

João Miguel Charrua de Sousa

## Caracterização horária de consumos de energia elétrica e previsão de curto prazo em mercados liberalizados

Tese de Doutoramento na área científica de Engenharia Eletrotécnica, especialidade em Sistemas de Energia, orientada pelo Senhor Professor Doutor Humberto Manuel Matos Jorge e pelo Senhor Professor Doutor Luís Miguel Pires Neves e apresentada ao Departamento de Engenharia Eletrotécnica e de Computadores da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

2013

• U • C •



UNIVERSIDADE DE COIMBRA



UNIVERSIDADE DE COIMBRA  
FACULDADE DE CIÊNCIAS E TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA ELETROTÉCNICA E DE COMPUTADORES

**Caracterização horária de consumos de  
energia elétrica e previsão de curto prazo  
em mercados liberalizados**

João Miguel Charrua de Sousa

Tese apresentada no âmbito do curso de Doutoramento  
em Engenharia Eletrotécnica  
na área de especialização em Sistemas de Energia

Trabalho realizado sob orientação de:

Professor Doutor Humberto Manuel Matos Jorge,  
Departamento de Eng<sup>a</sup>. Eletrotécnica e de Computadores  
Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Professor Doutor Luís Miguel Pires Neves,  
Departamento de Eng<sup>a</sup>. Eletrotécnica  
Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Leiria

Coimbra, março de 2013



## RESUMO

O setor elétrico tem sentido alterações muito significativas nas últimas décadas. Com a reestruturação do setor elétrico tem-se assistido a um crescente envolvimento e a uma participação ativa de diferentes agentes, próprio de um setor que passou a ter as suas atividades integrantes juridicamente separadas e com condições criadas para que exista concorrência na produção e comercialização. A previsão de consumos é tradicionalmente imperiosa para o equilíbrio entre a oferta e a procura, bem como para uma rigorosa gestão e planeamento das redes elétricas de transporte e distribuição. A sua pertinência é ainda reforçada com a liberalização dos mercados, uma vez que os comercializadores pretendem dispor de ferramentas que lhes permitam estimar com acuidade a curva de procura agregada do conjunto de consumidores com quem contratualizam. Deste modo, terão uma participação mais conscienciosa em regime de mercado, procedendo a uma ajustada oferta pela quantidade de energia que satisfaça as necessidades de consumo da sua carteira de consumidores.

Com a liberalização do setor, a caracterização de consumos tornou-se também uma necessidade legal para permitir o acerto de contas entre o distribuidor e os comercializadores de eletricidade, sempre que as instalações consumidoras não disponham de sistemas de medição com apropriados intervalos de integração. O uso de perfis de consumo é seguido para diferentes classes de consumidores, de modo a que o consumo total medido nos contadores tradicionais possa ser distribuído de forma coerente por diferentes períodos horários.

A perspetiva de integração da informação proveniente dos perfis de consumo numa ótica de previsão tem sido pouco explorada, pelo que esta tese procura avaliar a pertinência desta informação para melhorar indicadores de previsão. O estudo explora recentes metodologias de previsão que são também bem reconhecidas na comunidade científica, como são as redes neuronais artificiais e as máquinas de suporte vetorial.

O estudo é validado através da previsão de consumos para as diferentes horas do dia seguinte e recorre a dois cenários distintos: no primeiro caso em estudo é considerado um diagrama de carga simulado resultante da agregação de consumidores individuais, enquanto o segundo estudo de caso é relativo a uma saída específica de uma subestação de distribuição de 60/15kV. Os resultados de previsão apresentam erros percentuais absolutos médios da ordem dos 5% em amostras de teste para os dois estudos de caso

considerados. A integração de informação proveniente da classificação de consumidores resulta no aumento da qualidade de previsão, nomeadamente em dias especiais, o que é também confirmado através da elevada dependência que os modelos usados têm a estas novas entradas.

Palavras-Chave: Previsão de Consumos de Eletricidade; Perfis de Consumo; Redes Neurais; Máquinas de Suporte Vetorial; Análise de Sensibilidades.

## **ABSTRACT**

The electricity sector has been subjected to significant changes in the last decades. The sector deregulation has prompted an increasing involvement and a more active participation of the different agents, as a result from the legal split of the different activities and the establishment of the necessary conditions to enable competition at the production and retail levels. Load forecasting is not only traditionally fundamental to provide a balance between the supply and demand, but also valuable to the electric grid management and planning. The sector liberalization reinforces the importance of load forecasting, because the different suppliers intend to employ useful tools to accurately estimate the aggregated load curve relative to their group of customers. This way, the suppliers may have a more conscientious participation in an open market in order to make reasoned offers for the amount of energy that meets the consumers' requirements.

Also associated with the liberalization of the energy markets, the characterization of consumption became mandatory in order to enable the settlement between the distributor and the different suppliers, whenever the consuming facilities are not equipped with metering systems with appropriate integration intervals. The use of load profiles for different consumer classes allows for a coherent distribution of the total consumption recorded in the traditional energy meters by different hourly periods.

The perspective of integrating information derived from load profiles in load forecasting has not yet being well exploited, so this work intends to evaluate the relative importance that this kind of information could have in guaranteeing an improvement in forecasting indicators. The study explores actual and well accepted forecasting methodologies as the artificial neural networks and support vector machines.

The study is validated through the load forecasting for the following day in two different scenarios: in the first case study a simulated load diagram resulting from individual consumers' aggregation is considered, while the second case study is related to a specific output of a 60/15kV distribution substation.

The forecasting results show mean absolute percentage errors of approximately 5% in testing samples for both case studies. It was proven that integrating information derived from consumers' classification raises the forecasting quality, mainly on special days,

which is also confirmed by the high dependence of the different models used to these new inputs.

Keywords:

Electric Load Forecasting; Load Profiling; Artificial Neural Networks; Support Vector Machines; Sensitivity Analysis.



## **AGRADECIMENTOS**

Este trabalho só foi possível com uma expressiva contribuição dos orientadores na parte científica. A experiência operacional do Professor Humberto Jorge no setor energético foi determinante para validar alguns pressupostos, mas também para evidenciar algumas lacunas do trabalho na sua fase de desenvolvimento. O meu reconhecimento pela sua ajuda, valiosa disponibilidade e pela sua afabilidade característica.

O meu agradecimento também ao Professor Luís Neves, pela constante partilha de informação e estímulo para explorar diferentes metodologias. A sua perseverança e determinação nos momentos mais críticos do trabalho foram também um excelente exemplo que guardarei.

Agradeço a professores, colegas e amigos que contribuíram de forma técnica e científica, partilhando a sua experiência e trabalhos desenvolvidos ou debatendo temas relacionados. Este contributo serviu indiscutivelmente para impulsionar este estudo.

Um agradecimento ao Gabinete de Estudos Empresariais da EDP pela cedência de dados de consumo monitorizados em subestações de distribuição circundantes da cidade de Leiria.

Uma palavra de apreço também ao Ministério da Defesa Nacional - Estado-Maior da Força Aérea, e mais particularmente ao Gabinete de Comando da Base Aérea nº 5 pela cedência de dados meteorológicos registados na Base Aérea de Monte Real.

Aos meus pais agradeço o exemplo de disciplina, esforço, dedicação e simplicidade para lidarem com os desafios, bem como toda a confiança que sempre mantiveram. A vossa “rede” foi sempre sentida e muito tem jogado a meu favor.

Aos familiares mais próximos por me manterem com uma boa disposição nos encontros que tivemos nos últimos anos. Muito agradeço também todo o suporte possibilitado para que conseguisse dedicar mais tempo ao trabalho em horários muito pouco compatíveis com o tempo de família. Neste grupo, dedico a maior consideração aos meus irmãos, sogros, cunhados e bem-dispostos sobrinhos.

Aos responsáveis do Instituto Politécnico de Leiria pelas condições facultadas, materializadas através de regimes de redução de serviço docente e apoio no pagamento de propinas.

Aos colegas de trabalho na ESTG - IPEiria pelos exemplos de profissionalismo e dedicação e pelas palavras encorajadoras. Ao vasto grupo de envolvidos, pretendo

destacar o apoio do colega Romeu Vitorino com quem foi partilhada muita da vivência durante este desafio, do colega Licínio Moreira com quem tenho tido uma próxima experiência docente e de orientação de projetos e à Lídia Paiva pelo apoio nas revisões do Inglês para os artigos de revista.

Ao INESC Coimbra por todo o apoio concedido e pelas proveitosas jornadas técnicas que tem organizado.

Uma última referência, mas por ventura a mais importante, um agradecimento muito especial à Sofia que sempre esteve do meu lado e que sempre soube escolher as palavras apropriadas para moralizar, reconhecer o esforço e ajudar, quando necessário, a relativizar e a “levantar a cabeça fora de água”. À Sofia e aos meus preciosos filhos, Miguel e Pedro, as minhas desculpas pelo tempo que não consegui acompanhar-vos, pelo tempo que, apesar de estar presencialmente convosco, estive demasiado ocupado mentalmente com outros assuntos e um muito obrigado pelos momentos que me consegui entregar e gozar realmente o prazer da vida na vossa companhia. Por muitas arquiteturas de redes neuronais que venham a ser exploradas, bem como algoritmos de aprendizagem mais robustos que sejam validados, nada me surpreenderá mais do que a oportunidade que tenho tido em assistir às “pequenas” redes neuronais dos meus filhos a desenvolverem-se e a criarem as sinapses necessárias para se tornarem mais ajustados à realidade que os rodeia.

## ÍNDICE

1. Introdução.....	1
1.1. Perspetiva sobre evolução de mercados de energia elétrica.....	1
1.2. Caracterização horária de consumos de energia elétrica.....	6
1.3. Previsão de consumos de energia elétrica.....	8
1.4. Organização do documento.....	9
2. Caracterização horária de consumos de energia elétrica.....	11
2.1. Perfis de consumo.....	11
2.2. Critérios usados na classificação de consumidores.....	13
2.2.1. Classificação baseada no setor de atividade económica.....	13
2.2.2. Classificação baseada em informação comercial.....	17
2.2.3. Classificação baseada em reconhecimento de padrões e em métodos de <i>clustering</i> .....	17
2.2.4. Metodologia de classificação de consumos adotada em Portugal.....	20
3. Métodos de previsão de consumos.....	23
3.1. Abordagem às Redes Neurais Artificiais.....	27
3.1.1. Processo de treino.....	29
3.1.2. Arquitetura da rede.....	34
3.1.3. Pré-processamento dos dados.....	35
3.1.4. Escolha de entradas a usar no processo de previsão.....	36
3.1.5. Considerações adicionais acerca das redes neuronais.....	41
3.2. Previsão de consumos com recurso a máquinas de suporte vetorial.....	43
3.3. Relação entre caracterização horária de consumos e previsão de curto prazo.....	46
4. Caracterização dos dados disponíveis e das séries temporais resultantes.....	49
4.1. Dados disponíveis no primeiro estudo de caso considerado.....	49
4.1.1. Análise da série temporal resultante do primeiro estudo de caso.....	53
4.1.2. Processo de elaboração do diagrama sintetizado associado ao primeiro estudo de caso a partir de perfis de consumo.....	55
4.1.3. Correlação cruzada entre o diagrama de carga real e o diagrama de carga sintetizado associados ao primeiro estudo de caso.....	59
4.2. Dados disponíveis no segundo estudo de caso considerado.....	60
4.2.1. Análise da série temporal resultante do segundo estudo de caso.....	62

4.2.2. Processo de elaboração do diagrama de carga sintetizado associado ao segundo estudo de caso a partir de perfis de consumo .....	64
4.2.3. Correlação cruzada entre o diagrama de carga real e o diagrama de carga sintetizado associados ao segundo estudo de caso .....	67
4.2.4. Análise dos dados climatéricos disponíveis na área geográfica coincidente com a série temporal do segundo estudo de caso .....	69
5. Aplicação de métodos de previsão aos estudos de caso considerados .....	73
5.1. Uso de redes neuronais em previsões de consumos de energia elétrica.....	73
5.1.1. Análise de sensibilidades das redes neuronais baseadas em derivadas parciais. .....	81
5.2. Uso de máquinas de suporte vetorial em previsões de consumo de eletricidade ....	83
5.2.1. Descrição da meta-heurística Simulated Annealing.....	83
5.2.2. Recurso a Simulated Annealing para permitir a previsão baseada em máquinas de suporte vetorial .....	85
5.2.3. Adaptação da metodologia para inclusão de estágio associado à seleção de parâmetros explicativos .....	89
5.3. Apuramento da importância relativa das diferentes entradas a partir de modelos baseados em máquinas de suporte vetorial.....	91
6. Resultados obtidos a partir dos modelos de previsão adotados.....	93
6.1. Resultados de previsão com redes neuronais relativos ao primeiro estudo de caso considerado.....	94
6.1.1. Análise de sensibilidades da rede neuronal adotada no primeiro estudo de caso considerado.....	101
6.2. Aplicação da metodologia baseada em máquinas de suporte vetorial para a previsão de consumos relativos ao primeiro estudo de caso considerado .....	104
6.2.1. Descrição do processo de criação de diferentes modelos horários otimizados .. .....	104
6.2.2. Apresentação de resultados de previsão com modelos baseados em máquinas de suporte vetorial .....	106
6.2.3. Análise de importância das entradas mais relevantes no modelo de previsão baseado em máquinas de suporte vetorial .....	111
6.3. Resultados de previsão com redes neuronais aplicados ao segundo estudo de caso considerado.....	117

6.3.1. Análise de sensibilidades da rede neuronal adotada no segundo estudo de caso considerado .....	122
6.4. Aplicação da metodologia baseada em máquinas de suporte vetorial para a previsão de consumos relativos ao segundo estudo de caso considerado .....	124
6.4.1. Descrição do processo de criação de diferentes modelos horários otimizados .. .....	124
6.4.2. Apresentação de resultados de previsão com modelos baseados em máquinas de suporte vetorial .....	126
6.4.3. Análise de importância das entradas mais relevantes no modelo de previsão baseado em máquinas de suporte vetorial .....	131
6.5. Avaliação de modelo apenas baseado em informação de consumo resultante de perfis e em variáveis meteorológicas para previsão aplicada ao segundo estudo de caso .....	136
7. Conclusões e pistas de trabalho futuro .....	141
Anexo A. – Segmentação dos dados em distintos períodos de treino, de validação e de teste.....	A-1
Anexo B. – Detalhe da análise de sensibilidades às diferentes horas do dia seguinte em modelos RN com DC sintetizado.....	B-1
Anexo C. – Descrição do processo de otimização associado aos modelos MSV adotados .....	C-1
Anexo D. – Detalhe da análise de sensibilidades às diferentes horas do dia seguinte em modelos MSV com DC sintetizado .....	D-1
Referências bibliográficas .....	



## LISTA DE FIGURAS

Figura 1.1 – Esquema de reorganização do setor elétrico .....	1
Figura 1.2 – Curvas agregadas típicas de oferta e de procura de energia elétrica.....	4
Figura 2.1 – Perfis de consumo relativos a consumidores BTN residenciais.....	14
Figura 2.2 – Perfis de consumo relativos a consumidores BTN comerciais.....	14
Figura 2.3 – Perfis de consumo relativos a consumidores BTN industriais.....	15
Figura 2.4 – Perfis de consumo relativos ao setor industrial (consumidores BTE) por dia de semana .....	15
Figura 2.5 – Perfis de consumo relativos a consumidores BTN do setor hoteleiro/restauração.....	16
Figura 2.6 – Perfis de consumo relativos a consumidores BTN de serviços.....	16
Figura 2.7 – Perfis de consumo aprovados pela ERSE (semana de Inverno) .....	20
Figura 3.1 – Modelo genérico de um neurónio artificial.....	27
Figura 3.2 – Exemplo de uma arquitetura de rede neuronal com 2 neurónios na camada intermédia.....	27
Figura 3.3 – Exemplos de funções de ativação típicas em redes neuronais .....	28
Figura 3.4 – Exemplo de validação cruzada com curvas de evolução de erro.....	32
Figura 3.5 – Evolução da procura de energia com a temperatura média registada .....	38
Figura 3.6 – Regressão com base em máquinas de suporte vetorial .....	44
Figura 4.1 – Distribuição do número total de consumidores monitorizados por área geográfica .....	50
Figura 4.2 – Evolução do número de consumidores com dados de consumo disponíveis ao longo do período que reporta às monitorizações efetuadas .....	50
Figura 4.3 – Diagrama de carga resultante da agregação dos consumos medidos em consumidores individuais .....	51
Figura 4.4 – Diagrama de carga resultante da agregação aleatória de 740 consumidores . .....	52
Figura 4.5 – Distribuição do consumo mensal agregado.....	52
Figura 4.6 – Diagrama de carga semanal resultante da agregação aleatória de 740 consumidores.....	53
Figura 4.7 – Análise de autocorrelação da série temporal considerada no primeiro estudo de caso.....	54

Figura 4.8 – Coeficientes de correlação identificados no primeiro estudo de caso para diferentes atrasos impostos (em dias).....	55
Figura 4.9 – Repartição do consumo total pelas diferentes classes de consumo consideradas durante o período estudado .....	57
Figura 4.10 – Diagrama de carga sintetizado evidenciando a contribuição de cada classe .....	58
Figura 4.11 – Comparação entre o diagrama de carga real e o diagrama de carga sintetizado obtido para o primeiro estudo de caso.....	59
Figura 4.12 – Correlograma entre registos horários de potência ativa do diagrama de carga real e registos horários de potência ativa do diagrama de carga sintetizado .....	59
Figura 4.13 – Diagrama de carga da saída de Monte Real na subestação de Ortigosa ..	61
Figura 4.14 – Distribuição do consumo mensal na saída de Monte Real da subestação de Ortigosa .....	62
Figura 4.15 – Diagrama de carga semanal da saída de Monte Real da subestação de Ortigosa .....	62
Figura 4.16 – Análise de autocorrelação da série temporal considerada no segundo estudo de caso.....	63
Figura 4.17 – Coeficientes de correlação identificados no segundo estudo de caso para diferentes atrasos impostos (em dias).....	63
Figura 4.18 – Procedimento ilustrativo do processo de inferência dos pesos assumidos por cada classe no DC sintetizado .....	65
Figura 4.19 – Evolução da função objetivo durante o processo de otimização.....	66
Figura 4.20 – Diagrama de carga sintetizado evidenciando a contribuição de cada classe .....	67
Figura 4.21 – Comparação entre o diagrama de carga real e o diagrama de carga sintetizado obtido para o segundo estudo de caso .....	67
Figura 4.22 – Correlograma entre registos horários de potência ativa do diagrama de carga real e registos horários de potência ativa do diagrama de carga sintetizado .....	68
Figura 4.23 – Análise de correlação do consumo diário de energia do segundo estudo de caso com a temperatura máxima e mínima local.....	69
Figura 4.24 – Análise de correlação do consumo diário de energia do segundo estudo de caso com a temperatura máxima local para épocas frias.....	70
Figura 4.25 - Análise de correlação do consumo diário de energia do segundo estudo de caso com a humidade relativa e o nível de insolação do local .....	71



Figura 5.1 – Ilustração da estratégia de segmentação dos dados disponíveis no primeiro estudo de caso.....	76
Figura 5.2 – Ilustração da estratégia de segmentação dos dados disponíveis no segundo estudo de caso.....	76
Figura 5.3 – Modelo I inspirado em redes neuronais para a previsão de consumos do dia seguinte .....	78
Figura 5.4 – Modelo II inspirado em redes neuronais para a previsão de consumos do dia seguinte.....	79
Figura 5.5 – Modelo III inspirado em redes neuronais para a previsão de consumos do dia seguinte.....	80
Figura 5.6 – Descrição sumária do algoritmo associado ao Simulated Annealing .....	85
Figura 5.7 – Descrição sumária da metodologia de previsão baseada em máquinas de suporte vetorial .....	86
Figura 6.1 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) numa semana típica de novembro de 2004 (período de treino) – 1º Estudo de Caso (1º EdC).....	95
Figura 6.2 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) numa semana típica de março de 2005 (período de treino) (1º EdC) .....	95
Figura 6.3 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) numa semana típica de fevereiro de 2006 (período de teste) (1º EdC).....	96
Figura 6.4 – Distribuição dos erros percentuais por modelo (RN) durante o período de treino considerado (1º EdC) .....	96
Figura 6.5 – Distribuição dos erros percentuais por modelo (RN) durante o período de teste considerado (1º EdC) .....	97
Figura 6.6 – Histograma dos erros percentuais por modelo (RN) para o período de treino considerado (1º EdC).....	97
Figura 6.7 – Histograma dos erros percentuais por modelo (RN) para o período de teste considerado (1º EdC).....	98
Figura 6.8 – Comparação na distribuição de erros entre modelos de previsão (RN) (boxplot) para os dados de treino (1º EdC) .....	98
Figura 6.9 – Comparação na distribuição de erros entre modelos de previsão (RN) (boxplot) para os dados de teste (1º EdC) .....	99

Figura 6.10 – Distribuição de erros em modelos de previsão (RN) para as diferentes horas do dia (dados de treino) (1º EdC) .....	99
Figura 6.11 – Distribuição de erros em modelos de previsão (RN) para as diferentes horas do dia (dados de teste) (1º EdC) .....	100
Figura 6.12 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) numa semana atípica com dia de Carnaval – 8 de fevereiro de 2005 (1º EdC).....	101
Figura 6.13 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) numa semana atípica com feriado de 5 de outubro de 2005 (1º EdC).....	101
Figura 6.14 – Médias das sensibilidades horárias de cada saída do Modelo (RN) relativamente às 24 h de cada vetor de entrada (1º EdC).....	102
Figura 6.15 – Sensibilidades manifestadas pelas saídas do modelo RN referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (1º EdC) .....	103
Figura 6.16 – Evolução do mínimo resultado de EPAM para o conjunto de validação – modelo referente à hora 0 (1º EdC).....	105
Figura 6.17 – Evolução do resultado EPAM verificado no conjunto de validação durante o processo de otimização – modelo referente à hora 0 (1º EdC).....	105
Figura 6.18 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana típica de novembro de 2004 (período de treino) (1º EdC) .....	107
Figura 6.19 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana típica de fevereiro de 2006 (período de teste) (1º EdC).....	107
Figura 6.20 – Distribuição dos erros percentuais por modelo (MSV) durante o período de treino considerado (1º EdC).....	107
Figura 6.21 – Distribuição dos erros percentuais por modelo (MSV) durante o período de teste considerado (1º EdC).....	108
Figura 6.22 – Histograma dos erros de previsão com modelos MSV para os períodos de treino (1º EdC).....	108
Figura 6.23 – Histograma dos erros de previsão com modelos MSV para os períodos de teste (1º EdC).....	108

Figura 6.24 – Comparação na distribuição de erros entre modelos de previsão (MSV) (boxplot) para os dados de treino (1º EdC) .....	109
Figura 6.25 – Comparação na distribuição de erros entre modelos de previsão (MSV) (boxplot) para os dados de teste (1º EdC) .....	109
Figura 6.26 – Distribuição de erros em modelos de previsão (MSV) para as diferentes horas do dia (dados de treino) (1º EdC) .....	109
Figura 6.27 – Distribuição de erros em modelos de previsão (MSV) para as diferentes horas do dia (dados de teste) (1º EdC) .....	110
Figura 6.28 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana atípica com dia de Carnaval – 8 de fevereiro de 2005 (1º EdC).....	111
Figura 6.29 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana atípica com feriado de 5 de outubro de 2005 (1º EdC) .....	111
Figura 6.30 – Médias das sensibilidades horárias de cada saída de Modelo MSV (sem seleção de entradas) relativamente às 24 h de cada vetor de entrada (1º EdC).....	114
Figura 6.31 – Sensibilidades manifestadas pelas saídas do modelo MSV (sem seleção de entradas) referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (1º EdC) .....	115
Figura 6.32 – Médias das sensibilidades horárias de cada saída de Modelo MSV (com seleção de entradas) relativamente às horas contributivas de cada vetor de entrada (1º EdC) .....	115
Figura 6.33 – Sensibilidades manifestadas pelas saídas do modelo MSV (com seleção de entradas) referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (1º EdC) .....	116
Figura 6.34 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir dos diferentes modelos de previsão (RN) numa semana típica de fevereiro de 2007 (período de treino) (2º EdC).....	118
Figura 6.35 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir dos diferentes modelos de previsão (RN) numa semana típica de outubro de 2008 (período de treino) (2º EdC) .....	118
Figura 6.36 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir dos diferentes modelos de previsão (RN) numa semana típica de outubro de 2009 (período de teste) (2º EdC) .....	118

Figura 6.37 – Distribuição dos erros percentuais por modelo (RN) durante o período de treino considerado (2º EdC) .....	119
Figura 6.38 – Distribuição dos erros percentuais por modelo (RN) durante o período de teste considerado (2º EdC) .....	119
Figura 6.39 – Histograma dos erros percentuais por modelo (RN) para o período de treino considerado (2º EdC) .....	119
Figura 6.40 – Histograma dos erros percentuais por modelo (RN) para o período de teste considerado (2º EdC) .....	120
Figura 6.41 – Comparação na distribuição de erros entre modelos de previsão (RN) (boxplot) para os dados de treino (2º EdC) .....	120
Figura 6.42 – Comparação na distribuição de erros entre modelos de previsão (RN) (boxplot) para os dados de teste (2º EdC) .....	120
Figura 6.43 – Distribuição de erros em modelos de previsão (RN) para as diferentes horas do dia (dados de treino) (2º EdC) .....	121
Figura 6.44 – Distribuição de erros em modelos de previsão (RN) para as diferentes horas do dia (dados de teste) (2º EdC) .....	121
Figura 6.45 – Médias das sensibilidades horárias de cada saída do Modelo RN relativamente às 24 h de cada vetor de entrada (2º EdC) .....	122
Figura 6.46 – Sensibilidades manifestadas pelas saídas do modelo RN referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (2º EdC) .....	123
Figura 6.47 – Evolução do mínimo resultado de EPAM para o conjunto de validação – modelo referente à hora 0 (2º EdC) .....	125
Figura 6.48 – Evolução do resultado EPAM verificado no conjunto de validação durante o processo de otimização – modelo referente à hora 0 (2º EdC).....	125
Figura 6.49 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana típica de fevereiro de 2007 (período de treino) (2º EdC) .....	126
Figura 6.50 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana típica de outubro de 2009 (período de teste) (2º EdC) .....	127
Figura 6.51 – Distribuição dos erros percentuais por modelo (MSV) durante o período de treino considerado (2º EdC).....	127

Figura 6.52 – Distribuição dos erros percentuais por modelo (MSV) durante o período de teste considerado (2º EdC).....	128
Figura 6.53 – Histograma dos erros de previsão com modelos MSV para os períodos de treino (2º EdC).....	128
Figura 6.54 – Histograma dos erros de previsão com modelos MSV para os períodos de teste (2º EdC).....	129
Figura 6.55 – Comparação na distribuição de erros entre modelos de previsão (MSV) (boxplot) para os dados de treino (2º EdC) .....	129
Figura 6.56 – Comparação na distribuição de erros entre modelos de previsão (MSV) (boxplot) para os dados de teste (2º EdC) .....	130
Figura 6.57 – Distribuição de erros em modelos de previsão (MSV) para as diferentes horas do dia (dados de treino) (2º EdC) .....	130
Figura 6.58 – Distribuição de erros em modelos de previsão (MSV) para as diferentes horas do dia (dados de teste) (2º EdC) .....	130
Figura 6.59 – Médias das sensibilidades horárias de cada saída de Modelo MSV (sem seleção de entradas) relativamente às 24 h de cada vetor de entrada (2º EdC).....	133
Figura 6.60 – Sensibilidades manifestadas pelas saídas do modelo MSV (sem seleção de entradas) referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (2º EdC) .....	134
Figura 6.61 – Médias das sensibilidades horárias de cada saída de Modelo MSV (com seleção de entradas) relativamente às horas contributivas de cada vetor de entrada (2º EdC) .....	135
Figura 6.62 – Sensibilidades manifestadas pelas saídas do modelo MSV (com seleção de entradas) referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (2º EdC) .....	136
Figura 6.63 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) com recurso ao DC sintetizado e a variáveis meteorológicas (semana durante o período de teste).....	138
Figura 6.64 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (MSV) com recurso ao DC sintetizado e a variáveis meteorológicas (semana durante o período de teste) .....	139
Figura Anexo A.1 – 1º Cenário alternativo de validação cruzada para o primeiro estudo de caso .....	A-1

Figura Anexo A.2 – 2º Cenário alternativo de validação cruzada para o primeiro estudo de caso .....	A-2
Figura Anexo A.3 – 1º Cenário alternativo de validação cruzada para o segundo estudo de caso .....	A-3
Figura Anexo A.4 – 2º Cenário alternativo de validação cruzada para o segundo estudo de caso .....	A-4
Figura Anexo B.1 – Análise de sensibilidades horárias verificadas em modelos RN para o primeiro estudo de caso considerado.....	B-3
Figura Anexo B.2– Análise de sensibilidades horárias verificadas em modelos RN para o segundo estudo de caso considerado .....	B-5
Figura Anexo C.1 – Otimização do processo de previsão associado ao modelo horário referente à hora 8 do dia seguinte – sem seleção de entradas no 1º estudo de caso .....	C-1
Figura Anexo C.2 – Otimização do processo de previsão associado ao modelo horário referente à hora 16 do dia seguinte – sem seleção de entradas no 1º estudo de caso ...	C-1
Figura Anexo C.3 – Otimização do processo de previsão associado ao modelo horário referente à hora 0 do dia seguinte – com seleção de entradas no 1º estudo de caso ....	C-3
Figura Anexo C.4 – Otimização do processo de previsão associado ao modelo horário referente à hora 8 do dia seguinte – com seleção de entradas no 1º estudo de caso ....	C-3
Figura Anexo C.5 – Otimização do processo de previsão associado ao modelo horário referente à hora 16 do dia seguinte – com seleção de entradas no 1º estudo de caso...	C-3
Figura Anexo C.6 – Otimização do processo de previsão associado ao modelo horário referente à hora 8 do dia seguinte – sem seleção de entradas no 2º estudo de caso .....	C-5
Figura Anexo C.7 – Otimização do processo de previsão associado ao modelo horário referente à hora 16 do dia seguinte – sem seleção de entradas no 2º estudo de caso ...	C-5
Figura Anexo C.8 - Otimização do processo de previsão associado ao modelo horário referente à hora 0 do dia seguinte – com seleção de entradas no 2º estudo de caso ....	C-7
Figura Anexo C.9 – Otimização do processo de previsão associado ao modelo horário referente à hora 8 do dia seguinte – com seleção de entradas no 2º estudo de caso ....	C-7
Figura Anexo C.10 – Otimização do processo de previsão associado ao modelo horário referente à hora 16 do dia seguinte – com seleção de entradas no 2º estudo de caso...	C-7
Figura Anexo D.1 – Análise de sensibilidades horárias verificadas em modelos MSV sem seleção de entradas para o primeiro estudo de caso considerado.....	D-3
Figura Anexo D.2 – Análise de sensibilidades horárias verificadas em modelos MSV com seleção de entradas para o primeiro estudo de caso considerado .....	D-5

Figura Anexo D.3 – Análise de sensibilidades horárias verificadas em modelos MSV sem seleção de entradas para o segundo estudo de caso considerado .....	D-7
Figura Anexo D.4 – Análise de sensibilidades horárias verificadas em modelos MSV com seleção de entradas para o segundo estudo de caso considerado.....	D-9





## LISTA DE TABELAS

Tabela 2.1 – Descrição da classificação assumida pela ERSE para consumidores BT..	20
Tabela 4.1 – Coeficientes de correlação linear entre registos horários de potência ativa do DC real e outros registos analisados no primeiro estudo de caso.....	60
Tabela 4.2 – Contribuição apurada por cada classe de consumidores BTN para caracterização do consumo identificado na série temporal associada ao segundo estudo de caso .....	66
Tabela 4.3 – Coeficientes de correlação linear entre registos horários de potência ativa do DC real e outros registos analisados no segundo estudo de caso .....	68
Tabela 5.1 – Valores máximos por parâmetro nas máquinas de suporte vetorial .....	87
Tabela 5.2 – Parâmetros e critérios associados ao processo de <i>Simulated Annealing</i> ...	88
Tabela 6.1 – Resultados de previsão baseada em redes neuronais para o primeiro estudo de caso .....	94
Tabela 6.2 – Comparação entre os modelos de previsão (RN) aplicados ao primeiro estudo de caso para dias atípicos (como feriados e pontes) .....	100
Tabela 6.3 – Descrição de alguns dos modelos MSV obtidos (para horas específicas) (1º EdC) .....	104
Tabela 6.4 – Resultados de previsão baseada em máquinas de suporte vetorial para o primeiro estudo de caso .....	106
Tabela 6.5 – Erros de previsão (modelos MSV) verificados no primeiro estudo de caso para dias atípicos (como feriados e pontes).....	110
Tabela 6.6 – Efeito de remoção de diferentes vetores de entrada usando o modelo MSV sem seleção de entradas (no primeiro estudo de caso) .....	112
Tabela 6.7 – Efeito de remoção de diferentes vetores de entrada usando o modelo MSV com seleção de entradas (no primeiro estudo de caso).....	113
Tabela 6.8 – Resultados de previsão baseada em redes neuronais para o segundo estudo de caso .....	117
Tabela 6.9 – Descrição de alguns dos modelos MSV obtidos (para horas específicas) (2º EdC) .....	124
Tabela 6.10 – Resultados de previsão baseada em máquinas de suporte vetorial para o segundo estudo de caso .....	126
Tabela 6.11 – Efeito de remoção de diferentes vetores de entrada usando o modelo MSV sem seleção de entradas (no segundo estudo de caso).....	131

Tabela 6.12 – Efeito de remoção de diferentes vetores de entrada usando o modelo MSV com seleção de entradas (no segundo estudo de caso) .....	133
Tabela 6.13 – Resultados de previsão com redes neuronais baseadas apenas em informação de perfis de consumo e em variáveis climáticas .....	137
Tabela 6.14 – Resultados de previsão com máquinas de suporte vetorial baseadas apenas em informação de perfis de consumo e em variáveis climáticas.....	138
Tabela Anexo A.1 – Resultados de erros de previsão com modelos RN para o primeiro cenário alternativo de validação cruzada aplicado ao primeiro estudo de caso .....	A-1
Tabela Anexo A.2 – Resultados de erros de previsão com modelos RN para o segundo cenário alternativo de validação cruzada aplicado ao primeiro estudo de caso .....	A-2
Tabela Anexo A.3 – Resultados de erros de previsão com modelos RN para o primeiro cenário alternativo de validação cruzada aplicado ao segundo estudo de caso.....	A-3
Tabela Anexo A.4 – Resultados de erros de previsão com modelos RN para o segundo cenário alternativo de validação cruzada aplicado ao segundo estudo de caso.....	A-4
Tabela Anexo C.1 – Parâmetros MSV identificados e número de vetores de suporte resultantes após otimização – sem seleção de entradas no 1º estudo de caso .....	C-2
Tabela Anexo C.2 – Parâmetros MSV identificados e número de vetores de suporte resultantes após otimização – com seleção de entradas no 1º estudo de caso.....	C-4
Tabela Anexo C.3 – Parâmetros MSV identificados e número de vetores de suporte resultantes após otimização – sem seleção de entradas no 2º estudo de caso .....	C-6
Tabela Anexo C.4 – Parâmetros MSV identificados e número de vetores de suporte resultantes após otimização – com seleção de entradas no 2º estudo de caso.....	C-8

## **LISTA DE ACRÓNIMOS**

AT – Alta Tensão

BTE – Baixa Tensão Especial

BTN – Baixa Tensão Normal

DC- Diagrama de Carga

DSM – Demand Side Management

EdC – Estudo de Caso

EDP – Energias de Portugal

EE- Energia Elétrica

EPAM – Erro Percentual Absoluto Médio

REPQM – Raiz quadrada do Erro Percentual Quadrático Médio

ERSE – Entidade Reguladora dos Serviços Energéticos

MAT – Muito Alta Tensão

MIBEL – Mercado Ibérico de Eletricidade

MSV – Máquinas de Suporte Vetorial

MT – Média Tensão

PCT – Perfil de Carga Típico

REN – Rede Elétrica Nacional

RN – Redes Neurais

RRC – Regulamento de Relações Comerciais



# 1. Introdução

O capítulo introdutório desta tese visa dar uma perspetiva das principais alterações sentidas nos últimos anos na estrutura do setor elétrico, descrevendo o crescente envolvimento de um maior número de agentes económicos interessados em participar em regime de mercado liberalizado. É também pretendido descrever-se de forma sumária o novo paradigma em que o consumidor final toma um papel mais ativo no setor, sendo esperado que passe a estar mais exposto e, conseqüentemente, mais sensível à variação real dos preços de energia numa lógica de mercado e com o objetivo último de se promover uma almejada gestão da procura. A motivação para o uso de perfis de consumo como recurso para viabilizar a caracterização horária de consumos é também abordada neste capítulo, quer numa ótica de funcionamento de mercado concorrencial, quer na possível aplicação destes perfis no favorecimento de processos de previsão de consumos. A reorganização do setor elétrico reforça a importância da previsão de consumos, tema que merecerá um enquadramento numa secção integrada no presente capítulo.

## 1.1. Perspetiva sobre evolução de mercados de energia elétrica

A evolução mais recente e significativa da estrutura de organização dos sistemas elétricos de energia nos EUA e nos países europeus mais desenvolvidos desencadeou-se entre as décadas de 80<sup>[1]</sup> e 90 com uma separação clara das diferentes atividades da cadeia de valor <sup>[2]</sup>. O modelo proposto que demonstra a reorganização do setor apresenta-se na Figura 1.1.

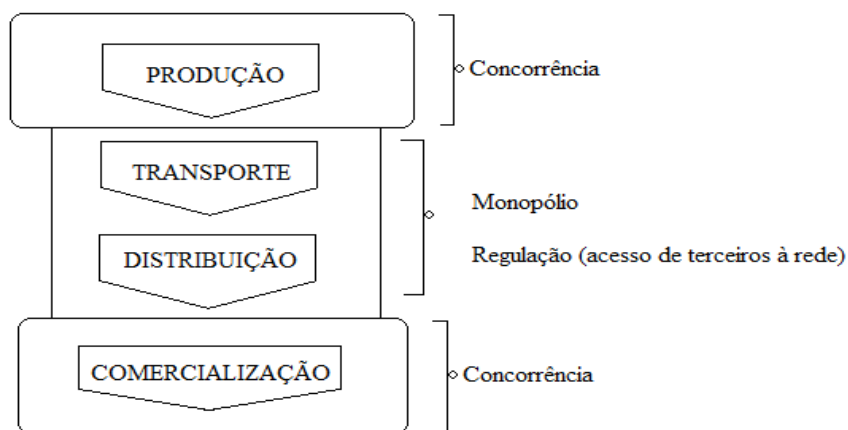


Figura 1.1 – Esquema de reorganização do setor elétrico

[1] A lei de 1982 exercida no Chile para a reestruturação do sistema elétrico foi pioneira a nível mundial a criar condições de mercado (Rudnick, 1994).

[2] O processo de reestruturação é normalmente conhecido na comunidade académica como *Unbundling*.

A Comunidade Europeia aprovou no ano de 1996 a diretiva 96/92/CE de modo a permitir uma evolução gradual para um maior grau de abertura e concorrência dos mercados de energia. Esta última foi revogada em 2003 pela diretiva 2003/54/CE, sendo reforçados os benefícios obtidos com o mercado interno de eletricidade, nomeadamente através de um aumento de eficiência nas diferentes atividades da cadeia de valor, aumento de competitividade, a melhoria de níveis de qualidade de serviço e o conseqüente efeito esperado de reduções de preços. A mais recente diretiva europeia 2009/72/CE perspetivou essencialmente criar regras apropriadas que permitam uma independência plena entre os operadores das redes e as atividades de produção e comercialização (evitando os indesejáveis conflitos de interesse), bem como reforçar a importância das conexões transfronteiriças para assegurar o abastecimento de energia elétrica proveniente das mais diversas formas possíveis e aos preços mais competitivos.

A liberalização dos mercados de energia viabiliza assim a competição do lado da geração e também no lado da comercialização, passando a energia elétrica a ser encarada como uma *commodity* sujeita às normais regras de mercado. As redes de transporte e distribuição de energia são encaradas como um meio físico para garantir o abastecimento de qualquer consumidor, justificando-se deste modo que estas duas atividades (transporte e distribuição) sejam exploradas sem qualquer concorrência (sendo também justificável pela incoerência que seria a duplicação da infraestrutura já existente) e que sejam devidamente reguladas por uma entidade externa e idealmente isenta de qualquer interesse económico ou político. Esta regulação é complexa já que deve atender a interesses porventura conflitantes como o equilíbrio económico-financeiro das empresas do setor, a satisfação do consumidor, a uniformidade tarifária (missão do serviço público), a pressão dos produtores independentes (no sentido de aumento de preços) e dos grandes consumidores (no sentido oposto) (Paiva, 2005). Uma adequada regulação deve atuar no sentido de garantir um acesso livre e não discriminatório das redes de energia aos consumidores e a qualquer produtor vinculado que pretenda injetar energia elétrica nas redes, bem como determinar uma tarifa de acesso às redes que simultaneamente permita o retorno do custo real de exploração das redes e admita determinadas estratégias de investimento para expansão ou adaptação das mesmas. As tarifas de acesso e uso da rede elétrica em Portugal são assim obtidas pela adição de diferentes componentes que visam a recuperação dos proveitos relacionados com a operação do sistema (tais como os custos com a Entidade Reguladora dos Serviços Energéticos – ERSE, sobrecustos de produção a partir de energias renováveis e outros custos associados à política energética), a recuperação dos proveitos relacionados com o estabelecimento, a operação e a manutenção das redes de transporte, a recuperação dos proveitos relacionados com o estabelecimento, operação e manutenção das redes de distribuição (possibilitando veicular a

energia elétrica dos seus pontos de receção até aos consumidores finais) e a recuperação dos proveitos da atividade de comercialização (contratação, leitura, faturação e cobrança dos serviços associados ao uso de redes e outros serviços regulados) (ERSE, 2011).

A concorrência no mercado grossista altera as transações para uma base horária (ou para intervalos de meia hora ou de quartos de hora) que é coordenada por um Operador de Mercado de modo a possibilitar o ajuste das flutuações intra-diárias da procura de energia. Os preços intra-diários são alterados de acordo com a interação entre os diferentes agentes de mercado e o tipo de tecnologia usada na produção de eletricidade de modo a satisfazer a procura (Soares e Sarmiento, 2010). Em paralelo deve existir, associada à atividade de transporte, a responsabilidade assumida pelo Operador do Sistema em coordenar as transações de eletricidade que garantam um equilíbrio físico em tempo real entre a oferta e a procura e a consequente garantia de abastecimento.

A tradicional estrutura vertical que caracterizava a organização das atividades dos serviços de energia, com um comprador único que celebrava contratos de aquisição de energia a longo prazo com companhias produtoras (partilhando assim os riscos inerentes a investimentos na construção de centrais e sua exploração), passou deste modo gradualmente a dismantelar-se pretendendo-se tendencialmente adotar o paradigma de compra a quem vende ao melhor preço e/ou pratica os serviços de valor acrescentado mais vantajosos. Os objetivos essenciais passam a ser a descentralização associada à procura de uma maior eficiência em cada uma das atividades e à concorrência e participação privada no setor. Cada comercializador poderá tomar posição no mercado quer pelo preço de venda de energia elétrica (*business as usual*), quer pela oferta de determinados serviços apelativos para o consumidor final. Uma questão problemática em alguns mercados de energia (como aliás sucede também em Portugal) tem sido efetivamente a persistência de preços regulados de eletricidade para o consumidor final <sup>[3]</sup>. Apesar de ser uma medida de transição recomendável <sup>[4]</sup>, existe o risco de limitar a concorrência, reduzir o investimento e contrapor o efeito esperado de *Unbundling*. O consumidor final deve passar a beneficiar de um ambiente de mercado, passando a sentir redução do preço final de energia <sup>[5]</sup> e a ter mais cenários de escolha quer de diferentes comercializadores a operar, quer de distintas opções tarifárias ou serviços de valor acrescentado associados e propostos pelos comercializadores.

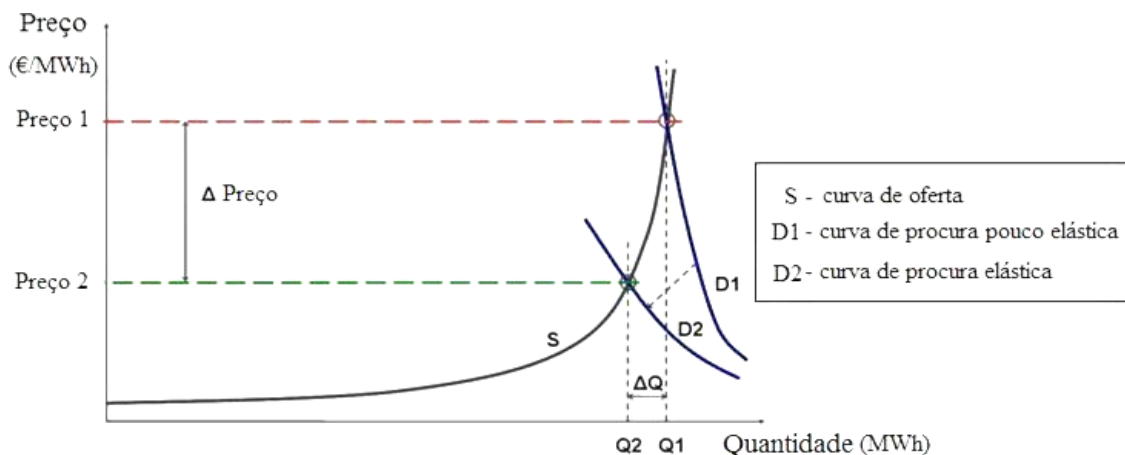
---

[3] À data de escrita desta tese as tarifas reguladas foram substituídas por tarifas transitórias, apesar de ainda se caracterizarem por serem fixadas pelo poder político.

[4] Coexistindo com os acordos de longo prazo de aquisição de eletricidade e com mercados ditos concorrenciais.

[5] Existem outros fatores não relacionados que podem distorcer o efeito esperado de redução de preço tais como a escassez dos combustíveis fósseis ou os custos de interesse económico geral onde se incluem os recentes investimentos em energias renováveis devido à crescente preocupação com o desenvolvimento sustentável.

O funcionamento saudável de um mercado competitivo deverá pressupor ainda uma maior exposição do consumidor final às variações reais do preço de energia, sentidas para cada período do dia. Deste modo, espera-se que o consumidor esteja mais sensível aos preços reais da energia em cada instante (período de integração) e assim adaptar o seu comportamento, despoletando-se uma desejada resposta da procura (*demand response*). Se as tarifas forem efetivamente variáveis segundo as horas do dia, espera-se uma maior elasticidade da procura, sendo esta uma forma de acautelar aumentos abruptos de preços e poder de mercado das companhias produtoras. Numa *pool* os preços e as quantidades de energia para cada período de tempo são apresentados previamente ao fornecimento efetivo de energia na lista de potenciais compras e vendas. Na Figura 1.2 é apresentado um gráfico com a variável preço unitário em função da quantidade de energia, evidenciando-se as curvas agregadas de oferta (curva *S*) e de procura (curvas *D1* e *D2*). O preço de mercado <sup>[6]</sup> é determinado pelo preço do grupo gerador mais oneroso, sendo encontrado no ponto de interseção entre as curvas de oferta e de procura do bem. Uma procura demasiado rígida (curva *D1*) aproxima-se de uma reta vertical com preços elevados (preço *P1*) (consumidores pouco sensíveis ao preço), resultando num mercado estrangulado pelas características da própria procura. Com uma maior elasticidade na procura (curva *D2*), os preços podem reduzir substancialmente (preço *P2*) <sup>[7]</sup>, possibilitando ao mercado recompor-se melhor a ofertas que se revelem elevadas, artifício que poderá por vezes ser criado intencionalmente quando existe elevada concentração no lado da produção. A adoção de uma estrutura de preços dinâmica e a aplicação de programas de redução de carga são estratégias usadas para o aumento de elasticidade da procura <sup>[8]</sup> e incentivo à participação no mercado retalhista.



**Figura 1.2 – Curvas agregadas típicas de oferta e de procura de energia elétrica**  
**Figura adaptada de IEA, 2003**

[6] Também denominado como preço marginal do sistema é o preço pago a todos os produtores.

[7] Estima-se que uma redução de 5% na procura de eletricidade nas horas de ponta, poder-se-ia ter refletido em reduções de preço da ordem dos 50% durante a crise energética sentida no estado da Califórnia em 2000/2001 (IEA, 2003).

[8] Aumento da elasticidade da procura tem sido nos últimos anos uma preocupação crescente devido à volatilidade dos preços praticados.



