010	0.408
269	Um Al Qiwayn	12	AEQIW	8.013	0.408
270	Redcar	110	GBRER	8.025	0.407
271	Port Walcott (Cape Lambert)	19	AUPWL	8.036	0.406
272	Port Rashid	5	AEPR	8.046	0.405
273	Dubai	6	AEDXB	8.046	0.405
274	Umm Said (Mesaieed)	279	QAUMS	8.054	0.405
275	Ras Al Ghar	292	SA001	8.094	0.401
276	Mina Sulman (Al Manamah)	47	BHMIN	8.097	0.401
277	Novorossiysk, Russia	284	RUNVS	8.113	0.400
278	Cochin	139	INCOK	8.121	0.399
279	Bushehr	156	IRBUZ	8.168	0.396
280	New York New York (New Jersey)	324	USNYC	8.218	0.392
281	Port Harcourt	252	NGPHC	8.229	0.391
282	Haldia	140	INHAL	8.236	0.390
283	Midia	283	ROMID	8.257	0.388
284	Port Kelang	245	MYPKG	8.322	0.383
285	Constanta	281	ROCND	8.356	0.380
286	Mangalia	282	ROMAG	8.411	0.376
287	Sharjah	14	AESHJ	8.412	0.376
288	Kanda Fukuoka	187	JPKND	8.419	0.375
289	Kuwait (Shuwaikh; KWSWK)	236	KWKWI	8.430	0.375
290	Ilyichevsk	320	UAILK	8.431	0.375
291	Dunkerque	99	FRDKK	8.441	0.374
292	Le Havre	104	FRLEH	8.451	0.373
293	Shuaiba	240	KWSAA	8.453	0.373
294	Sikka	149	INSIK	8.462	0.372
295	Amamapare Irian Jaya	134	IDAMA	8.463	0.372
296	Varna, Bulgaria	45	BGVAR	8.470	0.371
297	Mina Al Ahmadi	237	KWMAA	8.476	0.371
298	Moneypoint	135	IEMOT	8.508	0.368
299	Lae	264	PGLAE	8.510	0.368
300	Beaumont	337	USBPT	8.562	0.364
301	Bangkok	304	THBKK	8.566	0.364
302	Bremen	78	DEBRE	8.669	0.355
303	Vladivostok	286	RUVVO	8.694	0.353
304	Rotterdam	256	NLRTM	8.719	0.351
305	Khark Island	157	IRKHK	8.723	0.351
306	Huangpu Guangdong	66	CNHUA	8.733	0.350
307	Roberts Bank (British Columbia)	62	CARBK	8.758	0.348
308	Bourgass	44	BGBOJ	8.763	0.348
309	Pusan	232	KRPUS	8.784	0.346
310	Halifax Nova Scotia	59	CAHAL	8.826	0.343
311	Malta (Valletta)	242	MTMLA	8.864	0.340



Posição	Porto	No.	Código ONU	Distância Euclidiana	Coefficiente de Similaridade
312	Vancouver (British Columbia)	61	CAVAN	8.912	0.336
313	Kunsan	229	KRKUV	8.919	0.335
314	Odessa	318	UAODS	8.983	0.330
315	Kandla	142	INIXY	8.986	0.330
316	Port Talbot	109	GBPTB	9.053	0.325
317	Karumba	40	AUKRB	9.078	0.323
318	Sevastopol	322	UASVP	9.248	0.309
319	Baltimore Maryland	327	USBAL	9.252	0.309
320	Burry Port (Llanelly)	108	GB001	9.297	0.305
321	Damietta	86	EGDAM	9.328	0.303
322	Istanbul	308	TRIST	9.348	0.301
323	Sirri Island Oil Terminal	159	IRSXI	9.358	0.300
324	Lavan Island	158	IRLVP	9.369	0.299
325	Immingham	107	GBIMM	9.377	0.299
326	Wilmington Delaware	326	USILG	9.410	0.296
327	Hamburg	79	DEHAM	9.425	0.295
328	Jubail	289	SAJUB	9.446	0.293
329	Davant	333	USDVT	9.468	0.291
330	Yarimca	313	TRYAR	9.519	0.287
331	Dnepro-Bugsky (Ochakov)	319	UADNB	9.534	0.286
332	Come By Chance	57	CACBC	9.660	0.276
333	Sept-Iles (Pointe Noire) Quebec	58	CASEI	9.694	0.273
334	Philadelphia Pennsylvania (Port Richmond)	325	USPHL	9.713	0.272
335	Izmit (Tutuncifilik Oil Terminal)	310	TRIZT	9.888	0.258
336	Ashdod	136	ILASH	9.901	0.257
337	La Have	60	CALHA	9.922	0.255
338	Amsterdam	258	NLAMS	9.935	0.254
339	Chittagong	41	BDCGP	9.964	0.252
340	Bandar Mushar (Mushahr)	154	IRBMR	9.969	0.251
341	Ghent (Gent)	43	BEGNE	10.041	0.245
342	Shanghai Shanghai	69	CNSHA	10.072	0.243
343	Antwerpen	42	BEANR	10.136	0.238
344	Port Said	88	EGPSD	10.210	0.232
345	Shanghai Baoshan	70	CNSHB	10.304	0.224
346	Bandar Imam Khomeyni	153	IRBKM	10.327	0.223
347	Nicolayev	321	UANIK	10.470	0.211
348	Campana	17	ARCMP	10.544	0.205
349	Hafnarfjordur	160	ISHAF	10.549	0.205
350	New Orleans	334	USMSY	10.563	0.204
351	Straumsvik	161	ISSTR	10.573	0.203
352	Vancouver Washington	343	USBCC	10.580	0.202
353	Portland Oregon	342	USPDX	10.617	0.199
354	Calcutta	138	INCCU	11.167	0.155
355	Al Juaymah Terminal	290	SAJUT	11.843	0.101
356	Cilacap Java	129	IDCXP	11.979	0.090
357	Anchorage Alaska	341	USANC	12.481	0.050

**Anexo 5 - Lei Nº 11.200, de 24 de novembro de 2005**

**LEI Nº 11.200, DE 24 DE NOVEMBRO DE 2005.**

Altera a denominação do Porto de Sepetiba, no Estado do Rio de Janeiro, para Porto de Itaguaí.

**O PRESIDENTE DA REPÚBLICA** Faço saber que o Congresso Nacional decreta e eu sanciono a seguinte Lei:

Art. 1º O Porto de Sepetiba, localizado na baía de mesmo nome, no Município de Itaguaí, Estado do Rio de Janeiro, passa a denominar-se "Porto de Itaguaí".

Art. 2º Esta Lei entra em vigor na data de sua publicação.

Brasília, 24 de novembro de 2005; 184º da Independência e 117º da República.

LUIZ INÁCIO LULA DA SILVA  
*José Alencar Gomes da Silva*  
*Alfredo Nascimento*

Publicada no D.O.U. de 25.11.2005