É atualmente assumido um risco pelas empresas comercializadoras de eletricidade, já que compram a preços sujeitos à volatilidade do mercado e vendem a preços acordados com os consumidores. O tipo de acordo entre o fornecedor e o utilizador final de eletricidade pode tomar diferentes formatos, sendo comum encontrar-se diversos modelos de tarifas (Kirschen, 2003) enunciados a seguir segundo a sua evolução histórica <sup>[9]</sup>:

– Contratos adotando tarifas fixas – o consumidor paga sempre um preço acordado, independentemente da procura real de energia elétrica vista pelo sistema para cada período do dia e das conseqüentes variações de preço real. A tarifa definida prevê naturalmente um prémio de seguro que compensa o fornecedor do risco associado;

– Contratos adotando tarifas variáveis segundo as horas do dia (como a adoção de tarifas bi-horárias ou tri-horárias) – parte do risco é assim transmitido ao consumidor, sendo este incentivado a deslocar parte do seu consumo para as horas de vazio quando o preço de energia é mais reduzido;

– Contratos que preveem eventuais interrupções de serviço, permitindo ao fornecedor diminuir a quantidade de energia a fornecer nos períodos de maior preço com contrapartida para o consumidor ao nível das tarifas praticadas;

– Contratos com tarifas associadas ao mercado *spot*, nos quais o consumidor é exposto aos preços de mercado <sup>[10]</sup>, servindo o fornecedor como intermediário <sup>[11]</sup> e sendo recompensado através de um prémio. Torna-se necessário por vezes estipular um preço máximo (*cap*) e um preço mínimo (*floor*) a pagar pelo consumidor para limitar o risco assumido pelas duas partes. Este tipo de contratos permite que o consumidor decida o nível de risco a assumir.

Os fornecedores de energia elétrica devem por esta razão dispor de informação considerada relevante sobre cada consumidor, de modo a acordar com estes o tipo de contrato que melhor sirva a ambas as partes e, para fazer face à concorrência, devem também diferenciar a sua oferta, não somente através do preço e/ou tipo de tarifa, mas também através dos serviços de valor acrescentado que propõem. Destes serviços destacam-se os programas de aconselhamento sobre utilização racional de energia, a divulgação de meios para redução da fatura através da escolha

---

[9] Apesar de se pretender evidenciar uma evolução histórica dos tipos de contrato adotados, não é lícito considerar-se que todos os mercados concorrenciais tenham já um grau de maturidade que permita encontrar tarifas associadas ao mercado *spot* para todo o universo de consumidores (como é o caso do MIBEL).

[10] Destinada apenas a grandes consumidores quando têm contratos bilaterais físicos com os produtores.

[11] A figura do fornecedor como intermediário deixa de existir nos contratos bilaterais celebrados entre produtores e consumidores considerados elegíveis (com elevados consumos ou características distintas).

apropriada de opções tarifárias <sup>[12]</sup> ou a oferta de produtos especializados como a proposta de venda de energia elétrica que seja produzida a partir de energias renováveis, ainda que a um preço superior, mas que poderá cativar um determinado nicho de mercado caracterizado por consumidores mais sensíveis a questões ambientais.

O conhecimento acerca dos padrões de consumo de diferentes grupos de consumidores torna-se naturalmente uma mais-valia para um adequado funcionamento de mercado em regime liberalizado, sendo exatamente nesta temática que reside um dos tópicos do estudo que será introduzido na secção seguinte.

## 1.2. Caracterização horária de consumos de energia elétrica

De forma histórica, a identificação de perfis de consumo tem sido usada de modo a favorecer a gestão e planeamento das redes elétricas e apoiar a fixação de tarifas adequadas, sendo esta estratégia atualmente imposta pela legislação para apurar o valor apropriado que os comercializadores devem pagar pelo acesso às redes elétricas para satisfazer a procura do seu conjunto de consumidores. A utilização da infraestrutura pode ser cobrada pelo distribuidor ao comercializador, tendo em conta o número e o(s) perfil(is) de consumo do conjunto dos seus clientes. A aplicação de perfis de consumo destina-se a apurar a forma como a energia é consumida por todos aqueles consumidores que não disponham de sistemas de telecontagem <sup>[13]</sup>, sendo encarada como alternativa à instalação massiva destes sistemas para todo o universo de consumidores <sup>[14]</sup>, o que acarretaria custos consideráveis e exigiria elevada coordenação técnica para abranger diferentes áreas geográficas. As vantagens dos sistemas de telecontagem reportam à caracterização real dos consumos horários dos consumidores finais, à possibilidade de uma maior sensibilização do consumidor para a *demand response*, a um funcionamento adequado de mercado liberalizado já que são favorecidas as intrincadas relações comerciais entre os agentes, evitando as leituras periódicas no local de consumo e reduzindo os processos fraudulentos que têm decorrido com alguns contadores eletrodinâmicos ainda instalados.

No processo de elaboração de perfis deve recorrer-se a um conjunto de dados de monitorização de uma amostra de diferentes consumidores, de forma a existir uma representatividade adequada

---

[12] Estes tipos de estratégias têm sido divulgadas pelas operadoras de fornecimento através de simuladores de potência a contratar ou simuladores de tarifas e de comparação de preços.

[13] Em Portugal são os consumidores de Baixa Tensão com uma tensão entre fases inferior a 1kV que não são obrigados a dispor de sistemas de telecontagem.

[14] Na fase em que este texto é redigido encontra-se em processo de análise e discussão da proposta de substituição dos contadores de energia elétrica no âmbito do plano de compatibilização regulatória (MIBEL).

ao processo de estimação. Os perfis obtêm-se através de uma estimativa do diagrama de carga médio de uma classe de consumidores durante um determinado período, pressupondo a normalização de cada diagrama em particular para que o resultado final se baseie apenas na forma assumida pelos diversos diagramas de carga individuais, contrariando contribuições significativas dos consumidores com maiores consumos, que sem uma adequada normalização poderiam inadvertidamente impor uma determinada forma no diagrama global. Os diferentes perfis a adotar devem ser tomados em número reduzido para tornar o processo de tratamento de dados mais expedito, representando grupos relativamente homogêneos de consumidores, que sejam claramente distintos entre si e cada perfil deve cobrir uma considerável percentagem de consumidores. Os perfis caracterizam diferentes hábitos de consumo (segundo diferentes particularidades dos consumidores), contemplando ainda uma análise sazonal para diferentes épocas do ano que permitem distinguir principalmente variáveis meteorológicas como temperatura, humidade ou efeitos de radiação solar. A caracterização deve também acautelar uma fundamental distinção entre dias de semana, dias de fim de semana e dias especiais como feriados.

A utilização de perfis de consumo possibilita a distribuição do registo acumulado de energia no final de um determinado período <sup>[15]</sup> por intervalos horários (ou intra-horários), sendo desta forma que se elabora o designado diagrama de carga reconstruído. Este diagrama reconstruído permite servir como aproximação para discriminar consumos por períodos horários de medição. De realçar que se os contadores instalados possibilitarem o registo de consumo para diferentes períodos tarifários (como o caso dos contadores bi-horários por exemplo), o diagrama de carga reconstruído basear-se-á em registos acumulados de energia para diferentes intervalos respeitantes ao horizonte temporal a estimar, logo é expectável que este diagrama acompanhe com maior precisão o diagrama de carga real, que na ausência de monitorização não é verdadeiramente conhecido e terá de ser estimado.

Os perfis são atualmente adotados de acordo com a legislação Portuguesa de modo a possibilitarem trocas comerciais entre os diferentes agentes envolvidos, procurando-se também evidenciar neste estudo a aplicabilidade destes perfis para efeitos de previsão de consumos.

---

[15] O intervalo entre duas leituras efetuadas pelo distribuidor não pode nunca ser superior a três meses (ERSE, 2012).

### 1.3. Previsão de consumos de energia elétrica

A previsão de consumos é uma tarefa de reconhecida importância histórica no que concerne ao planeamento dos sistemas de energia. Se por um lado se torna providencial no equilíbrio físico entre a oferta e a procura, torna-se também vantajosa numa ótica de médio/longo prazo para possibilitar uma análise de necessidades de reforço ou expansão da infraestrutura existente ou no planeamento de ações de manutenção de modo a garantir uma operação fiável do sistema. Numa perspetiva de curto prazo, a previsão de consumos pode tornar-se importante para encontrar uma otimizada configuração da rede (tendo como objetivo a minimização de perdas e a estabilidade de tensão), para possibilitar a decisão sobre as unidades de geração que devem satisfazer a procura (de acordo com a disponibilidade e custos de produção) e também para facilitar decisões relacionadas com a integração da geração distribuída. Com um adequado planeamento poderá consumir-se o intuito de ter eletricidade disponível em qualquer instante para satisfazer a procura solicitada, ao mais baixo custo e sem pôr em risco a normal operação do sistema.

O processo de liberalização de mercados de energia elétrica acrescenta ainda uma maior importância à temática de previsão de consumos. Com mais agentes económicos envolvidos, a pertinência económica acentua-se e, como tal, os métodos de previsão tornam-se uma ferramenta essencial para simplificar e fundamentar as trocas comerciais que venham a ser praticadas. É reconhecido que previsões acima ou abaixo do valor real de procura conduzem a perdas de proveitos e/ou a custos operacionais mais avultados (Fan e Hyndmand, 2012), pelo que se torna vital a adoção de métodos apropriados que minimizem os erros de previsão <sup>[16]</sup>. Como exemplos de que a liberalização do setor acarreta riscos económicos substanciais para a atividade de comercialização, pode supor-se que um fornecedor que subestime a quantidade de energia elétrica a adquirir para satisfazer a procura da sua carteira de clientes, vê-se forçado a recorrer ao mercado *spot* <sup>[17]</sup> adquirindo a energia em défice a preços muito provavelmente superiores para evitar a pressão dos consumidores ou as penalizações aplicadas pela entidade reguladora. Por outro lado, se o mesmo fornecedor previr por excesso o consumo de energia existe o risco de se produzir mais energia do que a necessária ou de se alocar um nível excessivo de reserva, desperdiçando-se energia e recursos, havendo também neste caso redução de proveitos e de valor acrescentado nesta atividade (Casazza e Delea, 2003, Fan e Hyndmand, 2012, Hippert e Pedreira, 2004). É imperativo que os comercializadores que atuam no mercado liberalizado

---

[16] Um estudo realizado em 1984 no mercado britânico de energia elétrica aponta que um erro de previsão de 1% resulta em aumentos nos custos anuais de operação na ordem dos 10 milhões de libras (Bunn e Farmer, 1985).

[17] Tipo de mercado em que as transações são feitas praticamente em simultâneo (apenas a minutos) com o efetivo fornecimento de energia elétrica, sendo que os preços procuram espelhar o preço real de energia elétrica no instante da transação, muito embora existe o risco de “algum” tipo de especulação.

negociem a quantidade necessária de eletricidade de modo a satisfazer as necessidades dos consumidores com quem contratualizam, sendo determinante disporem de informação fiável de procura de energia futura para melhorarem a sua participação no mercado.

São comumente distinguidas três categorias de previsão de consumo baseadas na duração da previsão para o planeamento e operação dos sistemas de energia. A primeira categoria, denominada como previsão de longo prazo, em que os horizontes temporais são previstos com uma antecedência superior a um ano, é tipicamente pensada no ponto de vista do planeamento dos sistemas de energia elétrica. Este tipo de previsão serve assim de referência a investimentos futuros no setor elétrico e a determinadas decisões comerciais. A segunda categoria, denominada como previsão de médio prazo prevê para períodos de meses até um ano. Este tipo de previsão torna-se necessário para a gestão e programação do recurso a fontes de energia primária, operações de manutenção e planeamento de intercâmbio através das interligações com outros mercados (ou países <sup>[18]</sup>). Em suma, este tipo de previsão é idealizado para uma gestão otimizada dos recursos disponíveis. A última categoria denomina-se como previsão de curto prazo e pretende prever para períodos horários seguintes (horas seguintes ou dias seguintes) sendo usada para uma operação apropriada do sistema de energia. A previsão de curto prazo assegura o compromisso assumido na exploração das unidades de geração para satisfazer a procura dos consumidores e compensar perdas associadas ao transporte e distribuição, facilita a coordenação entre diferentes tipos de energia primária (por exemplo entre centrais térmicas e hídricas), promove a normal operação do sistema possibilitando análises de rede e estudos de fluxos de carga, permite a elaboração de planos de contingência e gera informação para um despacho ótimo.

#### **1.4. Organização do documento**

Após o enquadramento que se pretende transmitir no presente capítulo acerca do conteúdo e dos propósitos desta tese, os restantes capítulos estão organizados da seguinte forma:

- o capítulo 2 descreve de uma forma detalhada a temática associada à caracterização horária de consumos de energia elétrica, referindo as vantagens e a motivação para esta estratégia dando especial enfoque aos diferentes métodos usados para se classificarem os consumidores a partir do seu padrão típico de consumo. O modelo de caracterização horária de consumos de energia elétrica adotado em Portugal é ilustrado já que será escrutinado em capítulos seguintes.

---

[18] O exemplo do Mercado Ibérico de Eletricidade – MIBEL, trata-se efetivamente de um mercado comum que serve de plataforma às transações de energia entre Portugal e Espanha.

- o capítulo 3 aponta alguns dos métodos aplicados na previsão de consumos. O capítulo inicia-se com um levantamento dos métodos tipicamente seguidos em estratégias de médio/longo prazo, centrando-se posteriormente em métodos aplicados a cenários de curto prazo, como o estudo pretende explorar. De entre os métodos de curto prazo, as redes neuronais e as máquinas de suporte vetorial serão os modelos que merecerão maior destaque, pelo que os modelos matemáticos providenciados por estas duas estratégias serão detalhados no decorrer deste capítulo.
- o capítulo 4 apresenta os dados disponíveis no âmbito deste estudo, como dados de consumos de energia elétrica monitorizados ao nível de consumidores individuais e ao nível de uma subestação de distribuição que abastece uma determinada área geográfica do distrito de Leiria, bem como dados climatéricos que possam ter alguma influência nos consumos registados. Neste capítulo são ainda descritas as séries temporais e estabelecidas análises de correlação entre as mesmas para se despistarem possíveis relações causa-efeito nos estudos de caso adotados.
- no capítulo 5 são descritas as metodologias que recorrem a alguns dos métodos de previsão apresentados e sistematizados no capítulo 3 para aplicação aos estudos de caso considerados. O capítulo pretende elucidar o leitor acerca das opções tomadas e de processos de otimização seguidos para que os modelos de previsão tivessem a aplicabilidade desejada. Neste capítulo são também indicadas e fundamentadas as formas usadas para avaliar a importância de determinadas variáveis como entradas em modelos de regressão.
- no capítulo 6 são apresentados resultados dos métodos explorados no capítulo 5, com uma cuidada comparação dos diferentes modelos criados, acompanhada de uma análise crítica e comentada de todas as ilações a reter.
- no capítulo 7 são apontadas as principais conclusões deste estudo, procurando avaliar-se a pertinência e rigor do mesmo para os objetivos que se pretendem alcançar (investigação da forma como a previsão de consumos é favorecida com informação proveniente de perfis de consumo e validação deste pressuposto a partir de diferentes métodos de previsão). Nesta secção serão ainda apresentadas pistas para trabalho futuro.

## *2. Caracterização horária de consumos de energia elétrica*

Este capítulo incide na caracterização de consumidores de energia elétrica e apresenta uma descrição detalhada de metodologias que podem ser seguidas para se proceder a uma classificação dos consumidores. Os perfis de carga (também denominados como perfis de consumo), são adotados como alternativa aos sistemas de telecontagem para caracterização horária dos consumos e resultam naturalmente do processo de classificação de consumidores adotado. Neste capítulo são apresentados diferentes perfis que podem ser assumidos consoante o processo de classificação escolhido e será apresentada em maior detalhe a estratégia adotada para a classificação de consumidores em Portugal e os perfis que daí resultam e que são indispensáveis para viabilizar o funcionamento do mercado liberalizado.

### **2.1. Perfis de consumo**

O consumo de energia elétrica, quer em termos de quantidade total consumida, quer em termos da sua distribuição temporal, pode variar consideravelmente estando sujeito a fatores como o tipo de atividade económica desenvolvida, o tipo de equipamento elétrico utilizado, os hábitos de consumidores, influência de condições meteorológicas, efeitos de sazonalidade, os tipos de dia considerados (dias de semana, fins de semana ou dias especiais como feriados), e/ou a maior ou menor resposta dos consumidores à volatilidade de preços de eletricidade. A caracterização de consumos pretende criar diferentes grupos de consumidores, consolidando-se a necessidade de se efetivar uma análise mais pormenorizada à forma como a energia se distribui para consumidores que pertençam a um mesmo grupo e de se estimar também a evolução de consumo ou a sensibilidade manifestada por cada grupo perante diferentes cenários de alteração consoante as alterações meteorológicas, os efeitos de sazonalidade, o desenvolvimento económico sentido, a atualização de tarifas ou o efeito de dias especiais.

Historicamente a utilização dos perfis favorecia já determinados processos associados às diferentes atividades da cadeia de valor, atividades que se encontravam integradas e tipicamente exploradas por uma mesma entidade <sup>[19]</sup>. Estes processos incluem o planeamento da rede, o

---

[19] Exceção feita a companhias de distribuição que poderiam ser locais comumente exploradas por entidades autárquicas.

investimento em capacidade de produção, gestão da procura (*Demand Side Management - DSM*), bem como a determinação das tarifas adequadas ao sistema elétrico. Em ambiente de mercado liberalizado, o uso de perfis torna-se obrigatório para permitir o acerto de contas entre o distribuidor e o comercializador sempre que as instalações consumidoras não disponham de um sistema de telecontagem <sup>[20]</sup> e os consumos horários tenham de ser estimados a partir de registo(s) acumulado(s) de energia lido(s) no contador tradicional (Eurelectric, 2000).

Os métodos de classificação aplicados a diagramas de carga de consumidores pressupõem um período de monitorização dos consumidores que compõem a amostra original considerada. Para cada consumidor deve ser encontrado um diagrama de carga diário típico, assumindo-se como a média dos registos disponíveis para cada período horário (intervalos de 15 em 15 minutos para o caso português), devendo na análise distinguir-se os diferentes tipos de dias e épocas do ano, de modo a obter-se para cada consumidor um diagrama de carga diário de referência para dias de semana (ou particularizando ainda diferentes tipos de dia em cada semana), sábados, domingos e diferentes estações do ano. Este efeito garante por um lado o apuramento de atividade económica e hábitos de consumo diferenciados para cada tipo de dia e, ao considerar diferentes épocas do ano, permite também distinguir diferentes necessidades de utilização de eletricidade para aquecimento, arrefecimento, ventilação, ar condicionado, bem como o recurso a sistemas de iluminação.

A normalização dos diagramas de carga típicos torna-se igualmente importante, já que a classificação deve ser baseada na forma assumida pelos diagramas, sendo evitável que dependa dos fatores de escala assumidos pelos registos. De entre os fatores de normalização encontrados em diversos trabalhos podem ser apresentados os seguintes:

– o valor da potência máxima registada no período de monitorização (Gerbec *et al.*, 2005) – este fator de normalização tem como desvantagem o facto da potência máxima poder ocorrer de forma imprevista e numa situação anómala, podendo implicar imprecisão no processo de normalização. No caso dos consumidores de menor consumo, que podem ainda ser uma parte considerável da amostra, os valores normalizados podem com este equívoco vir a assumir valores muito baixos, perdendo-se alguma informação acerca da forma do diagrama. Na verdade, os agrupamentos seriam feitos tendo também em linha de conta o valor de consumo e não somente a forma do diagrama;

---

[20] Instalações referentes a contratos BT com valores de tensão eficaz inferior ou igual a 1 kV (para contratos BTE com potências contratadas superiores a 41,4 kVA e contratos BTN com potências contratadas inferiores ou iguais a 41,4 kVA).



– o valor da potência média registada no período de monitorização (Jardini *et al.*, 2000), que permite manter a informação relativa ao consumo de energia, manifestando contudo o risco de se perder informação acerca da forma do diagrama de carga entre consumidores com valores de potência mais afastados do valor da potência média utilizada na normalização;

– a potência máxima registada no diagrama de carga representativo como fator de normalização tem-se revelado cada vez mais comum (Chicco *et al.*, 2003, Gerbec *et al.*, 2005), uma vez que os diagramas de cargas representativos de consumidores passam a variar assim numa gama de valores compreendidos entre [0;1], assumindo a mesma ordem de grandeza quando comparados entre si (Chicco *et al.*, 2003). Na sua essência, esta normalização preserva as relações iniciais entre os valores de dados, permitindo manter a informação relativa à forma da curva assumida por cada diagrama representativo.

## **2.2. Critérios usados na classificação de consumidores**

### **2.2.1. Classificação baseada no setor de atividade económica**

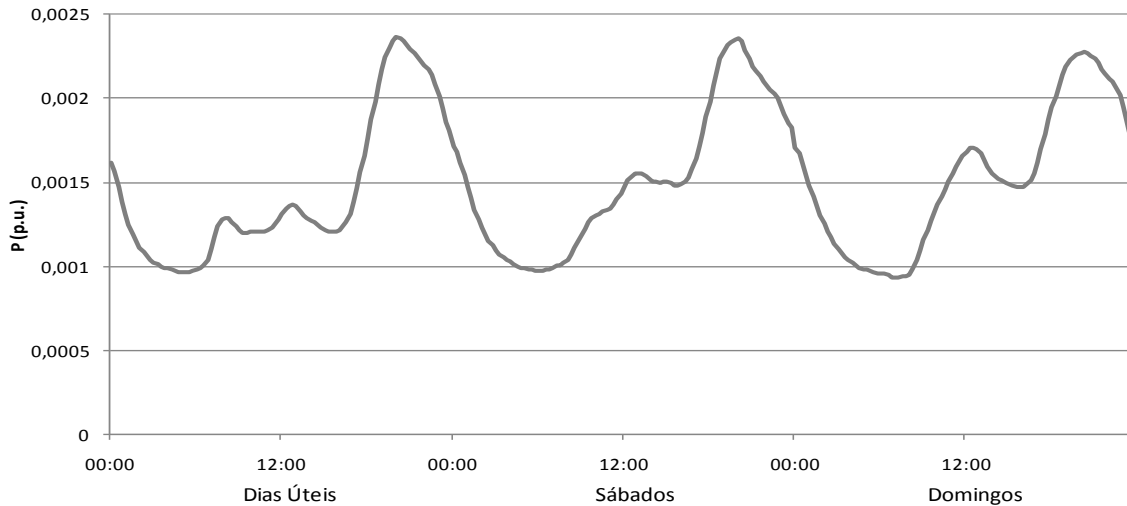
Um dos critérios que pode ser escolhido para a segmentação de consumidores baseia-se nos padrões de utilização comuns, discriminando os consumidores pela atividade económica desenvolvida. Deste modo são criados grupos de consumidores domésticos (Paatero e Lund, 2006), do setor comercial, do setor industrial, do setor hoteleiro/restauração e relativos a outros serviços, tais como escolas, companhias de seguros, agências bancárias, pavilhões desportivos, teatros,... (Jardini *et al.*, 2000).

Nas Figuras 2.1 a 2.6 são apresentados modelos de perfis que permitem descrever a forma como a energia elétrica é utilizada em distintos tipos de dias por diferentes grupos de consumidores de baixa tensão <sup>[21]</sup>. Nas análises que se seguem devem ser tidas em linha de conta as distribuições temporais do consumo pelas diferentes horas do dia já que os diagramas se apresentam normalizados numa base semanal e o total de consumo não é considerado.

Para os consumidores residenciais e durante os dias úteis, o período de maior consumo (ponta do diagrama de carga) ocorre entre as 18h30 e as 20h30, quando os residentes regressam às habitações, sendo natural considerar-se algum efeito de simultaneidade de diferentes cargas elétricas de maior potência, como sistemas de aquecimento ou ar condicionado e máquinas de lavar e fornos elétricos.

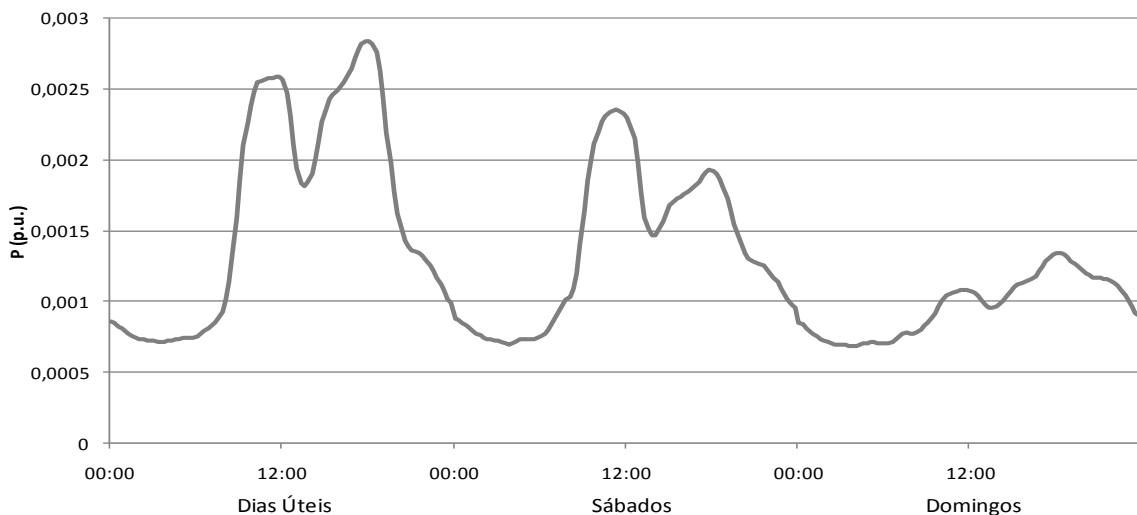
---

[21] Perfis apurados com base numa amostra representativa de diagramas de carga obtidos a partir de campanhas de monitorização que será detalhadamente descrita no Capítulo 4. Os perfis apresentados reportam ao período de Inverno.



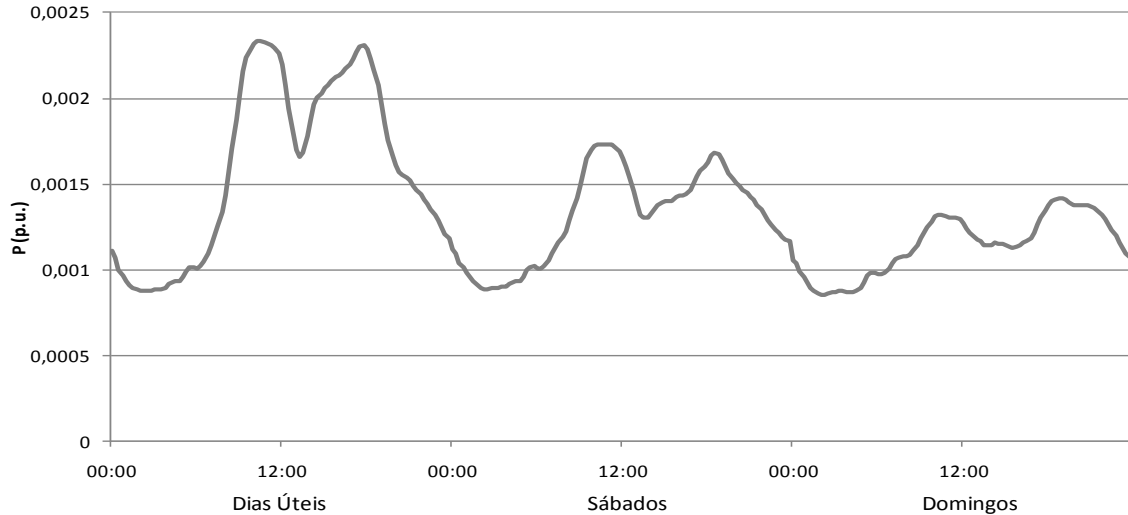
**Figura 2.1 – Perfis de consumo relativos a consumidores BTN residenciais**

A maior diferença que se pode verificar entre dias úteis e fins de semana está essencialmente relacionada com o início do período de consumo durante a manhã, já que a responsabilidade no cumprimento de horários obriga a que o consumo se verifique mais cedo em dias laborais.



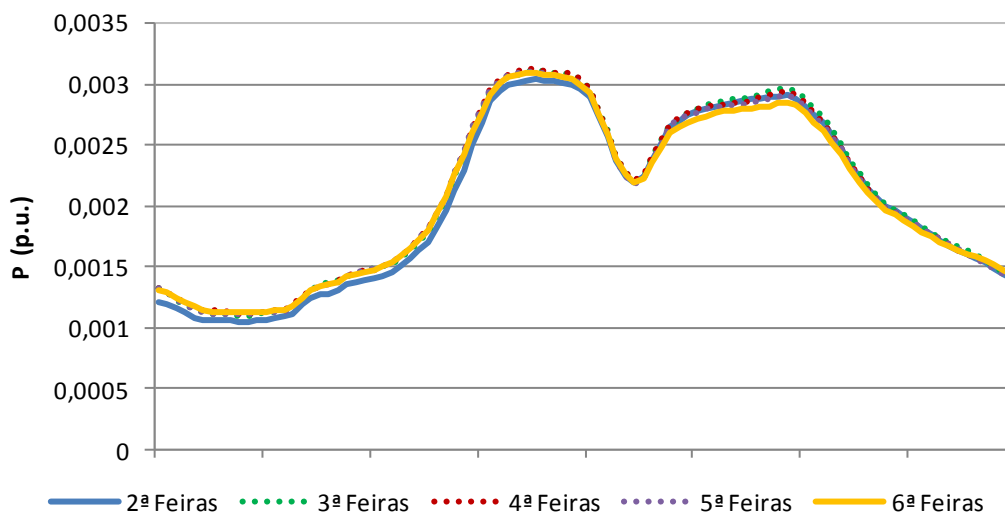
**Figura 2.2 – Perfis de consumo relativos a consumidores BTN comerciais**

No setor comercial, caracterizado na Figura 2.2, pode verificar-se que o consumo nos dias úteis acompanha o normal horário de expediente (sensivelmente entre as 9h e as 18h), com uma quebra de consumo associada à interrupção laboral durante as horas de almoço. Os sábados são caracterizados por terem um consumo mais pronunciado durante o período da manhã, com quebra da atividade económica a partir das 12 horas. Esta quebra na atividade económica é também visível durante os domingos que apresentam um consumo praticamente residual (a forma do diagrama apresenta-se mais plana).



**Figura 2.3 – Perfis de consumo relativos a consumidores BTN industriais**

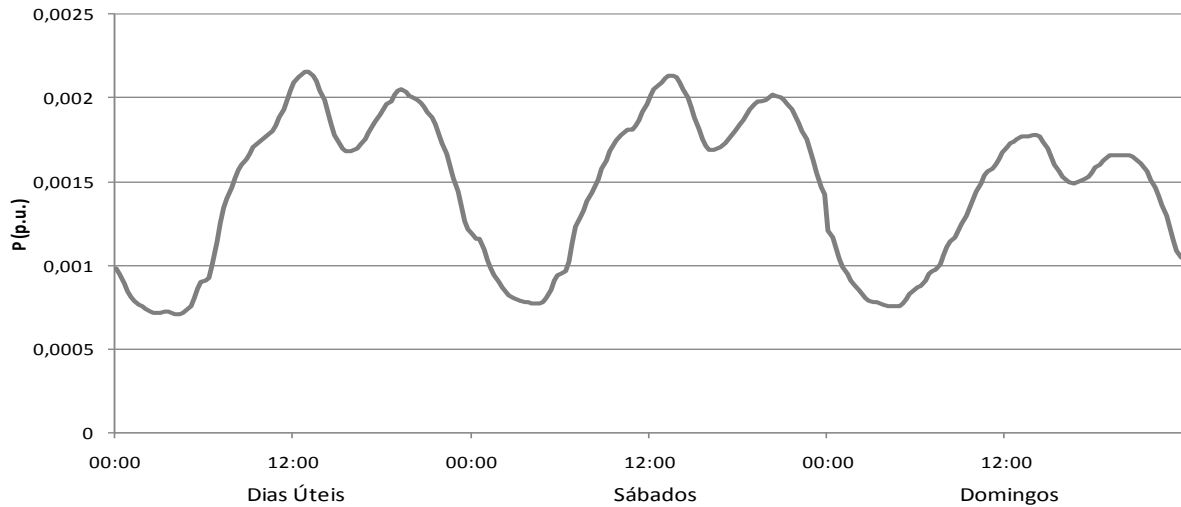
O setor industrial caracteriza-se por apresentar padrões de consumo que são ilustrados na Figura 2.3. Estes padrões evidenciam uma utilização de energia mais uniforme durante todo o dia (o fator de carga do diagrama é mais elevado do que no caso do perfil comercial). Ainda a respeito dos consumidores industriais pretende elucidar-se na Figura 2.4 como a utilização de energia sofre variações em diferentes dias da semana, com as segundas-feiras e as sextas-feiras a apresentarem um padrão um pouco distinto, essencialmente nas manhãs de segunda-feira e no final do dia útil de sexta-feira, correspondendo respetivamente ao início e final de atividade económica da semana.



**Figura 2.4 – Perfis de consumo relativos ao setor industrial (consumidores BTE) por dia de semana**

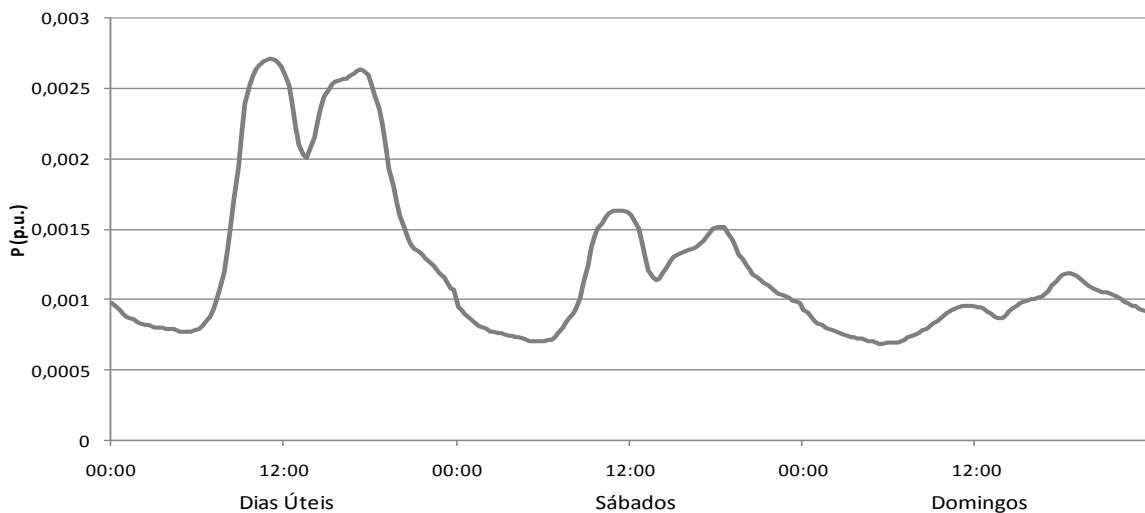
A agregação dos consumidores que pertencem aos setores hoteleiros e de restauração resulta em perfis de consumo (ilustrados na Figura 2.5) que mantêm uma forma análoga entre dias de

semana e fins de semana, com um peso considerável de consumo verificado no período noturno e com o efeito de picos correspondentes naturalmente aos horários típicos das refeições.



**Figura 2.5 – Perfis de consumo relativos a consumidores BTN do setor hoteleiro/restauração**

Os consumidores que apresentam uma atividade económica relacionada com o setor de serviços apresentam um padrão de consumo muito idêntico aos do setor comercial. A distribuição temporal, sendo concordante com o típico horário de expediente das instalações revela novamente uma quebra significativa a partir das 12h de Sábado, sendo que o consumo durante as manhãs de Sábado não se apresenta tão significativo como no caso dos utilizadores comerciais.



**Figura 2.6 – Perfis de consumo relativos a consumidores BTN de serviços**

A apresentação dos perfis por setor de atividade visa dar uma perspetiva mais detalhada do modo como é utilizada a energia elétrica por cada setor e para permitir uma comparação posterior com outros critérios usados para segmentação de classes de consumo. Com a apresentação destas distintas formas assumidas pelos diagramas de carga típicos pode também assumir-se que se torna expectável que o consumo agregado do sistema ou de uma determinada área geográfica de

dimensão regional (ao nível por exemplo da distribuição) espelhe as diferentes contribuições individuais de cada setor económico para cada hora do dia.

### 2.2.2. Classificação baseada em informação comercial

Um processo considerado expedito é o de formação de grupos de consumidores através de atributos comerciais. Entre estes, destacam-se algumas características que são comuns encontrarem-se nas faturas de energia elétrica, tais como: o nível de tensão <sup>[22]</sup>, a potência contratada, o consumo médio de energia ou o tipo de tarifa aplicada (tarifa simples, bi-horária ou tri-horária). Outros atributos podem ser derivados destes já apontados, dando-se como exemplo a utilização anual da potência contratada (como a relação entre o consumo de energia e a potência contratada). O risco associado a esta estratégia é que exista uma grande diversidade de padrões de consumo que podem caracterizar uma única classe e que as diferenças entre classes podem não ser tão substanciais como o desejável.

### 2.2.3. Classificação baseada em reconhecimento de padrões e em métodos de *clustering*

Uma forma alternativa de segmentar os consumidores pode resultar de uma classificação dos mesmos de acordo com um critério de reconhecimento de padrões aplicado aos diagramas de carga típicos analisados por cada consumidor individual (Chicco e Ilie, 2009, Fidalgo *et al.*, 2012, Gavrilas *et al.*, 2010, Jota *et al.*, 2011). Esta estratégia procura a formação de grupos de consumidores garantindo que cada grupo contemple padrões de consumo com elevada similitude e que exista uma clara distinção face a padrões admitidos em outros grupos criados. A abordagem evita uma classificação de consumidores que se baseie simplesmente em critérios como a atividade económica desenvolvida ou em dados comerciais, logrando de uma seriação estabelecida e fundamentada através de métodos *clustering*.

Estes métodos são usados para discriminar estruturas em dados sem necessidade de se proceder a uma interpretação formal desta disposição, baseando-se apenas em análises de proximidade verificada entre padrões. De entre os métodos *clustering* mais explorados destacam-se os algoritmos hierárquicos, os métodos partitivos e os métodos recorrendo a inteligência artificial. Nos métodos hierárquicos são encontrados sucessivos grupos (*clusters*) a partir de *clusters* previamente estabelecidos. De entre estes, os algoritmos hierárquicos aglomerativos são os mais comuns considerando-se numa fase inicial cada padrão de entrada como um *cluster* procedendo ao agrupamento destes em *clusters* sucessivamente maiores seguindo sempre uma estratégia de

---

[22] No caso do Regulamento Tarifário aplicado em Portugal, apenas faz sentido a discriminação entre consumidores BTE e BTN.

agrupamento baseado nas menores distâncias encontradas (Everitt *et al.*, 2001, Gerbec *et al.*, 2004) . Nos métodos de partição, logo na primeira iteração é assumida uma partição dos dados em diferentes grupos, sendo especificado *a priori* o número de grupos a considerar. O processo parte efetivamente de uma consideração aleatória inicial, sendo avaliada a função objetivo (função que determina a variância intra-classe) que se pretende minimizar. Em cada iteração são atualizados os protótipos (centros dos *clusters*) e avaliada a função objetivo, sendo que o processo iterativo termina quando o número máximo de iterações especificado é atingido ou a alteração sentida na função objetivo é inferior a um determinado valor limiar também especificado *a priori*. Uma variante deste que é denominado como método *k-means* tradicional é o método difuso *k-means* que inclui um grau de pertença de cada padrão aos diferentes grupos, sendo este grau de natureza difusa.

Em alternativa, os métodos baseados em inteligência artificial mereceram uma atenção muito considerável na última década do século XX (Everitt *et al.*, 2001). O recurso a redes neuronais é uma prática comum a uma vasta gama de aplicações para as quais métodos estatísticos convencionais são já sobejamente usados. As redes neuronais podem ser encaradas como algoritmos que pretendem simular as capacidades computacionais de extensas redes com relações entre elementos simples, como os neurónios que compõem um cérebro humano. O interesse neste tipo de técnicas aplicadas à classificação deriva do facto de se pretender imitar alguns dos tipos de tarefas desejáveis de reconhecimento de padrões para as quais o cérebro humano está adequadamente dotado.

Para além dos métodos *clustering* apresentados, existem outras variantes que pretendem beneficiar das vantagens oferecidas por métodos evolucionários. Ivanov e Gavrilas (2007) propõem o recurso a algoritmos genéticos como meta-heurística para apurar perfis de carga resultantes de grupos de consumidores. O algoritmo codifica possíveis soluções do problema, correspondendo no caso em concreto a cadeias de números (cromossomas) com um comprimento equivalente ao número de consumidores a classificar e em que cada número (gene) toma um valor no intervalo limitado pelo número total de grupos a considerar e especificado *a priori*. Diferentes combinações são geradas e o processo é avaliado recorrendo a uma função objetivo que incorpora o erro quadrático médio entre os perfis de carga típicos obtidos e os diagramas efetivamente monitorizados (que se pretende minimizar) e o erro quadrático médio entre os perfis obtidos (que se pretende maximizar). Deste modo, o objetivo da otimização é efetivamente determinar o melhor compromisso entre o número de perfis de carga típicos e a precisão na representação de consumidores.

É também proposto por Gavrilas *et. al* (2010) um método baseado no processo de acasalamento em colmeias, considerando que o processo de otimização deriva de meta-heurísticas como algoritmos genéticos (no processo de exploração de soluções) e de *Simulated Annealing* (que determinam a probabilidade de existirem cruzamentos na população existente).

Os perfis de consumo devem estar associados a cada grupo criado e resultam assim da média dos diagramas típicos dos consumidores que fazem parte de cada grupo. Após a formação dos grupos através de um processo de classificação não supervisionado (sem classes predefinidas *a priori*), deve ser estabelecida uma análise cruzada entre os grupos obtidos e atributos disponíveis dos consumidores, como sejam os atributos de índole comercial (potência contratada, consumo de energia ou o tipo de tarifa adotada), de natureza discriminatória do setor de atividade económica (consumidores residenciais, comerciais, industriais ou de serviços) ou resultantes de inquéritos direcionados aos consumidores. Esta análise cruzada visa identificar uma consistente correlação entre grupos formados e tais atributos, de modo a possibilitar aos consumidores que não sejam alvo de monitorização a atribuição expedita e sem ambiguidades de um perfil de consumo, apenas a partir de informação disponível numa base de dados que se pretende completa e atualizada com informação detalhada sobre os consumidores.

A problemática associada à utilização de perfis de consumidores reside, na grande maioria dos casos, no processo de afetação de perfis a consumidores não monitorizados (Gerbec *et al.*, 2005, Matos *et al.*, 2005). Usando uma estratégia de definição prévia das classes com base em dados comerciais ou do setor de atividade económica, pode originar-se uma diversidade considerável de padrões de consumo no seio de cada classe, com os diagramas de carga médios a afastarem-se consideravelmente dos protótipos (centros ideais dos grupos formados com base em padrões de proximidade). Se em alternativa se optar por recorrer aos métodos de *clustering* apresentados anteriormente, tem-se a garantia de rigor na identificação dos protótipos, muito embora a relação destes com informação comercial seja parca, o que dificulta a classificação de consumidores não monitorizados.

Uma metodologia proposta (Fidalgo *et al.*, 2012, Matos *et al.*, 2005) para lidar com este desafio baseia-se no recurso a uma classificação prévia que maximize a homogeneidade em cada classe através da criação de células de classificação do espaço da informação comercial disponível. Estas células são então agregadas através de um método de otimização que permita garantir a maior homogeneidade de diagramas em cada classe. A definição destas células de classificação consiste na seleção das variáveis a usar na classificação de consumidores (variáveis comerciais ou relativas à utilização de energia) e as gamas das variáveis são divididas em intervalos

discretos, de acordo com as suas distribuições na amostra. O número de intervalos deverá ser suficiente para garantir uma adequada separação mas simultaneamente limitado para evitar a geração de um número elevado de células.

#### 2.2.4. Metodologia de classificação de consumos adotada em Portugal

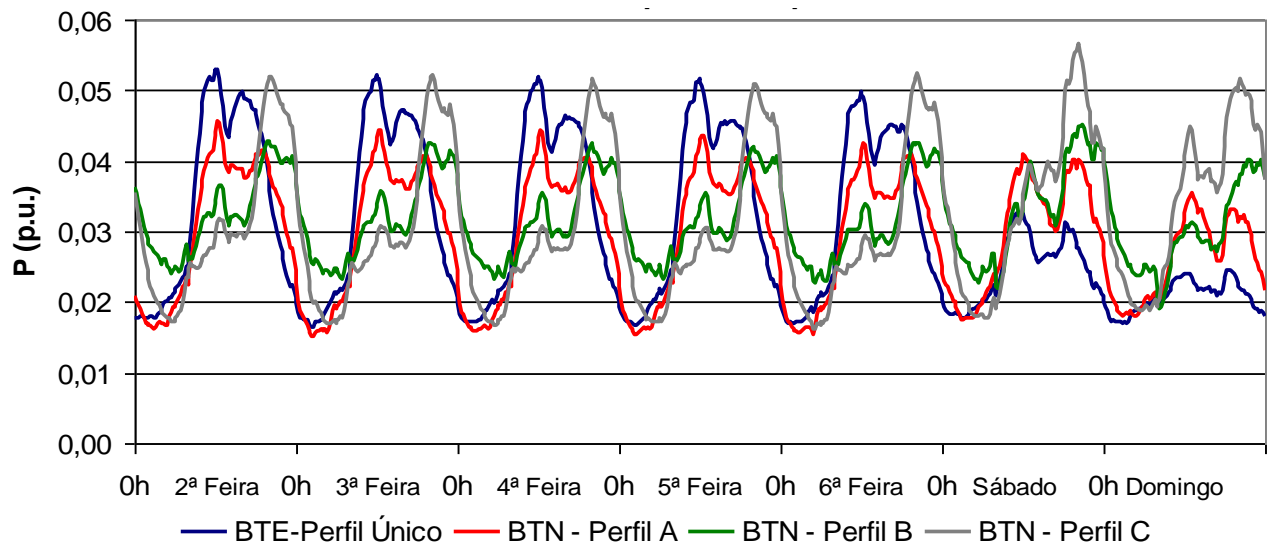
Esta última metodologia apresentada (Fidalgo *et al.*, 2012, Matos *et al.*, 2005) esteve na génese do processo atualmente adotado para viabilizar o adequado funcionamento do mercado liberalizado (*MIBEL*), permitindo também garantir o princípio imposto pela Entidade Reguladora dos Serviços Energéticos (*ERSE*) de não discriminação do tipo de utilização de energia. Por este motivo, as células de classificação que com os devidos agrupamentos geram as classes de consumidores, são diferenciadas com base em dados comerciais como a potência contratada e os consumos anuais de energia (de acordo com a descrição apresentada na Tabela 2.1).

**Tabela 2.1 – Descrição da classificação assumida pela ERSE para consumidores BT**

CLASSE	Descrição
BTE	Aplicável a qualquer consumidor BTE (com tensão entre fases $\leq 1\text{kV}$ e potência contratada $> 41,4\text{kVA}$ )
BTN – Classe A	Aplicável a consumidores do tipo BTN com potência contratada superior a $13,8\text{kVA}$
BTN – Classe B	Aplicável a consumidores do tipo BTN com potência contratada $\leq 13,8\text{kVA}$ e consumo anual $> 7140\text{kWh}$
BTN – Classe C	Aplicável a consumidores do tipo BTN que apresentem uma potência contratada $\leq 13,8\text{kVA}$ e consumo anual $\leq 7140\text{kWh}$

BT – Baixa Tensão reporta a consumidores com uma tensão entre fases inferior ou igual a  $1\text{kV}$

As formas assumidas pelos perfis resultantes, durante uma semana típica de Inverno são apresentadas na Figura 2.7.



**Figura 2.7 – Perfis de consumo aprovados pela ERSE (semana de Inverno)**



Os perfis aprovados e divulgados numa base anual pela ERSE (denominados perfis iniciais com divulgação no início de cada ano civil) fazem-se sempre acompanhar de um diagrama de carga de referência que se baseia no histórico de anos antecedentes do diagrama de carga global. Ao verificar-se a evolução efetiva do diagrama de carga do sistema podem ser atenuados pequenos desacertos suscitados por fatores imponderáveis tais como a temperatura ou luminosidade. Os perfis iniciais são então adaptados segundo a relação que se verifica entre o diagrama de carga do sistema e o diagrama de carga de referência dando origem aos perfis finais. O cálculo dos perfis finais do setor elétrico em Portugal (divulgados mensalmente pela REN <sup>[23]</sup>) processa-se de forma semelhante para todas as classes de consumidores consideradas, pressupondo que o padrão de consumo para cada uma destas classes reflete de forma semelhante as alterações meteorológicas sentidas, o que é na realidade pouco realista. O estudo de Chen *et al.* (2001) propõe uma metodologia de elaboração de perfis de diferentes classes de consumidores (residencial, comercial e industrial), evidenciando análises de regressão para apurar o consumo de energia de cada grupo de consumidores em função de variáveis como a temperatura e a humidade. Com base nas primeiras derivadas destas funções podem ser determinados os denominados índices de sensibilidade às condições meteorológicas, que serão indicadores consideravelmente diferenciados para cada classe. A título exemplificativo, verifica-se no estudo que o consumo de energia elétrica resultante da atividade industrial é, como esperado, o menos sensível a questões climatéricas.

A delimitação de diferentes tipos de consumidores revela-se inquestionavelmente um benefício no planeamento dos serviços de energia, já que as diferentes classes têm diferentes padrões de utilização de energia com impactos diferenciados nos custos de exploração e operação. Usando o exemplo do tradicional mercado regulado, aos consumidores eram imputadas taxas pelo seu uso de energia baseado no custo do seu fornecimento, provando que a classificação permitia apurar diferentes tipos de tarifas a aplicar a cada classe em particular (Casazza e Delea, 2003).

A análise de diferentes tipos de consumidores facilita também a previsão de alterações nas necessidades de consumo por parte dos utilizadores finais. Estas previsões favorecem consideravelmente o planeamento a longo prazo sendo também importantes nas estratégias de operação no curto prazo. Na verdade, até aos finais dos anos 70, a ênfase era dada à garantia de capacidade de geração de energia para existir um equilíbrio físico com a procura nos períodos de potência máxima (ponta do diagrama). Esta estratégia, sendo apenas de acompanhamento,

---

[23] REN - Redes Energéticas Nacionais.

poderia revelar-se extremamente dispendiosa devido aos custos elevados associados ao aumento do número de centrais de produção (ou expansão das existentes), com algumas destas a serem necessárias apenas para assegurar determinados períodos mais críticos do ano. Os programas de gestão de procura foram por isso idealizados como um meio para permitir simultaneamente a redução da ponta, através do desfasamento temporal de utilização de cargas elétricas (reduzir a coincidência temporal de cargas elétricas) e/ou desviando consumo para períodos de menor solicitação, através de uma utilização mais racional de energia. A contribuição dos perfis de consumo pode servir de perspetiva para previsões de potência máxima em períodos futuros, agregando as curvas típicas para cada classe de consumidores que estejam alimentadas num determinado nível de abastecimento (ao nível da distribuição por exemplo). É comum apresentar-se na bibliografia o fator de coincidência temporal que evidencia um decréscimo do eventual efeito associado à potência máxima por consumidor, quando cada vez mais consumidores são adicionados a um grupo. Estas pontas assumidas por cada consumidor são raramente coincidentes, o que origina a que a ponta do grupo se verifique quando a combinação das curvas de carga individuais atingem um máximo, e esta ponta do grupo é por norma substancialmente mais reduzida do que a soma das pontas individuais (Willis, 2004).

$$\text{Fator de coincidência} = \frac{\text{ponta registada pelo grupo de consumidores}}{\text{soma das pontas individuais registadas por consumidor}} \quad (2-1)$$

As redes de distribuição estendem-se por toda a localização geográfica coberta pelos distribuidores, providenciando a ligação última entre o distribuidor e os clientes. Deve ser encarado como o sistema mais adequado para capturar necessidades localizadas dos consumidores, elaborar perfis de carga para caracterização de consumos e estimar padrões de evolução de carga. É facilmente aceite que diferentes áreas geográficas possam evidenciar distintas evoluções da procura e/ou apresentarem preocupações especiais com a fiabilidade de serviço. O processo efetivo de planeamento deve ter início nos sistemas de distribuição, sendo a procura do sistema global apurada num sentido ascendente (*bottom-up approach*), desde a identificação de necessidades de reforço e melhoramento das redes de distribuição, à expansão/criação de subestações de modo a satisfazer os requisitos do sistema de distribuição. A um nível superior deve também acautelar-se o desenvolvimento da rede de transporte para os requisitos das subestações, para que globalmente se convirja para atender os objetivos finais de satisfação das necessidades dos consumidores e assegurar uma otimização técnico-económica.

### *3. Métodos de previsão de consumos*

Este capítulo apresenta um enquadramento acerca dos métodos de previsão que são recorrentemente aplicados para diferentes horizontes temporais em causa, procurando-se sintetizar a evolução sentida nas técnicas de previsão de médio/longo prazo, dando-se um enfoque maior na previsão de curto prazo. No que concerne aos métodos de previsão de curto prazo serão descritas distintas técnicas utilizadas, tais como, os tradicionais métodos regressivos, métodos baseados na análise da própria série temporal bem como os mais recentes métodos baseados em Inteligência Artificial.

Os modelos de usos finais e a modelação econométrica, ou a combinação destes dois tipos de modelos são recorrentemente usados para previsões de carga de médio/longo prazo. As descrições do tipo de equipamento usado pelos consumidores, as áreas das suas habitações, a idade do equipamento, as alterações sentidas na tecnologia, o comportamento dos consumidores (padrões de utilização) e a dinâmica populacional são parâmetros usados nos modelos estatísticos e de simulação baseados na intitulada aproximação por usos finais (Feinberg e Genethliou, 2005). A aproximação estabelecida por meio de usos finais permite estimar diretamente o consumo de energia usando informação diversificada acerca do equipamento de uso final e também acerca do perfil dos consumidores finais. Deste modo, com perspetivas de tipo de equipamento elétrico utilizado, hábitos de consumo, informação relevante sobre a atividade desenvolvida pelos consumidores ou o padrão do agregado familiar, o objetivo será o delineamento de um perfil de consumo. A informação estatística sobre os consumidores a par com as dinâmicas de alteração é encarada como a base para a previsão. Os modelos de usos finais focam os vários tipos de utilização, como são os setores residenciais, comerciais e industriais. Estes modelos são baseados no princípio que a procura de energia deriva da procura dos consumidores para fins específicos como são a iluminação, o arrefecimento, o aquecimento, a ventilação, a refrigeração, ou a força-motriz,... Os modelos explicam a procura de energia como uma função do número de aplicações (equipamentos elétricos) existentes no sistema. Este tipo de abordagem pode apresentar uma grande precisão, muito embora esteja muito dependente e sensível à quantidade e qualidade dos dados disponíveis. Nesta previsão, mais importante que um grande conjunto de dados históricos de consumo é necessário possuir vasta informação acerca dos consumidores e equipamento que possuem.

Na modelação econométrica são incorporados fatores económicos como índices de preços ao consumidor, rendimentos *per capita*, taxas de desemprego ou a inflação verificada. A modelação

econométrica procura combinar a teoria económica e as técnicas estatísticas para prever a procura de eletricidade. É estimada a relação entre consumo de energia (variáveis dependentes) e fatores que influenciam este consumo. Estas relações podem ser estimadas através do método de mínimos quadrados ou por meio de métodos baseados em séries temporais. Uma das opções nesta estrutura é a agregação da abordagem econométrica quando o consumo nos diferentes setores (residencial, comercial, industrial,...) é calculado em função de variáveis meteorológicas, índices económicos ou outras variáveis, e posteriormente as estimativas são geradas recorrendo a dados históricos recentes.

A aplicação de métodos de usos finais e econométricos requerem uma grande quantidade de informação e participação humana especializada no processo. Esta informação pode ser desconhecida para alguns consumidores e são deste modo usados perfis para diferentes classes de consumidores. A previsão de longo prazo deve incluir adicionalmente estudos de migração demográfica e índices de desenvolvimento económico, industrial e tecnológico.

Em previsões de médio/longo prazo, a par da estimativa de quantidade e distribuição temporal de energia consumida, é normal pretender incluir-se o fator localização para favorecer o planeamento de redes de energia elétrica. Deste modo, é recorrente proceder-se a estudos de evolução da procura em determinadas regiões, consoante o seu nível de desenvolvimento económico e social, e prever esta evolução da procura num determinado horizonte temporal. É atualmente esperado que a potência máxima solicitada e a utilização de energia numa determinada área geográfica (podendo esta área ser delineada de acordo com a distribuição das subestações ou de forma a garantir que as diferentes áreas geográficas tenham a mesma dimensão) possa sofrer variações de acordo com o aumento demográfico e/ou novas formas de utilização de energia (como é exemplo o uso em crescimento do Ar Condicionado em Portugal).

A previsão de evolução de consumos para um sistema elétrico global apresenta uma curva de crescimento praticamente linear, quando se excluem fatores atípicos como condições meteorológicas extremas ou efeitos de *booms*/recessões económicas. Numa perspetiva de alta resolução (analisando áreas geográficas de menor dimensão) esta curva de evolução da procura apresenta formas que podem ser consideravelmente diferentes, sendo normalmente assumida uma forma em “S”. Esta forma em “S” evidencia uma evolução brusca de procura numa determinada zona que inicialmente estava praticamente desocupada e que tende a ser explorada de uma forma relativamente rápida (com consumos associados a processos de construção e à atração de novos e diferentes tipos de consumidores) até se atingir uma zona de maturação na curva de crescimento da procura de energia. Em trabalhos dedicados à previsão espacial de carga (Miranda e Monteiro, 2000, Miranda e Monteiro, 1999, Willis e NorthCote-Green, 1983, Wu *et*

*al.*, 1999) é sugerida a aplicação de regras de inferência difusa que evidenciem influências de determinados fatores geográficos e que desta forma possibilitem criar um mapa demonstrativo da potencial evolução de procura de energia e da potência máxima solicitada. Os trabalhos analisados visam assim prever a evolução do consumo de energia com base em fatores estruturais (altitude, inclinação, classificação do local em termos de população e desenvolvimento), distâncias a certas referências geográficas (como proximidades a estradas ou a centros urbanos), efeitos de vizinhança (pretendem representar a influência de entidades ou especificidades na área adjacente ou na própria área geográfica através de curvas de saturação). Estes estudos podem também contemplar análises multi-cenário. Contudo, devem ser encarados numa perspetiva de longo prazo. De registar que a maturidade atualmente verificada dos sistemas de informação geográfica tem permitido um aperfeiçoamento das técnicas de previsão espacial de carga.

No que concerne a previsões de curto prazo, tipicamente previsões para horas (ou partes de hora) seguintes, são usadas categorias distintas de métodos de previsão (Feinberg e Genethliou, 2005):

- abordagem a dias similares <sup>[24]</sup> – baseada na pesquisa de dados históricos (até um horizonte temporal máximo de 2 a 3 anos) com características semelhantes ao dia a prever (Mandal *et al.*, 2006). Estas características verificadas podem ser condições meteorológicas, tipo de dia da semana e/ou uma dada data específica. Para tornar a previsão mais completa, pode ainda basear-se numa combinação linear (processo de regressão) de um conjunto de dias similares.

- métodos de regressão – modelos que pretendem modelar as relações entre o consumo de energia e outros fatores como variáveis meteorológicas, tipo de dia e classe de consumidores (Feinberg e Genethliou, 2005). Os métodos regressivos que são recorrentemente utilizados podem ser de regressão linear simples, regressão linear múltipla, regressão polinomial, exponencial ou logarítmica.

- séries estocásticas temporais – modelos baseados na consideração que os dados têm uma estrutura interna, como uma autocorrelação, tendência ou variação sazonal. Os métodos de previsão com base em séries temporais detetam e exploram este tipo de estrutura, sendo comum distinguirem-se os seguintes métodos (Espinoza *et al.*, 2005, Feinberg e Genethliou, 2005):

- modelo *AR* – *Autoregressive* – o modelo pretende prever uma determinada variável a partir de uma função linear com dados transatos desta mesma variável.

---

[24] Denominados na comunidade científica como *proxy day*.

– modelo *MA* – *Moving Average*– o modelo considera a média da série adicionando uma combinação linear de erros anteriores (cada termo com determinada ponderação).

– modelo *ARMA* – *Autoregressive Moving Average* – para além de registos anteriores da variável a prever, inclui-se no modelo uma combinação linear de erros anteriormente verificados. A estratégia permite reter alterações abruptas manifestadas no passado.

– modelo *ARIMA* <sup>[25]</sup> – *Autoregressive Integrated Moving Average* – pretende aperfeiçoar o modelo *ARMA*, permitindo a correção de eventuais séries não estacionárias <sup>[26]</sup>. Deste modo despista-se algum efeito de tendência, a partir das diferenças encontradas entre registos sucessivos. É usada uma diferenciação na série temporal com base num parâmetro  $d$  (número inteiro positivo) que controla o número de diferenças não sazonais assumidas (com  $d=0$  assume-se que o modelo é equivalente ao modelo *ARMA*). O modelo *ARIMA* deve ser estudado para cada caso, devendo ser evidenciado o efeito de sazonalidade na série temporal. É também comum a adoção de técnicas focadas na redução de irregularidades ou flutuações aleatórias nas séries temporais, como o método de decaimento exponencial de pesos (sendo tipicamente atribuídos maiores pesos aos registos mais recentes).

– modelo *ARMAX* – *Autoregressive moving average with exogenous inputs* – contempla ainda possíveis contribuições de variáveis externas (no caso de previsão de consumos de energia elétrica poderá por exemplo considerar-se a contribuição de variáveis meteorológicas para otimizar o processo).

– métodos baseados em inteligência artificial - procuram aprender com a experiência adquirida e desta forma modelar dados futuros, com recurso a redes neuronais ou a regras de inferências difusas, revelando-se métodos mais apropriados para previsão de dias atípicos (como feriados, ou a ocorrência de eventos casuais) já que lidam com relações não lineares entre a procura de energia e fatores que lhe estejam relacionados (Ferreira e da Silva, 2007, Hippert *et al.*, 2001). Os trabalhos de investigação mais recentes têm apontado para regras de inferência difusas e modelos mistos difuso-neuronais mas é indubitavelmente nas redes neuronais artificiais que se tem explorado mais e reportado resultados bastante satisfatórios (Campbell e Adamson, 2006, Cavallaro, 2005, Chang *et al.*, 2011, Hippert e Taylor, 2010). As primeiras abordagens a este tema foram publicadas no final da década de 80 e desde essa altura que os trabalhos desenvolvidos neste campo têm ganho grande aceitação (Hippert *et al.*, 2001).

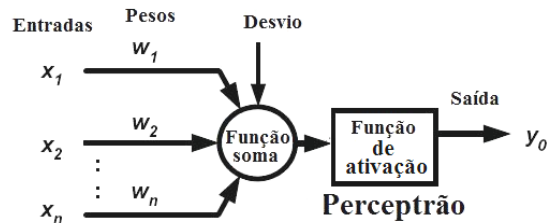
---

[25] Também descrito como modelo de Box-Jenkins por ter sido apresentado por estes dois estatísticos em 1967.

[26] As séries estacionárias caracterizam-se por apresentarem uma média, variância e estrutura de autocorrelação constantes no tempo.

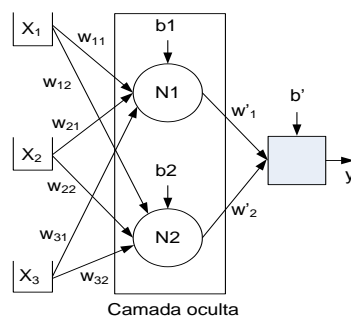
### 3.1. Abordagem às Redes Neurais Artificiais

As redes neurais artificiais são ferramentas matemáticas originalmente inspiradas na forma como o cérebro humano processa informação. A sua unidade básica é considerada o neurónio artificial, que recebe informação numérica através de um conjunto de nós de entrada, processando internamente esta informação e fornecendo uma determinada resposta (saída). Os valores de entrada são linearmente combinados e o resultado é usado como argumento de uma função de ativação. A combinação usa os pesos  $w_i$  atribuídos a cada conexão e um termo constante que pretende funcionar como desvio (*offset* <sup>[27]</sup>) e que pode ser representado por um peso de uma ligação cuja entrada assume um valor fixo unitário (como se ilustra na Figura 3.1).



**Figura 3.1 – Modelo genérico de um neurónio artificial**  
 Figura adaptada de Palit e Popovic, 2005

Na Figura 3.2 é apresentada uma rede neuronal comportando mais do que um simples neurónio na camada intermédia, para exemplificar as ligações (com pesos associados) entre elementos da rede (entre entradas e neurónios e entre neurónios e saídas). A apresentação deste exemplo torna também mais sugestiva a descrição mais detalhada do modelo matemático que caracteriza a arquitetura usada.



**Figura 3.2 – Exemplo de uma arquitetura de rede neuronal com 2 neurónios na camada intermédia**

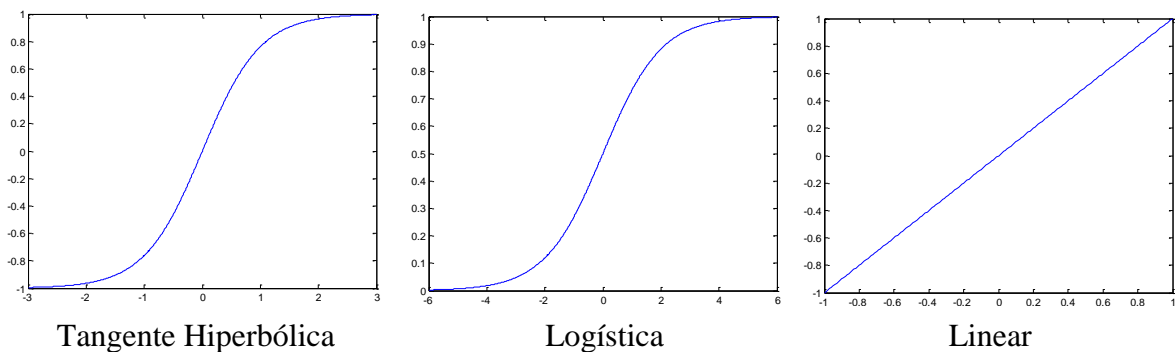
As camadas intermédias entre as camadas de entrada e de saída são designadas por camadas ocultas, uma vez que estas podem estar apenas acessíveis a partir da camada de entrada e/ou da camada de saída mas não são diretamente observadas na estrutura de rede neuronal. Os neurónios em cada camada podem partilhar as mesmas entradas, mas não estão interligados entre

[27] Na comunidade científica internacional também é muito adotado o termo “bias”.

si. Cada ligação tem um determinado peso associado, como é o exemplo da ligação entre a entrada  $x_1$  e o segundo neurónio que é definida pelo peso  $w_{12}$ . Cada neurónio como elemento básico da arquitetura processa a informação capturada, que tanto pode resultar das entradas do processo (no caso de se tratar de um neurónio na 1ª camada oculta tal como descreve a Figura 3.2) ou antes resultar das saídas de neurónios que compõem as camadas intermédias precedentes, e aplica uma função à soma ponderada das entradas adicionando ainda um desvio (Demuth *et al.*, 2007, Hagan *et al.*, 1996). Dando como exemplo o primeiro neurónio, este produziria uma saída calculada matematicamente através da Equação 3-1.

$$\text{saída}(n_1) = f_1(x_1 \cdot w_{11} + x_2 \cdot w_{21} + x_3 \cdot w_{31} + b_1) \quad (3-1)$$

O termo  $f_1$  reporta à função de ativação adotada. A função de ativação deve ser não decrescente e diferenciável (Hippert *et al.*, 2001) e depende do tipo de previsão que se pretenda. As funções de ativação mais recorrentes são: tangente hiperbólica, logística e função linear. Estes tipos de funções são apresentados na Figura 3.3.



**Figura 3.3 – Exemplos de funções de ativação típicas em redes neuronais**

O recurso a funções não lineares é essencial em casos para os quais se verificam relações não lineares entre saída(s) e entradas. A saída  $y$  é obtida através da mesma metodologia mas aplicada à camada de saída, de acordo com a Equação 3-2.

$$y = f_2[\text{saída}(n_1) \cdot w'_1 + \text{saída}(n_2) \cdot w'_2 + b'] \quad (3-2)$$

A função de ativação da saída (neste caso  $f_2$ ) deve também ter as mesmas características (não decrescente e diferenciável) e pode ser distinta das funções de ativação usadas nas camadas anteriores.



### 3.1.1. Processo de treino

A perspetiva das redes neuronais no âmbito da previsão pode ser interpretada como um método guiado pelos próprios dados históricos usados. Apenas com a descrição do vetor de entrada e da(s) saída(s) desejada(s), as redes automaticamente mapeiam as relações encontradas através do registo da aprendizagem efetuada nos parâmetros da rede. A utilização de redes é particularmente útil quando se dispõe de uma grande quantidade de dados e para os quais se tem um parco conhecimento prévio sobre a forma como as entradas e as alterações nas entradas se refletem nos dados de saída.

Os dados históricos usados assumem um papel muito importante, já que a rede é adaptada tendo em conta estes dados que lhe são apresentados. Este processo é designado como processo de treino, dando à rede a possibilidade de ajuste dos seus parâmetros (pesos e desvios) de forma a minimizar os erros entre os valores reais e as estimativas realizadas. Diferentes tipos de métodos de treino podem ser encarados, baseados em processos iterativos de otimização. O objetivo é minimizar a função que avalia o desempenho da rede (função esta que se baseia tipicamente no erro quadrático médio).

Um dos métodos de treino inicialmente proposto em 1960 por *Widrow and Hoff* (Palit e Popovic, 2005) para ajuste dos pesos associados a um perceptrão <sup>[28]</sup> baseia-se na regra de delta que adiciona o valor de peso corrente  $w(k)$  a um termo de compensação que gera o valor do peso seguinte, de acordo com a Equação 3-3.

$$w(k+1) = w(k) + \eta \cdot \varepsilon(k) \cdot x(k) \quad (3-3)$$

em que:  $\eta$  corresponde a um termo de proporcionalidade;

$\varepsilon(k)$  corresponde a um erro no instante  $k$ ;

$x(k)$  corresponde a um valor do sinal de entrada no instante corrente  $k$ .

Apesar da sua simplicidade, a regra de aprendizagem delta tem na maioria dos casos demonstrado um excelente desempenho e simultaneamente uma elevada rapidez na convergência durante o treino do perceptrão. Contudo, um único perceptrão está limitado para resolver problemas mais complexos devido ao raio de ação computacional estar muito restringida à simplicidade da sua estrutura.

Para as redes multicamada foi reformulado por *Rumelhart and McClelland* (Palit e Popovic, 2005) um algoritmo de aprendizagem que é no plano atual consideravelmente disseminado,

---

[28] Perceptrão corresponde ao tipo de rede neuronal mais simplificado que se encontra ilustrado na Figura 3.1.

denominado por algoritmo de treino por retro-propagação <sup>[29]</sup>. Este algoritmo calcula o gradiente do erro <sup>[30]</sup> ( $E$ ) em relação aos pesos e desvios que são ajustáveis, o que justifica a escolha de funções de ativação diferenciáveis. Os pesos e desvios são deste modo atualizados na direção para a qual a função objetivo decresce mais rapidamente, usando um processo iterativo de gradiente descendente (Hagan *et al.*, 1996), de acordo com as expressões 3-4 e 3-5.

$$\Delta w_{ij} = -\varepsilon \cdot \frac{\partial E}{\partial w_{ij}} \quad (3-4) \quad ; \quad \Delta b_j = -\varepsilon \cdot \frac{\partial E}{\partial b_j} \quad (3-5)$$

com  $\varepsilon$  a expressar a taxa de aprendizagem que multiplica pelo gradiente do erro para determinar as alterações nos pesos e desvios, sendo um termo que controla a velocidade de treino.

Uma taxa de aprendizagem mais elevada pode tornar o algoritmo célere, embora podendo comprometer a desejada convergência. Por outro lado, valores reduzidos para a taxa de aprendizagem tornam o processo de treino muito moroso em termos computacionais.

Algumas variantes deste método foram entretanto adotadas, como o uso de taxas de aprendizagem distintas e ajustáveis para cada peso/desvio. Esta alternativa, também denominada como algoritmo de retro-propagação adaptativo permite que se aumente a taxa de aprendizagem quando a superfície de erro se torna muito plana, e diminuir esta taxa quando o declive se torna considerável, com o objetivo primordial de se acelerar a convergência (Hagan *et al.*, 1996). Pode assumir-se que cada taxa de aprendizagem é incrementada ou decrementada dependendo da superfície de erro nesse particular peso ser ou não monótona. Deste modo a taxa de aprendizagem pode aumentar consideravelmente ao longo das iterações para alguns pesos e decrescer até se anular para outros pesos. O tempo requerido por esta variante é geralmente inferior ao tradicional algoritmo de retro-propagação (Fidalgo e Matos, 2007, Hagan *et al.*, 1996).

Um outro algoritmo de treino é intitulado de gradiente descendente com momento, baseando-se não somente na resposta da rede neuronal ao gradiente local, mas também nas tendências recentes na superfície do erro, por forma a amenizar as oscilações sentidas na trajetória assumida. Atuando como um filtro passa-baixo, o momento permite à rede ignorar pequenas variações na superfície de erro, permitindo à rede evitar prender-se a mínimos locais.

[29] Denominado na comunidade científica como “Backpropagation training algorithm”.

[30] Sendo considerado o erro quadrático médio.

Um método de treino mais recente é o algoritmo de Levenberg-Marquardt que é uma aproximação ao método de Newton. Para o método de Newton é obtida uma matriz Jacobiana que contém as primeiras derivadas da função de desempenho em relação a todos os parâmetros ajustáveis da rede. A atualização dos pesos e desvios respeita a Equação 3-6.

$$\Delta \mathbf{x} = \left[ \mathbf{J}^T \cdot \mathbf{J} \right]^{-1} \cdot \mathbf{J}^T \mathbf{e}(\mathbf{x}) \quad (3-6)$$

com:

$\Delta \mathbf{x}$  – vetor com atualizações para os pesos e desvios da rede neuronal;

$\mathbf{J}$  – matriz Jacobiana;

$\mathbf{e}(\mathbf{x})$  – vetor com os erros identificados.

O algoritmo de Levenberg-Marquardt incorpora um parâmetro adaptativo (expresso na Equação 3-7) de modo a manter como objetivo a busca de boas soluções de uma forma expedita mas também preservando a convergência (Hagan *et al.*, 1996, Hagan e Menhaj, 1994).

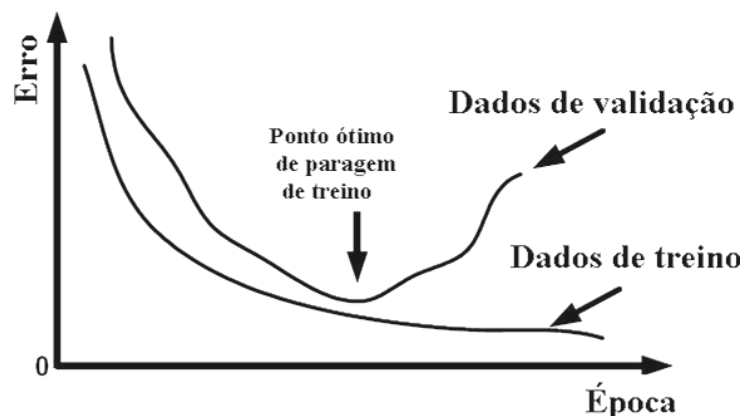
$$\Delta \mathbf{x} = \left[ \mathbf{J}^T \cdot \mathbf{J} + \mu \cdot \mathbf{I} \right]^{-1} \cdot \mathbf{J}^T \mathbf{e}(\mathbf{x}) \quad (3-7)$$

com  $\mathbf{I}$  a representar a matriz identidade.

O parâmetro adaptativo  $\mu$  é multiplicado por um fator  $\beta$  sempre que uma iteração resultar num aumento da função de desempenho baseada no erro de previsão. Por outro lado, quando uma iteração resultar na redução da função de desempenho, o parâmetro adaptativo  $\mu$  é dividido por  $\beta$ . Com um elevado parâmetro adaptativo, o algoritmo torna-se similar a um algoritmo de gradiente descendente com pequenas alterações sentidas na superfície de erro. Considerando um reduzido valor para o parâmetro adaptativo, o algoritmo tende a aproximar-se do método de Newton configurando um método mais rápido de treino. O algoritmo de Levenberg-Marquardt pretende ser um compromisso entre a celeridade garantida pelo método de Newton e a garantia de convergência providenciada pelo método de gradiente descendente. Para este algoritmo, as especificações para o período de treino são relativas ao número máximo de épocas (iterações) de treino, um valor objetivo para a função de desempenho (até se atingir um valor de erro de previsão igual ou inferior a um valor pré-definido), um valor mínimo de gradiente do erro, o tempo máximo de treino, um valor inicial para o parâmetro adaptativo, os fatores de aumento ou decréscimo do parâmetro adaptativo, bem como um valor máximo para este parâmetro (Demuth *et al.*, 2007).

Como pode ser verificado pelas especificações que são requeridas *a priori* nos diferentes algoritmos de treino, diferentes estratégias podem ser adotadas para interromper o algoritmo de

treino. Todavia, um processo que tem sido amplamente reconhecido pela comunidade científica é o recurso a validação cruzada. Durante o processo de treino é expectável que o modelo baseado em redes neuronais se adapte convenientemente aos dados para os quais a rede é efetivamente treinada (dados de treino), satisfazendo deste modo os propósitos dos algoritmos de treino subjacentes que visam a minimização do erro. No entanto, pode não ser conseguido um desempenho desejado da rede quando aplicada a dados de teste (dados que não façam parte da amostra de treino), comprometendo a capacidade de generalização da rede que é modelada. As duas causas mais comuns para esta limitação é o excessivo período de treino e/ou a complexidade do modelo criado. De forma a limitar o excesso de treino é assim recomendável proceder-se à técnica de validação cruzada, separando os dados da amostra original em dados de treino e em dados de validação. Os parâmetros são assim estimados com recurso aos dados de treino e o desempenho é avaliado para cada iteração no conjunto de dados de validação. No instante em que o desempenho se começa a deteriorar (de acordo com o que se pretende ilustrar na Figura 3.4) subentende-se que a rede se encontra com um treino exacerbado que deve ser interrompido, admitindo-se como parâmetros da rede aqueles que conduziram a um menor erro de previsão no conjunto de dados de validação. Outro parâmetro que é tipicamente ajustável pelo utilizador no processo de treino é o número máximo de iterações para o qual se permite a degradação do indicador de erro para dados de validação. Assim que este limite é atingido, o treino é imediatamente interrompido, tomando-se como parâmetros da rede aqueles que conduziram ao menor erro de validação.



**Figura 3.4 – Exemplo de validação cruzada com curvas de evolução de erro para dados de treino e de validação em função do número de iterações consideradas no processo de treino – Figura adaptada de Palit e Popovic, 2005 –**

Uma outra abordagem para evitar demasiado treino poderá ser a aplicação de técnicas de regularização (Hippert *et al.*, 2001) que incluem na função objetivo a minimizar (tipicamente dependente do erro de precisão encontrado) um termo que penalize a complexidade do modelo. Este termo pode por exemplo penalizar a elevada curvatura do modelo considerando a segunda

derivada da saída em função das entradas. Os modelos mais complexos apresentam normalmente uma curvatura pronunciada já que procuram contemplar todo e qualquer ponto relativo a cada um dos dados de treino. As segundas derivadas são nestes casos elevadas e o termo de regularização aumenta, condicionando o resultado da função objetivo.

O excesso de treino pode também estar diretamente dependente da sobre-parametrização da rede, devido ao elevado número de entradas e de camadas intermédias que se incluem na arquitetura da rede e para as quais se torna necessário estimar um conjunto vasto de parâmetros.

A avaliação do desempenho da rede deve também contemplar um subconjunto de dados que não tenham sido usados nem para treino nem para validação, também denominados de dados de teste e para os quais se procede a uma comparação dos resultados reais de consumo com os resultados obtidos a partir do modelo de previsão com redes neuronais. Por uma questão de rigor na apresentação dos erros para possível comparação e discussão entre a comunidade científica, devem ser inequivocamente indicados os tipos de erro analisados, se estes se reportam aos dados de treino ou aos dados de teste, sendo que a implementação é adequada quando os níveis de erros de cada uma destas espécies de amostras são comparáveis <sup>[31]</sup>. É também aconselhável usarem-se diferentes amostras de teste impondo alguma rotatividade para se poder exercer uma avaliação coerente dos erros verificados. Os erros podem ser apresentados em diferentes formatos tais como:

- o erro absoluto médio (em valor ou em percentagem) – apropriado nos casos em que se verifica que os valores de módulo dos erros de previsão têm registos próximos. Quando em determinados casos se encontram elevados erros individuais, estes podem acarretar repercussões desastrosas para um operador de sistema ou outro agente económico que dependa de uma previsão;

- o erro quadrático (erro quadrático médio ou erro percentual quadrático médio), adotado com o intuito de penalizar elevados erros individuais.

- a perspetiva da distribuição do erro pode também ser um útil indicador, podendo apresentar-se uma distribuição cumulativa, por percentis, ou através de histogramas.

Por um lado, pode averiguar-se se a rede terá capacidade suficiente (com um suficiente número de neurónios na camada intermédia ou mesmo se terá um suficiente número de camadas intermédias na estrutura) para modelar os dados apresentados, quando os erros associados a dados de treino não se manifestam elevados. O confronto entre os erros verificados nas amostras

---

[31] Erros reduzidos nos dados de treino não implicam diretamente erros reduzidos na amostra de teste.

de dados de treino e no subconjunto de teste permitirá averiguar se a rede tem boa capacidade de generalização. Esta comparação entre erros de treino e erros de teste revela-se efetivamente de extrema utilidade prática para apoiar a definição de uma arquitetura adequada para a rede.

### 3.1.2. Arquitetura da rede

A arquitetura de uma rede neuronal pressupõe uma adequada decisão acerca do número de nós de entrada, número de camadas ocultas, número de neurónios por camada, número de nós de saída e tipos de função de ativação entre camadas.

A decisão do número de nós de entrada depende exclusivamente do número de entradas que se pretendem incorporar. Na Secção 3.1.4 é apresentado com maior detalhe o processo de escolha do número e tipo de entradas que são passíveis de se utilizar em previsão de consumos de energia.

Já a seleção do número de camadas ocultas deve ter em vista a inclusão de tantas camadas quanto as necessárias para criar alguma flexibilidade na modelação dos dados usados, evitando ainda que sejam em número demasiado elevado que possa suscitar um dispendioso período de treino e tornar o modelo de complexa interpretação. Uma opção válida e apresentada por diversos autores (Hippert *et al.*, 2001, Sarle, 2000) é a comparação de diferentes alternativas para se apurar o melhor desempenho. A capacidade de aprendizagem da rede é essencialmente devida ao mapeamento interno das características essenciais na camada intermédia. A aprendizagem acumulada durante o processo de treino da rede pode ser automaticamente aproveitada durante a fase da aplicação em processamentos seguintes. Apesar de parecer que as capacidades para capturar características da rede aumentam com eventual inclusão de mais camadas ocultas, o tempo de treino passa a estar sacrificado e um compromisso deverá ser assim encontrado, dependendo também da complexidade do problema a modelar.

O número de nós de saída depende do tipo de previsão a estabelecer-se, podendo-se distinguir diferentes categorias. Numa primeira categoria pode ser interpretada uma única saída, tipicamente para previsão de consumo para o período de integração seguinte (hora seguinte ou fração de hora seguinte), ou para previsão da potência máxima ou mesmo do consumo total de energia no dia seguinte. Esta previsão pode ser encarada como iterativa quando é prevista uma variável relacionada com o consumo num determinado período e agregada à série, possibilitando que as previsões futuras sejam baseadas em previsões anteriores. Na categoria de previsão do tipo multimodelo usam-se diferentes redes em paralelo dedicadas a distintas horas do dia garantindo que as redes individuais são relativamente pequenas, não existindo um risco de sobre-ajustamento (demasiados parâmetros da rede a configurar face ao número de diferentes

amostras existentes na entrada). A última categoria considera um modelo único multivariável, já que são considerados múltiplos nós de saída de modo a prever um conjunto de consumos horários, enquadrando-se nos exemplos de previsão de consumos para o dia seguinte (perfil de consumo para o dia seguinte). Uma solução híbrida foi também já proposta (Alfuhaid *et al.*, 1997), recorrendo a pequenas redes que processam previamente alguns dados e estimam potências máximas, potências mínimas e consumo total, sendo estes dados fornecidos em conjunto com outros dados adicionais que se considerem relevantes a redes consideravelmente maiores que estimam o perfil de consumo do dia seguinte.

### 3.1.3. Pré-processamento dos dados

O pré-processamento permite reduzir a dimensão do vetor de entrada que tende a afetar exponencialmente a complexidade do modelo e eliminar ou corrigir dados atípicos ou registos em falta, aos quais as redes são muito sensíveis. O pré-processamento pode implicar uma partição do espaço de entrada, através por exemplo da classificação prévia dos dados de entrada (perfis de carga anteriores ou dados meteorológicos disponíveis) e usar posteriormente redes separadas para modelar dados das diferentes classes. Uma classificação comum à maioria das abordagens diz respeito à separação dos dados em dias de semana e fins de semana. Um fator importante na previsão diz respeito às condições meteorológicas que podem ser consideradas como entradas, podendo ser categorizados os dias através, por exemplo, de medidas estatísticas de similaridade entre variáveis relacionadas com condições meteorológicas (temperatura, humidade, radiação solar, ...).

O processo de normalização deve também ser adotado, permitindo que haja maior facilidade em lidar com diferentes escalas possíveis em variáveis distintas, garantindo assim que todas as variáveis se centrem numa gama similar de valores, para além de permitir atenuar o efeito de possíveis desvios sentidos em determinadas variáveis. A normalização pode estabelecer-se de modo a que as entradas e saídas tenham uma média nula e uma variância unitária ou, em alternativa, recorrer aos valores máximos e mínimos das variáveis de modo a que os registos das diferentes variáveis estejam incluídos numa gama  $[-1, 1]$ . Devido às funções de ativação que são recorrentemente usadas na(s) camada(s) intermédia(s) e na camada de saída assumirem valores entre  $-1$  e  $1$ , os registos de entrada devem ser normalizados previamente para respeitar este mesmo intervalo. Algumas variantes podem ainda ser assumidas, procedendo-se a uma normalização do conjunto de dados de aprendizagem numa escala entre  $0,15$  e  $0,85$ , por forma a evitar a saturação de funções sigmoide nas vizinhanças de  $0$  e de  $1$  e também de modo a que

permita valores previstos superiores aos valores máximos históricos que formam o conjunto de dados de treino (Gavrilas, 2002).

#### 3.1.4. Escolha de entradas a usar no processo de previsão

Um processo de alguma complexidade é a seleção do número e do tipo de entradas a usar (Drezga e Rahman, 1998, Hippert e Taylor, 2010). Esta temática é conhecida como o processo de seleção de parâmetros explicativos <sup>[32]</sup> e, de entre as variáveis a selecionar a que manifestamente deve ser usada diz respeito aos próprios dados de consumos em intervalos temporais anteriores (ou de potência ativa em instantes anteriores), já que a série temporal de consumo como variável endógena é naturalmente autocorrelacionável com a(s) variável(is) de saída. Na previsão de curto prazo torna-se necessário especificar os valores de consumo em atraso (dados históricos) que devem servir para o vetor de entrada, tendo como certo que deve ser obtido um franco compromisso entre a precisão do modelo e a complexidade do mesmo, já que um elevado número de dados de entrada pode acarretar a sobre-parametrização da rede sem necessariamente se minorarem os erros associados à previsão de novos dados que sejam simulados e que não constem na amostra de treino de rede. A análise de autocorrelação (ou a análise de funções de autocorrelação parcial) pode servir para identificar registos apropriados para integrar o processo de previsão, contudo, existe o risco de descartar registos (ou variáveis) que não apresentem uma correlação linear significativa com a carga mas que todavia podem ser francamente correlacionáveis de forma não linear com a mesma. Alguns autores (Santos *et al.*, 2007) recorrem a técnicas como a análise de entropia da série temporal (forma de aferir o grau de aleatoriedade de um sinal), complementando com uma análise de tendência, de forma a reduzir a discricionariedade associada ao estabelecimento de variáveis na composição do vetor de entrada, em modelos baseados em redes neuronais.

Outras técnicas são também exploradas, enquadrando-se no conceito de condensação de parâmetros explicativos <sup>[33]</sup>, como a análise de componentes principais (Ashfin *et al.*, 2007, Guo *et al.*, 2004). Esta estratégia visa identificar as entradas com maior influência no contexto do modelo de previsão, transformando o conjunto de variáveis de entrada original. A análise permite a redução do número inicial de variáveis correlacionadas a um reduzido número de variáveis que representará o conjunto de fatores principais que contém a informação considerada essencial, resultante do conjunto de variáveis iniciais e sem perda de informação relevante. Este processo tem como vantagem a redução do esforço computacional necessário para posterior

---

[32] Denominado na comunidade científica de *Feature Selection*.

[33] Denominado na comunidade científica de *Feature Extraction*.



processamento de dados (Palit e Popovic, 2005). A técnica transforma linearmente as entradas originais em novos parâmetros não correlacionados entre si e que são concretamente a transformação ortogonal das entradas originais baseadas na análise dos vetores próprios. No final, são selecionadas as componentes principais correspondendo aos valores próprios obtidos de forma decrescente.

O recurso à transformada *Wavelet* <sup>[34]</sup> (Guan *et al.*, 2009, Zhao *et al.*, 2009) é também apontado como um meio para corrigir a série temporal original, tipicamente caracterizada por ter algum ruído, numa série que apresente melhor comportamento. O processo passa por decompor o vetor com os dados de entrada a partir da escolha de uma função base (família de funções *Wavelet*) e através da escolha do número de níveis para os quais se deseja que a série seja decomposta. Em cada nível de decomposição é considerado um sinal de aproximação (correspondendo a baixas frequências) e um sinal de detalhe (visto como a diferença entre duas aproximações sucessivas, podendo ser encarado como resultante de um filtro passa-alto) (Guan *et al.*, 2009). A remoção do ruído é conseguida através da eliminação dos coeficientes menores que um determinado valor limiar na série detalhada. Esta série corrigida é a que efetivamente é utilizada no método de previsão, sendo necessário numa fase posterior reconstruir o sinal a partir dos coeficientes de decomposição usados (Guan *et al.*, 2009).

A seleção de variáveis exógenas como variáveis relacionadas com condições meteorológicas é também uma prática comum, quando se demonstra que a procura de energia é influenciada por esta natureza de variáveis (Chen *et al.*, 2001, Falvo *et al.*, 2006, Gerse, 2007, Hippert e Pedreira, 2004, Xia *et al.*, 2010). Atualmente a inclusão deste tipo de variáveis toma ainda um particular interesse devido à crescente penetração dos sistemas de ar condicionado que proporcionam um substancial aumento do consumo de energia predominantemente nas épocas quentes, enquanto nas épocas frias é também expectável o aumento do consumo de energia, mas nem sempre se pode atribuir diretamente a um aumento de energia elétrica (dependendo da fonte de energia primária usada, sendo o gás natural a forma alternativa mais utilizada em Portugal).

No que diz respeito ao impacto das condições meteorológicas em previsões de médio prazo da potência máxima e/ou consumo de energia, podem ser assumidos diferentes critérios (Willis, 2004):

---

[34] Transformada Wavelet é encarada como uma extensão da Transformada de Fourier (Palit e Popovic, 2005)

– abordagem sem critérios – não é contemplada a influência das condições meteorológicas no estudo de previsão de consumos – normalmente adotada por sistemas com elevadas margens de capacidade na geração;

– abordagem arbitrária – modelo tradicional baseado na análise de padrões meteorológicos e seus efeitos na procura;

– abordagem reacionária – baseada num cenário extremo de condições meteorológicas ocorrido no passado, tendo como prioridade a garantia de abastecimento.

– abordagem baseada no risco – recorrendo a um método probabilístico que analisa como as probabilidades de falha do equipamento podem afetar a fiabilidade – baseia-se num planeamento tendo em vista o cumprimento de determinados objetivos de fiabilidade.

A função que relaciona a temperatura com a procura de energia é claramente não linear, o que se torna uma motivação adicional para o uso de redes neuronais, já que estas redes têm uma capacidade intrínseca de modelar relações não lineares entre variáveis de saída e variáveis de entrada. A função que modela o consumo de energia (ou registos de potência) com a temperatura verificada apresenta uma curva característica em forma de “U”, enunciando uma elevada dependência de energia elétrica para satisfazer necessidades de aquecimento em épocas frias e húmidas, ou para satisfazer necessidades de arrefecimento em épocas quentes, contrastando com épocas de meia estação em que o consumo tende a reduzir substancialmente (como pretende demonstrar a Figura 3.5.

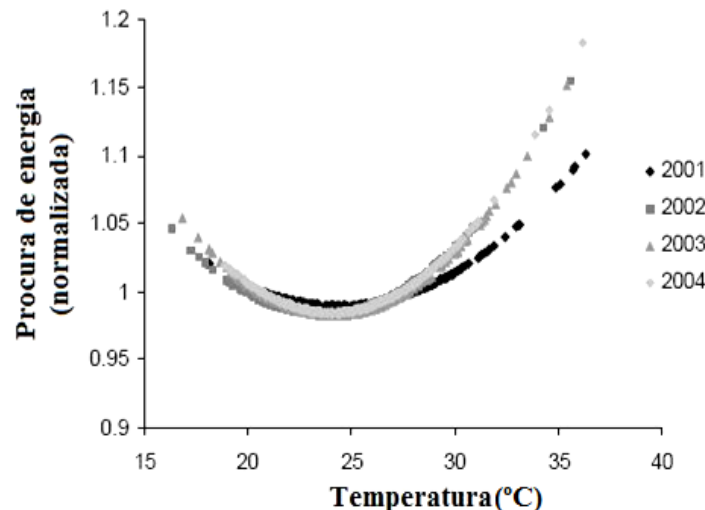


Figura 3.5 – Evolução da procura de energia com a temperatura média registada  
 Figura adaptada de Ziser *et al.*, 2005

Uma opção pode passar por uma análise separada de diferentes épocas do ano, particionando desta forma a curva consumo médio/temperatura ou potência ativa/temperatura em várias gamas de temperatura por forma a aperfeiçoar funções lineares/quadráticas em diferentes intervalos. Devem ser tidos em consideração os denominados efeitos de inércia térmica, já que alterações

bruscas que se sintam nas condições meteorológicas não se refletem instantaneamente nos consumos de energia. A sensibilidade associada ao conforto térmico dos consumidores difere temporalmente destas variações de temperatura com determinado atraso (que pode variar entre horas a dias) e o próprio isolamento térmico dos edifícios tende a amenizar o efeito destas ocorrências. Neste sentido, e tipicamente quando se relaciona a potência máxima com a temperatura média diária, podem usar-se índices climáticos que correspondem a uma média ponderada (Santos, 2005) dos registos de temperatura média verificada nos dias anteriores.

Para além da temperatura também podem ser usadas outras variáveis como a humidade relativa e a velocidade do vento já que podem ter um considerável efeito na sensação humana de conforto, justificando o uso de sistemas elétricos para aquecimento ou arrefecimento. A radiação solar pode também ser avaliada como válida no processo de previsão de consumo, podendo ser estipulado um determinado valor limiar que diferencie o período do dia solar do período da noite, para desta forma se determinar o número de horas diárias de radiação solar. É expectável que este número de horas diárias de radiação solar possa ser um indicador importante a ter em conta na estimação de consumos de energia elétrica, uma vez que os hábitos de consumo podem estar diretamente relacionados com esta variável, no que diz respeito à utilização de energia para climatização e iluminação. Determinadas simulações operadas na previsão de consumo adotando redes neuronais usam dados de previsão de condições meteorológicas futuras nos parâmetros de entrada. Outra hipótese pode passar pelo recurso aos valores efetivamente verificados nos períodos usados de previsão em vez dos valores de meteorologia previstos, de forma a evitar que os erros decorrentes da previsão das condições meteorológicas se transmitam e condicionem a previsão de consumo. Uma abordagem alternativa poderá ser o uso de dados reais de meteorologia, impondo ainda assim posteriormente um ruído aleatório definido previamente numa gama limitada por margens de erro (Fidalgo e Matos, 2007, Hippert e Pedreira, 2004), de forma a serem simuladas imprecisões na previsão dos dados climáticos. Esta última metodologia para além de permitir contornar possível falta de informação, permite a distinção clara entre os erros provenientes do modelo de previsão adotado e os erros originados pela incerteza nos dados de entrada (Hippert e Pedreira, 2004). Nas previsões de médio e de longo prazo já não é dado tanto relevo ao efeito das variáveis meteorológicas no consumo, uma vez que os efeitos das variáveis meteorológicas no consumo tendem a dissipar-se ao longo do tempo, acrescentando o facto da previsão meteorológica para horizontes temporais mais dilatados se tornar indisponível ou menos fiável (Gavrilas *et al.*, 2001). No contexto de previsão de muito curto prazo (da ordem

dos minutos, previsão tipicamente idealizada para garantia da estabilidade da frequência) a escala temporal é por sua vez encarada como demasiado curta para os efeitos climáticos se poderem fazer sentir. De salientar que, por norma, os serviços meteorológicos disponibilizam dados médios ou cenários extremos sem indicação de horas de ocorrência, sendo eventualmente definidos apenas dois a três períodos do dia mas sem serem disponibilizados perfis meteorológicos para cada dia.

Como exemplo de outros tipos de variáveis exógenas pode referir-se o recurso a registos de potência reativa como forma de distinguir diferentes períodos do dia (essencialmente a potência máxima verificada no período da manhã da potência máxima verificada no período da tarde) e também revela informação sobre consumo nos fins de semana já que o decréscimo da atividade industrial origina a redução de consumo de energia reativa (Santos, 2005). De realçar que o estudo de previsão elaborado ao nível de subestações de distribuição pode ficar extremamente dependente da compensação de fator de potência a um nível local (por consumidor) ou a um nível global por subestação o que, a rigor, tornaria esta entrada pouco determinante sempre que fossem adotadas com acuidade estratégias de compensação do fator de potência.

Outras variáveis podem também ser incluídas como variáveis auxiliares com indicação de periodicidades, para discriminar o tipo de dia da semana e também as horas do dia. O uso de funções trigonométricas é assim idealizado para a definição de variáveis cíclicas, usando entradas baseadas nas funções *seno* e *coseno*. O tipo de dia ou a hora a prever é interpretada de uma forma periódica e o uso das duas funções trigonométricas justificam-se por assim permitirem especificar cada tipo de dia ou hora de forma unívoca (Drezga e Rahman, 1999, Fidalgo e Peças Lopes, 2003, Ramezani *et al.*, 2005). Apresentam-se alguns exemplos:

$$\text{sen}\left(\frac{2 \cdot \pi \cdot d}{7}\right); \cos\left(\frac{2 \cdot \pi \cdot d}{7}\right) \Rightarrow \text{funções trigonométricas para distinguir o tipo de dia da semana}$$

(com  $d=1, \dots, 7$ )

$$\text{sen}\left(\frac{2 \cdot \pi \cdot h}{24}\right); \cos\left(\frac{2 \cdot \pi \cdot h}{24}\right) \Rightarrow \text{funções trigonométricas para distinguir a hora do dia}$$

com ( $h=1, \dots, 24$ )

Outras abordagens podem ser seguidas por forma a diferenciar tipos de dia e meses do ano recorrendo a variáveis binárias (Lamedica *et al.*, 1996). Os dias especiais como feriados são também recorrentemente discriminados a partir de variáveis binárias. Diversos trabalhos focam esta problemática associada à previsão reportada a dias especiais, ou em torno de dias especiais (Bakirtzis *et al.*, 1996, Chicco *et al.*, 2001, Fidalgo e Matos, 2007, Fidalgo e Peças Lopes, 2005,

Lamedica *et al.*, 1996, Srinivasan *et al.*, 1999,). O efeito dos feriados pode ser visto como um fator de perturbação em processos de previsão já que os valores conjecturados para estes dias tendem a ser superiores aos valores reais registados, uma vez que o consumo decresce de forma considerável devido essencialmente à redução da atividade económica. Por outro lado, se as previsões se basearem em registos de dias anteriores, sempre que ocorre um feriado este efeito irá repercutir-se para as previsões dos dias posteriores que, apesar de não serem feriados, são previstos com base em dados históricos de dias especiais e por conseguinte apresentam previsões abaixo dos valores reais efetivos. O efeito das pontes deve também ser tido em conta, com as considerações que se revelem úteis para este tipo de dias de carácter especial. Chicco *et al.* (2001) apresenta uma fase inicial do estudo em que considera métodos de *clustering* de forma a classificar diferentes tipos de dia, assumindo-se quatro grupos que coincidentemente correspondem de forma quase inequívoca a grupos de dias de semana, sábados, domingos e outros casos especiais como dias feriados. O propósito desta fundamentada classificação é determinar classes de padrões similares de consumo nos dados históricos, que através de uma transposição adequada de calendário (no caso em questão transpondo para o ano seguinte), permita a seleção de um adequado vetor de treino para o período de previsão em análise.

Bakirtzis *et al.* (1996) resumem dois métodos distintos para previsão relativa a dias feriados. Uma primeira abordagem pressupõe o treino de uma rede neuronal relativa apenas a dias especiais (e eventualmente englobando ainda o efeito dos fins de semana), separando estes dados de treino de dados relativos a dias normais de semana. De referir que é expectável que os erros resultantes da previsão de dias atípicos sejam superiores face aos erros verificados nos dias normais, já que a amostra de dias atípicos e fins de semana será forçosamente inferior, qualquer que seja o horizonte temporal disponível em histórico. No segundo método a procura de energia relativa a dias especiais é encarada como tendo duas componentes distintas: uma componente básica de dia normal e uma componente de ajustamento do efeito de feriado. Esta componente de ajustamento é determinada usando dados históricos dos anos anteriores que relatam a média do desvio da componente básica face aos valores de carga medidos em cada dia especial.

### 3.1.5. Considerações adicionais acerca das redes neuronais

O ceticismo subjacente ao uso de redes neuronais na previsão de consumos fundamenta-se especialmente em argumentos relacionados com a sua sobre-parametrização, já que podem estruturar-se redes demasiado complexas com demasiados parâmetros a definir em função da pequena amostra de treino que se pretende modelar. Outra limitação que é comum reportar

relaciona-se com a separação dos dados em período de treino/validação e período de teste. Esta segmentação pode revelar-se em alguns estudos pouco sistematizada e nem sempre apresentada da forma mais transparente e credível como é desejável.

Hippert *et al.* (2001) apresentam um levantamento exaustivo de vários trabalhos na área com a descrição de pontos fortes e fracos das redes neuronais, resumindo e identificando diversas incertezas em relação à modelação por redes neuronais para previsão de consumos. Um trabalho mais recente desta equipa de investigação (Hippert e Pedreira, 2004) descreve uma metodologia com vista a prever perfis de temperatura para posteriormente se integrar esta informação nos perfis de consumo. Nesse âmbito, e tendo em conta que determinadas previsões de consumo de curto prazo recorrem a variáveis meteorológicas, toda a precisão obtida nos perfis meteorológicos pode ser fundamental para beneficiar a precisão dos perfis de consumo. O estudo apresenta um modelo de redes neuronais que representa os perfis como um vetor de dimensão dependente do número de horas do dia, com as previsões a serem feitas num único passo para todas estas horas (modelo único multivariável). Esta estratégia é justificada pelo facto de se combinarem dados que resultam de previsões (normalmente associados à temperatura mínima e à temperatura máxima) com dados efetivamente observados. As redes neuronais podem lidar com os perfis num formato completo em apenas um único passo nas suas entradas e saídas, em vez de modelar valores isolados e, deste modo, não tendem a perder informação de possíveis autocorrelações da série e da forma dos perfis diários. Por outro lado é reconhecido que permitem introduzir mais facilmente o efeito de variáveis exógenas, tendo em consideração que, em comparação com modelos regressivos, existirá uma menor sensibilidade aos efeitos de colinearidade<sup>[35]</sup> que caracterizam estas séries de dados.

O resultado das comparações efetuadas permitiram concluir que as redes demonstram ser mais robustas do que os tradicionais modelos lineares relativamente à sensibilidade manifestada à incerteza e ruído que possam caracterizar as entradas usadas (Campbell e Adamson, 2006). O risco que se prevê é o de sobreparametrização e excessivo treino. Todavia, a sua adoção tem sido uma prática corrente com grandes provas de sucesso demonstrado. A forma como a sobreparametrização pode condicionar o período de treino não pode ser descrito de forma completamente esclarecedora, sendo argumentado por alguns autores (Hippert *et al.*, 2001) que o número de parâmetros efetivos da rede não corresponde ao mesmo número dos seus pesos e desvios, uma vez que alguns destes se tornam praticamente irrelevantes. Alguns estudos procuraram demonstrar que o modelo de redes neuronais aplicado à previsão de uma variável

---

[35] Colinearidade ocorre sempre que se verificam relações lineares entre as variáveis explanatórias.

dependente a partir de uma variável explanatória recorre apenas a um sub-espaço dimensional efetivo da variável explanatória que retém as relações estatísticas entre a variável de saída a prever e a variável explanatória. Esta explicação pode fundamentar que a sobreparametrização pode nem sempre acarretar um excessivo período de treino. Em problemas de classificação foram já apresentados comentários que provam que a complexidade da rede não deve estar intimamente ligada ao número de pesos, mas antes à norma do vetor de pesos, o que em muitos casos justifica que simulações com alterações no número de neurónios (logo nos pesos associados) não afetam significativamente o desempenho da rede.

O processo de treino permite a atualização dos parâmetros da rede de forma a serem minimizados os erros de previsão, onde os parâmetros que vão sendo experimentados ao longo do algoritmo de treino dependem necessariamente dos parâmetros para os quais a rede é inicializada.

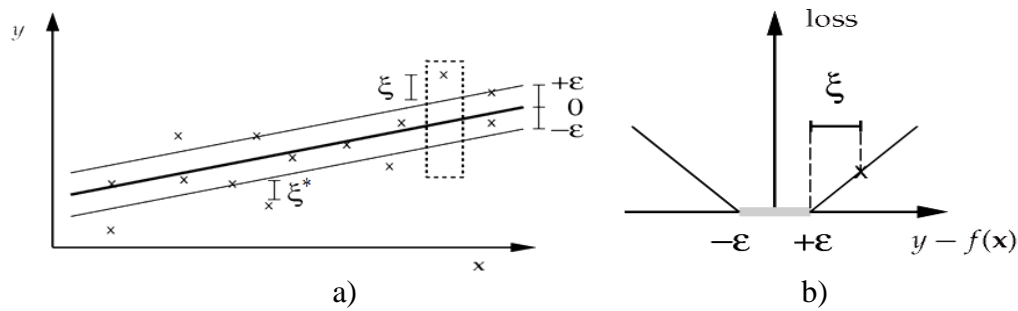
### **3.2. Previsão de consumos com recurso a máquinas de suporte vetorial**

As máquinas de suporte vetorial baseiam-se numa técnica de aprendizagem supervisionada que implementa um princípio estruturado de minimização do risco de pouca generalização, em vez de um princípio empírico de minimização do risco, que é no fundo o princípio subjacente às redes neuronais (Hamel, 2009, Pai e Hong, 2005, Scholkopf e Smola, 2002, Scholkopf *et al.*, 1998). A grande vantagem que tem sido apontada ao uso de regressão baseada em suporte vetorial é o conceito de minimização de uma tolerância máxima para o erro de generalização ao invés do processo de treino de redes neuronais que apenas permite minimizar o erro associado ao conjunto de dados segundo os quais a rede é efetivamente treinada. Deste modo, as máquinas de suporte vetorial podem ser entendidas como um problema de programação quadrática de restrição linear que garante uma solução que será sempre única e globalmente ótima, como é apresentado posteriormente (Pai e Hong, 2005).

As máquinas de suporte vetorial tiveram as primeiras aplicações em problemas relativos a reconhecimento de padrões, para o qual o método apenas lida com classes de valores, como por exemplo saídas que pudessem tomar valores -1 ou +1. Contudo, a contribuição de *Vapnik* <sup>[36]</sup> (Schölkopf *et al.*, 2002) foi decisiva para estender a metodologia à estimação de funções com valores reais. Esta estratégia passa por adotar uma função que permite quantificar as perdas incorridas com erros de previsão, ainda que a mesma apresente um determinado grau de insensibilidade. Na Figura 3.6 torna-se possível entender o conceito descrito.

---

[36] Vladimir Vapnik é um reconhecido matemático de nacionalidade Russa, tendo sido um dos investigadores pioneiros na área das máquinas de suporte vetorial.



**Figura 3.6 – Regressão com base em máquinas de suporte vetorial**

**a) Margem de tolerância associada à regressão; b) Função de penalização associada aos erros identificados**  
 Figura adaptada de Schölkopf, 2002

A função que regula as perdas incorridas pelos erros de previsão apresenta-se na Equação 3-8.

$$\text{Função associada às perdas} = |y - f(x)|_{\varepsilon} = \max \{0, |y - f(x)| - \varepsilon\} \quad (3-8)$$

A variável  $\varepsilon$  define a tolerância ao erro que se assume como aceitável na regressão, devendo este grau de precisão desejado ser especificado *a priori*. De modo a exemplificar com um modelo de regressão linear, pode ser considerada uma função  $f$  dada por  $f(x) = \langle \mathbf{w} \cdot \mathbf{x} \rangle + b$ . Se para a estimativa se contar com uma precisão  $\varepsilon$ , a função objetivo expressa na Equação 3-9 deve ser minimizada.

$$\frac{1}{2} \cdot \|\mathbf{w}\|^2 + C \cdot \sum_{i=1}^m |y_i - f(x_i)|_{\varepsilon} \quad (3-9)$$

O termo  $\frac{1}{2} \cdot \|\mathbf{w}\|^2$  que deriva já de propósitos de classificação quando se pretende encontrar o ótimo hiperplano que separa duas classes distintas deve ser minimizado reduzindo por esta via a complexidade do modelo (Pai e Hong, 2005). O segundo termo  $C \cdot \sum_{i=1}^m |y_i - f(x_i)|_{\varepsilon}$  é usado com o intuito de precaver que os erros associados à regressão, sendo inevitáveis em casos práticos, sejam penalizados através de um fator  $C$ . Este fator de penalização é também um parâmetro que deve ser especificado *a priori* incidindo nos erros que ultrapassem a tolerância considerada, concluindo-se que a penalização será diretamente relacionada com a dimensão do erro. O parâmetro  $C$  é por isso usado para especificar o compromisso entre o risco empírico e a simplicidade do modelo. Duas variáveis auxiliares são introduzidas para os dois casos possíveis, de acordo com as Equações 3-10 e 3-11.

$$y_i - f(x_i) \leq \varepsilon + \xi_i \quad (3-10); \quad f(x_i) - y_i \leq \varepsilon + \xi_i^* \quad (3-11)$$



Adotando estas variáveis auxiliares, um problema de otimização com restrições pode ser encarado segundo a expressão apresentada em 3-12.

$$\min \left[ \frac{1}{2} \cdot \|w\|^2 + C \cdot \sum_{i=1}^m (\xi_i + \xi_i^*) \right] \quad (3-12)$$

sujeito às restrições expressas em 3-13 e 3-14.

$$y_i - f(x_i) \leq \varepsilon + \xi_i \quad (3-13)$$

$$f(x_i) - y_i \leq \varepsilon + \xi_i^* \quad (3-14)$$

$$\text{com } \xi_i, \xi_i^* \geq 0$$

Deste modo, verifica-se que a qualquer erro inferior a  $\varepsilon$  não está associado nenhum dos termos  $\xi_i$  ou  $\xi_i^*$ , não contribuindo desta feita para a função objetivo.

Generalizando para a regressão não linear, a ideia subjacente é mapear o conjunto de dados para outro espaço dimensional  $F$  através de um mapeamento não linear e proceder a uma regressão linear em  $F$  (Schölkopf *et al.*, 1998). O espaço dimensional original de entradas é tipicamente não relacionável linearmente com a(s) variável(is) a prever, por isso é usada uma função *kernel* que transforma num espaço de maior dimensão de modo a se poder contornar as não linearidades encontradas no espaço original. O espaço dimensional que está implicitamente embebido na matriz *kernel* será usado como novas entradas para o problema de regressão (Üstün *et al.*, 2007). Diferentes funções *kernel* podem ser usadas, sendo as mais comuns as funções lineares, polinomiais e de base radial. Para os propósitos de previsão de consumos, a função *kernel* que tem tido maior aceitação na comunidade científica é a função de base radial (Abbas e Arif, 2006, Afshin *et al.*, 2007, Espinoza *et al.*, 2007, Guo *et al.*, 2006, Pai e Hong, 2005), devido essencialmente à sua aptidão para lidar com não linearidades. A função de base radial é definida de acordo com a Equação 3-15.

$$K(x, y) = \exp^{-\gamma \|x - y\|^2} \quad (3-15)$$

em que o parâmetro  $\gamma$  deve ser visto como uma constante que será determinante para o processo e que deve, à semelhança de outros parâmetros já apresentados, ser especificado *a priori*.

De modo a respeitar a formulação primal-dual e a introduzir os multiplicadores de Lagrange (representados com os termos  $\alpha_i$  e  $\alpha_i^*$ ), o problema de otimização é equivalente a obter-se a expressão 3-16 (Hamel, 2009).

$$\max W(\alpha, \alpha^*) = \sum_{i=1}^m (\alpha_i - \alpha_i^*) \cdot \mathbf{y}_i - \varepsilon \sum_{i=1}^m (\alpha_i + \alpha_i^*) - \frac{1}{2} \cdot \sum_{i=1}^m \sum_{j=1}^m (\alpha_i - \alpha_i^*) \cdot (\alpha_j - \alpha_j^*) \cdot k(\mathbf{x}_i, \mathbf{x}_j) \quad (3-16)$$

sujeita a:

$$0 \leq \alpha_i \leq C, 0 \leq \alpha_i^* \leq C \text{ para } i = 1, \dots, m \text{ e } \sum_{i=1}^m (\alpha_i - \alpha_i^*) = 0 \quad (3-17)$$

A função que define o valor que resulta da regressão apresenta-se em 3-18.

$$f(x) = \sum_{i=1}^m (\alpha_i - \alpha_i^*) \cdot k(\mathbf{x}_i, \mathbf{x}) + b \quad (3-18)$$

em que  $b$  pode ser obtido considerando os casos em que a inequação em 3-10 tornar-se-ia uma igualdade caso  $\xi_i = 0$  e com  $0 < \alpha_i < C$ ; por outro lado a inequação 3-11 tornar-se-ia uma igualdade no caso de  $\xi_i^* = 0$  e com  $0 < \alpha_i^* < C$ .

O processo de seleção dos três parâmetros não negativos ( $\varepsilon$ ,  $C$  e  $\gamma$ ) é de extrema importância para se providenciar uma previsão com um desejado grau de precisão (Abbas e Arif, 2006, Afshin *et al.*, 2007, Elattar *et al.*, 2010, Guo *et al.*, 2006, Hamel, 2009, Lee, 2009, Mohandes, 2002, Pai e Hong, 2005, Üstün *et al.*, 2007). Contudo, são escassos os métodos estruturais para confirmar eficientemente a seleção de parâmetros. Para otimizar o procedimento associado à seleção destes parâmetros é recorrente o uso de métodos combinatórios (Abbas e Arif, 2006, Afshin *et al.*, 2007, Pai e Hong, 2005).

Pela conveniência associada à fase de codificação de soluções, é explorado neste estudo o método de otimização *Simulated Annealing*, como heurística para a determinação do conjunto de parâmetros ajustáveis <sup>[37]</sup>. O objetivo primordial é efetivamente garantir a melhor combinação de parâmetros de entrada nos modelos de máquinas de suporte vetorial que garantam um mínimo erro de previsão.

### 3.3. Relação entre caracterização horária de consumos e previsão de curto prazo

Espinosa *et al.* (2005) apresenta um artigo que relaciona os temas de perfis de consumo (*Load Profiling*) com previsão de carga (*Load Forecasting*), dois temas que têm sido caracteristicamente abordados de forma independente. O estudo baseia-se em dados de monitorização de 245 subestações e a metodologia apresentada é orientada de forma a lidar com

[37] O processo de otimização *Simulated Annealing* a ser aplicado a modelos de regressão inspirados em máquinas de suporte vetorial que serão usados neste estudo, será descrito com maior detalhe no Capítulo 5.

o desafio de previsão de consumos de curto prazo e a segmentação de perfis de consumidores numa ótica unificada. Numa primeira fase recorrem-se a séries temporais para criar um modelo de previsão de consumos. Estas séries temporais contemplam informação acerca do consumo nas últimas 48 horas, estendendo também a análise para as variações sazonais (mês do ano e dia de semana), contemplando ainda a influência de variáveis exógenas relacionadas com condições meteorológicas (necessidades de aquecimento ou arrefecimento). Posteriormente, o estudo propõe a conversão do modelo de autoregressão periódica num vetor de autoregressão (usando os coeficientes já estimados na regressão) que anule as contribuições de influências meteorológicas e sazonais, de modo a obter-se assim o denominado perfil diário típico de carga. Um processo de *clustering* não supervisionado é então aplicado de forma a identificar diferentes grupos ou padrões no conjunto de dados. O objetivo será assim reconhecer a quantidade e os diferentes tipos de perfis de carga que devem ser considerados, desconhecendo-se previamente a composição particular de procura de cada subestação. A informação relativa a estes perfis poderá ser então usada posteriormente para um refinamento específico dos modelos de previsão usados, através por exemplo de modelos *ad-hoc* para cada cluster específico. O trabalho proposto incide essencialmente no aproveitamento do conceito de *Load Forecasting* para auxiliar o *Load Profiling*. Apesar de ser comentada uma possível abordagem *Bottom-Up Approach*, o trabalho não explora a informação proveniente dos perfis de consumo de diferentes grupos de consumidores para apoiar a previsão, pelo que se pretende investigar essa opção no decorrer deste estudo.

Chen *et. al* (2001) aborda no seu estudo um método que compreende uma fase de segmentação de grupos de consumidores em classes (residencial, comercial e industrial), e de acordo com a seleção de um conjunto de dados de monitorizações que componha cada amostra, são elaborados os perfis de consumo para cada grupo. A referida seleção depende do número total de consumidores efetivamente monitorizados em cada amostra e da variância sentida no consumo de energia para cada classe. Com base nos padrões diários de carga normalizados e no consumo total de energia que resulta de todos os consumidores de cada classe (pode ser um valor conhecido ou baseado em estimativas), pode obter-se o diagrama de carga reconstruído e através de uma agregação dos diagramas reconstruídos para cada classe resulta o diagrama global do sistema. O trabalho apresenta ainda uma análise regressiva para relacionar o consumo de energia em função de variáveis meteorológicas (temperatura e humidade) para cada classe em particular (como já foi anteriormente abordado no Capítulo 2).



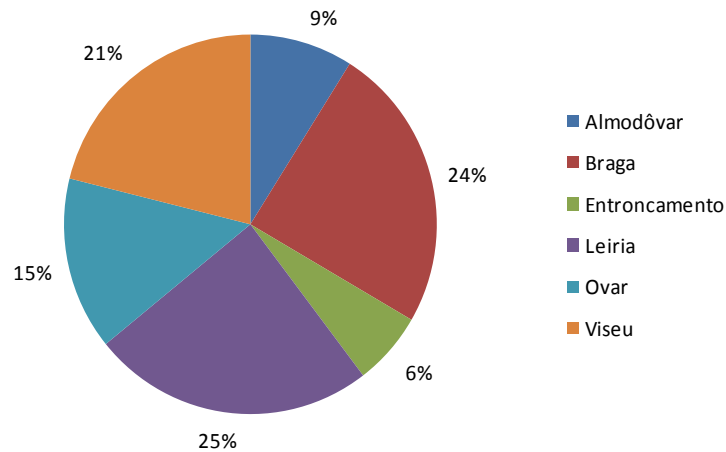
## ***4. Caracterização dos dados disponíveis e das séries temporais resultantes***

Neste capítulo serão inicialmente apresentados os dados originais que foram considerados no âmbito deste estudo. Para validar a proposta que é apresentada nesta tese, foram considerados dois estudos de caso consideravelmente distintos. No primeiro caso em estudo são usados dados de consumo que resultam da agregação de um conjunto de consumidores sujeitos a campanhas de medição nas suas instalações. O segundo caso reporta a uma monitorização efetuada numa saída de uma subestação de distribuição, tornando a análise mais realista.

### **4.1. Dados disponíveis no primeiro estudo de caso considerado**

O primeiro conjunto de dados utilizado foi obtido no âmbito de um projeto coordenado pelas unidades de investigação *INESC Coimbra* e *INESC Porto*, correspondendo a dados de consumo monitorizados, com uma resolução de quinze minutos, numa amostra de consumidores de Média e Baixa Tensão (BTE e BTN). A motivação para esse projeto resultou da necessidade de serem garantidas algumas condições essenciais à criação do mercado ibérico de eletricidade (*MIBEL*). Deste modo, o projeto visou apoiar o processo de caracterização de consumidores (através da elaboração de perfis de consumo de Baixa Tensão para diferentes classes consideradas) e a estimação da contribuição das cargas elétricas para as perdas da rede. Os dados disponíveis podem também vir a ser uma fonte de informação relevante para o estudo de procura de energia nomeadamente, e no que respeita a este estudo, auxiliar estudos relativos à previsão de carga elétrica. Para além dos dados de consumo monitorizados a esta amostra de consumidores, do conteúdo do projeto faz também parte um conjunto de dados comerciais de faturação e resultados de inquéritos realizados que permitem conhecer, entre outros atributos, o setor de atividade económica. A amostra de consumidores monitorizados contempla um total de 1158 consumidores, entre os quais 1080 consumidores BTN, 64 consumidores BTE e 14 consumidores MT. As monitorizações foram efetuadas maioritariamente nas regiões Norte e Centro de Portugal, envolvendo áreas geográficas como Braga, Ovar, Viseu, Coimbra, Leiria e Entroncamento. No Sul do país, apenas na região de Almodôvar se verificaram campanhas de

monitorização. A distribuição do número de consumidores monitorizados por área geográfica apresenta-se na Figura 4.1.



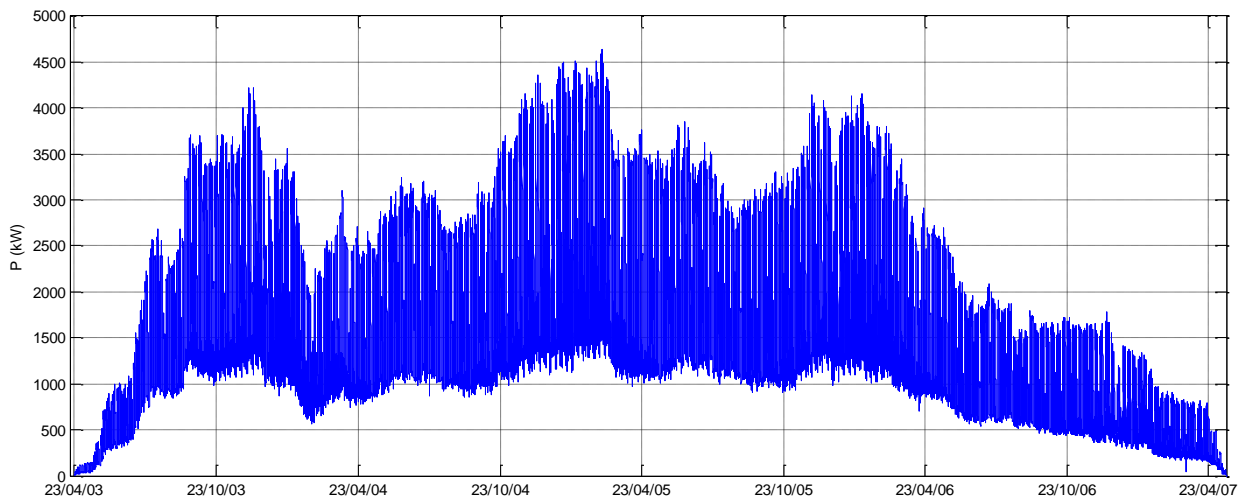
**Figura 4.1 – Distribuição do número total de consumidores monitorizados por área geográfica**

As primeiras monitorizações tiveram início a 23 de abril de 2003 enquanto os dados mais recentes que se encontram disponíveis no âmbito do projeto em causa reportam a 16 de maio de 2007. É de realçar que não se verifica coincidência temporal dos diagramas de carga obtidos e, durante este horizonte temporal, os consumidores envolvidos não estiveram sempre a ser sujeitos a monitorizações. Por um lado, o início do período das referidas campanhas de monitorizações foi efetivado de forma faseada por zonas geográficas, sendo também possível verificar que cada consumidor pode ter mais do que um diagrama de carga disponível em diferentes períodos. Esta constatação justifica-se pelas normais dificuldades técnicas associadas a um projeto desta envergadura, em que os equipamentos usados nas campanhas se encontravam consideravelmente dispersos e na ocorrência de registos erróneos nas leituras, ou mesmo interrupções de leitura, ser necessário desencadear a programação de novos intervalos de medição. Na Figura 4.2 é apresentada uma evolução do número de consumidores para os quais existem registos de leitura de consumo ao longo do horizonte temporal definido.



**Figura 4.2 – Evolução do número de consumidores com dados de consumo disponíveis ao longo do período que reporta às monitorizações efetuadas**

Neste primeiro caso em estudo procedeu-se a uma agregação de consumidores, por forma a simular um diagrama de carga agregado. Apesar da dispersão geográfica do consumo medido, a pertinência desta investigação prende-se com a possibilidade de adoção posterior da informação proveniente dos perfis de consumo (que na verdade foram projetados com base nesta mesma amostra de consumidores) em estratégias de previsão de consumos. No caso de se proceder a uma agregação dos diagramas de carga disponíveis que resultaram das campanhas de monitorização a esta amostra de consumidores, obtém-se o diagrama de carga apresentado na Figura 4.3.



**Figura 4.3 – Diagrama de carga resultante da agregação dos consumos medidos em consumidores individuais**

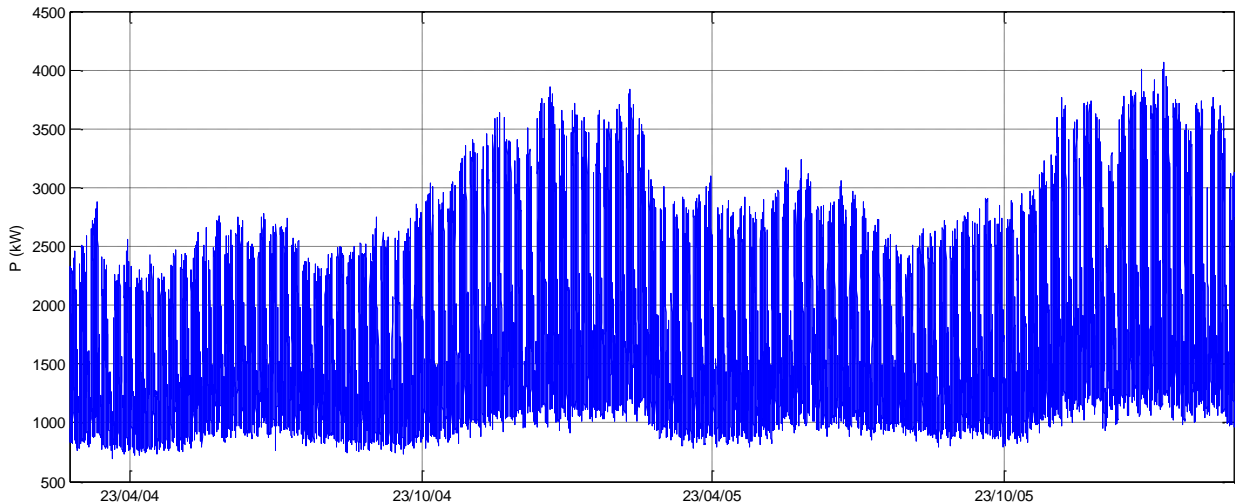
Por forma a encarar-se um critério que permita despistar evoluções de consumo que fossem apenas dependentes do número de consumidores contribuintes para este diagrama agregado, foi estabelecido um número fixo de consumidores, o que conduziu à agregação de 740 consumidores escolhidos de forma aleatória. Para que esta contribuição mínima de 740 consumidores fosse possível e também para que na janela temporal escolhida houvesse uma estabilidade desejada nos dados de consumo (sem haver interrupções de medição e/ou a presença de dados erróneos), foi considerado o período entre 16 de março de 2004 e 16 de março de 2006 (correspondendo a 2 anos) e, durante este período algumas considerações foram tomadas durante o processo de escolha aleatória de consumidores contribuintes:

- no instante em que se verifica alguma interrupção na medição de um determinado consumidor, procede-se à escolha de um novo consumidor contribuinte com características semelhantes (mantendo sempre fixo o número de 740 consumidores), o que neste caso se subentende um consumidor com o mesmo nível de tensão e com a mesma atividade económica;

- neste processo de escolha, é sempre guardada informação relevante que permita identificar posteriormente cada consumidor contribuinte. Deste modo, pode-se classificar em

cada período de monitorização (intervalos de 15 minutos) o tipo de consumo que contribui para o consumo acumulado, desagregando por setor de atividade, por nível de tensão ou por outro atributo comercial que se revele útil.

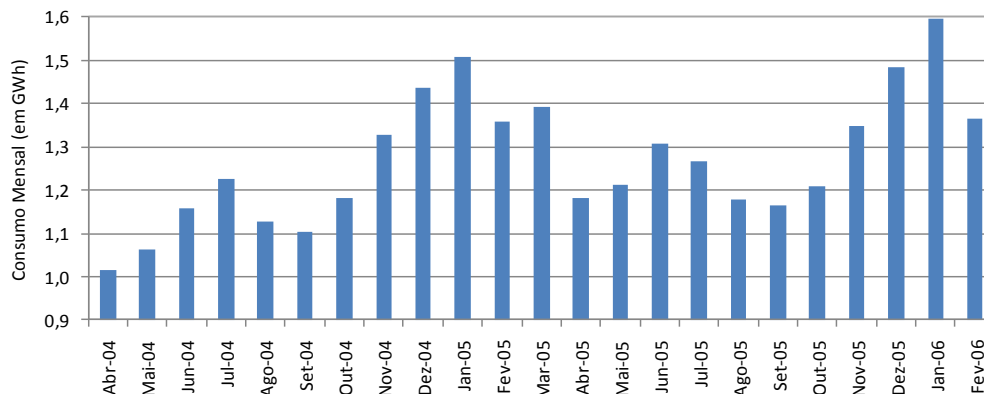
Da agregação aleatória de consumos resulta o diagrama de carga global que é apresentado na Figura 4.4.



**Figura 4.4 – Diagrama de carga resultante da agregação aleatória de 740 consumidores**

Apesar da fraca resolução que caracteriza os diagramas anteriormente apresentados, é possível comparar parte do diagrama da Figura 4.3 (restringindo ao período entre 16 de março de 2004 e 16 de março de 2006) com o diagrama da Figura 4.4 e verificar que há uma maior uniformidade na distribuição de consumo neste último caso que considera um número fixo de consumidores contribuintes. O efeito de sazonalidade é também notório, com o consumo a ser maior nos períodos de Inverno.

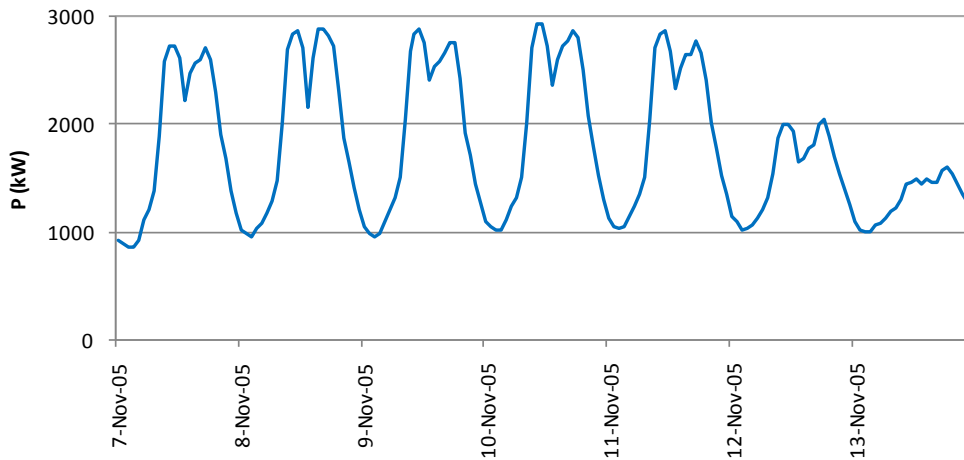
Na Figura 4.5 podem verificar-se os valores acumulados de energia consumida pelo conjunto de consumidores agregados distribuídos pelos meses completos do horizonte temporal considerado, de forma a evidenciarem-se os efeitos sazonais e a garantia de alguma estabilidade de tendência, apesar de eventuais alterações no conjunto de consumidores contribuintes.



**Figura 4.5 – Distribuição do consumo mensal agregado**



Por forma a aumentar-se a resolução do diagrama de carga já apresentado na Figura 4.4 e a tornar mais fácil a análise do padrão de consumo que resulta desta agregação, na Figura 4.6 é apresentado o diagrama de carga semanal referente a este estudo de caso.



**Figura 4.6 – Diagrama de carga semanal resultante da agregação aleatória de 740 consumidores**

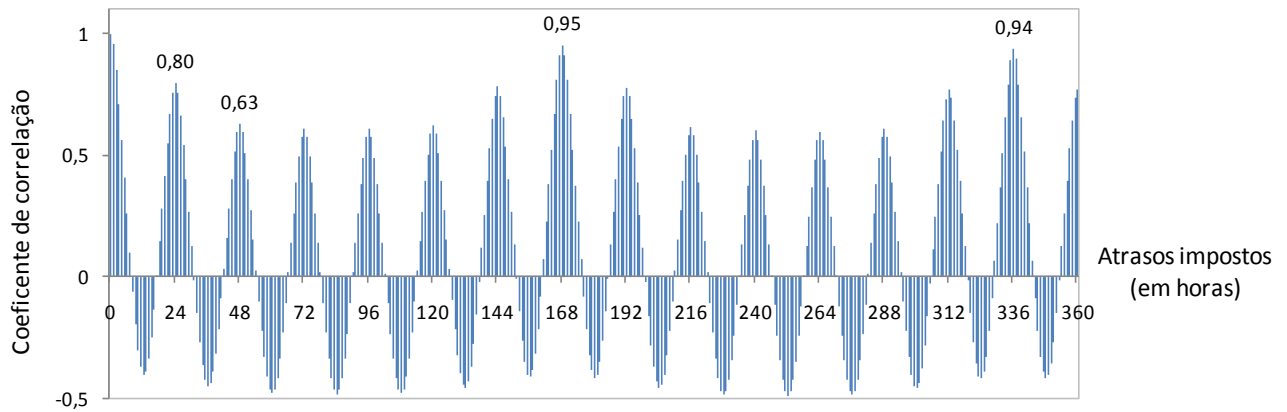
#### 4.1.1. Análise da série temporal resultante do primeiro estudo de caso

Conhecida a série temporal que caracteriza o primeiro estudo de caso, esta deve ser sujeita a uma análise mais pormenorizada, identificando os eventuais efeitos de autocorrelação (registos anteriores que terão um maior impacto no registo atual/futuro), tendências ou efeitos sazonais.

Uma vez que os modelos de previsão abordados no Capítulo 5 terão uma base horária, a série temporal foi ajustada de modo a serem considerados registos horários, em vez de registos com um período de monitorização de 15 minutos. Estes registos quarto-horários que compõem cada hora foram sujeitos a uma média simples para se encontrar o registo da hora correspondente. O risco da série temporal perder alguma resolução é de alguma forma atenuado pelo tamanho da amostra de dados considerado e o nível de consumo envolvido. O nível de agregação considerado possibilita algum amortecimento a variações bruscas no padrão de consumo que determinados consumidores individuais possam impor ao diagrama aglomerado. Uma análise de autocorrelação à referida série temporal apresenta-se na Figura 4.7. O coeficiente de correlação adotado nas análises subsequentes é o coeficiente de correlação de Pearson, obtido a partir da expressão 4-1.

$$\text{coeficiente de correlação} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4-1)$$

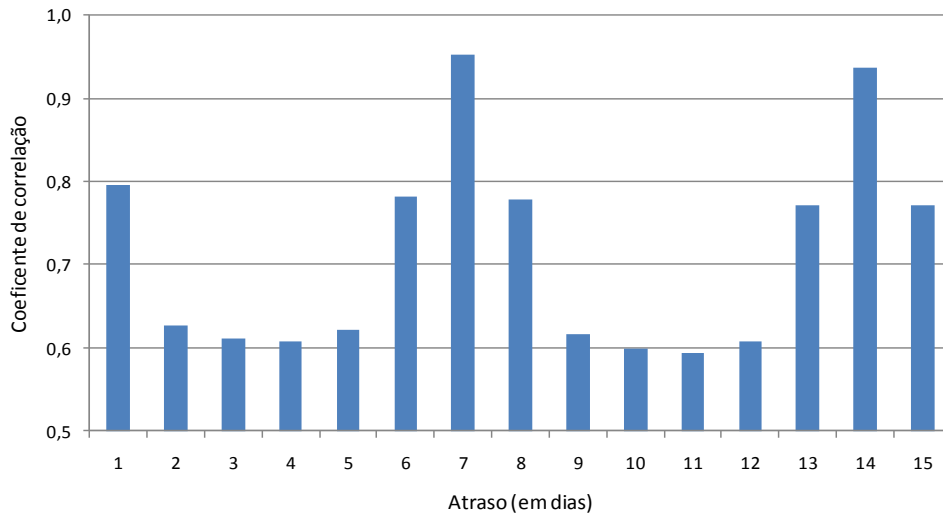
em que os valores  $x_i$  e  $y_i$  são os valores medidos e  $\bar{x}$  e  $\bar{y}$  são as médias aritméticas de ambas as variáveis.



**Figura 4.7 – Análise de autocorrelação da série temporal considerada no primeiro estudo de caso**

Da análise pode facilmente constatar-se que os registos identificados nas duas semanas anteriores à mesma hora (com atrasos de 168 horas e 336 horas respetivamente) têm uma forte correlação com o registo atual de consumo. A evolução sentida neste gráfico de autocorrelação é típica, já que as horas anteriores múltiplas do inteiro 24 apresentam um peso mais significativo, com uma correlação positiva com o registo atual. Para determinados registos passados a correlação toma valores negativos, facto justificado pelos consumos de energia típicos associados aos horários. Dando como exemplo o início do dia, por volta das 9h da manhã, o consumo de energia tende a aumentar devido à atividade económica e aos hábitos de consumo nas residências, enquanto às 21h o consumo diminui já que os horários laborais cessam e o consumo das famílias começa também a ser mais modesto. No gráfico da Figura 4.7 são ainda rotulados alguns valores de correlação que serão úteis para análises subsequentes que se apresentam no Capítulo 5.

Como se irá verificar posteriormente, os modelos de previsão visam estimar valores de potência ativa para os registos horários do dia seguinte. Neste sentido, os modelos têm de estar aptos a lidar com múltiplas saídas por forma a permitir prever em bloco um conjunto de 24 valores de potência correspondentes aos registos horários do dia seguinte. Focando no processo de escolha dos dias anteriores que podem estar mais associados ao dia que se pretende prever, é apresentada a Figura 4.8 que pretende sintetizar a informação detalhada na Figura 4.7.



**Figura 4.8 – Coeficientes de correlação identificados no primeiro estudo de caso para diferentes atrasos impostos (em dias)**

É atualmente aceite na comunidade científica que a seleção de entradas baseada apenas na análise de correlação pode revelar-se falível (Fidalgo, 2001, Hippert *et al.*, 2001). A análise de correlação permite por um lado identificar registos anteriores na série mais diretamente relacionados com a(s) saída(s). Esta análise de correlação é notável quando existem relações lineares entre entradas e saída(s). No entanto, quando há relações não lineares tal estratégia pode revelar-se insuficiente. Para além deste aspeto, a correlação analisa cada registo anterior individualmente, não procurando antever se uma eventual combinação destes valores antecedentes poderá ter também um poder discriminatório determinante para além das variáveis individuais.

A perspetiva de confrontar esta série temporal que caracteriza um consumo agregado com uma variável exógena não é neste primeiro caso em estudo aplicável, já que não se dispõe de dados suficientes para o efeito. No caso de se pretender integrar no estudo variáveis exógenas relativas a condições meteorológicas, o estudo viria consideravelmente distorcido já que a série resulta da agregação de um vasto conjunto de consumidores dispersos por diferentes regiões do território nacional e que apresentam registos meteorológicos muito diferenciados.

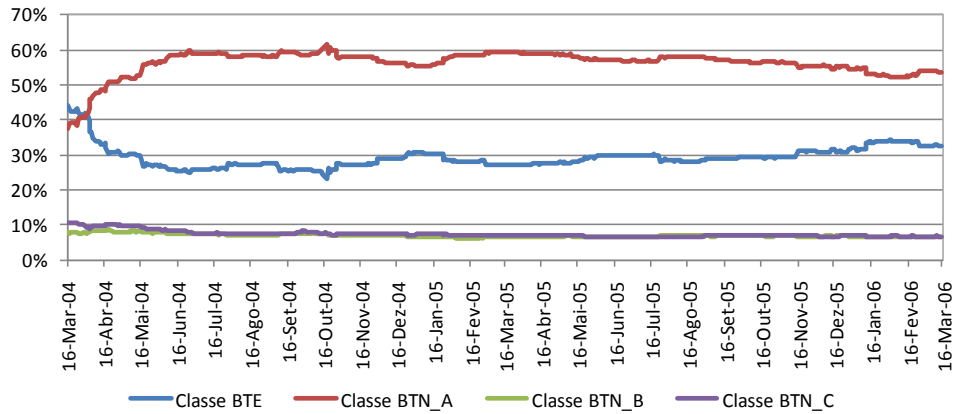
#### 4.1.2. Processo de elaboração do diagrama sintetizado associado ao primeiro estudo de caso a partir de perfis de consumo

Um dos objetivos deste estudo é a integração de informação resultante da caracterização de consumos de energia elétrica para fins de previsão. Tendo sido já assumido que a previsão de consumos e a caracterização de consumos por meio de perfis têm sido duas temáticas tratadas de forma independente (Espinoza *et al.*, 2005), é expeável que do ponto de vista de gestão esta

separação seja irrelevante e mesmo desnecessária. A caracterização de consumos em Portugal tem sido vista como um meio para viabilizar a participação de todos os consumidores num mercado liberalizado de energia, e é particularmente dedicada a pequenos consumidores que não dispõem de sistemas de medição horária. A sua adoção, pelo menos de uma forma formal, desencadeou-se aquando da necessidade de garantir um método para o acerto de contas entre o distribuidor e comercializadores presentes no mercado. Uma vez que a utilização de eletricidade passa a ser caracterizada e, por inerência estimada, este conhecimento pode ser transposto para favorecer a previsão de consumos (baseada numa abordagem *bottom-up*).

Adotando um determinado critério de segmentação de consumos é possível estruturar informação histórica para a elaboração de padrões que determinadas classes apresentam para distintas épocas do ano, tipos de dia e diferentes horas do dia. No Capítulo 5 serão apresentados modelos de previsão que incorporam informação proveniente de perfis de consumo como variáveis de entrada. Para este estudo de caso em específico, é determinante a existência de informação comercial, dados de faturação e setor de atividade acerca da amostra de consumidores que contribuíram para o diagrama de carga agregado, como foi descrito anteriormente aquando da apresentação dos critérios subjacentes a esta agregação. Neste âmbito existiriam distintos critérios para classificação de consumidores que poderiam ser adotados neste estudo (como foi aliás reportado no Capítulo 2). Contudo, por uma questão de simplicidade e aproveitamento da informação resultante dos perfis atualmente aprovados pela ERSE e a vigorar em Portugal para efeitos de faturação, foi seguida a estratégia de classificação de consumidores de Baixa Tensão segundo os atributos de índole comercial: potência contratada e consumo anual de energia (segundo os critérios apresentados na Tabela 2.1).

Para cada hora, uma descrição dos consumidores contribuintes para o diagrama de carga agregado foi analisada e confrontada com uma base de dados comercial que contempla um histórico de consumos anuais e, conseqüentemente, durante o período considerado o consumo total pode ser distribuído pelas diferentes classes de acordo com a Figura 4.9.



**Figura 4.9 – Repartição do consumo total pelas diferentes classes de consumo consideradas durante o período estudado**

Analisando esta distribuição pode concluir-se que as maiores contribuições resultam dos consumidores com maior potência contratada e consumo de energia, normalmente associados a um setor de atividade económica, sendo muito modesta a contribuição acumulada dos consumidores com potência contratada inferior a 13,8 kVA (Classes BTN - B e C), que nunca alcança os 19% do consumo total. Nesta distribuição também se denotam alguns períodos para os quais há uma considerável simultaneidade de interrupções de monitorizações, havendo por vezes necessidade de se selecionarem consumidores com outras características comerciais (este efeito está patente na segunda quinzena de outubro de 2004). Com a informação resultante desta distribuição de consumo pelas diferentes classes e tendo também os perfis de consumo associados a cada classe, torna-se possível elaborar o que se designará, de ora avante, como diagrama de carga sintetizado. Os perfis de consumo anuais foram diretamente importados da informação disponibilizada pela ERSE, sendo considerados os perfis iniciais (antes de contemplarem efeitos meteorológicos) normalizados de modo a que a soma dos diferentes registos horários que compõem um ano seja unitária. A composição do diagrama de carga sintetizado seguiu a metodologia que está expressa na Equação 4-2.

$$DC_{\text{sintetizado}_h} = W_i \cdot \left( \frac{\sum_{i=1}^4 PC_{\text{classe}_i_h} \cdot \text{percCons}_{\text{classe}_i_h}}{\sum_h \sum_{i=1}^4 PC_{\text{classe}_i_h} \cdot \text{percCons}_{\text{classe}_i_h}} \right) \quad i = 1, \dots, 4 \quad (4-2)$$

com:

$DC_{\text{sintetizado}_h}$  – registo horário associado ao diagrama de carga sintetizado;

$PC_{\text{classe}_i_h}$  – perfil de consumo da classe  $i$  à hora  $h$ ;

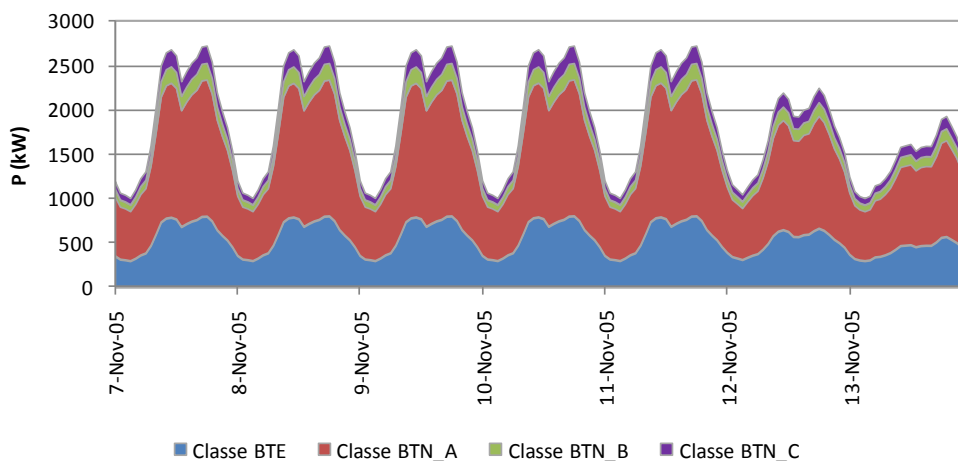
$\text{percCons}_{\text{classe}_i_h}$  – percentagem de consumo da classe  $i$  à hora  $h$ ;

$W_t$  – consumo total estimado da agregação considerada (numa base anual);

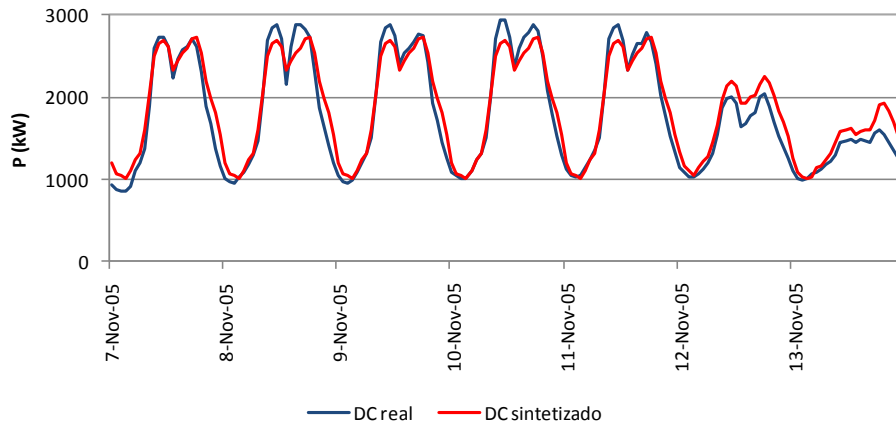
$h$  – cada hora considerada (entre a hora 0 do dia 16 de março de 2004 e a hora 23 do dia 16 de março de 2006).

O denominador da equação 4-1 visa uma nova normalização do diagrama que resulta da contribuição dos diferentes perfis, ficando-se a conhecer o peso que cada hora em específico terá no diagrama de carga sintetizado durante o horizonte temporal analisado. Com esta nova normalização a soma dos diferentes registos será equivalente à unidade, tornando-se deste modo possível multiplicar o diagrama pelo valor total de consumo durante o período em análise, garantindo-se que o integral do diagrama corresponda ao valor de consumo total utilizado. Este valor total de consumo pode ser encarado como uma estimativa, tendo como base a evolução típica do consumo anual. Como se irá facilmente comprovar no Capítulo 5, a estimativa do consumo horário dependerá, por esta via, do consumo total anual. Se este consumo total resulta de uma estimativa baseada na evolução típica anual, este facto pode ser uma fonte de propagação de erros quando se escrutina o consumo horário. Ainda assim, como os modelos de previsão analisados são horários, tão ou mais importante do que prever com acuidade o valor exato de cada registo, será prever com rigor a forma assumida pelo diagrama de carga (através dos pesos associados aos diferentes registos horários). Pode então concluir-se que qualquer erro que exista na estimativa do consumo total anual será diluído pelos diferentes registos horários, sem comprometer em grande escala a metodologia adotada.

A título exemplificativo é apresentado um diagrama de carga sintetizado na Figura 4.10 que evidencia a contribuição cumulativa das quatro classes que foram adotadas. Na Figura 4.11 é apresentada uma comparação entre o diagrama de carga real e o diagrama que é obtido a partir das contribuições das diferentes classes.



**Figura 4.10 – Diagrama de carga sintetizado evidenciando a contribuição de cada classe – efeito de agregação durante uma semana de novembro de 2005**

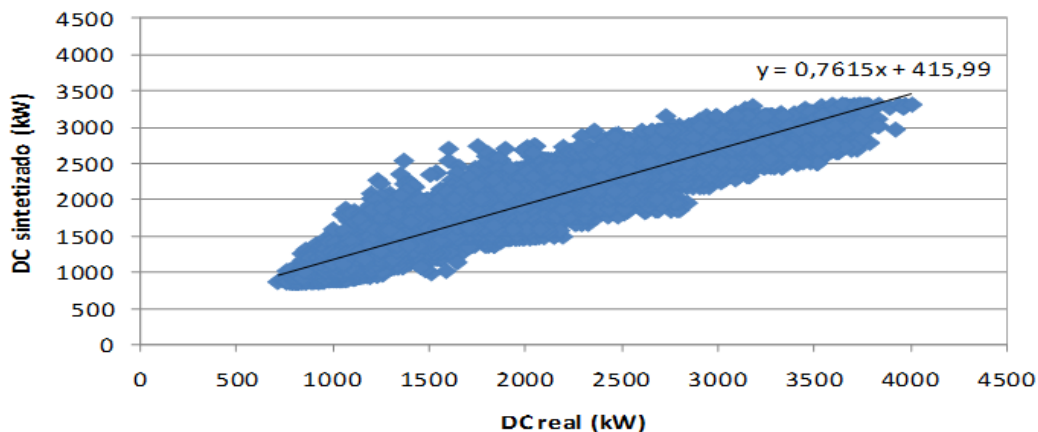


**Figura 4.11 – Comparação entre o diagrama de carga real e o diagrama de carga sintetizado obtido para o primeiro estudo de caso**

Esta última comparação permite verificar que no efeito da elaboração dos perfis consideram-se todos os dias úteis com um similar padrão de consumo, o que pode acarretar alguns desvios não desprezáveis face ao consumo verificado na realidade. Na secção seguinte é verificada a correlação cruzada que resulta da comparação entre estes dois diagramas.

#### 4.1.3. Correlação cruzada entre o diagrama de carga real e o diagrama de carga sintetizado associados ao primeiro estudo de caso

Como foi já elaborada uma análise de autocorrelação da série temporal associada ao diagrama de carga real (na Secção 4.1.1), torna-se aconselhável proceder a uma análise da correlação entre esta mesma série temporal e a série que resulta da composição do diagrama de carga sintetizado. A Figura 4.12 apresenta um correlograma entre os registos horários de potência ativa do diagrama de carga real que resulta da agregação aleatória de consumidores monitorizados e os registos horários de potência ativa que caracterizam o correspondente diagrama de carga sintetizado (relativo à mesma amostra de consumidores).



**Figura 4.12 – Correlograma entre registos horários de potência ativa do diagrama de carga real e registos horários de potência ativa do diagrama de carga sintetizado**

O correlograma da Figura 4.12 apresenta a linha de tendência definida por um tipo de regressão linear e também a equação que descreve esta linha. A regressão em causa é estabelecida através da técnica de mínimos quadrados e com esta regressão linear permite obter-se um coeficiente de correlação de 0,936 que, em termos quantitativos, pode ser comparado na Tabela 4.1 com os coeficientes de autocorrelação já analisados em 4.1.1.

**Tabela 4.1 – Coeficientes de correlação linear entre registos horários de potência ativa do DC real e outros registos analisados no primeiro estudo de caso**

	Confronto dos registos horários de potência ativa do DC real à hora $h$ com registos resultantes de:	Coeficiente de correlação linear
Registos prévios do DC real	- 1 hora de atraso no DC real ( <i>hora h-1</i> )	0,954
	- 2 horas de atraso no DC real ( <i>hora h-2</i> )	0,848
	- 1 dia de atraso no DC real ( <i>hora h-24</i> )	0,796
	- 2 dias de atraso no DC real ( <i>hora h-48</i> )	0,627
	- 1 semana de atraso no DC real ( <i>hora h-168</i> )	0,953
	- 2 semanas de atraso no DC real ( <i>hora h-336</i> )	0,937
Registo do DC sintetizado	- hora $h$ do DC sintetizado	0,936

## 4.2. Dados disponíveis no segundo estudo de caso considerado

O segundo caso em estudo pretende criar um cenário mais realista, já que recorre a dados obtidos a partir de monitorizações numa subestação de distribuição na vizinhança de Leiria. Os dados foram gentilmente cedidos pelo Gabinete de Estudos Empresariais da EDP para efeitos deste estudo, correspondendo à subestação de distribuição de Ortigosa que transforma o nível de alta tensão a 60 kV em média tensão a 15 kV, abastecendo uma zona com qualidade de serviço C (localidade com um número inferior a 2500 clientes). De entre um conjunto de 6 saídas que esta subestação alimenta (entre as quais Monte Redondo, Monte Real, Base Aérea, Amor e Gândara), a análise deste estudo de caso debruçar-se-á na saída de Monte Real <sup>[38]</sup>. Os dados de monitorização disponíveis reportam a um período entre 15 de dezembro de 2006 e 30 de novembro de 2009, o que configura um horizonte temporal que se aproxima a três anos completos. Durante o período de monitorização denotaram-se alguns dados erróneos, associados a reais quebras de energia, mas também a falhas de monitorização ou até mesmo a eventuais erros associados à transmissão de dados. Cada período alvo de correção nunca teve uma duração

[38] A base aérea de Monte Real é alimentada a partir de uma saída específica, pelo que a saída analisada no âmbito deste estudo não é influenciada pelo padrão de consumo desta instituição.



superior a 1 semana, tendo sido adotada uma estratégia de correção de acordo com o seguinte procedimento:

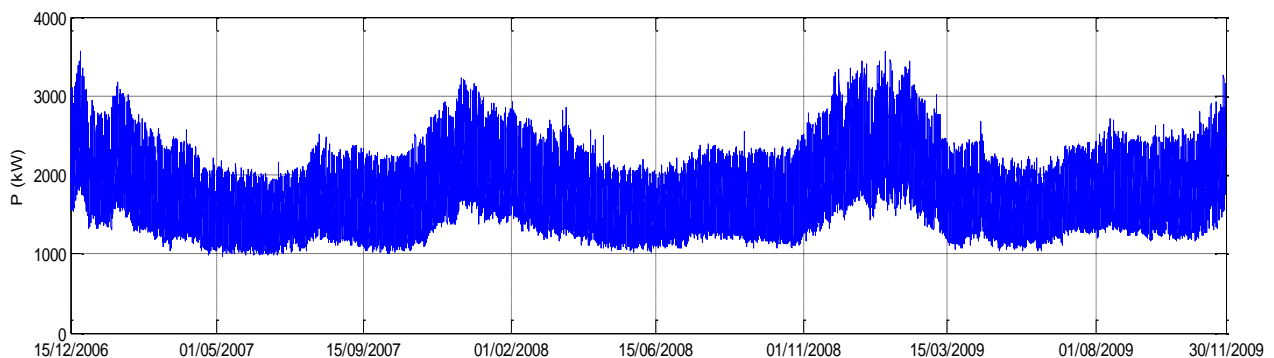
– quando a falha identificada se manifesta com uma duração inferior a um dia, cada dado erróneo é substituído por um registo resultante da média ponderada de dados congéneres (ocorridos à mesma hora) identificados em dias similares das duas semanas vizinhas, ou seja na semana precedente e na semana seguinte (com um peso de 30%), e de dados decorridos à mesma hora no dia útil imediatamente anterior e no dia útil imediatamente posterior (com um peso de 70%). Em termos matemáticos, esta metodologia respeita a expressão 4-3.

$$P_{\text{dia } d \text{ _hora } h} = 0,3 \times \left( \frac{P_{\text{dia } d-7 \text{ _hora } h} + P_{\text{dia } d+7 \text{ _hora } h}}{2} \right) + 0,7 \times \left( \frac{P_{\text{dia útil } d-1 \text{ _hora } h} + P_{\text{dia útil } d+1 \text{ _hora } h}}{2} \right) \quad (4-3)$$

– quando a falha identificada se manifesta com uma duração compreendida entre um dia e uma semana, cada dado erróneo é substituído por um registo resultante da média de dados congéneres identificados em dias similares das duas semanas vizinhas, ou seja na semana precedente e na semana seguinte. Em termos matemáticos, esta metodologia respeita a expressão 4-4.

$$P_{\text{dia } d \text{ _hora } h} = \frac{P_{\text{dia } d-7 \text{ _hora } h} + P_{\text{dia } d+7 \text{ _hora } h}}{2} \quad (4-4)$$

Apesar da fraca resolução associada, a Figura 4.13 permite dar uma perspetiva do diagrama de carga corrigido correspondente a esta saída para o período de monitorização em causa.



**Figura 4.13 – Diagrama de carga da saída de Monte Real na subestação de Ortigosa**

A sazonalidade está patente com os maiores consumos a ocorrerem durante os invernos. Apesar dos dados disponibilizados pela EDP terem um período de integração de 15 minutos, os dados foram transformados, numa fase inicial, para garantir um período de integração de 1 hora (adotando para cada hora um registo de potência ativa que resulta da média de 4 registos quarto-horários). Esta resolução será a mais indicada para o tipo de previsão que se irá

estabelecer no Capítulo 5. O consumo mensal distribui-se durante o período disponível segundo a Figura 4.14.

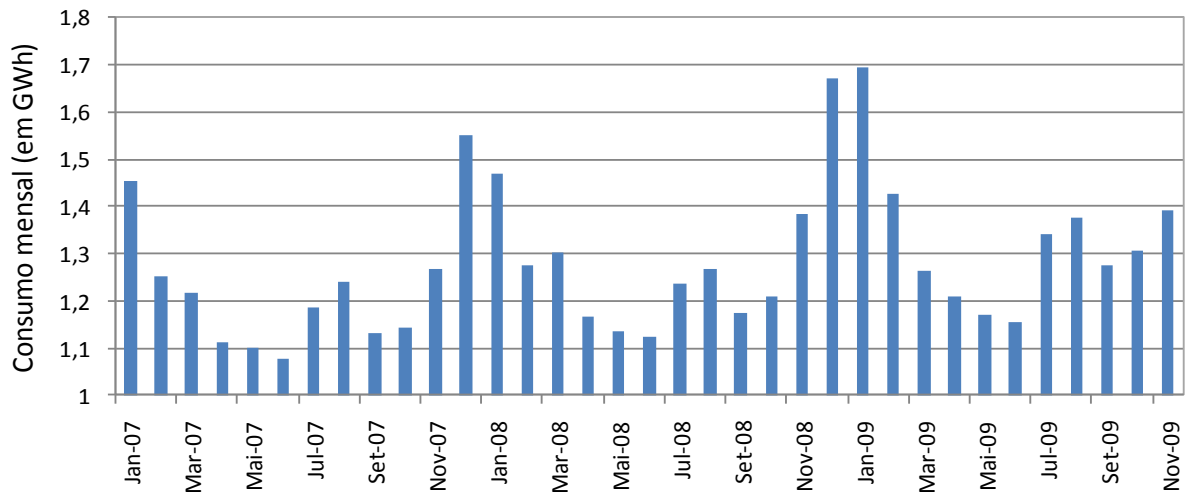


Figura 4.14 – Distribuição do consumo mensal na saída de Monte Real da subestação de Ortigosa

Apesar de não ter sido facultada informação complementar de qualquer género acerca da amostra de consumidores abastecidos por esta saída, é razoavelmente seguro admitir que o consumo agregado é predominantemente influenciado por consumidores de baixa tensão, em que a utilização de energia é essencialmente destinada a fins domésticos, como se evidencia na Figura 4.15. O diagrama de carga semanal ilustra incontestavelmente que o padrão de consumo é semelhante quando se confrontam dias laborais com dias associados a fins de semana.

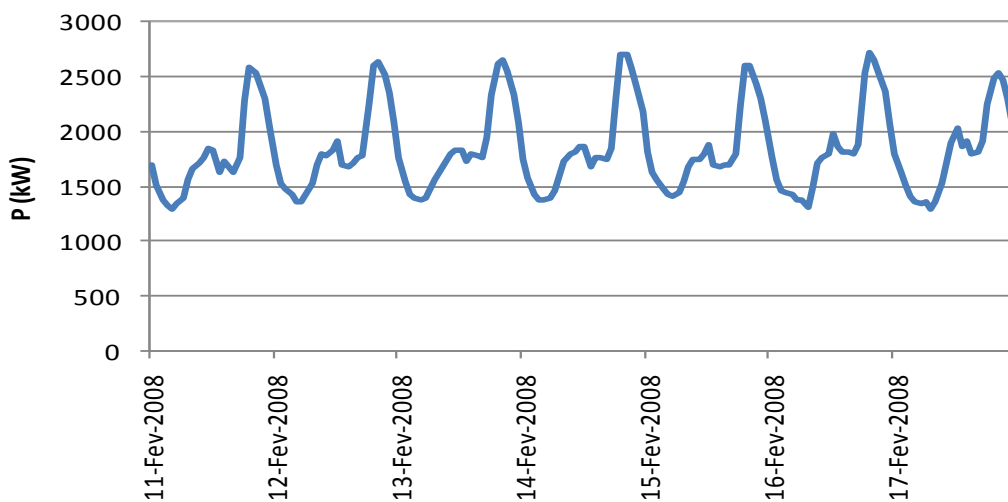
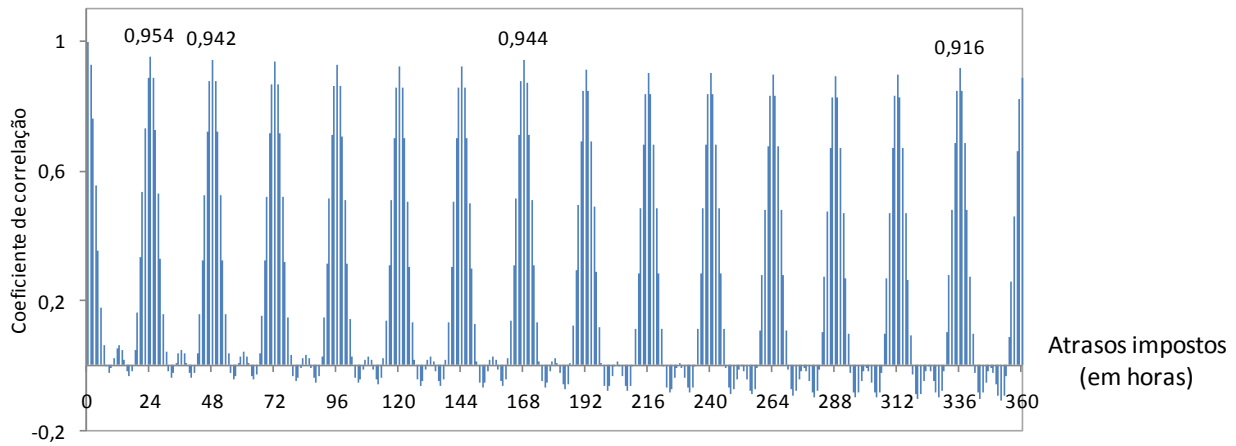


Figura 4.15 – Diagrama de carga semanal da saída de Monte Real da subestação de Ortigosa

#### 4.2.1. Análise da série temporal resultante do segundo estudo de caso

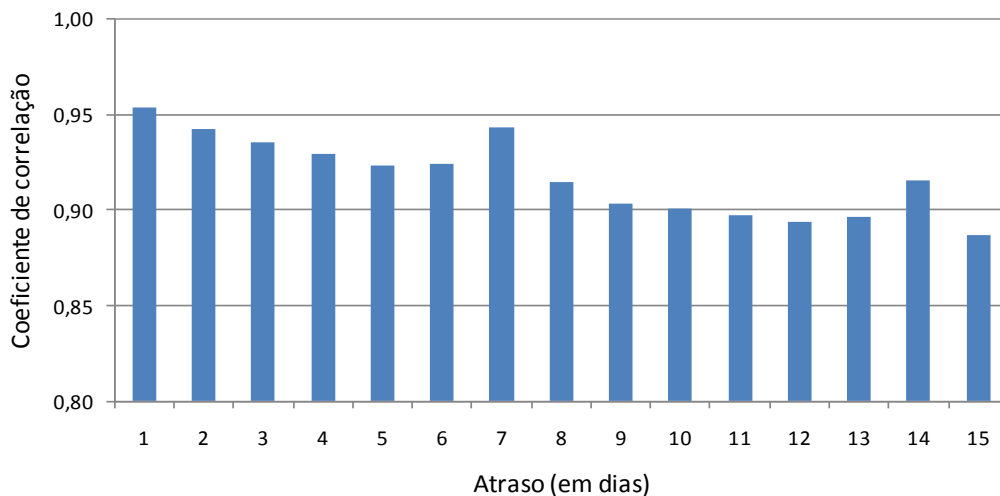
Estendendo a análise já tomada para o primeiro estudo de caso, torna-se também relevante verificar os eventuais efeitos de autocorrelação sentidas nesta série temporal. Uma análise de autocorrelação à referida série temporal vem expressa na Figura 4.16.



**Figura 4.16 – Análise de autocorrelação da série temporal considerada no segundo estudo de caso**

Para esta série temporal verifica-se que a autocorrelação apresenta os valores máximos para horas anteriores múltiplas do inteiro 24, sendo que um maior atraso não origina um decréscimo significativo de coeficientes de correlação, como sucedia na série associada ao primeiro caso em estudo. Esta constatação justifica-se pelo padrão de consumo praticamente homogéneo entre diferentes tipos de dias de semana. Outra conclusão evidente é que os coeficientes de correlação negativos têm valores absolutos pouco significativos, uma vez que esta série temporal não depende tanto de consumos relativos à atividade económica e, como tal, os efeitos de início e fim de períodos de maior utilização de energia eléctrica não são tão demarcados. No gráfico da Figura 4.16 são ainda rotulados alguns valores de correlação que serão úteis para análises subsequentes que se apresentam no Capítulo 5, à semelhança do que foi analisado para o primeiro caso em estudo.

A Figura 4.17 apresenta uma súmula da informação disponibilizada na Figura 4.16, sintetizando apenas os coeficientes de correlação identificados na análise de autocorrelação com atrasos impostos em dias.



**Figura 4.17 – Coeficientes de correlação identificados no segundo estudo de caso para diferentes atrasos impostos (em dias)**

#### 4.2.2. Processo de elaboração do diagrama de carga sintetizado associado ao segundo estudo de caso a partir de perfis de consumo

O facto de não se dispor de informação comercial acerca dos consumidores abastecidos por esta saída condiciona o processo de elaboração do diagrama de carga sintetizado, já que este estágio pressupunha uma rigorosa segmentação do consumo total por classes para que posteriormente se elaborasse a composição (a partir de informações provenientes de perfis de consumo) de um diagrama de carga que se pretendia comparável ao diagrama de carga real. A solução para contornar esta lacuna foi encarar um método combinatório, neste caso baseado no *Simulated Annealing* [39], que explorasse diferentes pesos assumidos pelas diferentes classes de consumidores de Baixa Tensão. O modelo de otimização resume-se à busca de uma solução num espaço que envolve quatro variáveis (pesos associados a cada classe – BTE, BTN – Classes A, B e C). Em cada iteração o modelo de otimização gera quatro números no intervalo [0;1] que posteriormente são normalizados segundo a expressão 4-5 para se garantir que a soma dos mesmos é unitária, e quando multiplicados pelos perfis de consumo associados a cada classe, permitem obter um diagrama de carga sintetizado que resulta da agregação de cada classe individual.

$$percCons_{normalizada\_classe\_i} = \frac{percCons_{classe\_i}}{\sum_{i=1}^4 percCons_{classe\_i}}, i = 1, \dots, 4 \quad (4-5)$$

em que:

$percCons_{normalizada\_classe\_i}$  – percentagem de consumo da classe  $i$  normalizada;

$percCons_{classe\_i}$  – percentagem de consumo da classe  $i$  gerada aleatoriamente a partir do modelo combinatório de otimização.

A fórmula matemática que permite a obtenção dos registos horários de potência ativa do diagrama de carga sintetizado para este segundo caso em estudo apresenta-se na expressão 4-6.

$$DC_{sintetizado\_h} = W_t \cdot \left( \frac{\sum_{i=1}^4 PC_{classe\_i\_h} \cdot percCons_{normalizada\_classe\_i}}{\sum_h \sum_{i=1}^4 PC_{classe\_i\_h} \cdot percCons_{normalizada\_classe\_i}} \right) \text{ com } i = 1, \dots, 4 \quad (4-6)$$

em que:

---

[39] Esta meta-heurística será apresentada com maior detalhe no Capítulo 5.

$DC_{\text{ sintetizado}_h}$  – registo horário associado ao diagrama de carga sintetizado;

$PC_{\text{ classe}_i_h}$  – perfil de consumo da classe  $i$  à hora  $h$ ;

$\text{percCons}_{\text{normalizada\_classe}_i}$  – percentagem de consumo da classe  $i$  normalizada;

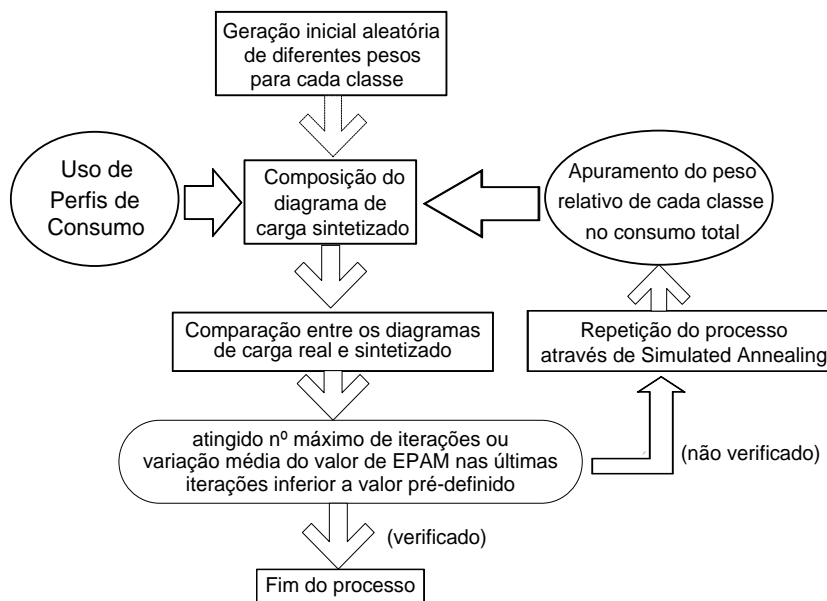
$W_t$  – consumo total estimado da agregação considerada (numa base anual <sup>[40]</sup>);

$h$  – cada hora considerada (entre a hora 0 do dia 15 de dezembro de 2006 e a hora 23 do dia 30 de novembro de 2009).

A função objetivo pelo qual o modelo se rege é determinada pelo erro percentual absoluto médio que se identifica na comparação entre o DC real e o DC sintetizado. O erro percentual absoluto médio é deste modo calculado em cada iteração a partir da expressão 4-7.

$$\text{Erro percentual absoluto médio (EPAM)} = \frac{\sum_h \left( \frac{|P_{DC\_real\_h} - P_{DC\_sintetizado\_h}|}{P_{DC\_real\_h}} \right)}{\text{Número total de horas no período considerado}} \quad (4-7)$$

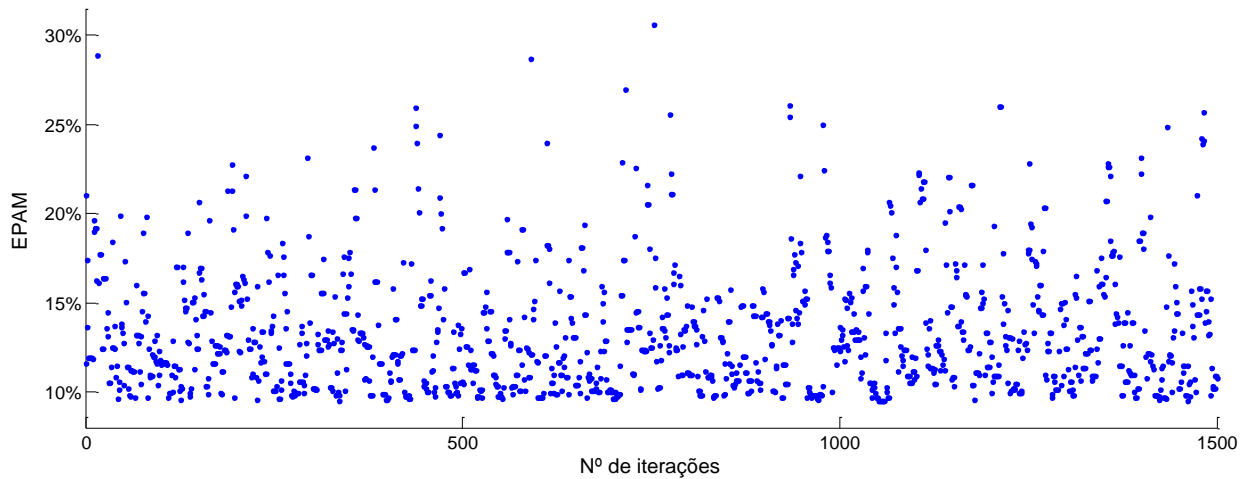
O processo de inferência das contribuições para o grupo de consumidores pretende deste modo verificar uma combinação que permite uma melhor aproximação do DC sintetizado ao DC real (através da minimização do EPAM). O processo sintetizado apresenta-se na Figura 4.18.



**Figura 4.18 – Procedimento ilustrativo do processo de inferência dos pesos assumidos por cada classe no DC sintetizado**

O processo de otimização resultante vem explícito na Figura 4.19 através da perspetiva de evolução da função objetivo a partir de diferentes combinações de pesos associados às diferentes classes que são exploradas.

[40] No estudo também foi adotada uma abordagem numa base mensal, que garantiu efetivamente uma aproximação com menor erro ao DC real, mas sem benefícios demonstrados na fase de previsão com os métodos utilizados.



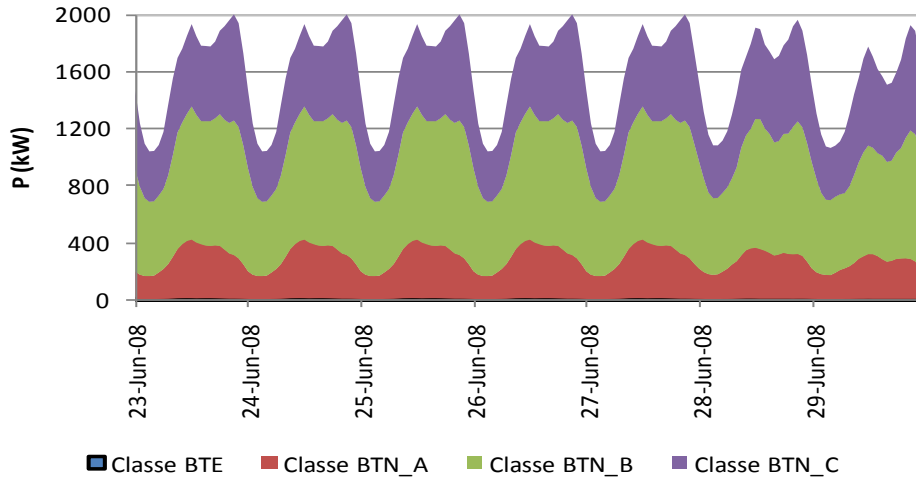
**Figura 4.19 – Evolução da função objetivo durante o processo de otimização**

Como se pode verificar, a aproximação fica consideravelmente dependente da distribuição do consumo total pelas diferentes classes adotadas, com erros percentuais absolutos médios associados à aproximação a variar entre os 9,5% e os 31%. A solução seguida foi a que conduziu ao menor erro percentual absoluto médio entre o diagrama de carga real e o diagrama de carga sintetizado, neste caso igual a 9,5%, com a distribuição do consumo total pelas diferentes classes consideradas segundo a Tabela 4.2.

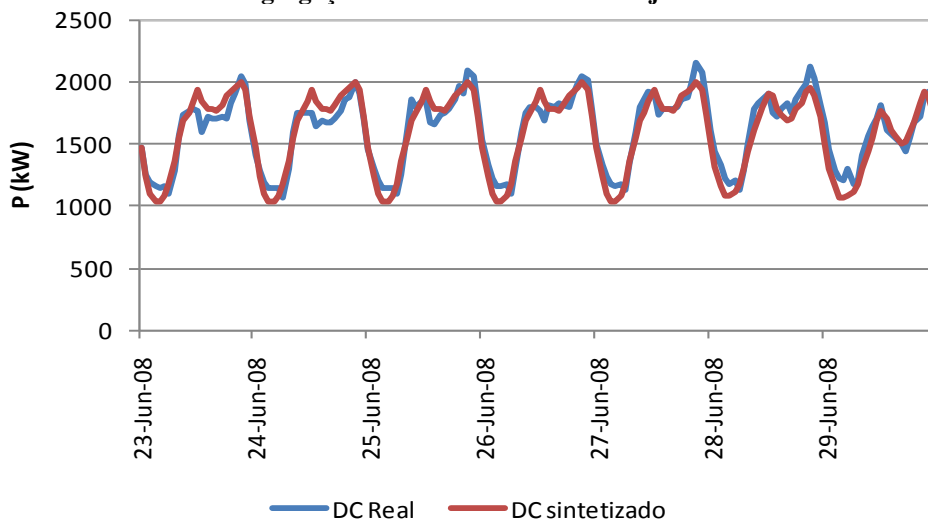
**Tabela 4.2 – Contribuição apurada por cada classe de consumidores BTN para caracterização do consumo identificado na série temporal associada ao segundo estudo de caso**

CLASSE	Peso assumido por cada classe na elaboração do DC sintetizado
BTE	0,33%
BTN – Classe A	15,80%
BTN – Classe B	49,47%
BTN – Classe C	34,40%

Com esta distribuição das diferentes classes de consumo, prosseguiu-se o estudo com a análise do diagrama de carga sintetizado resultante. Na Figura 4.20 é apresentada a contribuição cumulativa das diferentes classes para o diagrama sintetizado, para uma determinada semana do período em causa, que conduziu ao menor erro absoluto médio face ao diagrama de carga real. Uma comparação entre os dois diagramas de carga durante uma semana é referenciada na Figura 4.21.



**Figura 4.20 – Diagrama de carga sintetizado evidenciando a contribuição de cada classe – efeito de agregação durante uma semana de junho de 2008**



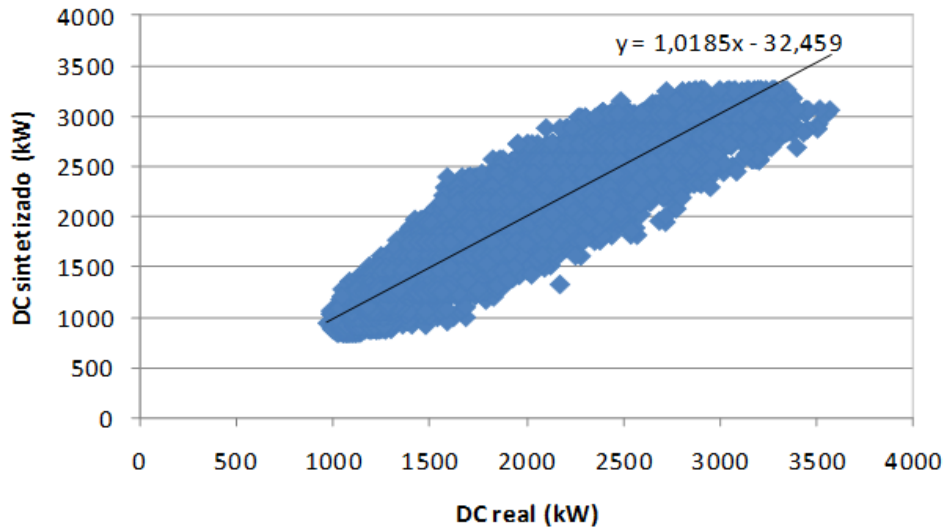
**Figura 4.21 – Comparação entre o diagrama de carga real e o diagrama de carga sintetizado obtido para o segundo estudo de caso**

Na secção seguinte é verificada a correlação cruzada que resulta da comparação entre estes dois diagramas.

#### 4.2.3. Correlação cruzada entre o diagrama de carga real e o diagrama de carga sintetizado associados ao segundo estudo de caso

Esta subsecção visa apresentar uma análise de correlação cruzada entre o diagrama de carga real e o diagrama de carga sintetizado, à semelhança da que foi desenvolvida para o primeiro estudo de caso.

A Figura 4.22 apresenta um correlograma entre os registos horários de potência ativa do diagrama de carga real que resulta da saída da subestação e os registos horários de potência ativa que caracterizam o correspondente diagrama de carga sintetizado.



**Figura 4.22 – Correlograma entre registos horários de potência ativa do diagrama de carga real e registos horários de potência ativa do diagrama de carga sintetizado**

O correlograma da Figura 4.22 apresenta a linha de tendência definida por uma regressão linear e também a equação que descreve esta linha. A regressão usando o critério dos mínimos quadrados apresenta uma correlação linear de 0,904 que, em termos quantitativos, pode ser comparada na Tabela 4.3 com os coeficientes de autocorrelação já analisados em 4.2.1.

**Tabela 4.3 – Coeficientes de correlação linear entre registos horários de potência ativa do DC real e outros registos analisados no segundo estudo de caso**

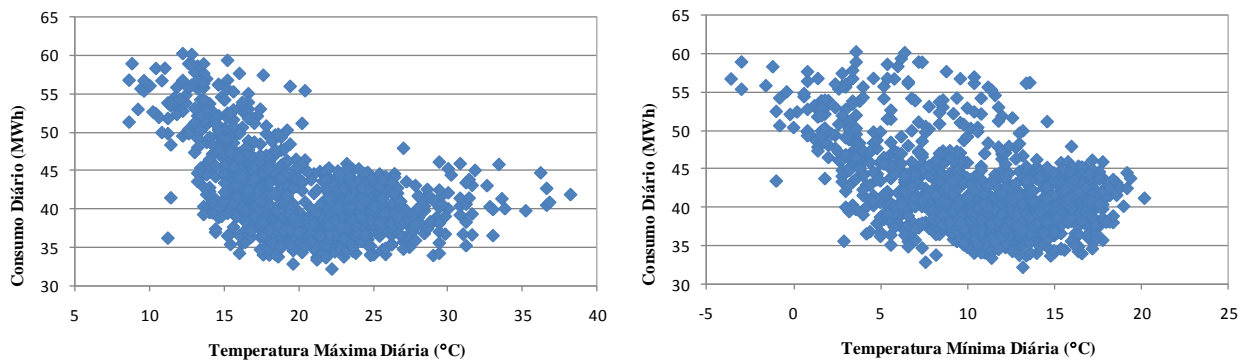
	Confronto dos registos horários de potência ativa do DC real à hora $h$ com registos resultantes de:	Coefficiente de correlação linear
Registos prévios do DC real	- 1 hora de atraso no DC real ( <i>hora h-1</i> )	0,926
	- 2 horas de atraso no DC real ( <i>hora h-2</i> )	0,761
	- 1 dia de atraso no DC real ( <i>hora h-24</i> )	0,954
	- 2 dias de atraso no DC real ( <i>hora h-48</i> )	0,942
	- 1 semana de atraso no DC real ( <i>hora h-168</i> )	0,944
	- 2 semanas de atraso no DC real ( <i>hora h-336</i> )	0,916
Registo do DC sintetizado	- hora $h$ do DC sintetizado	0,904

Apesar da correlação cruzada entre o diagrama de carga sintetizado e o diagrama de carga real se evidenciar menos relevante em comparação com autocorrelações para atrasos mais comumente adotados, a previsão poderá sair favorecida com possíveis relações não lineares entre as variáveis dependentes e variáveis independentes (que uma simples análise de correlação linear não permite despistar). Por conseguinte, no capítulo 5 será encarada a inclusão do diagrama de carga sintetizado com os seus registos horários a serem usados como variáveis de entrada para o processo de previsão.



#### 4.2.4. Análise dos dados climatéricos disponíveis na área geográfica coincidente com a série temporal do segundo estudo de caso

Para este estudo de caso foi possível ainda dispor de dados meteorológicos registados na Base Aérea de Monte Real. Estes dados foram gentilmente cedidos pelo Ministério da Defesa Nacional - Estado-Maior da Força Aérea após um pedido oficial dos mesmos ao Gabinete de Comando da Base Aérea nº 5. Os dados supracitados referem-se a temperaturas máximas e mínimas, humidade relativa e número de horas de insolação <sup>[41]</sup> em que os registos reportam a uma base diária. Os dados disponibilizados são relativos ao período entre 1 de dezembro de 2006 e 30 de novembro de 2009. O cruzamento entre a variável potência consumida e estas variáveis exógenas torna-se premente, procurando evidenciar-se possíveis efeitos de correlação entre as mesmas. Uma vez que os dados meteorológicos têm uma resolução diária, na análise de correlação foi considerado o consumo diário de energia. As Figuras 4.23 a 4.25 descrevem os correlogramas obtidos entre a variável de consumo energético e diferentes variáveis exógenas durante o período para o qual há coincidência de dados disponíveis (entre 15 de dezembro de 2006 e 30 de novembro de 2009).



**Figura 4.23 – Análise de correlação do consumo diário de energia do segundo estudo de caso com a temperatura máxima e mínima local**

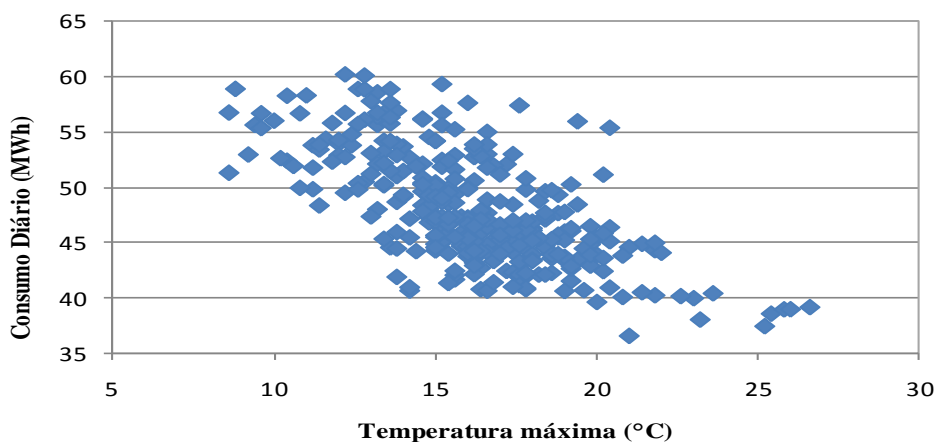
Do correlograma que relaciona o consumo diário e a temperatura máxima pode verificar-se que quando as temperaturas são mais reduzidas, o consumo de energia elétrica tende a aumentar, nomeadamente para fins de aquecimento e eventualmente de iluminação. Esta constatação é também consubstanciada pelo correlograma que confronta o consumo diário com a temperatura mínima, muito embora a correlação do consumo com a temperatura mínima não demonstre ser tão conclusiva, já que um considerável número de registos aparentam distorcer uma linha de tendência <sup>[42]</sup>. Um levantamento local dos equipamentos de uso final para aquecimento poderia

[41] Número de horas de insolação corresponde ao número de horas de sol descoberto.

[42] Registos comumente designados na comunidade científica como “outliers”.

manifestar-se neste estudo como um considerável contributo, para despistar se é de facto a eletricidade o recurso mais comum para esta finalidade (através de radiadores elétricos, sistemas de ar condicionado, piso radiante,...), ou antes o recurso a outras fontes de energia (como sistemas a gás natural ou a propano, recuperadores de calor ou sistemas de energia solar).

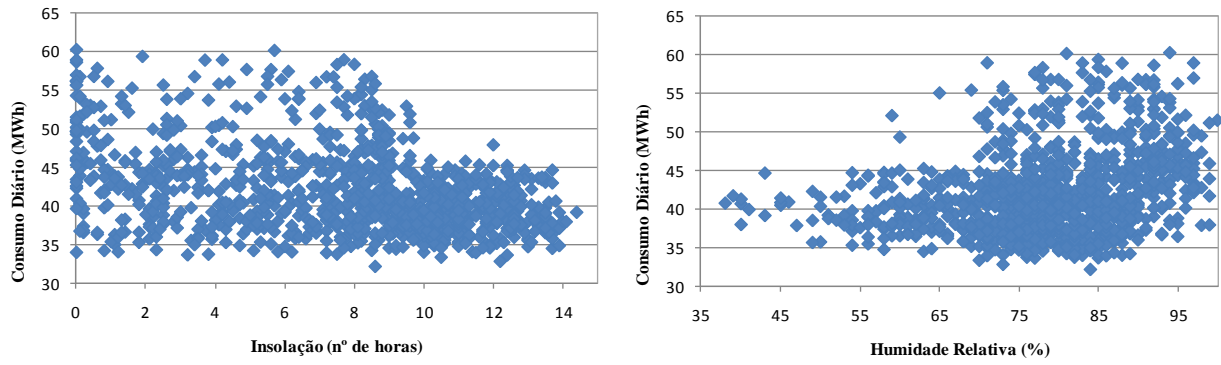
Por outro lado, com temperaturas mais elevadas a tendência típica em zonas desenvolvidas é a de um aumento de consumo de eletricidade que resulta maioritariamente de sistemas de ar condicionado. Para este caso em estudo o aumento de consumo de eletricidade em dias quentes, quando existe, é muito ténue. Esta constatação justifica-se pela proximidade da zona em questão à orla marítima (o que atenua alterações bruscas no registo de temperatura e reduz o desconforto associado a calor seco), bem como pelo reduzido nível de despesas de consumo das famílias residentes (quando comparável com grandes centros urbanos a nível nacional) que compromete o investimento inicial na aquisição de sistemas de ar condicionado e decorrente consumo associado à sua utilização. Se na análise apenas se considerar o período de época fria parece haver uma maior correlação entre consumo diário e a temperatura. Para tal, e usando apenas os meses mais frios do ano (novembro, dezembro, janeiro e fevereiro), podem verificar-se na Figura 4.24 os resultados obtidos.



**Figura 4.24 – Análise de correlação do consumo diário de energia do segundo estudo de caso com a temperatura máxima local para épocas frias**

O valor do coeficiente de correlação linear entre o consumo e a temperatura máxima passa de -0,533 quando se consideram todos os registos disponíveis para -0,671 quando passam a ser considerados apenas os referidos meses de época fria.

No que respeita à humidade relativa e ao nível de insolação (veja-se a Figura 4.25), os valores de correlação linear com a variável de consumo elétrico anunciam-se reduzidos (de 0,284 para a humidade e de -0,364 para a insolação) e estas duas variáveis apresentam alguns registos em que se identificam erros de medição e/ou de leitura, pelo que, combinando estas duas constatações se sugere o abandono destas variáveis como explicativas do valor do consumo a prever.



**Figura 4.25 - Análise de correlação do consumo diário de energia do segundo estudo de caso com a humidade relativa e o nível de insolação do local**

Em alguns dos modelos de previsão que serão apresentados nos seguintes capítulos, apenas se consideram os valores máximos e mínimos de temperatura diária como variáveis de entrada que pretendem ser, de alguma forma, justificativas do consumo de energia.



## *5. Aplicação de métodos de previsão aos estudos de caso considerados*

Neste capítulo serão apresentados modelos de previsão aplicados aos dois estudos de caso introduzidos no capítulo anterior <sup>[43]</sup>. Os modelos de previsão que serão explorados neste capítulo recorrem a alguns dos métodos já apresentados e descritos no capítulo 3, tratando-se de modelos inspirados em redes neuronais e de modelos baseados em máquinas de suporte vetorial. A escolha das redes neuronais deve-se essencialmente à grande aceitabilidade que têm atualmente na comunidade científica, o que sugere também uma análise pormenorizada do efeito da inclusão do diagrama de carga sintetizado resultante da informação de perfis de consumo no conjunto de variáveis determinantes para o processo. No caso das máquinas de suporte vetorial a escolha é justificada pelo carácter atual desta metodologia, pelos resultados satisfatórios que têm permitido quer no campo da classificação quer no campo da regressão, pela pertinência de uma comparação desta metodologia de previsão com as já sobejamente validadas redes neuronais e pela necessidade de ter um método complementar de validar a eventual importância que o diagrama de carga sintetizado tem na previsão.

### **5.1. Uso de redes neuronais em previsões de consumos de energia elétrica**

Os estudos que foram implementados neste âmbito incidiram no uso de redes de propagação multicamada, tendo-se considerado apenas uma camada intermédia com um número variável de neurónios. As entradas e saídas foram numa primeira fase normalizadas, de modo a terem uma média nula e uma variância unitária, acautelando escalas da mesma ordem de grandeza. Como foi abordado na secção 3.1.3, outras estratégias poderiam ser tomadas nesta fase de normalização, tal como uma abordagem baseada nos valores máximos e mínimos da série. No entanto, para que a normalização não ficasse dependente de eventuais registos atípicos da série (resultantes de pontas ou valores mínimos inesperados) que poderiam ocorrer, optou-se por se

---

[43] Os modelos de previsão são parcialmente baseados em Sousa *et al.* (2013) e Sousa *et al.* (2012).

considerar a média e o desvio padrão das séries como indicadores para a normalização. Desta forma, a normalização é estabelecida tendo em conta todos os registos identificados nas séries e não somente dois valores (máximo e mínimo) que poderiam distorcer características intrínsecas das séries <sup>[44]</sup>.

Uma única camada intermédia permite evitar estruturas muito pesadas com demasiados pesos e desvios que tornam o processo de treino consideravelmente dispendioso a nível computacional. Simultaneamente acautelam-se estruturas demasiado complexas que tornariam os modelos francamente dependentes dos processos de inicialização aleatória de pesos e desvios, evitando análises matemáticas mais intrincadas dos modelos gerados. Uma comparação entre diferentes arquiteturas revelou-se uma abordagem aconselhável, experimentando um número variável de neurónios e aferindo o efeito que essa alteração suscitaria nos erros de previsão, quer para os dados de treino, quer para os dados de teste. Exploraram-se redes com um número variável entre 3 e 15 neurónios na camada intermédia, com base na ideia de que um número inferior tornaria a previsão estabelecida pela rede pouco sensível às reais variações sentidas na série temporal em causa e um número superior a 15 neurónios fixaria demasiado a rede à série temporal disponível no período de treino (podendo mesmo passar à rede alguma informação acerca do ruído próprio existente na série) (Hippert *et al.*, 2001). Por influência de muitos trabalhos científicos encontrados na área da previsão de consumos (Chen *et al.*, 2001a, Hippert *et al.*, 2001, Ramezani *et al.*, 2005, Santos *et al.*, 2006, Xia *et al.*, 2010) foram consideradas funções de ativação baseadas em tangente hiperbólica na camada intermédia e funções lineares na camada de saída. A escolha da função tangente hiperbólica permite adaptar a rede a eventuais não linearidades entre entradas e saída(s), com uma aptidão para tornar o algoritmo de treino mais célere <sup>[45]</sup>. Já para a camada de saída, é recomendável a escolha da função linear sempre que as saídas sejam variáveis contínuas e não se conheçam *a priori* os seus valores limite.

As redes projetadas pretendem prever de uma só vez o conjunto de 24 registos horários de potência ativa para o dia seguinte, o que configura uma arquitetura com 24 saídas. O algoritmo de treino que se revelou mais apropriado foi o algoritmo de Levenberg-Marquardt, após comparação com outras variantes, tendo sido comprovada a sua eficácia no que respeita ao tempo consumido e à desejada convergência. De forma a limitar-se o processo associado ao treino, podem ser impostos diferentes critérios já apresentados no Capítulo 3, muito embora de

---

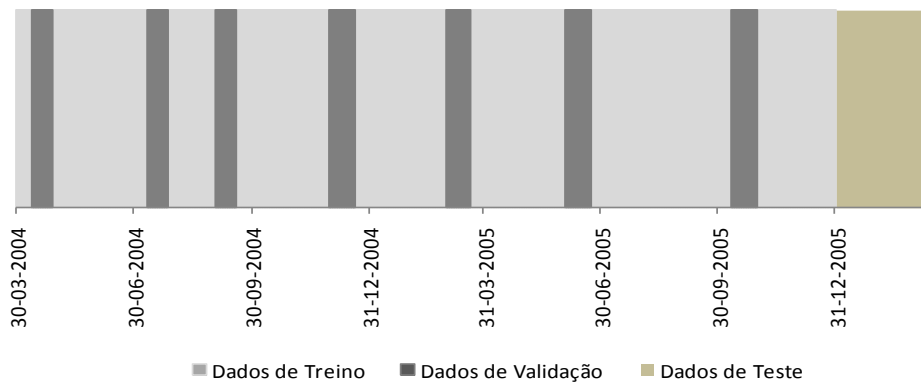
[44] Com esta normalização constatou-se que para o primeiro estudo de caso os dados normalizados de entrada ficaram compreendidos num intervalo [-2.82;5.86], os de saída num intervalo [-2.48;5.66] e para o segundo estudo de caso as entradas centraram-se numa gama [-2.59;4.86] e os de saída numa gama [-2.23;4.87].

[45] Uma vez que os parâmetros de entrada e de saída são normalizados, podendo tomar valores positivos e negativos, esta função de ativação torna-se mais ajustada para tornar o treino mais rápido.

entre estes critérios a estratégia de validação cruzada tende a ser unanimemente aceite para interrupção do período de treino quando se tende a verificar *overfitting* (perda de generalização associada à contínua melhoria de previsão nos dados associados ao treino, sem que esta melhoria se espelhe para períodos fora do conjunto de treino).

A validação cruzada adotada para os dois estudos de caso analisados pressupôs a separação do conjunto de dados em dados de treino, dados de validação e dados de teste. Para tornar o processo o mais realista possível, foi encarado como período de teste um subconjunto de dados que compreendiam o final do intervalo de dados completo disponível, já que desta forma a precisão das estimativas e a capacidade de generalização da rede neuronal é avaliada com os dados mais atuais e naturalmente mais próximos dos dados futuros que a rede pretende prever. No que respeita aos restantes subconjuntos de dados (dados de treino e dados de validação) a segmentação foi feita de forma intercalada para que não se desse um foco demasiado significativo apenas num determinado intervalo de tempo, correndo-se o risco de comprometer a validação e interromper prematuramente o treino ou, pelo contrário, necessitar de demasiadas iterações de treino pela natureza de dados de validação ser muito idêntica à característica dos dados de treino. Adicionalmente procurou-se definir intervalos com idênticas durações e evitar como dados de validação aqueles que dependeriam de dias feriados, épocas festivas ou outros eventos ocasionais, que pudessem distorcer os normais e expectáveis hábitos de consumo. A segmentação dos dados teve sempre uma divisão de 80% dos dados para conjunto de treino e de 20% dos dados para conjunto de validação.

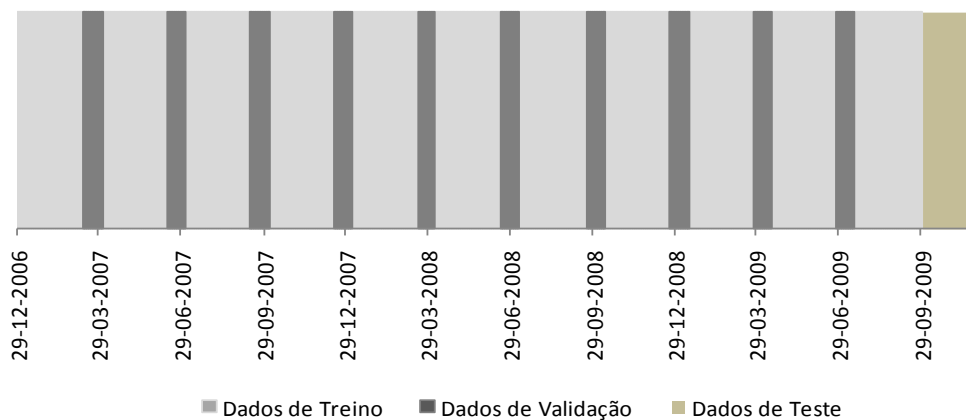
Para o primeiro estudo de caso os dados considerados reportaram ao período entre o dia 16 de março de 2004 e o dia 16 de março de 2006. Como tal, consideraram-se para período de teste os dados de consumo de 2006 (sensivelmente dois meses e meio) e como conjunto para treino e validação o período remanescente (aproximadamente 1 ano, 9 meses e 15 dias). Deste último conjunto, extraíram-se 7 sub-períodos com duração entre 16 e 21 dias para compor a amostra que irá ser determinante para interromper o processo de aprendizagem. Na Figura 5.1 pode verificar-se a segmentação usada, repartindo-se os dados disponíveis nos subconjuntos de treino, de validação e de teste.



**Figura 5.1 – Ilustração da estratégia de segmentação dos dados disponíveis no primeiro estudo de caso**

Esta segmentação serviu como estratégia base que terá também permitido alguma confiabilidade na comparação que se pretende entre modelos baseados em redes neuronais e baseados em máquinas de suporte vetorial. Outro aspeto importante é que, como se pretende aferir a importância de determinadas variáveis de entrada que serão introduzidas face a modelos já atualmente utilizados, reforça-se a necessidade de se utilizarem conjuntos de dados de treino, validação e de teste que sejam coincidentes entre simulações efetuadas. Não obstante, foram testadas outras segmentações <sup>[46]</sup> que interferem diretamente no processo de treino em comparação com esta estratégia base, com a apresentação de alguns resultados no Anexo A.

No que respeita ao segundo estudo de caso, os dados considerados reportaram ao período entre o dia 15 de dezembro de 2006 e o dia 30 de novembro de 2009. Mais uma vez foram considerados os últimos registos como dados de teste, neste caso os dois últimos meses disponíveis, entre o dia 1 de outubro de 2009 e o dia 30 de novembro de 2009. À semelhança do que sucedeu com o primeiro caso em estudo, consideraram-se dados de validação intercalados com dados de treino, de acordo com o que se pretende ilustrar na Figura 5.2. Os subconjuntos de dados de validação que foram intercalados com os dados de treino tiveram durações não superiores a 21 dias.



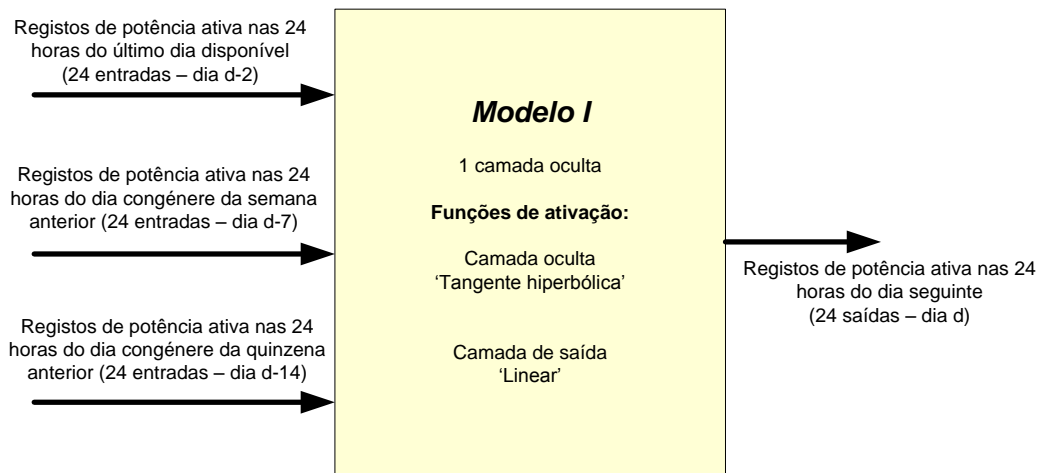
**Figura 5.2 – Ilustração da estratégia de segmentação dos dados disponíveis no segundo estudo de caso**

[46] Partindo sempre do pressuposto que dos dados usados para treino e validação, 80% destinaram-se ao treino e 20% à validação cruzada.



Para o segundo caso em estudo também foram testadas outras divisões de dados, que serão apresentadas e comentadas em maior pormenor no Anexo A.

Os modelos criados para a previsão dos registos horários de potência ativa tiveram como base a escolha de dias similares que tivessem alguma influência direta no dia a prever. Nesse sentido, considerou-se um primeiro modelo inspirado em outros trabalhos similares (Fidalgo e Matos, 2007, Fidalgo e Peças Lopes, 2005, Lamedica *et al.*, 1996, Ramezani *et al.*, 2005) com 72 entradas, contemplando os registos horários do último dia disponível, do dia congénere da semana anterior e o do dia congénere da quinzena anterior. O último dia disponível refere-se a dois dias anteriores ao da previsão, já que o dia atual não estaria inteiramente conhecido quando se pretendesse proceder à previsão do dia seguinte. Das análises de correlação efetuadas nas secções 4.1.1 e 4.2.1 aos dois estudos de caso verificou-se que a inclusão de informação relativa à semana anterior e quinzena anterior se revela de franca utilidade, já que se confirma que o dia a prever tem um padrão de consumo próximo de um dia de semana idêntico ocorrido recentemente. Já em relação ao vetor relativo ao último dia disponível a sua inclusão poderá tornar-se mais discutível, sobretudo no primeiro estudo de caso que apresenta uma correlação menos relevante, como se pode confirmar nos resultados da Tabela 4.1. O efeito de quebra de consumo durante os fins de semana que caracteriza esta série temporal condiciona a autocorrelação evidenciada quando se impõem atrasos de 48 horas. No entanto, como se irá demonstrar posteriormente, podem existir relações não lineares entre as entradas e as saídas que uma simples análise de correlação linear não permite despistar. Nesse sentido, foi adotada uma estratégia de seleção de parâmetros explicativos baseada na avaliação de parâmetros da rede após um período de treino bem-sucedido, por forma a avaliarem-se os potenciais benefícios das entradas escolhidas. Esta análise será descrita em maior detalhe na secção 5.1.1. No segundo estudo de caso o índice de correlação relativo a atrasos de 48 horas apresenta-se mais em linha com correlações identificadas com outros atrasos impostos na série temporal (como se pode confirmar na Tabela 4.3), devido essencialmente ao padrão de consumo muito homogêneo quando comparados dias úteis e dias de fins de semana. O Modelo I pretende ser mais facilmente sistematizado com a descrição da Figura 5.3.



**Figura 5.3 – Modelo I inspirado em redes neurais para a previsão de consumos do dia seguinte**

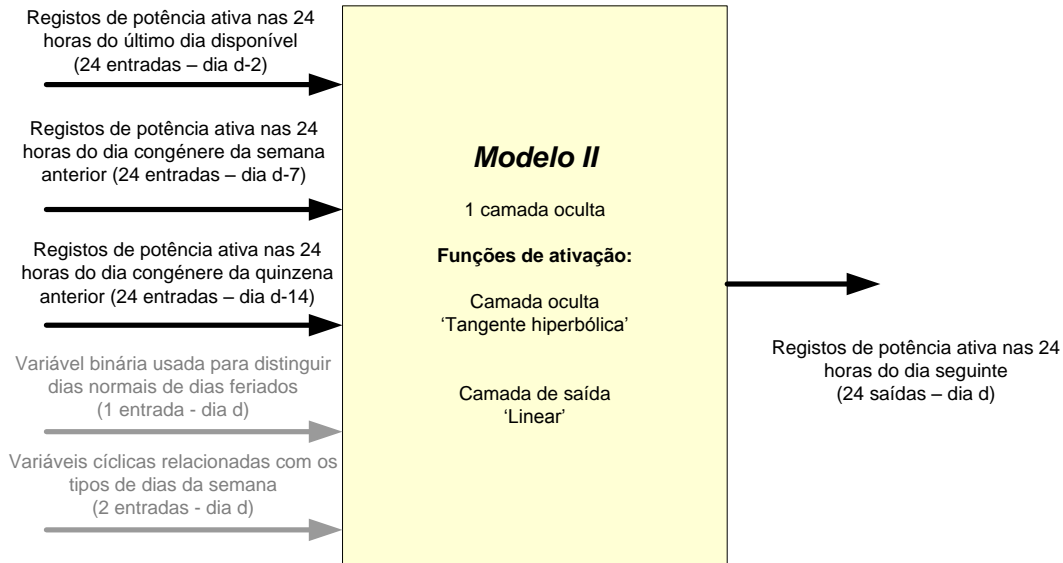
Este modelo multivariável recorre deste modo a 72 registos como entradas para estimar 24 registos de saída. A título exemplificativo, no caso de se pretender prever os registos horários de potência ativa de uma quinta-feira, 15 de setembro de 2005 (estimados durante a quarta-feira, 14 de setembro de 2005 e sem que este dia seja totalmente conhecido no que respeita ao consumo de energia para todas as horas), o modelo recorre aos registos horários de potência identificados na terça-feira, 13 de setembro de 2005, na quinta-feira, 8 de setembro de 2005 e na quinta-feira, 1 de setembro de 2005.

O segundo modelo criado procura contornar alguns problemas identificados na previsão estabelecida com o Modelo I, nomeadamente associados aos dias feriados. Nesse sentido, para além das 72 entradas a que o Modelo I recorre, são utilizadas mais três entradas:

- uma entrada binária para distinguir dias normais de dias feriados (a “0” quando se trata de prever um dia normal, e a “1” quando se pretende prever um dia “feriado”). Esta estratégia visa destacar a presença de dias feriados e dotar a rede de compreensão do efeito que este tipo de dias causa na previsão. Ao serem previstos dados futuros a reportarem a dias feriados, já é expectável que tenha havido alguma aprendizagem para lidar com este tipo de dias;

- duas entradas cíclicas para caracterizar o tipo de dia da semana. As funções periódicas *sen* e *cos* são usadas para garantir que cada dia tenha uma representação unívoca. Esta estratégia foi já descrita na secção 3.1.4.

Uma análise sumária do Modelo II pode ser verificada na Figura 5.4.

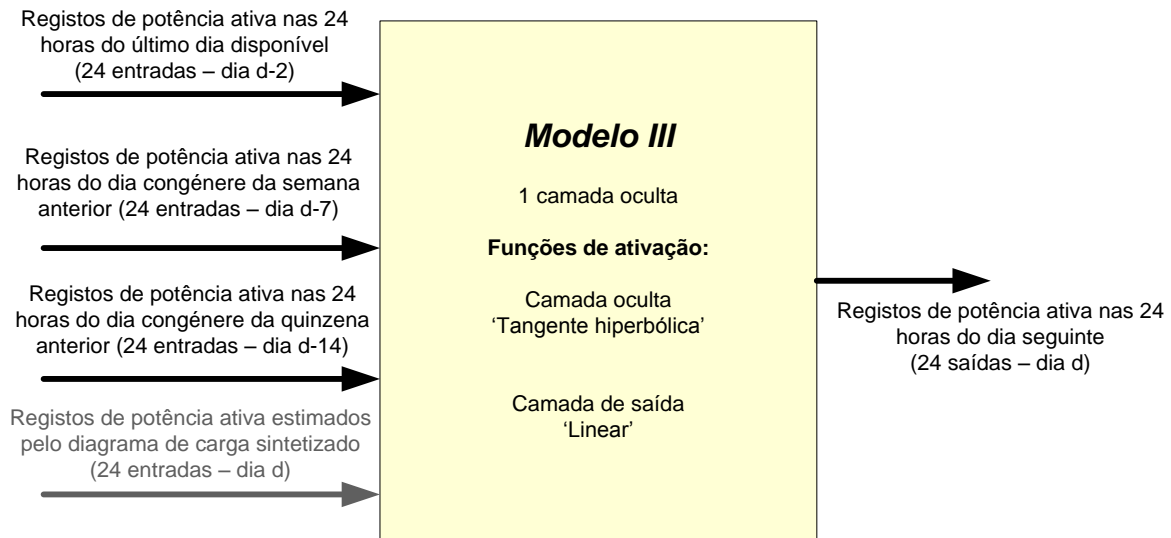


**Figura 5.4 – Modelo II inspirado em redes neurais para a previsão de consumos do dia seguinte**

Uma opção variante a este Modelo II foi ainda avaliada, tendo como base o estudo identificado em Fidalgo e Matos (2007), no qual se considerou um tratamento dos dados das entradas relativas ao passado recente (último dia disponível, dia congénere da semana anterior e dia congénere da quinzena anterior), de modo a que a previsão dos dias seguintes aos feriados não dependa destes dias, mas de dias vizinhos destes. Outro tratamento proposto diz respeito a pontes entre fins de semana e feriados, para os quais se sugere aplicar um fator de redução de consumo, tendo como base a relação média de consumo destes dias face aos seus dias vizinhos. O estudo em causa identifica uma relação da ordem dos 90% entre o consumo identificado em pontes relativamente ao consumo nos seus dias vizinhos. Para o primeiro estudo de caso analisado neste trabalho verificou-se também que em termos médios o consumo nos dias relativos a pontes é da ordem dos 90,52% face ao consumo verificado em dias vizinhos. Em todo o caso, este efeito de redução de consumo é mais denunciado em épocas de mudança de estação como os meses de fevereiro e de outubro, provando-se que a relação de consumo se torna muito dependente da época do ano em questão e também do tipo de feriado que se esteja a considerar. Por outro lado, o efeito de redução de consumo não deve ser aplicado a todas as entradas, já que para determinadas horas a diferença de consumo entre estes dias atípicos e dias vizinhos não é significativa, tornando-se necessária a identificação das horas para as quais se deve aplicar o fator de redução de consumo (tipicamente as horas de maior consumo e com atividade económica associada). Por estes motivos e uma vez que a informação do padrão histórico de comportamento de diferentes grupos de consumidores (proveniente dos perfis de consumo) já acarreta um tratamento de dias atípicos, esta informação acerca dos dias especiais foi explorada

na adoção do terceiro modelo o que em termos práticos tornou a abordagem mais simples e conduziu também a resultados mais satisfatórios.

Por conseguinte, o terceiro modelo de previsão baseado em redes neuronais explora novamente as mesmas entradas já usadas no Modelo I, adicionando como entradas os 24 registos horários de potência ativa obtidos a partir do diagrama de carga sintetizado para o dia a prever. O modelo recorre assim a 96 entradas para estimar 24 saídas, tal como pretende ser evidenciado na Figura 5.5.



**Figura 5.5 – Modelo III inspirado em redes neuronais para a previsão de consumos do dia seguinte**

Diversas análises de correlação entre o vetor de saída e os diferentes vetores de entrada foram averiguadas com elevado detalhe no capítulo 4, procurando antever a importância relativa de cada entrada na estimação das saídas. Estes valores de correlação resultaram simplesmente de uma análise de autocorrelação (usando diferentes atrasos na série temporal), ou de uma correlação cruzada entre os diagramas de carga reais e os diagramas de carga sintetizados. Nestes casos abordados, toda a série temporal disponível foi considerada nos cálculos de correlação.

É atualmente reconhecido que no caso de se optar por avaliar a importância de diferentes entradas aplicadas apenas a partir de análise de correlação, existirá um risco associado de descartar variáveis que não manifestem uma significativa correlação linear com a carga a prever, mas que podem estar fortemente relacionados de forma não linear com a mesma (Fidalgo, 2001, Hippert *et al.*, 2001). Nesse sentido, uma estratégia de seleção de parâmetros explicativos baseada na avaliação dos parâmetros da rede neuronal após esta ter sido efetivamente treinada, revelou-se um procedimento profícuo para avaliar os potenciais benefícios do recurso às variáveis de entrada adotadas. A análise da derivada parcial de cada saída individual em ordem a cada entrada usada, comumente designada como análise de sensibilidades, é usada no âmbito

deste trabalho de modo a aceder à relevância das entradas escolhidas nos diferentes modelos de rede neuronal implementados.

### 5.1.1. Análise de sensibilidades das redes neuronais baseadas em derivadas parciais

Uma primeira abordagem à estrutura de rede neuronal que contemple já um período de treino pode sugerir uma análise dos pesos entre entradas e os neurónios da camada intermédia. Após a rede ter sido treinada, com a subsequente adaptação “hipoteticamente” otimizada <sup>[47]</sup> dos pesos e desvios das ligações, uma análise dos pesos admitidos nas ligações evidenciadas parece sugerir uma importância relativa associada a cada entrada. Esta análise apesar de se tornar interessante pela sua simplicidade não será suficientemente robusta, já que podem existir relações não lineares existentes na arquitetura da rede e um aparente peso elevado associado a uma determinada entrada ser amortecido por uma zona de saturação da função de ativação. Em oposição, pesos modestos para uma determinada entrada e neurónios da camada intermédia podem não significar liminarmente que a entrada em questão terá um modesto contributo na saída, já que este contributo poderá vir a ser ampliado pelos pesos entre a camada intermédia e a camada de saída. Um método alternativo para procurar interpretar a pertinência das entradas usadas pode passar por uma análise de sensibilidade da rede, que permita perspetivar a taxa de atualização local da(s) saída(s) em relação a uma entrada individual, considerando que as restantes entradas permanecem inalteradas. A abordagem centra-se no cálculo de derivadas parciais da(s) saída(s) em função das entradas, pelo que é necessário introduzir o conceito matemático no qual esta metodologia se baseia. Como foi abordado na secção anterior, as redes neuronais utilizadas consideram apenas uma camada intermédia, com recurso a uma função de ativação tangente hiperbólica na camada oculta e a uma função linear na camada de saída. Usando o exemplo do Modelo III que pretende estimar um perfil de consumo para as 24 horas do dia seguinte recorrendo a 96 entradas e a uma camada oculta, no caso de se considerarem 3 neurónios nesta camada intermédia, a primeira saída (associada à hora 0) pode ser calculada de acordo com a Equação 5-1.

$$y'_1 = w'_{11} \cdot \tanh(w_{11} \cdot x_1 + w_{21} \cdot x_2 + \dots + w_{961} \cdot x_{96} + b_1) + w'_{21} \cdot \tanh(w_{12} \cdot x_1 + w_{22} \cdot x_2 + \dots + w_{962} \cdot x_{96} + b_2) + w'_{31} \cdot \tanh(w_{13} \cdot x_1 + w_{23} \cdot x_2 + \dots + w_{963} \cdot x_{96} + b_3) + b'_1 \quad (5-1)$$

com:

---

[47] “Hipoteticamente” otimizada, já que é impossível na prática explorar todas as combinações que os pesos e desvios das ligações possam tomar. O treino ajusta a rede para bons cenários de previsão em dados de treino e dados de teste, muito embora é inviável aferir se se tratam de mínimos locais ou de mínimos globais.

$w'_{11}$  – peso entre o neurónio 1 na camada oculta e a saída 1;

$b'_1$  – desvio encontrado na saída 1;

$w_{11}$  – peso entre a entrada 1 e o neurónio 1 na camada oculta;

$b_1$  – desvio encontrado no neurónio 1;

O conceito de derivada parcial implementado com o pretexto de uma análise de sensibilidade pretende, por definição, dar uma perspetiva da taxa de variação local da(s) saída(s) relativamente a cada entrada em particular, mantendo as restantes entradas fixas (Sarle, 2000). Uma vez que os modelos de redes neuronais utilizados recorrem a funções de ativação não lineares nas diferentes camadas, uma simples comparação de pesos não seria totalmente rigorosa. Em redes multicamada, a derivada parcial tem a mesma interpretação de um peso num simples modelo linear, mas ao invés de se estender a análise a todo espaço de entrada, apenas é focada a vizinhança em torno do ponto de entrada a ser considerado. A função de ativação considerada na camada intermédia é a função tangente hiperbólica, caracterizada matematicamente através da Equação 5-2.

$$\tanh(x) = \frac{2}{1+e^{-2x}} - 1 \quad (5-2)$$

A derivada da função tangente hiperbólica pode ser expressa através da Equação 5-3.

$$\frac{\partial \tanh(x)}{\partial x} = 1 - \tanh^2(x) \quad (5-3)$$

No caso da camada de saída a função de ativação usada é a função linear, tornando-se o cálculo matemático muito mais simplificado. A Equação 5-4 descreve matematicamente a função linear e a Equação 5-5 sintetiza a simplificação da sua derivada parcial.

$$\text{linear}(x) = x \quad (5-4)$$

$$\frac{\partial \text{linear}(x)}{\partial x} = 1 \quad (5-5)$$

Determinadas as implícitas expressões derivadas das funções de ativação contempladas nas arquiteturas de rede neuronal implementadas, torna-se possível calcular a derivada parcial de uma qualquer saída em função de determinada entrada. De modo a poder aferir-se a importância relativa de uma entrada  $x_I$  na saída  $y_I$  quando se recorre a um modelo com 96 entradas distintas e a uma camada intermédia com 3 neurónios, a derivada parcial pode ser calculada através da Equação 5-6 para cada padrão de dados disponível do conjunto de treino.

$$\left. \frac{\partial y'_l}{\partial x_l} \right|_{\text{padrão } p} = w'_{1l} \cdot w_{1l} \cdot \left[ I - \tanh^2(w_{1l} \cdot x_l + \dots + w_{96l} \cdot x_{96} + b_l) \right] + w'_{2l} \cdot w_{12} \cdot \left[ I - \tanh^2(w_{12} \cdot x_l + \dots + w_{962} \cdot x_{96} + b_2) \right] + w'_{3l} \cdot w_{13} \cdot \left[ I - \tanh^2(w_{13} \cdot x_l + \dots + w_{963} \cdot x_{96} + b_3) \right] \quad (5-6)$$

Este procedimento deve ser repetido para todos os padrões de treino disponíveis, de modo a poder-se proceder ao cálculo da média dos valores absolutos das derivadas parciais (Fidalgo, 2001) de acordo com a Equação 5-7.

$$\frac{\partial y'_l}{\partial x_l} = \frac{\sum_{\text{padrão}=1}^{\text{número total de padrões}} \left| \frac{\partial y'_l}{\partial x_l} \right|_{\text{padrão } p}}{\text{número total de padrões}} \quad (5-7)$$

A adoção dos valores absolutos das diferentes derivadas parciais por padrão prende-se com o risco de determinadas sensibilidades (derivadas parciais) poderem ser de elevada magnitude mas de sinais contrários, o que poderia suscitar um valor médio muito reduzido não espelhando a real importância que determinada entrada pode assumir no processo de previsão.

## 5.2. Uso de máquinas de suporte vetorial em previsões de consumo de eletricidade

O recurso a máquinas de suporte vetorial para efeitos de previsão de consumos de eletricidade apoiou-se na meta-heurística denominada *Simulated Annealing*, como aliás já havia sido relatado no final do Capítulo 3. Nesse sentido, o primeiro estágio desta subsecção do documento centra-se na descrição detalhada deste algoritmo de otimização, que terá sido também utilizado no Capítulo 4 mas com um propósito diferente, tendo sido remetida uma exposição pormenorizada do algoritmo para o presente capítulo. O propósito em causa referia-se à aferição da contribuição de diferentes classes de consumidores abastecidos pela saída de Monte Real da subestação de distribuição de Ortigosa.

### 5.2.1. Descrição da meta-heurística Simulated Annealing

*Simulated Annealing* é uma das meta-heurísticas mais antigas com uma estratégia explícita de contornar o problema associado a mínimos locais (de Vicente *et al.*, 2003) (Mathworks, 2009) (Metaheuristics, 2012). A ideia subjacente à técnica é permitir a aceitação de soluções que eventualmente deterioram a função objetivo, de modo a distanciar-se de mínimos locais. A

inspiração do método proveio do setor metalúrgico, já que a técnica envolve um estágio similar a um processo de aquecimento, seguido de um arrefecimento controlado a que determinados materiais são sujeitos, por forma a aumentarem a dimensão dos seus cristais e garantir a redução de defeitos. O aquecimento causa um estado de maior energia interna, com os átomos a “descolarem” de eventuais mínimos locais e, posteriormente o processo de arrefecimento pretende explorar configurações melhores, com menor energia interna do que o estado antecessor.

Deste modo, uma solução atual é substituída por uma nova solução numa vizinhança aleatória e a extensão da procura é baseada numa distribuição probabilística associada a uma escala proporcional ao conceito de temperatura do processo (Mathworks, 2009). O algoritmo aceita todas as novas soluções que resultem num decréscimo da função objetivo, tendo também a particularidade de aceitar, segundo uma determinada probabilidade, soluções que degradem a função objetivo. Desta forma, é possível o algoritmo contornar mínimos locais e explorar de forma mais global o espaço de soluções. O processo vai-se desenrolando com a temperatura a decair, de modo a que a extensão da procura vá também diminuindo e haja uma melhor aproximação e análise de mínimos da função. Todavia, é sempre difícil distinguir-se quando se trata de mínimo global ou de mínimos locais da função, por isso é frequentemente requerida a repetição do processo. O processo denominado de *reannealing* deve ser encarado como uma estratégia complementar de procura global no espaço de soluções, já que a temperatura em cada dimensão (para cada variável) é aumentada de acordo com as diferentes sensibilidades manifestadas pelas variáveis à “temperatura” (Mathworks, 2009), permitindo uma maior oportunidade de aceitar soluções dispersas das soluções atuais no espaço dimensional de soluções.

Uma descrição sumária do algoritmo é apresentada na Figura 5.6.



---

Definição de uma solução inicial  $X_0$  de forma aleatória;  
 Definição da temperatura inicial  $T_0$ ;  
 Cálculo da função objetivo  $f(X_0)$  para a solução inicial  $X_0$ ;  
 A primeira combinação será a solução inicial adotada  $\rightarrow X_t = X_0$ ;  
 A primeira função objetivo deriva da primeira solução adotada  $\rightarrow f(X_t) = f(X_0)$ ;  
 enquanto as condições de paragem não forem satisfeitas  
     escolha de uma solução aleatória vizinha  $X_{t+1}$ ;  
     cálculo da função objetivo para a nova solução  $f(X_{t+1})$ ;  
 se  
      $f(X_{t+1}) < f(X_t)$   
     a solução  $X_{t+1}$  passa a ser a nova solução  $X_t$ ;  
 ou se  
      $n^\circ$  aleatório no intervalo  $[0 ; 1] < 1 / [1 + \exp [f(X_{t+1}) - f(X_t) / T]]$   
     a solução  $X_{t+1}$  passa a ser a nova solução  $X_t$ ;  
 senão  
     é mantida a solução  $X_t$ ;  
      $T_{t+1} = T_0 * 0,95^{N^\circ \text{ de combinações avaliadas ;}}$   
 Fim de ciclo.

---

**Figura 5.6 – Descrição sumária do algoritmo associado ao Simulated Annealing**

Nesta descrição pode verificar-se que a temperatura do processo vai decaindo (enquanto não houver nova operação de *reannealing*) à medida que vai sendo avaliado um maior número de iterações e que é possibilitada a aceitação de soluções que degradem a função objetivo. A probabilidade desta aceitação é tanto maior quanto menor for a diferença entre soluções sucessivas e quando a temperatura do processo é ainda alta. Com valores mais reduzidos de temperatura, o termo  $\exp \frac{f(X_{t+1}) - f(X_t)}{T}$  tende para infinito, tornando cada vez menos provável a geração de um número aleatório que seja inferior à função de teste em causa, logo diminuindo a probabilidade de aceitação da nova solução. Para temperaturas reduzidas (após um número considerável de iterações de busca de soluções), esta estratégia pretende garantir uma fina aproximação, focando um determinado mínimo em estudo.

### 5.2.2. Recurso a Simulated Annealing para permitir a previsão baseada em máquinas de suporte vetorial

De entre os dados disponíveis para os dois estudos de caso adotados, os que foram usados na modelação do processo dizem respeito às amostras de treino e de validação. A amostra de treino foi então utilizada no processo de mapeamento do espaço de entrada num espaço dimensional superior, cumprindo o problema de otimização com restrições de acordo com formulação

primal/dual próprias do processo de treino de modelos de suporte vetorial. Para se evitar o efeito de *overfitting* é usada a amostra de validação que permite verificar a capacidade de generalização desta metodologia, sendo também tal amostra providencial neste estudo para identificar a combinação mais apropriada dos parâmetros das máquinas de suporte vetorial e que se entendeu corresponder ao valor mínimo do indicador de erro considerado. A função objetivo do processo de otimização analisa assim o erro percentual absoluto médio (*EPAM*) verificado na amostra de validação. As variáveis que são sujeitas ao método combinatório *Simulated Annealing* são as entradas dos modelos baseados em máquinas de suporte vetorial (Modelos *MSV*), já comentados anteriormente. O objetivo primordial é encontrar a melhor combinação de parâmetros de máquinas de suporte vetorial (para os diferentes modelos horários) que resulta na previsão mais rigorosa no conjunto de dados de validação. A Figura 5.7 dá uma perspetiva resumida do processo, com os termos  $\gamma_n$ ,  $\epsilon_n$  e  $C_n$  a corresponderem aos termos que são encontrados a partir da meta-heurística *Simulated Annealing*.

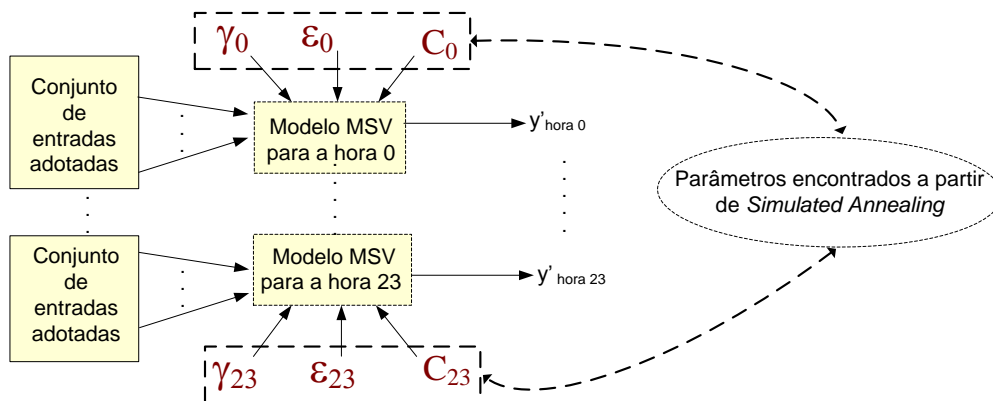


Figura 5.7 – Descrição sumária da metodologia de previsão baseada em máquinas de suporte vetorial

A metodologia idealizada visa uma aplicação semelhante ao recurso às redes neuronais artificiais, garantindo que sejam disponibilizadas as saídas relativas às 24 horas do dia seguinte. Por forma a tornar o processo comparável, são consideradas as mesmas entradas associadas ao Modelo III (que inclui a informação resultante do DC sintetizado), o mesmo critério de normalização dos dados e também a mesma estratégia de divisão dos dados. As máquinas de suporte vetorial foram adotadas complementarmente com o algoritmo de otimização *Simulated Annealing* por forma a identificar os parâmetros  $\epsilon$  (margem de tolerância do erro),  $C$  (o fator de penalização para as amostras existentes fora da margem de tolerância) e  $\gamma$  (constante usada na função de base radial). Após alguns testes efetuados e alguma análise crítica dos mesmos, optou-se por definir valores máximos para estes parâmetros de modo a limitar o espaço de

procura ocupado pelas variáveis contínuas. Na literatura é comum adotar-se uma tolerância ao erro de 0,5 (considerando as variáveis de entrada e de saída normalizadas) (Abbas e Arif, 2006, Chen *et al.*, 2004), pelo que este valor foi considerado como máximo deste parâmetro. Para o parâmetro  $\gamma$  foi seguido um valor máximo de 2 já que tendo diretamente influência na função de base radial, torna-se determinante para captar eventuais não linearidades. Os estudos prévios permitiram confirmar que valores mais elevados do que este máximo imposto aumentam a complexidade associada à distribuição de dados no espaço dimensional. Um limite para o fator de penalização foi também analisado após um número considerável de simulações, tendo sido reconhecido que valores de C superiores a 15 não melhoram a precisão dos modelos e podem aumentar consideravelmente o espaço de soluções possíveis. Estes valores máximos dos parâmetros foram impostos com a perceção que uma apropriada combinação destes parâmetros é o que se torna primordial identificar, ao invés de focar em valores ótimos para os parâmetros individuais (Guo, 2009). Os valores máximos estão sintetizados na Tabela 5.1.

**Tabela 5.1 – Valores máximos por parâmetro nas máquinas de suporte vetorial**

Parâmetro	Descrição	Valor máximo
$\varepsilon$	Margem de Tolerância	0,5
$\gamma$	Constante adotada na função de base radial	2
C	Fator de penalização para registos fora da margem de tolerância	15

Um interface de um software LIBSVM 2.89 disponibilizado online <sup>[48]</sup> para Matlab foi usado com o propósito de testar as máquinas de suporte vetorial no âmbito da previsão (Chang e Lin, 2011), tendo-se também recorrido em simultâneo e em paralelo à versão 2.2 de Genetic Algorithm and Direct Search Toolbox disponível na versão R2007b do Matlab para permitir o processo de otimização baseado em *Simulated Annealing*. A integração destas duas aplicações foi conseguida, tendo em consideração que diferentes combinações de parâmetros MSV devem ser analisadas (modelando a regressão baseada em suporte vetorial com recurso ao conjunto de dados de treino) e mantendo a combinação que garante o valor mínimo do erro percentual absoluto médio (*EPAM*) identificado no conjunto de dados de validação.

Após a avaliação de diferentes combinações de parâmetros MSV e com uma análise de diferentes efeitos causados na precisão associada à previsão dos dados de validação, o processo torna-se decisivo para definir a melhor combinação, sendo esperado que o modelo esteja assim

[48] <http://www.csie.ntu.edu.tw/~cjlin/libsvm>

apto para generalizar de forma adequada para novos dados apresentados (como seja a amostra de dados de teste). É reconhecido que o melhor resultado de previsão identificado no subconjunto de dados de validação não legitima de forma franca a hipótese de se ter identificado também o mínimo indicador de erro em dados de teste, no entanto o objetivo é obter-se maior confiança na estimativa de registos horários de consumo durante o período de teste. O fundamento é similar à estratégia de validação cruzada em redes neuronais, para o qual é sabido que um maior número de épocas de treino irá tender sempre a reduzir o erro no conjunto de treino, no entanto o processo é interrompido quando a previsão no subconjunto de validação tende a degradar-se. O que distingue o processo baseado em máquinas de suporte vetorial da metodologia alternativa é que, testando diferentes combinações de parâmetros das MSV durante o processo associado a *Simulated Annealing*, não se acarreta necessariamente um sucessivo decréscimo do erro na estimativa dos dados de treino, mas ao invés é possibilitada uma procura guiada no espaço de soluções.

Enquanto as redes neuronais podem acomodar múltiplas saídas (neste caso os 24 registos horários de potência ativa referentes ao dia seguinte), a regressão baseada em MSV adotada neste estudo é orientada por 24 modelos distintos (seguindo-se uma estratégia de modelos de previsão horários), cada um individualmente otimizado e caracterizado por diferentes parâmetros. Os diferentes modelos foram analisados seguindo os parâmetros de otimização e critérios apresentados na Tabela 5.2.

**Tabela 5.2 – Parâmetros e critérios associados ao processo de *Simulated Annealing***

Parâmetro	Valor / Função adotada
Temperatura inicial	1000 °
Intervalo entre operações de <i>Reannealing</i>	50 iterações
Número máximo de iterações	750 iterações
Valor limiar para ser comparado com a variação média na função objetivo	$1 \cdot 10^{-6}$
Fator de decréscimo da temperatura	Temperatura inicial $\cdot 0,95^{N^{\circ} \text{comb.} (*)}$
Critério associado a uma eventual aceitação de pior solução	$\frac{I}{I + \exp\left[\frac{f(X_{t+1}) - f(X_t)}{T}\right]}$

(\*) – N° comb. corresponde ao número de combinações já avaliadas

Após a perceção do tempo médio associado ao treino de cada modelo, foi imposto um número máximo de 750 iterações de modo a acautelar que, caso seja necessário, é possível ter um treino efetivo de todos os modelos para prever o perfil de carga do dia seguinte num intervalo inferior a 24 horas. De realçar que, uma vez que os modelos horários estão efetivamente treinados de

acordo com a melhor combinação de parâmetros encontrados, o processo de simular novos dados para prever consumos futuros tarda menos do que 2 minutos <sup>[49]</sup>.

### 5.2.3. Adaptação da metodologia para inclusão de estágio associado à seleção de parâmetros explicativos

Em trabalhos similares é consensual que alguns modelos podem apresentar um vasto número de parâmetros explicativos, o que pode possivelmente aumentar desmesuradamente o tempo associado ao processo de aprendizagem e degradar a capacidade de generalização do modelo (Lee, 2009, Tao *et al.*, 2004). Por esse motivo torna-se importante remover os parâmetros de entrada menos relevantes do conjunto de dados através de métodos apropriados de seleção. Todavia, esta escolha pode requerer uma exaustiva procura de todos os possíveis subconjuntos de dados de entrada que é computacionalmente dispendiosa. Como referenciado nos trabalhos de Guajardo *et al.* (2005) e de Tikka (2009), as técnicas de seleção de variáveis podem ser separadas em:

- métodos de filtragem – resultam de um estágio de pré-processamento dos dados disponíveis e é completamente independente do algoritmo de aprendizagem. Uma análise de correlação ou de correlação parcial entre diferentes entradas e/ou entre entradas e saída(s) é um exemplo deste tipo de método;

- métodos *wrapper* <sup>[50]</sup> – baseiam-se na seleção de diferentes subconjuntos de variáveis de entrada com diferentes impactos no algoritmo de aprendizagem. Com esta metodologia torna-se necessária a identificação de um critério para comparar o efeito de diferentes subconjuntos de entradas (neste caso o critério centra-se no indicador EPAM identificado no conjunto de dados de validação) e também uma estratégia de procura para guiar o processo. Uma possibilidade poderá passar por incluir de forma iterativa a variável de entrada com maior impacto, partindo de um conjunto vazio de dados (seleção *forward*) ou, em alternativa, partir de um conjunto de todas as variáveis de entrada e ir eliminando progressivamente a variável menos importante para o processo (seleção *backward*);

- métodos embebidos – incorporam uma seleção de parâmetros como parte do processo de treino, logo a seleção de variáveis de entrada é estabelecida durante a construção do modelo. Uma hipótese pode passar por incluir uma função objetivo que deve ser minimizada e que é

---

[49] Os tempos de computação são relativos a um processador de 1,87 GHz e a 2 GB de memória RAM.

[50] Foi usado o termo em Inglês por não ter sido identificado um termo apropriado na comunidade científica Portuguesa.

composta por dois termos. Um primeiro termo mede a eficiência da aplicação, tipicamente o erro de previsão, enquanto o segundo termo é usado para penalizar o número de entradas requeridas.

Este estudo inclui também uma estratégia para identificar um subconjunto apropriado de entradas, partindo novamente do recurso à otimização. Deste modo, a meta-heurística *Simulated Annealing* é também usada para a identificação do conjunto de entradas mais relevantes para inclusão no processo de previsão, avaliando-se diferentes combinações possíveis de subconjuntos de entrada por meio da função objetivo EPAM no conjunto de validação.

Das 96 entradas que compõem o conjunto total de entradas, foram considerados quatro blocos de 24 horas concordantes com cada vetor de entrada considerado e, conseqüentemente foram gerados aleatoriamente quatro números diferentes a variar entre 0 e  $2^{24}-1$ . Estes números foram seguidamente convertidos para um formato binário e os 24 bits que formam cada string utilizados para identificar cada entrada individual (relacionada com cada vetor de entrada individual) que entra no processo (correspondendo aos bits que se encontrem com valor lógico “1”) e aqueles que são descartados (correspondendo aos bits que se encontrem com valor lógico “0”) em cada iteração. O processo de otimização baseado em *Simulated Annealing* para identificação de três parâmetros MSV passa a ser estendido para procurar mais quatro entradas a variar entre 0 e  $2^{24}-1$ , fazendo com que o espaço de possíveis soluções aumente exponencialmente. Por este motivo o processo de otimização foi alterado para contemplar um número máximo de 5000 iterações. Deve ser salientado que este número máximo de iterações impostas pode ser efetivamente restrito conhecendo-se as dimensões reais do espaço de entradas implicado, mas assume-se como um adequado equilíbrio entre o espaço de procura e o tempo computacional despendido. Após testes efetuados verificou-se mesmo que em alguns modelos horários o processo de otimização foi interrompido antes do limite de 5000 iterações já que o critério baseado na variação média da função objetivo nas últimas iterações (neste caso, nas últimas 350 iterações) era inferior a  $1.10^{-6}$ . Para além desta justificação verifica-se que o risco de usar um número muito residual de combinações, face ao número total de todas as combinações de entradas possíveis, é tolerado já que a função objetivo a ser minimizada está relacionada com o EPAM do conjunto de validação. Deste modo, mesmo que se consiga um bom indicador de previsão em determinado modelo criado, este indicador pode dar resultados menos satisfatórios quando se usam dados de treino e de teste.

### 5.3. Apuramento da importância relativa das diferentes entradas a partir de modelos baseados em máquinas de suporte vetorial

Com os modelos de suporte vetorial já criados, torna-se premente recorrer a estratégias para aferir a importância dos diferentes registos de entrada usados, à semelhança do que foi já descrito nos modelos com redes neuronais. Uma das estratégias pode passar por remover individualmente cada um dos vetores de entrada que são utilizados, procurando quantificar o erro resultante e comparar com o erro calculado no cenário original em que os vetores considerados são todos incluídos. A verificação do(s) vetor(es) de entrada que tende(m) a degradar mais o erro aquando da sua remoção torna-se um indicador da importância que esse conjunto individual de entradas pode ter no processo de previsão, não sendo por isso aconselhável abdicar dessa informação.

Outra estratégia para reforçar a importância de determinados vetores de entrada nos diferentes modelos horários criados pode resultar de uma análise de sensibilidade de cada saída às diferentes entradas usadas (Cao e Tay, 2000). O processo é semelhante à análise de sensibilidades em redes neuronais, recorrendo à derivada parcial, e a ideia subjacente será partir da expressão matemática que permite obter cada saída. Procurando importar o conceito da referida função (já apresentada na Equação 3-18) e tendo por base uma adaptação da mesma à realidade da regressão efetuada neste estudo de caso, pode obter-se a Equação 5-8.

$$y_{\text{padrão}_p} = \sum_{j=1}^{\text{n}^\circ \text{ de vetores de suporte}} \alpha_j \cdot e^{-\gamma \cdot \|\mathbf{x}_p - \mathbf{sv}_j\|^2} + \mathbf{b} \quad (5-8)$$

em que

$\alpha_j$  – corresponde ao coeficiente relativo a cada vetor de suporte;

$\gamma$  – corresponde à constante da função de base radial;

$\mathbf{x}_p$  – corresponde a cada padrão individual de entrada;

$\mathbf{sv}_j$  – corresponde ao vetor de suporte de índice  $j$ ;

$\mathbf{b}$  – corresponde ao desvio (*bias*) a impor.

A referida equação pode vir a ser decomposta na estrutura identificada na Equação 5-9.

$$y_{i|\text{padrão}_p} = \sum_{j=1}^{\text{n}^\circ \text{ de vetores de suporte}} \alpha_j \cdot e^{-\gamma \cdot [(x_{p1} - \text{sv}_{j1})^2 + (x_{p2} - \text{sv}_{j2})^2 + \dots + (x_{p96} - \text{sv}_{j96})^2]} + \mathbf{b} \quad (5-9)$$

A sensibilidade manifestada por cada saída a cada entrada individual pode ser determinada a partir de derivadas parciais, tal como se exemplifica na Equação 5-10 com o cálculo da sensibilidade da saída  $y_l$  relativo à influência de  $x_l$  para o conjunto de padrões de treino.

$$\left. \frac{\partial y_l}{\partial x_l} \right|_{\text{padrão } p} = \sum_{j=1}^{\text{n}^\circ \text{ de vetores de suporte}} \alpha_j \cdot (-2 \cdot \gamma) \cdot (x_{p,l} - sv_{j,l}) \cdot e^{-\gamma[(x_{p,1} - sv_{j,1})^2 + (x_{p,2} - sv_{j,2})^2 + \dots + (x_{p,96} - sv_{j,96})^2]} \quad (5-10)$$

Por forma a tratar os resultados para todos os padrões, optou-se novamente por calcular a média dos valores absolutos das diferentes derivadas parciais por padrão. O procedimento é análogo ao que já foi tratado para as redes neuronais e rege-se pela aplicação da fórmula expressa na Equação 5.7.



## 6. Resultados obtidos a partir dos modelos de previsão adotados

Os resultados de previsão devem ser abordados numa ótica global, procurando-se indicadores que permitam aferir a acuidade de determinado método e comparar resultados entre diferentes modelos. Nesse sentido, este estudo apresenta resultados do erro percentual absoluto médio (*EPAM*) (indicador já apresentado na Equação 4-7) e da raiz do erro percentual quadrático médio (*REPQM*) que é descrito na equação 6-1.

$$\text{REPQM}(\%) = \sqrt{\frac{\sum_h \left( \frac{\text{Valor Real de Potência ativa}_h - \text{Valor Estimado de Potência ativa}_h}{\text{Valor Real de Potência ativa}_h} \right)^2}{\text{Tempo associado ao conjunto de dados (em horas)}}} \times 100\% \quad (6-1)$$

Os erros serão reportados quer para o conjunto de dados de treino/validação, quer para o conjunto de dados de teste, uma vez que é desejável que as redes neuronais treinadas sejam apropriadas para generalizar novos dados não usados durante a fase de treino. Na seleção da arquitetura mais vantajosa para cada modelo tornou-se possível aceitar que um reduzido número de neurónios origina elevados erros na previsão dos dados de teste mas também nos dados de treino (justificando-se pelo facto da rede não ter o número de unidades de processamento que lhe permitam aprender adequadamente as relações nos dados apresentados) e que um elevado número de neurónios tende genericamente a reduzir os erros nos dados de treino mas eleva os erros nos dados de teste. Este último efeito é consubstanciado pelo facto da rede ficar demasiado sensível aos padrões de treino apresentados, tornando-se muito rígida às relações que se evidenciam nestes padrões, perdendo consequentemente capacidade de generalização. Por este motivo, para cada modelo foi avaliado um número variável de neurónios na camada intermédia, que o estudo explora entre os 3 e os 15 neurónios, já que um número de neurónios fora deste intervalo não garante precisão suficiente no processo. Para cada número de neurónios diversas simulações foram efetuadas de modo a garantir diferentes inicializações aleatórias dos parâmetros ajustáveis das redes, podendo-se avaliar os diferentes resultados obtidos após processos de treino. De ressaltar que os resultados de previsão apresentados nas próximas

secções relativos aos dois estudos de caso considerados se encontram numa gama comparável com trabalhos que foram também desenvolvidos em Portugal e com um análogo nível de agregação de consumidores (Fidalgo e Peças Lopes, 2005, Gama e Rodrigues, 2008).

### 6.1. Resultados de previsão com redes neuronais relativos ao primeiro estudo de caso considerado

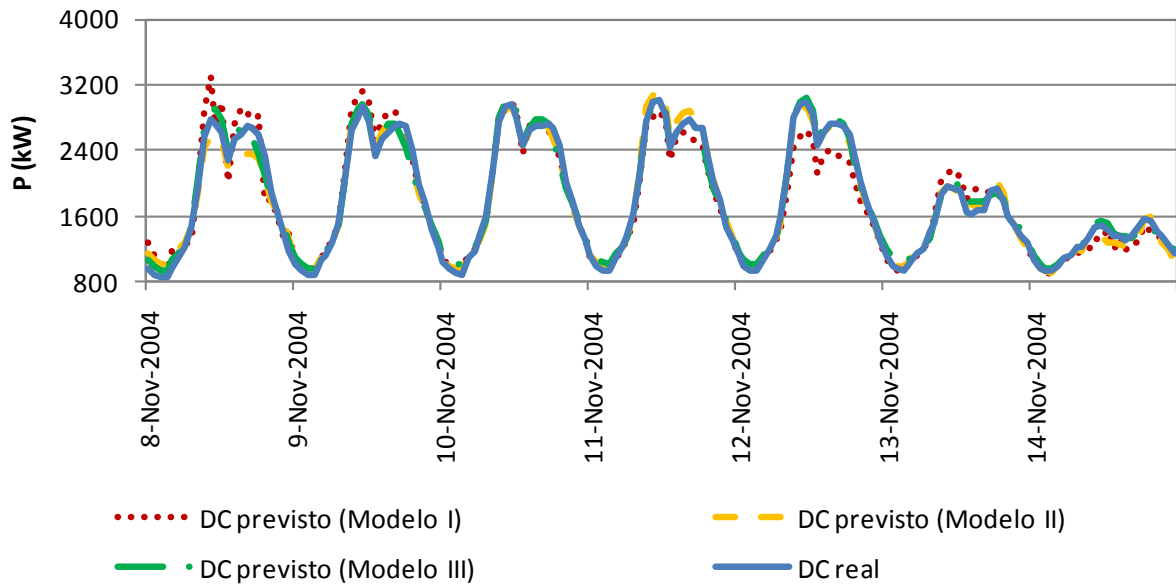
Por forma a sintetizar o estudo às previsões consideradas mais satisfatórias, os resultados apresentados para cada modelo serão exclusivamente relacionados com a melhor configuração encontrada para cada modelo após um leque de simulações analisadas.

**Tabela 6.1 – Resultados de previsão baseada em redes neuronais para o primeiro estudo de caso**

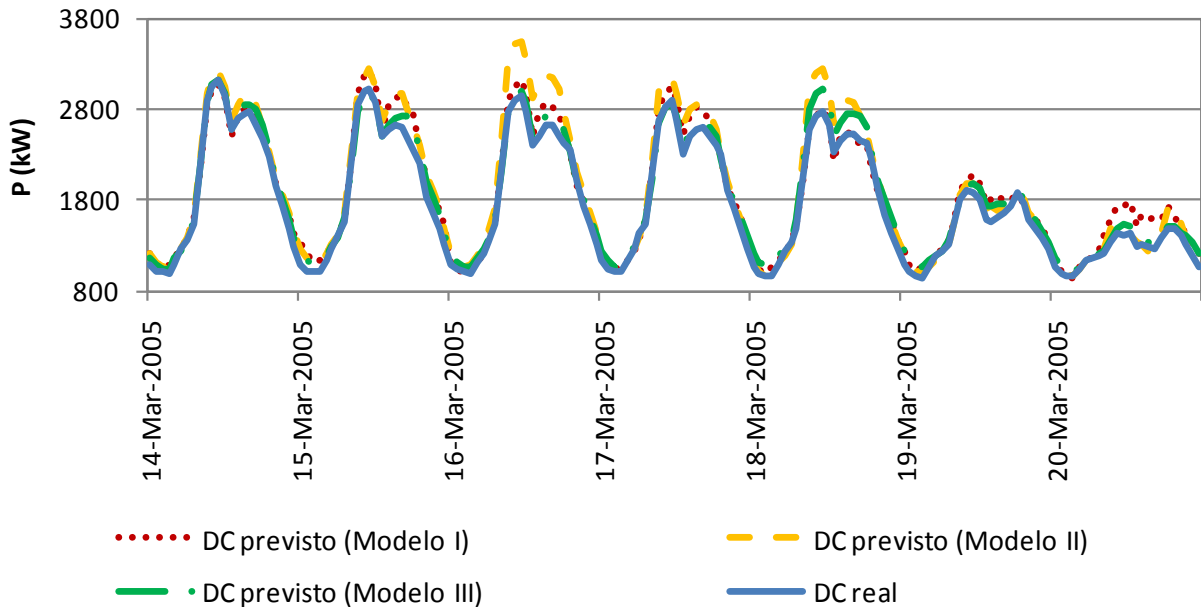
Modelo de previsão adotado	Arquitetura da rede com melhores resultados <sup>(*)</sup>	Medidas de Erro	EPAM	REPQM
Modelo I	72:7:24	Dados de treino	5,60%	9,00%
		Dados de teste	7,09%	11,47%
Modelo II	75:9:24	Dados de treino	4,94%	6,74%
		Dados de teste	5,54%	7,26%
Modelo III	96:3:24	Dados de treino	4,26%	5,69%
		Dados de teste	4,96%	6,76%

<sup>(\*)</sup> Número de entradas: Número de neurónios na camada intermédia: Número de saídas

Com a análise dos resultados da Tabela 6.1 é possível constatar-se que o Modelo I é o mais grosseiro, com resultados globais de erro superiores aos modelos concorrentes, demonstrando essencialmente a sua ineficiência para lidar com particularidades de distribuição de consumo em diferentes tipos de dia e na ocorrência de feriados. Esta constatação será abordada em maior detalhe em seguida, com uma análise mais pormenorizada aos casos com dias atípicos. Já o Modelo II apresenta uma melhoria em relação ao Modelo I, uma vez que incorpora variáveis que lhe permitem distinguir entre diferentes dias da semana, bem como uma variável binária que “alerta” a rede para a presença de dias atípicos, precavendo que o seu comportamento seja sensível à presença de tais dias. O Modelo III, que acarreta alguma informação de como diferentes classes de consumidores reagem em termos de comportamentos e hábitos de consumo em dias especiais, apresenta uma considerável redução dos erros de previsão, tornando o processo de estimativa mais rigoroso. As próximas figuras (Figura 6.1 e Figura 6.2) ilustram as diferentes precisões alcançadas com os diferentes modelos de previsão em semanas pertencentes ao período de treino. Outros resultados correspondendo a diferentes cenários de validação cruzada podem ser consultados no Anexo A.

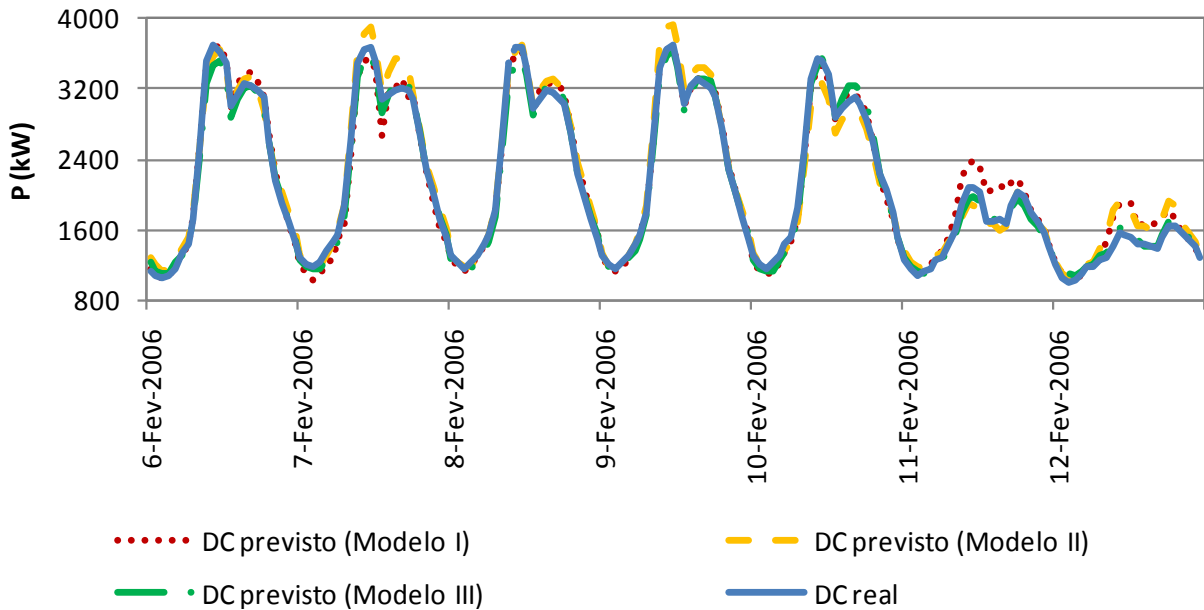


**Figura 6.1 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) numa semana típica de novembro de 2004 (período de treino) – 1º Estudo de Caso (1º EdC)**



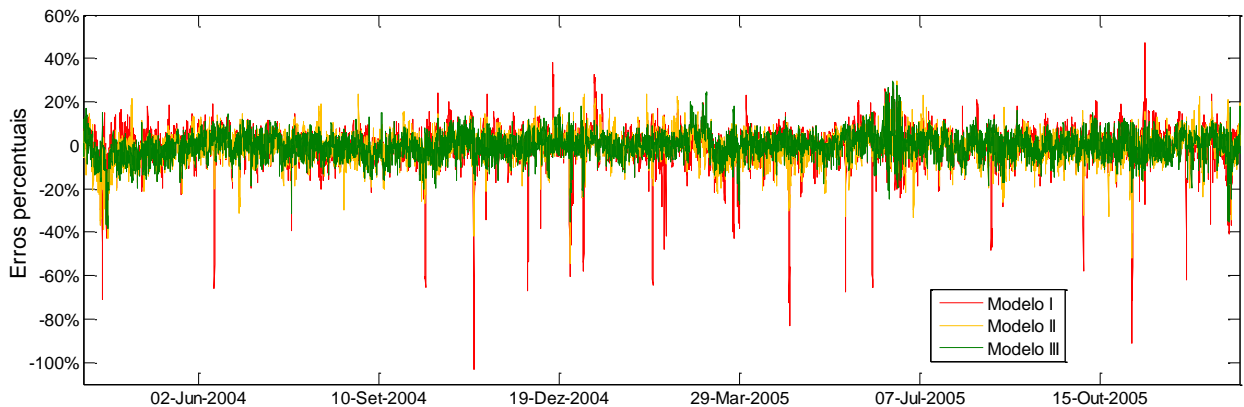
**Figura 6.2 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) numa semana típica de março de 2005 (período de treino) (1º EdC)**

A Figura 6.3 pretende também exemplificar a comparação entre os diferentes modelos prevendo o dia seguinte durante uma semana que pertence ao subconjunto de teste. A capacidade de generalização é assumidamente reconhecida quando está em causa a previsão de dias normais, sem características anómalas no padrão de consumo, uma vez que a precisão alcançada pelos modelos durante os dados de treino é extensível quando se aplicam estes modelos de previsão a dados de teste.



**Figura 6.3 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) numa semana típica de fevereiro de 2006 (período de teste) (1º EdC)**

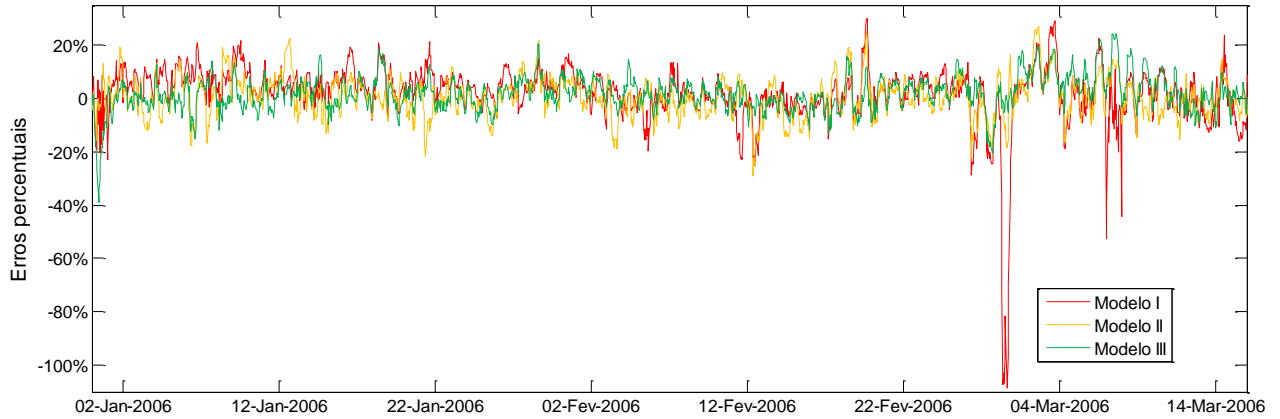
Por forma a possibilitar uma visão mais globalizada da distribuição dos erros relativos aos três diferentes modelos gerados, é apresentada na Figura 6.4 a forma como o erro evolui para o subconjunto de dados de treino.



**Figura 6.4 – Distribuição dos erros percentuais por modelo (RN) durante o período de treino considerado (1º EdC)**

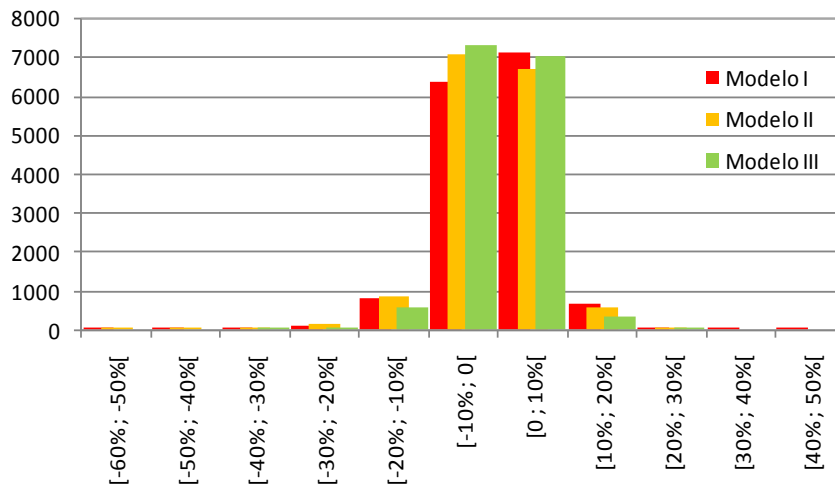
Da análise da Figura 6.4 sobressai a ideia que o Modelo I acarreta erros com maior projeção, enquanto o Modelo II já melhora substancialmente a acuidade nas estimativas de consumo do dia seguinte. O Modelo III revela indubitavelmente ser o mais apto para melhor centrar os erros em torno de zero. Os elevados erros do Modelo I não devem ser tolerados, pelo que se conclui que é premente a introdução de variáveis que retenham informação relevante acerca dos tipos de dia considerados, ou da presença de feriados. Em alternativa, alguma informação acerca de perfis de consumo típicos relativos aos dias que se pretendem estimar aparenta ser determinante para ser passada à arquitetura de rede neuronal adotada.

Na Figura 6.5 é apresentada uma distribuição dos diferentes erros por modelo para os dados que serviram de teste, sendo notória a falta de sensibilidade do Modelo I para lidar com o carácter atípico de consumo evidenciado no dia 28 de fevereiro de 2006, relativo a uma terça-feira de Carnaval.



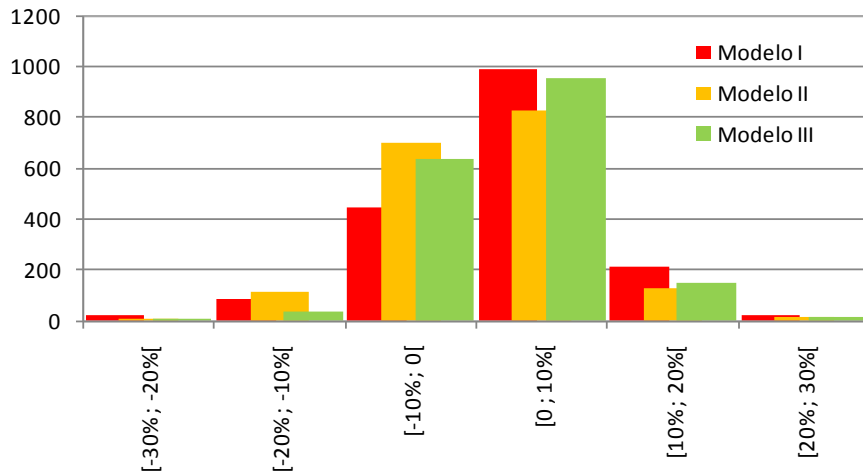
**Figura 6.5 – Distribuição dos erros percentuais por modelo (RN) durante o período de teste considerado (1º EdC)**

Nas seguintes figuras são apresentados histogramas relativos aos erros percentuais, neste caso subdivididos por escalões intercalados em 10%, para o período de treino (na Figura 6.6) e para o período de teste (na Figura 6.7). É verificado que o caso do Modelo III é o que garante um maior número de erros percentuais absolutos inferiores a 10%.



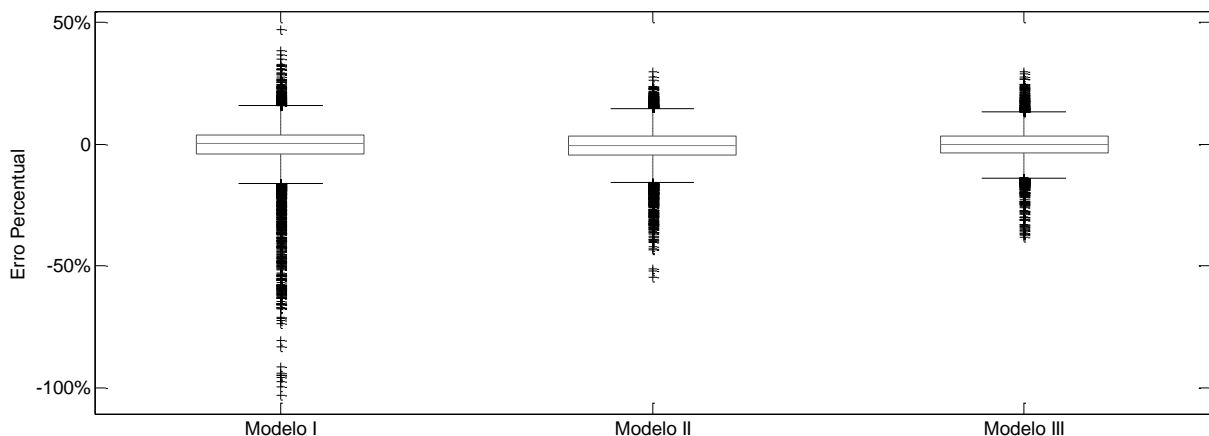
**Figura 6.6 – Histograma dos erros percentuais por modelo (RN) para o período de treino considerado (1º EdC)**

A mesma constatação pode ser apontada para o histograma que foca o período de teste, já que os Modelos I e II apresentam um número não negligenciável de erros percentuais absolutos acima dos 10%.



**Figura 6.7 – Histograma dos erros percentuais por modelo (RN) para o período de teste considerado (1º EdC)**

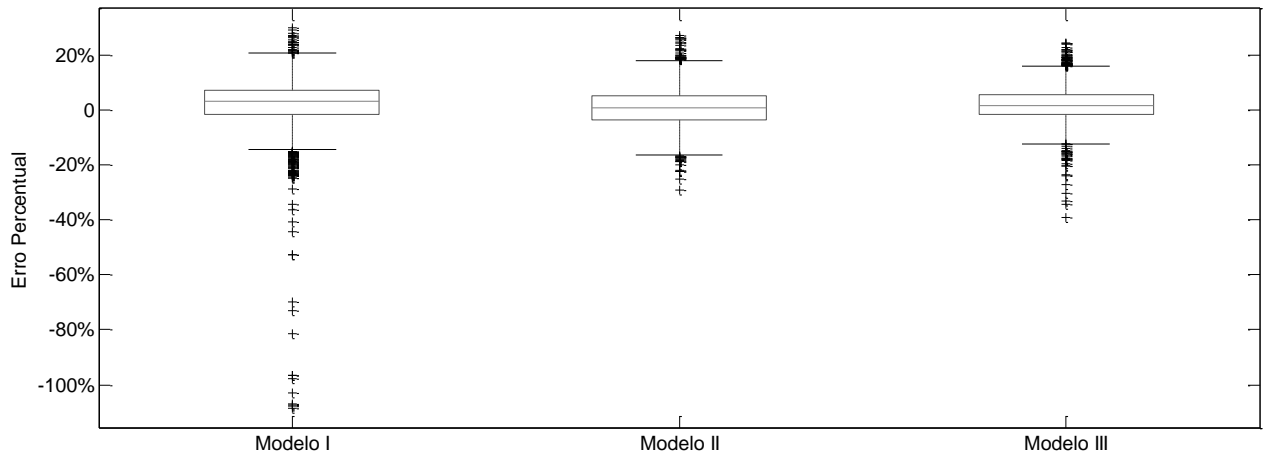
Uma distinta abordagem estatística para analisar a distribuição dos erros pode ser efetuada através de uma representação dos erros por meio de Boxplot <sup>[51]</sup>. Este tipo de representação evidencia o valor da mediana do conjunto de erros, e assume um formato de uma “caixa” com um limite superior afeto ao 3º quartil e um limite inferior afeto ao 1º quartil. É ainda considerada uma margem acima do 3º quartil e abaixo do 1º quartil, na qual se depreende que determinados erros se possam situar. Esta margem é definida a partir da amplitude inter-quartil e multiplicada por um parâmetro definido. Um elevado valor para este parâmetro pressupõe que elevados erros são tolerados, enquanto um reduzido valor deste parâmetro facilmente destaca os erros acima da banda aceitável, designados como *outliers*. Neste estudo foi considerado um parâmetro igual a 1,5 como fator multiplicador da amplitude inter-quartil. A Figura 6.8 incide no período de treino, sendo observável que o Modelo III apresenta os erros mais centrados em torno do zero, com um número mais reduzido de *outliers*, enquanto, por oposição, o Modelo I apresenta um número elevado de erros percentuais fora da gama expectável.



**Figura 6.8 – Comparação na distribuição de erros entre modelos de previsão (RN) (boxplot) para os dados de treino (1º EdC)**

[51] Neste documento adotar-se-á doravante o termo “Boxplot” por ser o termo internacionalmente reconhecido pela comunidade científica e por não ser um termo com uma tradução para Português direta e reconhecida pelo autor.

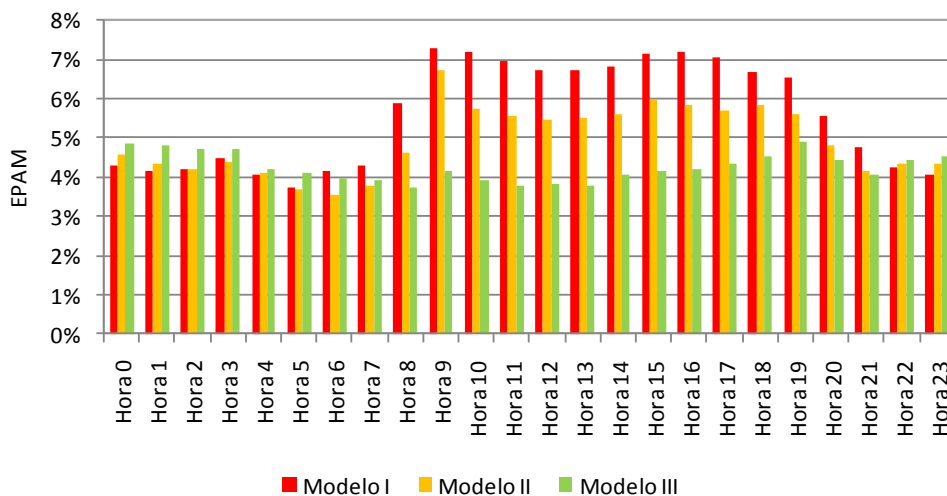
A Figura 6.9 apresenta o mesmo tipo de representação dos erros encontrados, mas relativo ao subconjunto de dados de teste.



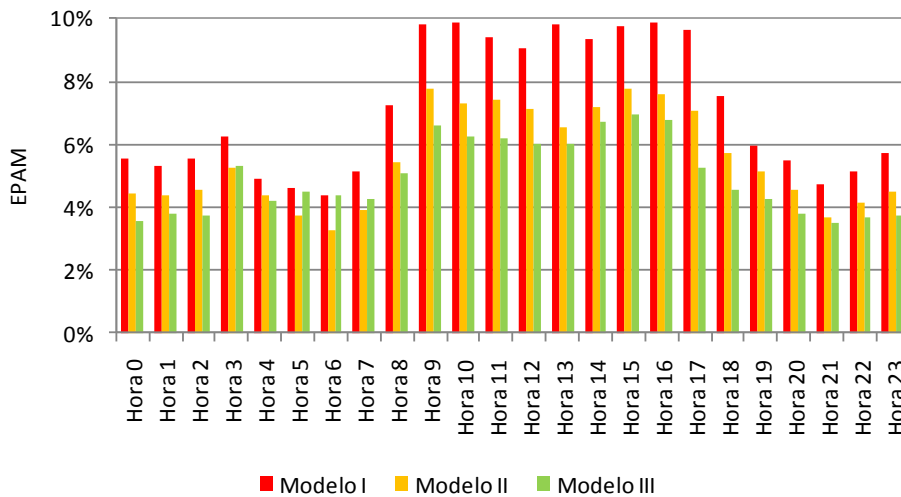
**Figura 6.9 – Comparação na distribuição de erros entre modelos de previsão (RN) (boxplot) para os dados de teste (1º EdC)**

Desta análise pode confirmar-se a inaptidão associada ao Modelo I, com elevados erros percentuais absolutos manifestamente assumidos como *outliers*. Estes erros são maioritariamente de sinal negativo, o que pode ser fundamentado pela ineficácia do Modelo I para lidar com dias feriados, nos quais o consumo real é sempre menor do que o consumo previsto por tal modelo. Apesar do Modelo III apresentar um valor relativo à mediana dos erros mais reduzido, assim como o primeiro e o terceiro quartis serem em módulo mais reduzidos do que nos modelos concorrentes, ainda assim é possível verificar-se que tal modelo não elimina o efeito de ocasionais *outliers* durante a fase de teste.

A forma como o erro se distribui pelas diferentes horas do dia pode suscitar algum interesse. Nas Figuras 6.10 e 6.11 são apresentados os valores de EPAM manifestados pelos diferentes modelos adotados para as distintas horas do dia, sem distinção entre dias de semana e fins de semana.



**Figura 6.10 – Distribuição de erros em modelos de previsão (RN) para as diferentes horas do dia (dados de treino) (1º EdC)**



**Figura 6.11 – Distribuição de erros em modelos de previsão (RN) para as diferentes horas do dia (dados de teste) (1º EdC)**

O adoção do Modelo III revela ser providencial para garantir que as previsões sejam mais rigorosas nos períodos mais críticos, normalmente associados às horas em que o consumo é maior e coincidente com atividades económicas.

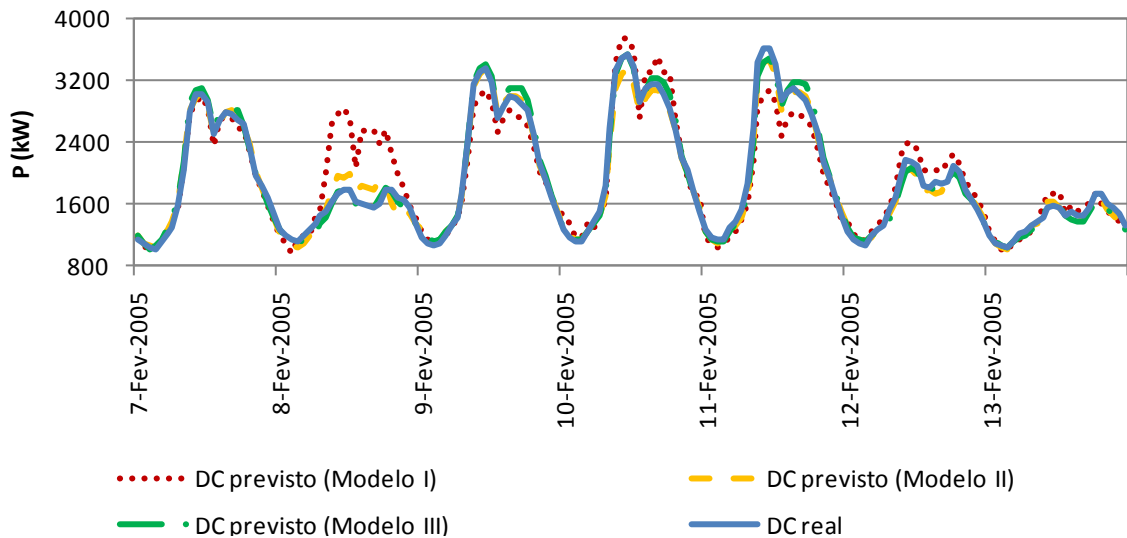
Como se pôde já concluir em resultados apresentados anteriormente, o efeito dos dias feriados revela ter um papel preponderante na comparação entre os modelos testados, pelo que uma atenção particular deverá ser dedicada neste primeiro caso em estudo à ocorrência destes dias atípicos. Deste modo, os resultados apresentados na Tabela 6.2 são obtidos a partir da aplicação de um filtro de modo a que só os dias atípicos sejam analisados (efeito de feriados e pontes). Os resultados dizem respeito a uma análise pormenorizada destes dias especiais englobando o período de treino e o período de teste, devido à pequena representatividade que os dias atípicos têm durante a fase de teste.

**Tabela 6.2 – Comparação entre os modelos de previsão (RN) aplicados ao primeiro estudo de caso para dias atípicos (como feriados e pontes)**

Modelo de previsão adotado	EPAM	REPQM
Modelo I	19,09%	29,14%
Modelo II	8,19%	11,11%
Modelo III	6,25%	8,64%

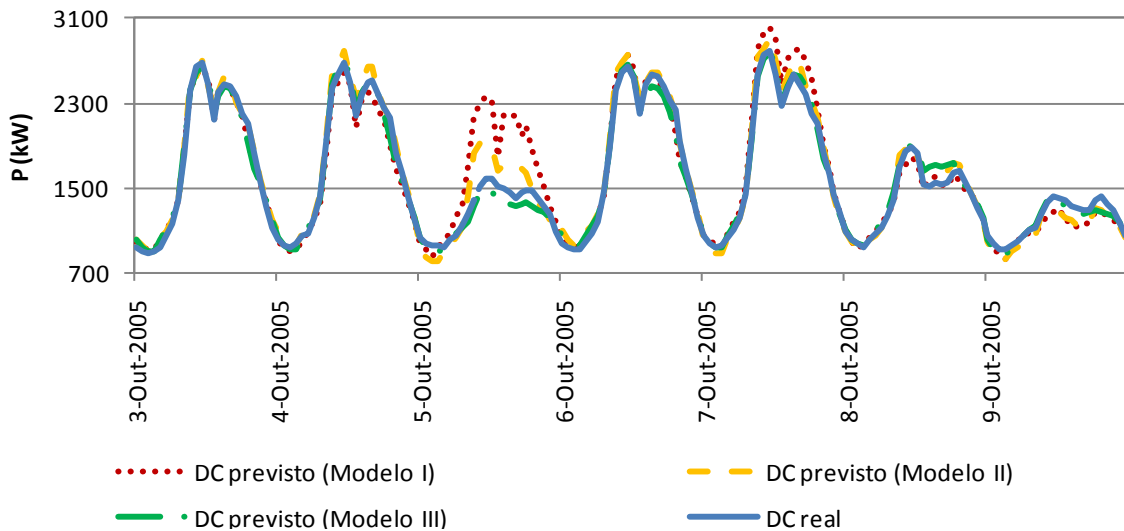
Por forma a evidenciar com maior detalhe a pertinência da escolha de modelos apropriados para a previsão de dias atípicos, são apresentadas comparações entre os diferentes modelos de previsão de dia seguinte, durante uma semana, nas Figura 6.12 e Figura 6.13.





**Figura 6.12 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) numa semana atípica com dia de Carnaval – 8 de fevereiro de 2005 (1º EdC)**

A Figura 6.12 permite verificar uma semana atípica que envolve uma terça-feira de Carnaval, bem como o efeito da ponte de segunda-feira. O Modelo III revela uma forte apetência para lidar com a problemática dos dias especiais, como se pode também confirmar com a previsão do feriado de 5 de outubro de 2005 na Figura 6.13. A informação do DC sintetizado aproveita a informação dos perfis de consumo por classe que compreende já um tratamento em relação a dias especiais, beneficiando por esta via os modelos de previsão.



**Figura 6.13 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) numa semana atípica com feriado de 5 de outubro de 2005 (1º EdC)**

6.1.1. Análise de sensibilidades da rede neuronal adotada no primeiro estudo de caso considerado

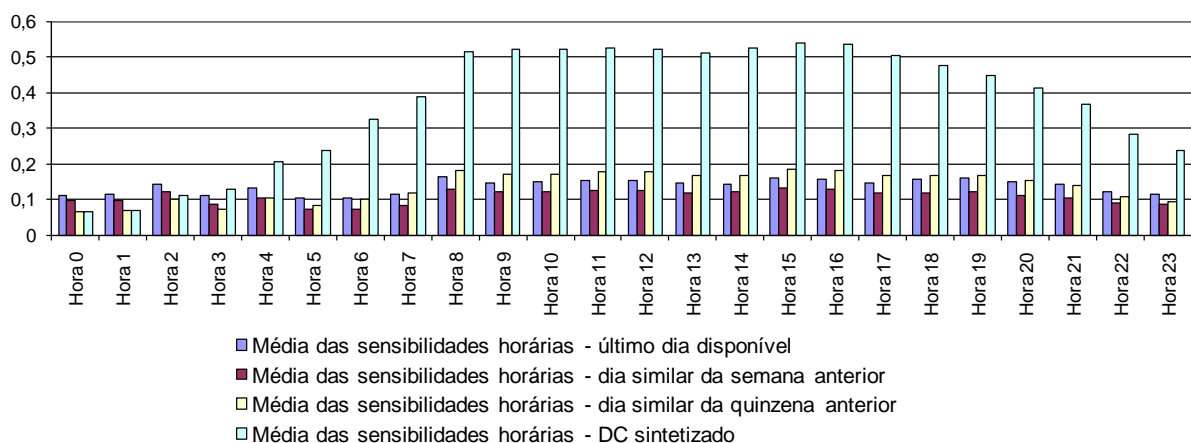
Esta subsecção visa avaliar a importância relativa que o diagrama de carga sintetizado toma relativamente a outros vetores de entrada usados (último dia disponível, dia similar da semana anterior e dia similar da quinzena anterior). O Modelo III será desta forma analisado em detalhe,

já que foi aquele que resultou em erros de previsão mais reduzidos e por ser aquele que inclui o diagrama de carga sintetizado. Este modelo recorre a 96 entradas distintas para prever os 24 registos horários de potência ativa do dia seguinte, podendo-se determinar as sensibilidades de cada saída em particular às diferentes entradas que são aplicadas. O processo pode tornar-se de difícil interpretação, já que resulta num elevado número de sensibilidades, neste caso 96 entradas para 24 saídas, ou seja 2304 diferentes sensibilidades. Para contornar este problema foram calculadas as médias das sensibilidades para cada saída relativamente às 24 horas que compõem cada dia do conjunto de dados de entrada. O processo pode ser descrito de forma mais detalhada:

– numa primeira fase é apurada a importância de cada entrada individual para cada saída, de acordo com a Equação 5-7, sendo apurados os descritores 2304 distintos valores.

– de forma a tratar e a analisar de forma mais coerente esta elevada quantidade de informação, optou-se por representar as sensibilidades relativas a cada vetor de entrada (cada vetor composto por 24 horas) através da média dos valores de sensibilidade de cada saída relativamente aos elementos individuais de um determinado vetor de entrada considerado. Por exemplo, para ser determinada a importância relativa do último dia disponível na previsão da primeira saída do modelo (registo horário de potência ativa referente à hora 0 do dia a prever), é calculada a média das 24 sensibilidades relacionadas com as contribuições de cada registo horário de potência do último dia disponível para o registo horário do dia seguinte à hora 0.

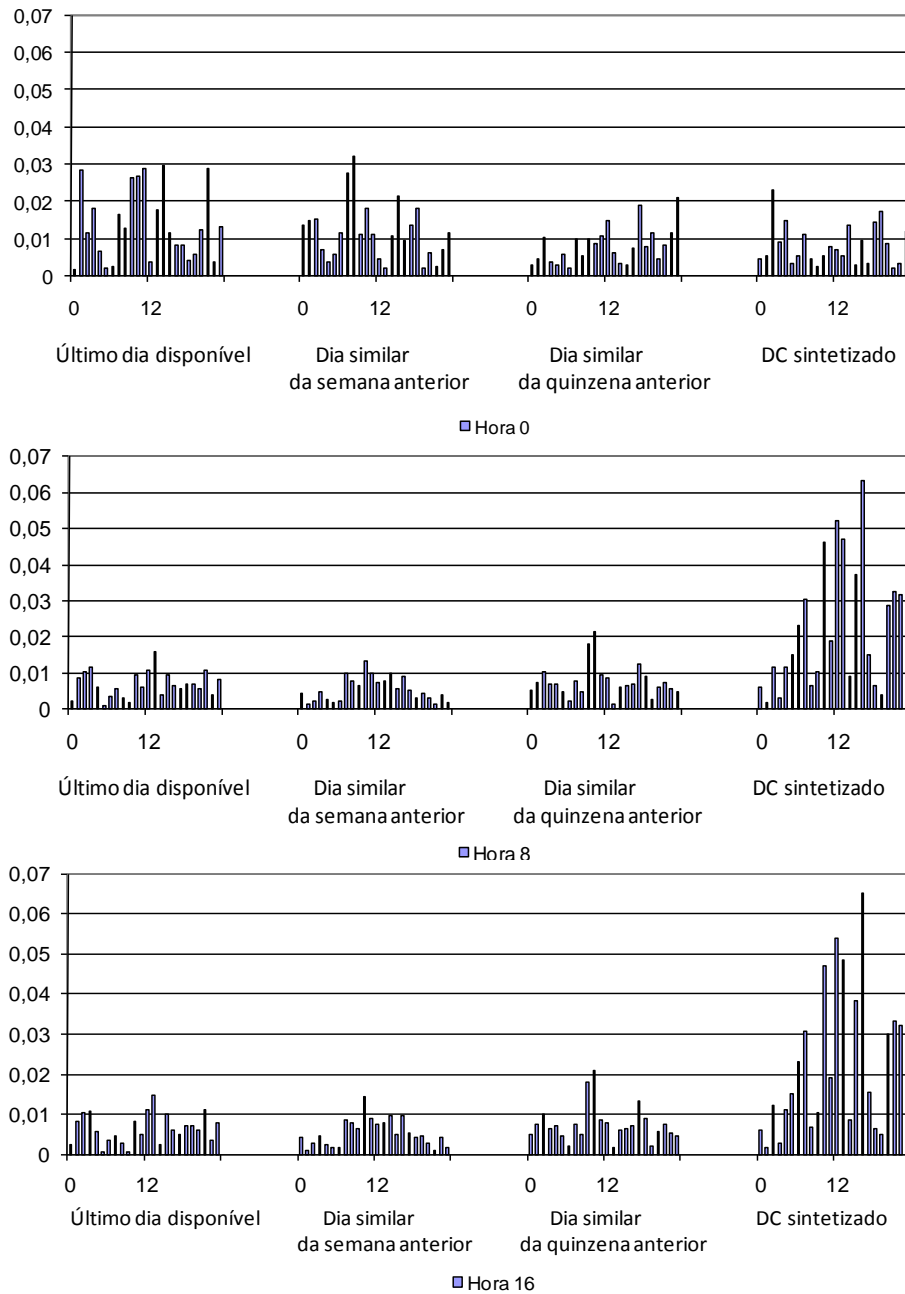
A Figura 6.14 ilustra os resultados obtidos, onde se evidencia uma relevância do vetor associado ao diagrama de carga sintetizado face às entradas que dependem do passado recente.



**Figura 6.14 – Médias das sensibilidades horárias de cada saída do Modelo (RN) relativamente às 24 h de cada vetor de entrada (1º EdC)**

Excetuando o período noturno (sensivelmente entre as 0h e as 2h), esta predominância permite validar que o Modelo terá maior sensibilidade ao conjunto de registos horários presente no diagrama de carga sintetizado.

De modo a dar-se uma perspetiva mais global da importância das diferentes entradas individuais para cada hora que se pretende prever, alguns exemplos são apresentados na Figura 6.15. De referir que as sensibilidades apresentadas nesta figura encontram-se normalizadas, garantindo que a soma das diferentes sensibilidades para cada saída é unitária.



**Figura 6.15 – Sensibilidades manifestadas pelas saídas do modelo RN referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (1º EdC)**

Com a apresentação destes exemplos é reforçada a ideia que o diagrama de carga sintetizado é genericamente uma contribuição relevante para prever os registos horários de potência ativa do dia seguinte. Ainda assim, as contribuições do último dia disponível e do dia similar da semana anterior não devem ser menosprezadas, já que, para prever algumas horas noturnas, o efeito de

tendência entre dias próximos manifesta-se determinante. Outra conclusão interessante a reter é que o modelo multivariável já treinado, quando visa prever uma hora específica do dia seguinte, não depende necessariamente apenas de horas congéneres dos dias anteriores e do DC sintetizado mas também de um conjunto de horas em torno desta hora congénere e de horas para as quais o consumo apresenta uma tendência de variação significativa. Uma abordagem mais detalhada acerca da aplicação da análise de sensibilidades ao modelo adotado apresenta-se no Anexo B.

## 6.2. Aplicação da metodologia baseada em máquinas de suporte vetorial para a previsão de consumos relativos ao primeiro estudo de caso considerado

Como foi já referido anteriormente, a previsão baseada em máquinas de suporte vetorial assumiu sempre um conjunto de entradas semelhantes ao Modelo III. Os registos horários de potência ativa referentes ao último dia disponível, dia similar da semana anterior, dia similar da quinzena anterior e o DC sintetizado foram novamente os parâmetros de entrada que serviram para a construção deste novo modelo.

### 6.2.1. Descrição do processo de criação de diferentes modelos horários otimizados

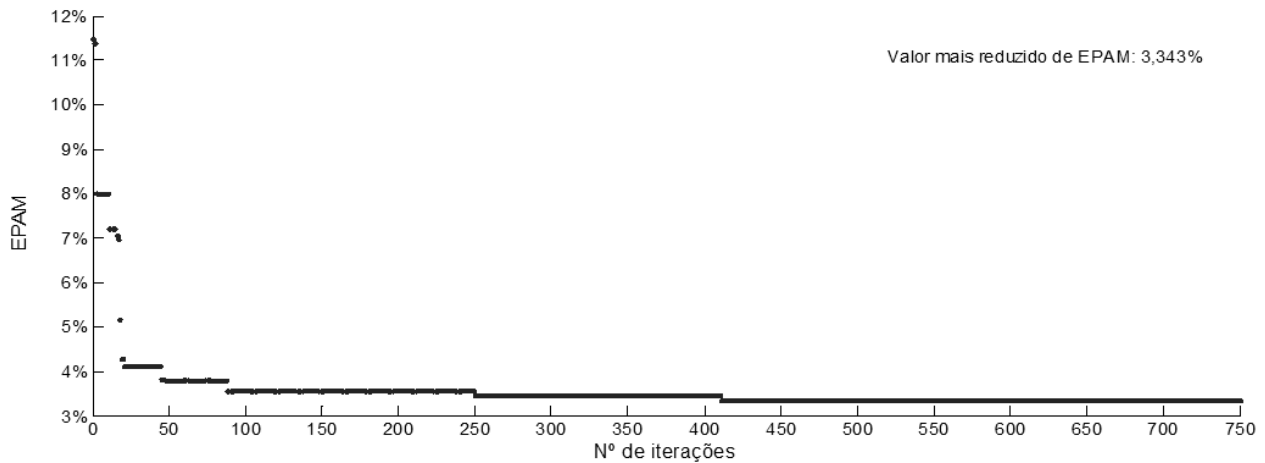
Os diferentes modelos MSV foram otimizados, de acordo com a informação já apresentada na Secção 5.2, e diferentes características foram originadas para cada hora que é prevista. Alguns exemplos são apresentados na Tabela 6.3, na qual é possível verificar-se o EPAM para os diferentes conjuntos de dados.

**Tabela 6.3 – Descrição de alguns dos modelos MSV obtidos (para horas específicas) (1º EdC)**

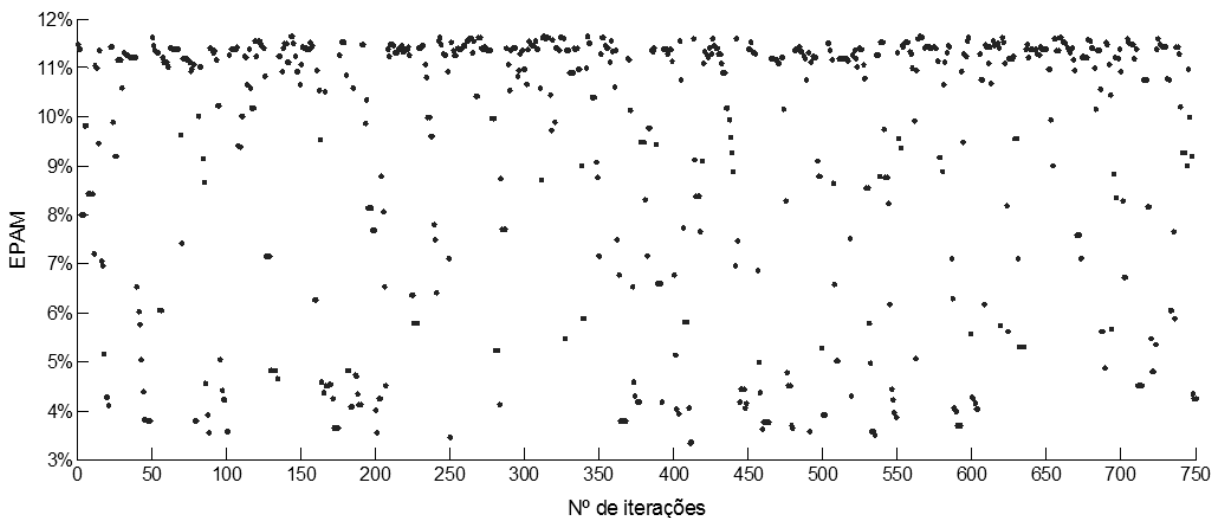
Modelo	Parâmetros MSV obtidos		Número de vetores de suporte	Dados de treino	Dados de validação	Dados de teste
				EPAM	EPAM	EPAM
Hora 0	$\gamma$	0,00138	246 <sup>(*)</sup>	2,64%	3,34%	3,18%
	C	14,972				
	$\varepsilon$	0,2063				
Hora 7	$\gamma$	0,00364	273 <sup>(*)</sup>	2,34%	3,58%	4,21%
	C	11,928				
	$\varepsilon$	0,1664				
Hora 18	$\gamma$	0,00197	213 <sup>(*)</sup>	3,85%	4,76%	4,80%
	C	4,981				
	$\varepsilon$	0,1574				

<sup>(\*)</sup> – Número de vetores de suporte reconhecidos de um total de 514 vetores considerados no período de treino

De modo a fornecer alguma informação adicional acerca do modo como o processo de otimização decorre para diferentes modelos horários, optou-se pela escolha do modelo referente à hora 0 do dia seguinte. Na Figura 6.16 explicita-se a evolução do valor mínimo do EPAM verificado nos dados de validação, enquanto na Figura 6.17 transparecem os resultados obtidos para diferentes combinações exploradas.



**Figura 6.16 – Evolução do mínimo resultado de EPAM para o conjunto de validação – modelo referente à hora 0 (1º EdC)**



**Figura 6.17 – Evolução do resultado EPAM verificado no conjunto de validação durante o processo de otimização – modelo referente à hora 0 (1º EdC)**

Para este modelo horário pode verificar-se que o indicador de previsão se situa entre os 3,34% e os 11,68%, demonstrando a importância da determinação dos parâmetros MSV ajustáveis para garantir rigor na previsão. A partir de outros modelos analisados e apresentados no Anexo C constata-se que é especialmente nas horas para as quais o consumo é superior que se identificam erros de previsão na ordem dos 30% durante o processo de otimização, tornando estes modelos muito dependentes de uma apropriada seleção de parâmetros MSV. Tendo os modelos horários otimizados e coerentes com esta metodologia apresentada, o passo seguinte deve incidir em

estender os resultados provenientes por modelo horário para todo o período considerado (para todos os subconjuntos de dados), tendo em consideração que cada hora tem um modelo correspondente.

### 6.2.2. Apresentação de resultados de previsão com modelos baseados em máquinas de suporte vetorial

Os erros de previsão são apresentados com uma divisão em dados de treino/validação e dados de teste, à semelhança dos resultados já apresentados para os modelos RN. A Tabela 6.4 apresenta os resultados obtidos com um total de 96 entradas semelhante ao Modelo III de redes neuronais que pode ser consultado na Figura 5.5 e também resultados congéneres alcançados após o recurso à seleção de parâmetros explicativos conforme foi apresentado na Secção 5.2.3.

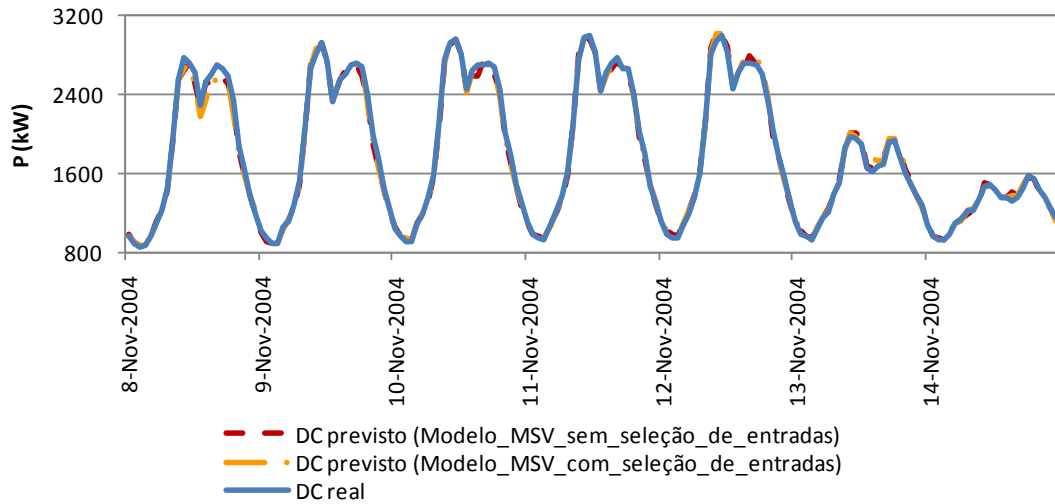
**Tabela 6.4 – Resultados de previsão baseada em máquinas de suporte vetorial para o primeiro estudo de caso**

Descrição do modelo adotado	Medidas de Erro	EPAM	REPQM
total de 96 entradas	Dados de treino	3,12%	5,02%
	Dados de teste	4,87%	6,24%
com seleção de parâmetros explicativos	Dados de treino	3,23%	5,18%
	Dados de teste	4,84%	6,26%

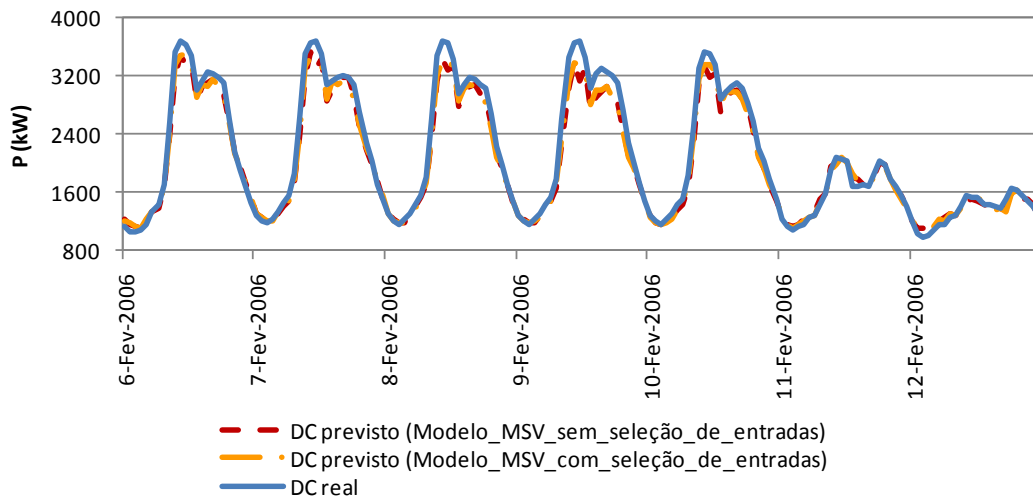
Dos resultados apresentados pode constatar-se que os modelos apresentam erros no conjunto de treino consideravelmente reduzidos, enquanto os erros no subconjunto de teste são já mais concordantes com os erros encontrados nos modelos RN. Apesar de poder parecer sugestiva, uma comparação direta entre as duas metodologias de previsão não se torna coerente, já que os pressupostos são substancialmente díspares. Enquanto, por razões de simplicidade, se recorreu a um modelo único multivariável no caso das redes neuronais (com vantagens associadas ao tempo de computação na fase de treino, posterior simulação de novos dados, envolvendo também uma análise crítica mais simplificada de um modelo único), nas máquinas de suporte vetorial são usados modelos diferenciados por hora.

Os dois modelos MSV apresentam indicadores de previsão muito semelhantes, pretendendo-se nesta comparação demonstrar que uma eventual redução do número de entradas pode ter um efeito de generalização do rigor obtido em estimativas de dados futuros, com modelos mais simplificados que acarretam um menor tempo associado ao treino e que tornam mais expeditas eventuais análises dos mesmos.

Alguns resultados comparativos com maior resolução (numa base semanal) são apresentados na Figura 6.18 e na Figura 6.19.

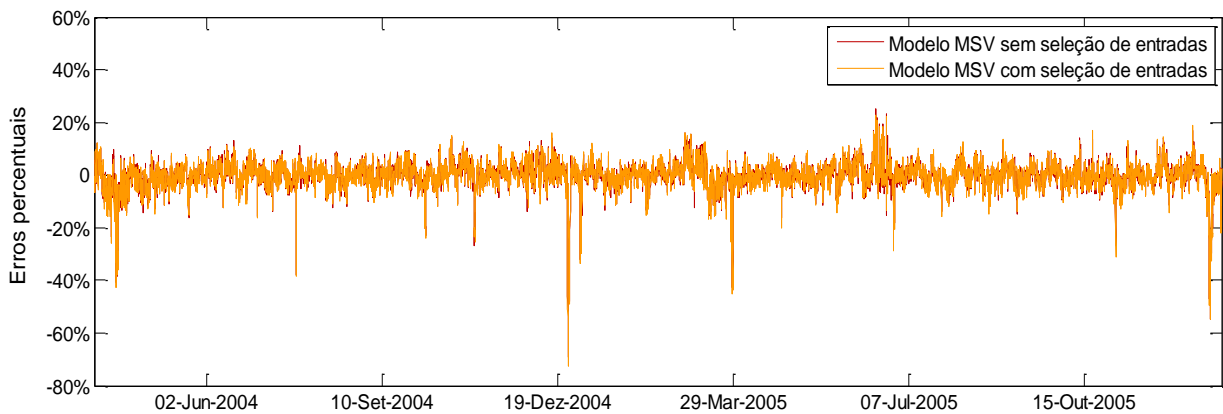


**Figura 6.18 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana típica de novembro de 2004 (período de treino) (1º EdC)**



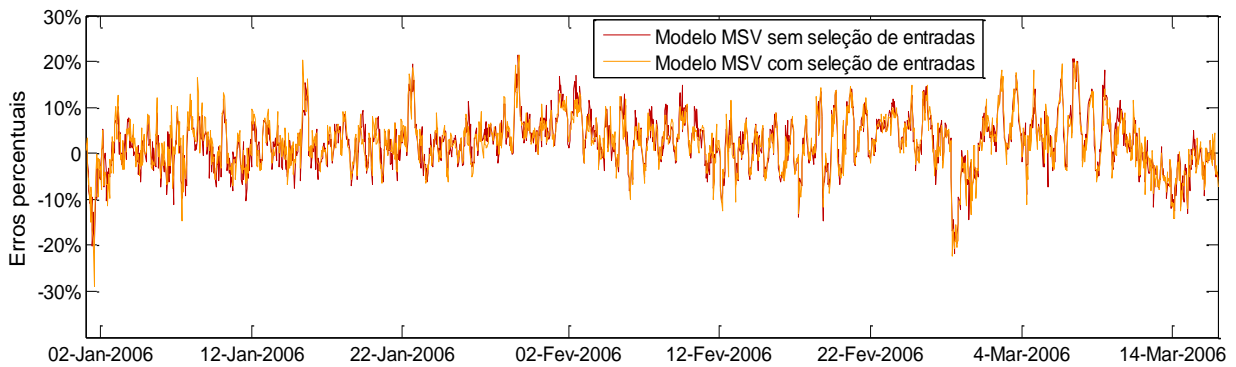
**Figura 6.19 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana típica de fevereiro de 2006 (período de teste) (1º EdC)**

A distribuição do erro para os dois subconjuntos de dados em análise encontra-se na Figura 6.20 (para os dados de treino) e na Figura 6.21 (para os dados de teste). Na distribuição de erros de treino existem registos de algum modo atípicos em torno da época natalícia e para outros dias especiais e que são comuns às duas variantes de modelos MSV.



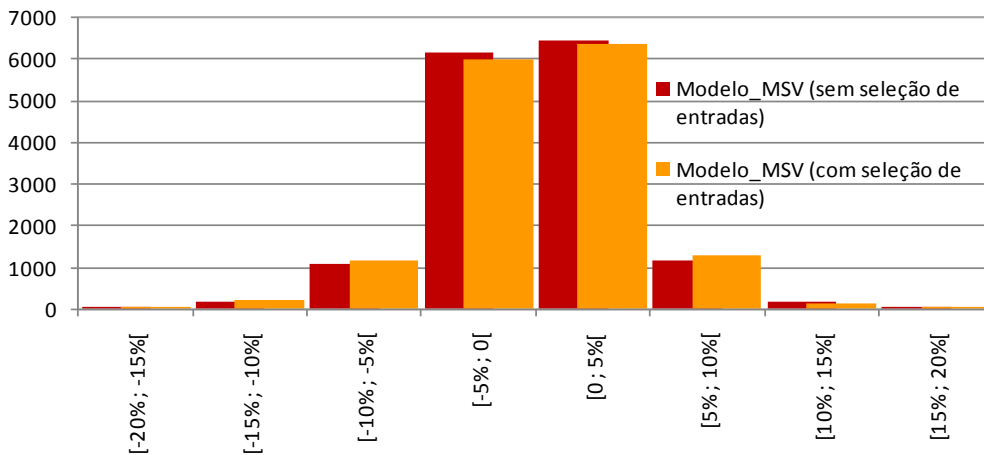
**Figura 6.20 – Distribuição dos erros percentuais por modelo (MSV) durante o período de treino considerado (1º EdC)**

Os erros de teste demonstram também um certo equilíbrio entre as duas variantes e os registos atípicos (embora numa gama menor do que o caso anterior) continuam a verificar-se em dias atípicos, como são o caso do feriado de 1 de janeiro e os dias em torno do Carnaval.

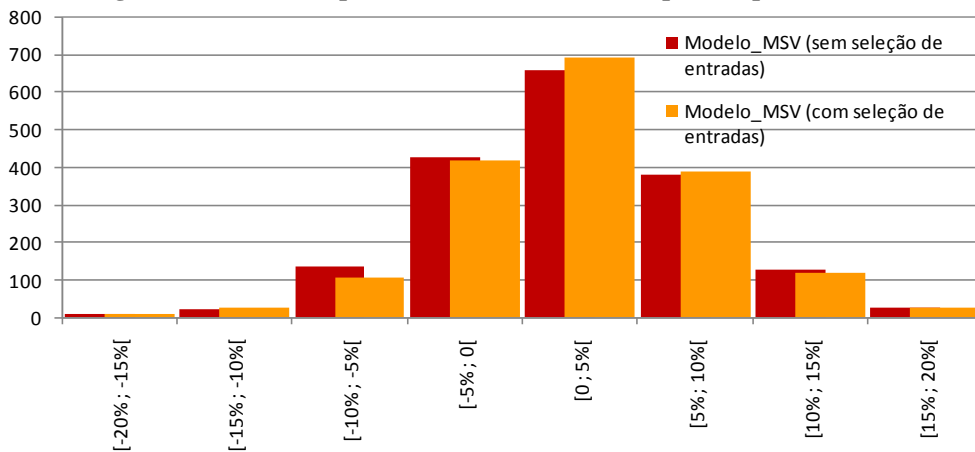


**Figura 6.21 – Distribuição dos erros percentuais por modelo (MSV) durante o período de teste considerado (1º EdC)**

Na Figura 6.22 pode analisar-se um histograma de distribuição dos erros no período de treino e na Figura 6.23 uma análise semelhante para dados de teste. Os resultados apontam novamente para distribuições semelhantes, com uma certa tendência para os erros de teste serem mais centrados em torno de zero no caso do modelo com seleção de parâmetros.



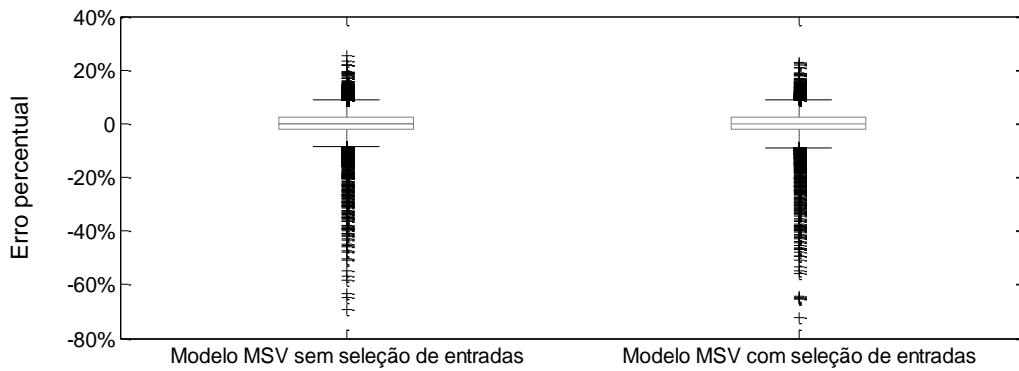
**Figura 6.22 – Histograma dos erros de previsão com modelos MSV para os períodos de treino (1º EdC)**



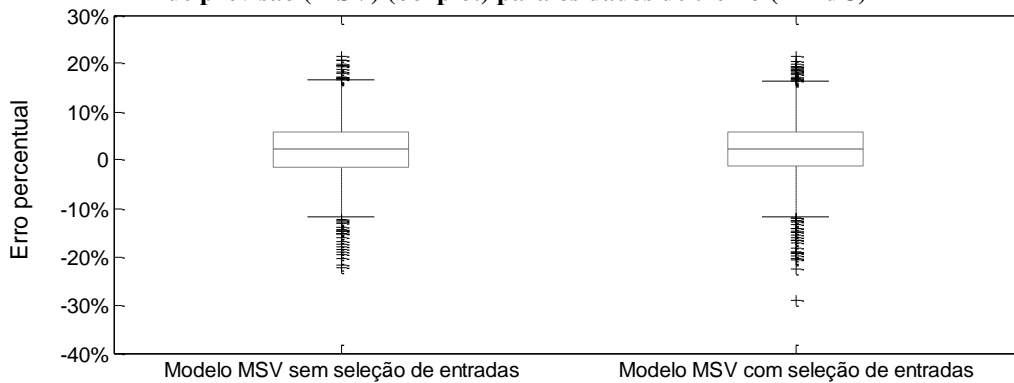
**Figura 6.23 – Histograma dos erros de previsão com modelos MSV para os períodos de teste (1º EdC)**



Como anteriormente foi verificada a existência de registos atípicos, apresenta-se também uma representação *boxplot* dos erros de treino e de teste na Figura 6.24 e Figura 6.25, respetivamente, por forma a destacar estes *outliers*. Novamente podem-se confirmar registos reduzidos de mediana dos erros, bem como dos quartis.

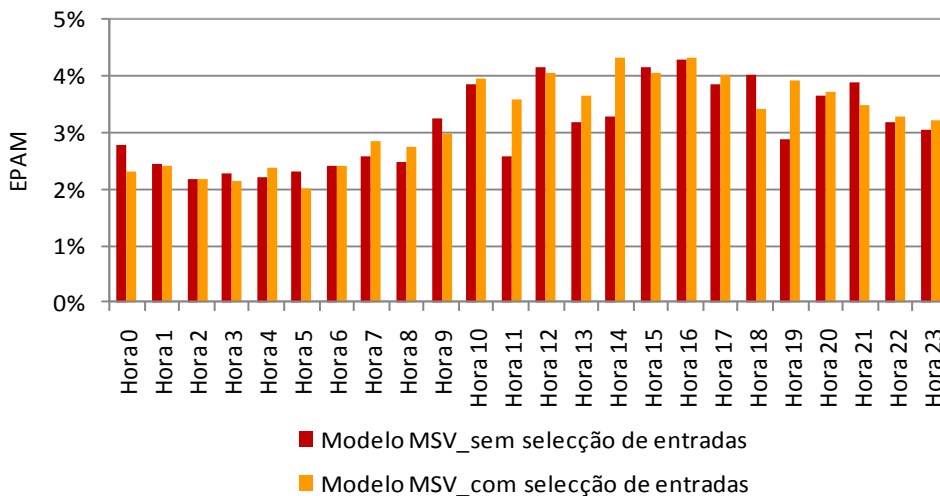


**Figura 6.24 – Comparação na distribuição de erros entre modelos de previsão (MSV) (boxplot) para os dados de treino (1º EdC)**

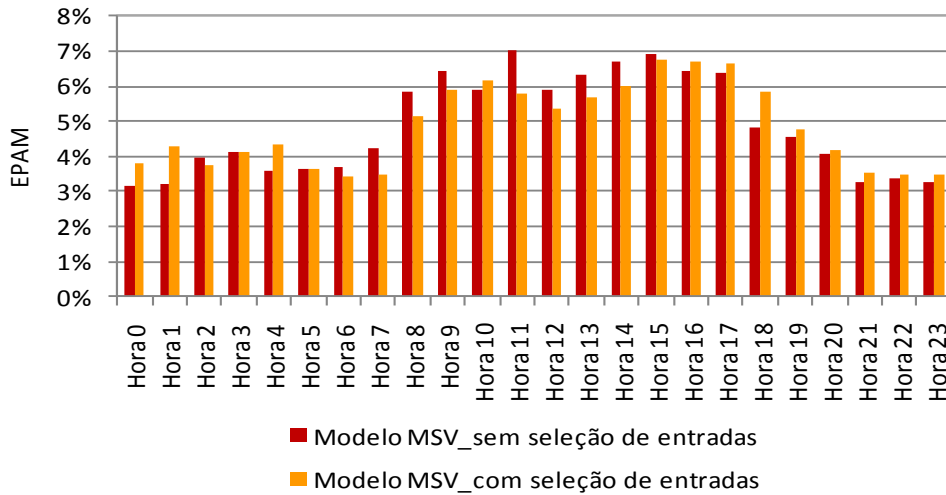


**Figura 6.25 – Comparação na distribuição de erros entre modelos de previsão (MSV) (boxplot) para os dados de teste (1º EdC)**

O indicador de erro EPAM distribui-se pelas diferentes horas do dia, segundo o que é apresentado na Figura 6.26 (para dados de treino) e na Figura 6.27 (para dados de teste).



**Figura 6.26 – Distribuição de erros em modelos de previsão (MSV) para as diferentes horas do dia (dados de treino) (1º EdC)**



**Figura 6.27 – Distribuição de erros em modelos de previsão (MSV) para as diferentes horas do dia (dados de teste) (1º EdC)**

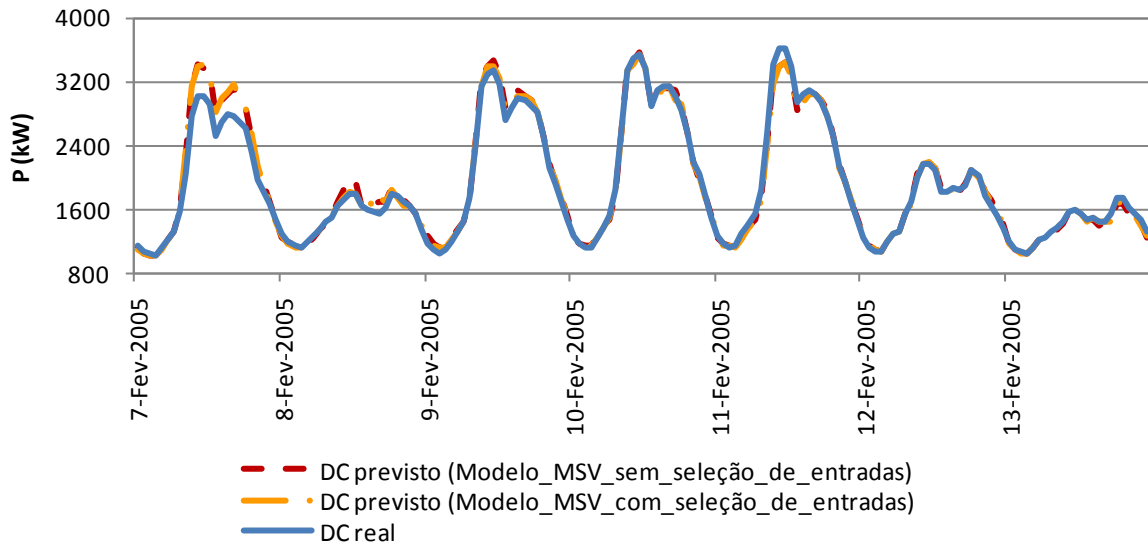
A tendência de distribuição do erro pelas diferentes horas é semelhante à já verificada nos modelos RN, denotando-se contudo algumas discrepâncias significativas para determinadas horas entre o modelo MSV sem seleção de entradas e o modelo MSV com seleção de entradas. Esta perceção justifica-se pelo facto dos diferentes modelos envolverem otimizações por hora, o que pode suscitar algumas variações significativas entre as duas estratégias seguidas.

De forma similar ao que foi analisado com os modelos baseados em redes neuronais, é analisado o impacto que os dias atípicos como feriados e pontes têm na previsão. Deste modo, filtrando os dados de modo a se considerarem apenas estes tipos de dias, são obtidos os erros de previsão indicados na Tabela 6.5.

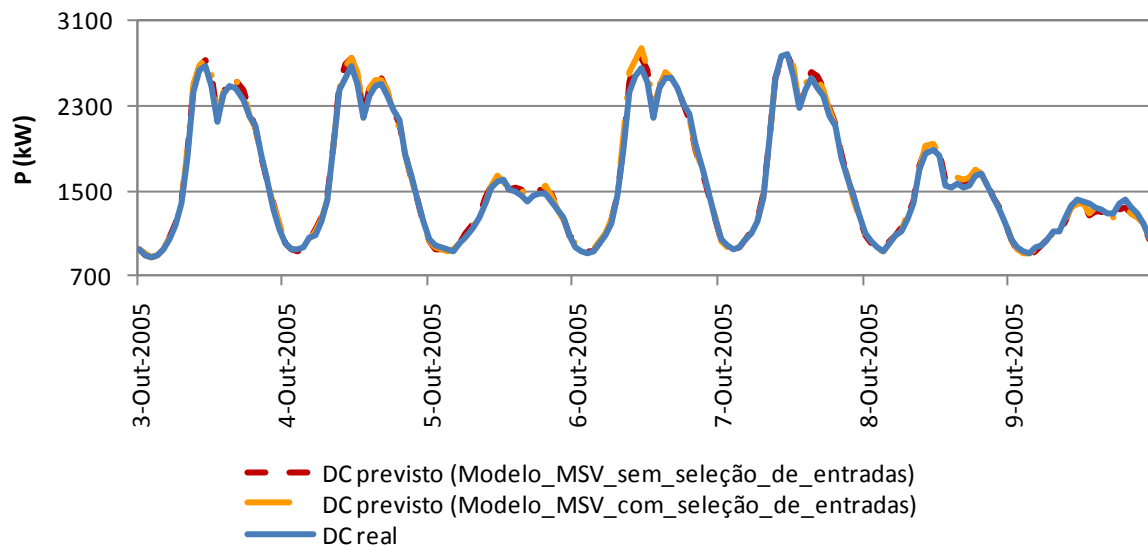
**Tabela 6.5 – Erros de previsão (modelos MSV) verificados no primeiro estudo de caso para dias atípicos (como feriados e pontes)**

Modelo de previsão MSV	EPAM	REPQM
sem seleção de entradas	6,18%	10,05%
com seleção de entradas	6,28%	10,38%

Para que a análise foque o efeito destes tipos de dias, são apresentados diagramas semanais que envolvem previsões de dia seguinte em época de Carnaval (na Figura 6.28) e em feriados fixos nacionais como foi o caso do 5 de outubro de 2005 (na Figura 6.29).



**Figura 6.28 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana atípica com dia de Carnaval – 8 de fevereiro de 2005 (1º EdC)**



**Figura 6.29 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana atípica com feriado de 5 de outubro de 2005 (1º EdC)**

### 6.2.3. Análise de importância das entradas mais relevantes no modelo de previsão baseado em máquinas de suporte vetorial

Uma das formas encontradas para apurar a importância das diferentes entradas usadas foi através da avaliação do impacto que cada vetor de entrada (cada bloco com registos de potência ativa compreendendo 24 horas distintas referentes a diferentes dias históricos, ou referentes à informação do DC sintetizado) poderá ter em cada modelo específico. Numa primeira fase é removido um vetor de entrada do conjunto de 4 vetores usados em cada modelo (num total de 96 diferentes entradas é removido um conjunto de 24 entradas) de cada vez, de modo a determinar qual o vetor de entrada que degrada mais a qualidade de previsão. Nestas diferentes análises

usaram-se sempre os parâmetros MSV ( $\gamma$ ,  $C$  e  $\varepsilon$ ) já encontrados para os 24 modelos horários no cenário base (incluindo os 4 vetores de entrada), sem que este processo contemplasse mais processos de otimização. Perante os modelos otimizados e já apresentados na secção anterior, o objetivo será apurar a sensibilidade que cada modelo manifesta aos diferentes conjuntos de entradas usadas. A Tabela 6.6 sintetiza os resultados obtidos e permite apurar a importância relativa que cada vetor de entrada tem no efeito global de previsão.

**Tabela 6.6 – Efeito de remoção de diferentes vetores de entrada usando o modelo MSV sem seleção de entradas (no primeiro estudo de caso)**

Descrição	Descrição do conjunto de dados	Medidas de Erro	
		EPAM	REPQM
4 vetores de entrada com 24 registos horários de potência incluídos (96 entradas)	Dados de treino	3,12%	5,02%
	Dados de teste	4,87%	6,24%
remoção do vetor de entrada associado ao último dia disponível (96-24= 72 entradas)	Dados de treino	3,88%	5,96%
	Dados de teste	5,15%	6,65%
remoção do vetor de entrada associado ao dia similar da semana anterior (96-24= 72 entradas)	Dados de treino	3,38%	5,34%
	Dados de teste	4,97%	6,37%
remoção do vetor de entrada associado ao dia similar da quinzena anterior (96-24= 72 entradas)	Dados de treino	3,39%	5,32%
	Dados de teste	4,61%	5,93%
remoção do vetor de entrada associado ao DC sintetizado (96-24= 72 entradas)	Dados de treino	4,54%	9,81%
	Dados de teste	5,87%	10,62%

Como se pode concluir, a informação proveniente do DC sintetizado demonstra um papel preponderante na metodologia adotada, sendo que a sua remoção tem um efeito consideravelmente nefasto na qualidade de previsão. Procurando de alguma forma seguir um *ranking* de pesos relativos, é possível verificar-se que o último dia disponível por contemplar dados de consumo decorridos no passado recente torna este vetor de entrada também relevante para o processo. O vetor seguinte no *ranking* é relativo ao dia similar da semana anterior e uma interessante conclusão é que a inclusão de informação do dia similar da quinzena anterior deve mesmo ser evitada, já que se prova que os modelos se tornam mais simples e também mais precisos quando aplicados a dados de teste.

Uma abordagem semelhante foi estabelecida para o modelo MSV que inclui no processo de otimização uma fase de escolha de entradas a usar. Neste caso, a estratégia passou por remover individualmente cada bloco de entradas (relacionadas com as horas contributivas de cada vetor de entrada cujo impacto é avaliado). No processo é reiterada a consideração de manter os parâmetros MSV ( $\gamma$ ,  $C$  e  $\varepsilon$ ) identificados no processo de otimização durante o cenário base, não

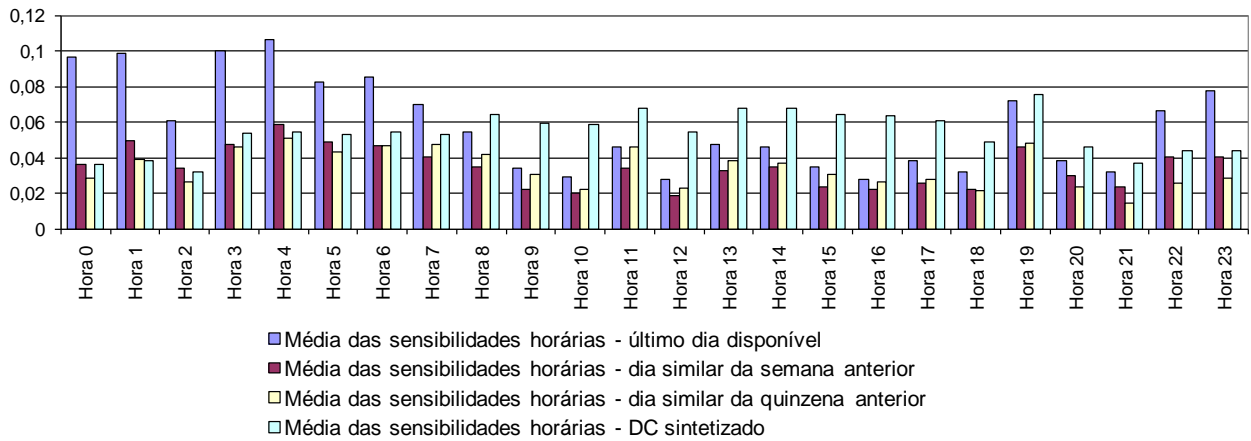
carecendo de novas iterações para nova procura de parâmetros. Os resultados obtidos apresentam-se na Tabela 6.7 podendo-se, em termos globais, retirar as mesmas ilações que já foram apontadas aquando da análise dos resultados da Tabela 6.6.

**Tabela 6.7 – Efeito de remoção de diferentes vetores de entrada usando o modelo MSV com seleção de entradas (no primeiro estudo de caso)**

Descrição	Descrição do conjunto de dados	Medidas de Erro	
		EPAM	REPQM
4 vetores de entrada parcialmente incluídos	Dados de treino	3,23%	5,18%
	Dados de teste	4,84%	6,26%
remoção do vetor de entrada associado ao último dia disponível	Dados de treino	4,03%	6,14%
	Dados de teste	5,16%	6,69%
remoção do vetor de entrada associado ao dia similar da semana anterior	Dados de treino	3,51%	5,52%
	Dados de teste	5,01%	6,43%
remoção do vetor de entrada associado ao dia similar da quinzena anterior	Dados de treino	3,50%	5,49%
	Dados de teste	4,65%	6,02%
remoção do vetor de entrada associado ao DC sintetizado	Dados de treino	4,68%	9,99%
	Dados de teste	5,84%	10,61%

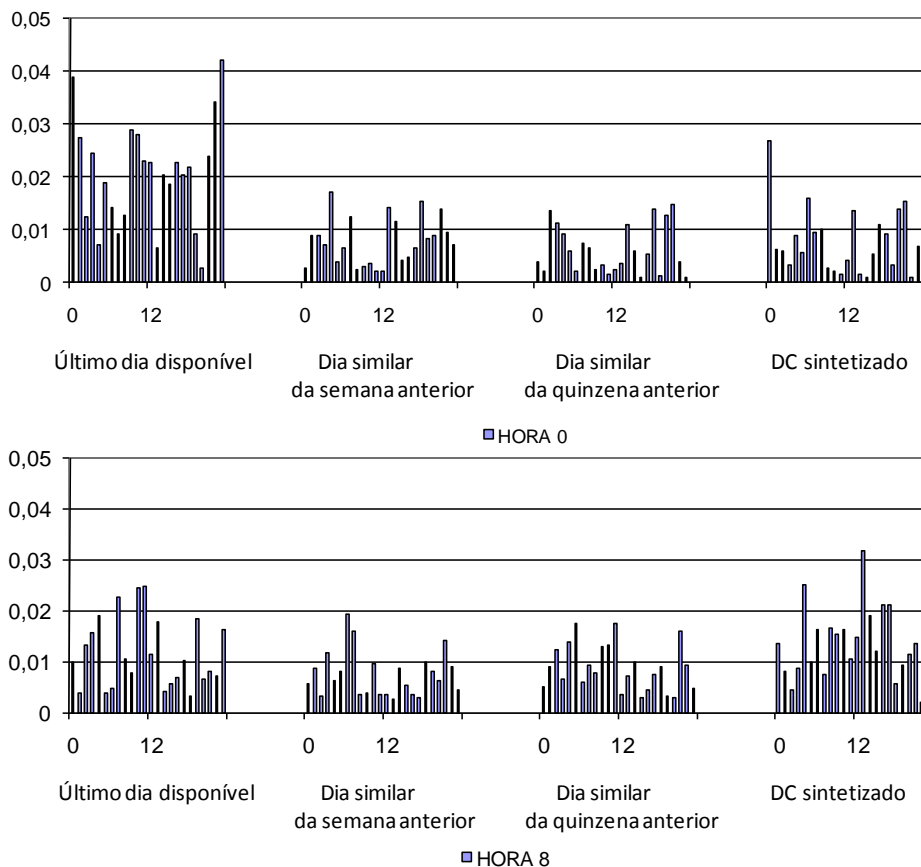
A conclusão mais premente reside novamente no facto da precisão alcançada com os diferentes modelos horários ser muito suscetível de se deteriorar quando se remove a informação do DC sintetizado. Com este exemplo é também consubstanciada a constatação que a informação da quinzena anterior pode vir a ser descartada do conjunto de entradas, já que não apresenta benefícios na estimativa de dados futuros.

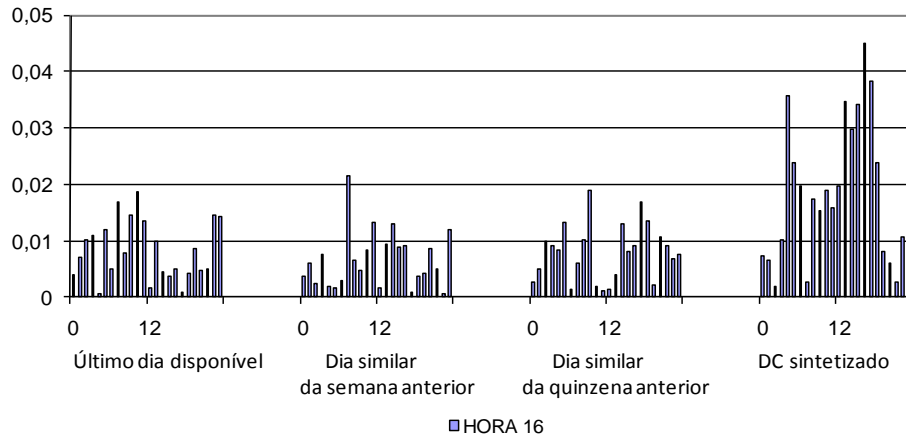
Para procurar validar os resultados anteriormente apresentados, recorreu-se a uma análise de sensibilidades das saídas identificadas em cada modelo horário relativamente às diferentes entradas aplicadas. Uma perspetiva da média das sensibilidades horárias de cada saída em particular a cada vetor de entrada aplicado é apresentada na Figura 6.30. Tal como no modelo de rede neuronal treinado, pode confirmar-se que as entradas referentes ao dia  $d-2$  têm um considerável impacto para prever as horas de menor consumo do dia a prever, enquanto no período entre a hora 8 e a hora 21 existe uma contribuição inequívoca do DC sintetizado. Como é exatamente neste período que existe um maior consumo e também maiores variações no padrão de consumo, é expetável que o indicador de erro se deteriore se a informação do DC sintetizado não estiver disponível ou não for a mais apropriada.



**Figura 6.30 – Médias das sensibilidades horárias de cada saída de Modelo MSV (sem seleção de entradas) relativamente às 24 h de cada vetor de entrada (1º EdC)**

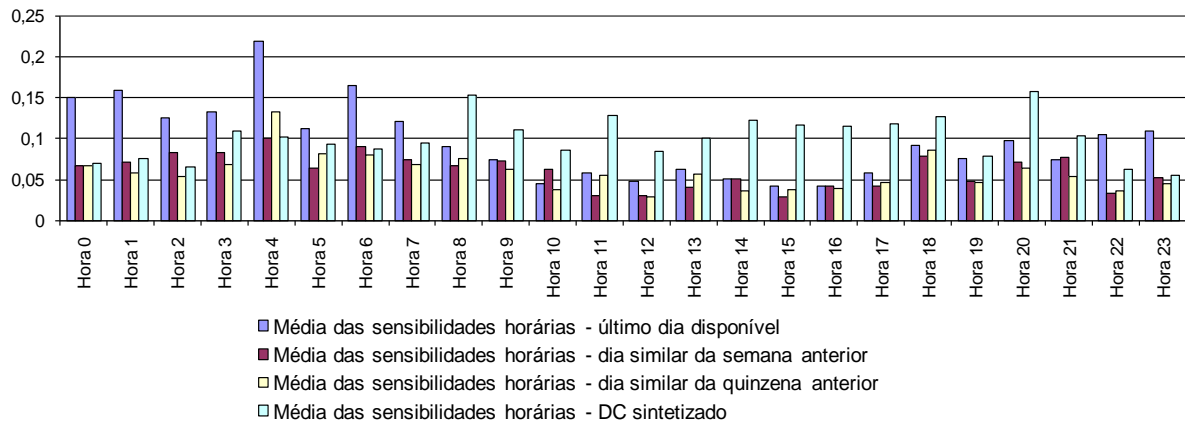
Para consubstanciar esta conclusão, apresenta-se na Figura 6.35 a distribuição de sensibilidades já normalizadas em diferentes modelos horários escolhidos como exemplo.





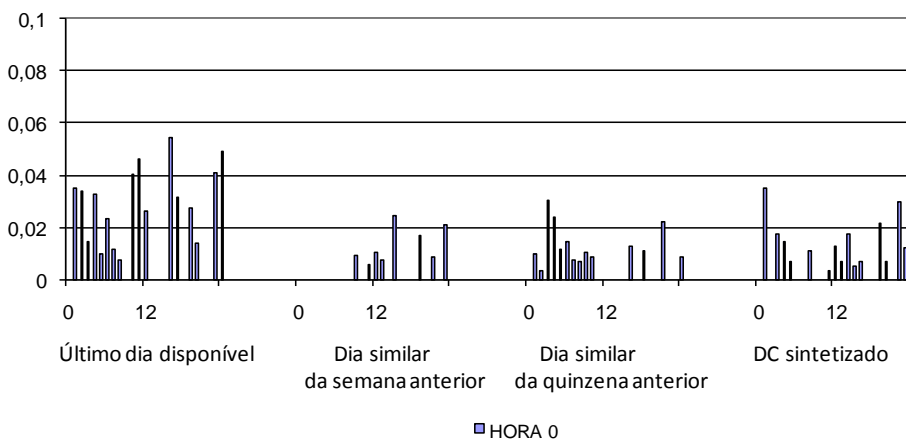
**Figura 6.31 – Sensibilidades manifestadas pelas saídas do modelo MSV (sem seleção de entradas) referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (1º EdC)**

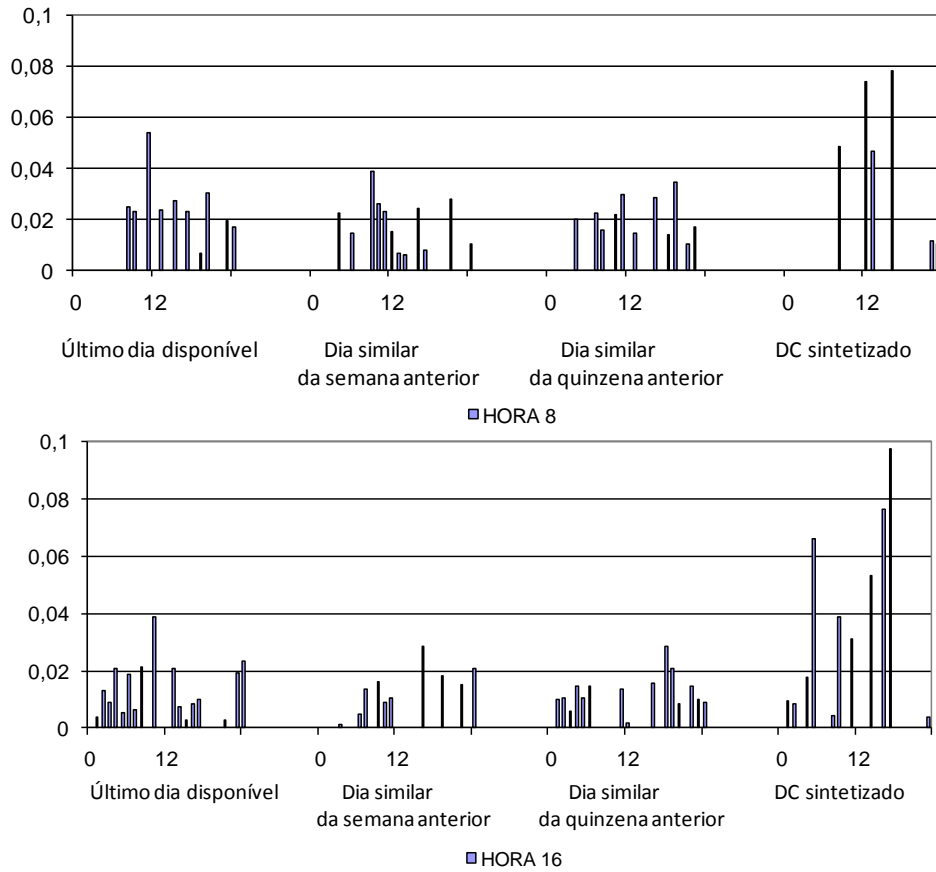
No caso de se considerar a estratégia de seleção de entradas na metodologia MSV podem verificar-se tendências muito aproximadas, como é espelhado na Figura 6.32.



**Figura 6.32 – Médias das sensibilidades horárias de cada saída de Modelo MSV (com seleção de entradas) relativamente às horas contributivas de cada vetor de entrada (1º EdC)**

As distribuições de sensibilidades normalizadas em diferentes modelos horários escolhidos como exemplo são apresentadas na Figura 6.33. De realçar que uma vez que o processo contempla uma fase prévia de seleção de entradas, denotam-se apenas algumas horas específicas de cada vetor de entrada que contribuem para cada modelo horário.





**Figura 6.33 – Sensibilidades manifestadas pelas saídas do modelo MSV (com seleção de entradas) referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (1º EdC)**

Novamente, e para novos modelos agora com seleção prévia de entradas, demarca-se o diagrama de carga sintetizado para favorecer a previsão, tipicamente contribuindo mais acentuadamente nas horas em que se verifica um maior consumo. Para permitir uma análise mais pormenorizada das sensibilidades para diferentes horas do dia, pode consultar-se o Anexo D.



### 6.3. Resultados de previsão com redes neuronais aplicados ao segundo estudo de caso considerado

O segundo caso em estudo é por inerência mais fácil de prever, uma vez que o padrão de consumo é mais homogéneo para diferentes tipos de dia. Os resultados obtidos para os modelos mais robustos de previsão apresentam-se na Tabela 6.8.

**Tabela 6.8 – Resultados de previsão baseada em redes neuronais para o segundo estudo de caso**

Modelo de previsão adotado	Arquitetura da rede com melhores resultados <sup>(*)</sup>	Medidas de Erro	EPAM	REPQM
Modelo I	72:7:24	Dados de treino	3,63%	4,85%
		Dados de teste	5,31%	6,66%
Modelo II	75:6:24	Dados de treino	3,63%	4,79%
		Dados de teste	5,35%	6,85%
Modelo III	96:7:24	Dados de treino	3,61%	4,79%
		Dados de teste	5,14%	6,71%

<sup>(\*)</sup> Número de entradas: Número de neurónios na camada intermédia: Número de saídas

Numa análise comparativa dos resultados para os diferentes modelos, constata-se que o Modelo II, face ao Modelo I, não apresenta uma melhor habilidade para generalização neste estudo de caso, já que os erros nos dados de teste são superiores. Deste modo, para esta série temporal em particular, o processo de descartar variáveis auxiliares usadas para distinguir diferentes dias da semana e a presença de dias especiais é passível de ser seguido, sem que implique perda de acuidade na previsão. Esta conclusão é facilmente suportada pela similaridade demonstrada nos diferentes padrões de carga e também porque a introdução de variáveis auxiliares para além de não serem discriminatórias no processo, tornam o modelo consideravelmente mais complexo sem benefícios práticos. Um estudo particular do efeito dos feriados como foi feito no primeiro estudo de caso também não se justifica para o segundo caso em estudo, já que o comportamento associado ao consumo de eletricidade não é muito distinto entre dias normais laborais e dias feriados. O uso do diagrama de carga sintetizado (como é característico do Modelo III) aparenta ser recomendável, apesar do aumento no erro percentual quadrático médio (face ao Modelo I) nos dados de teste, essencialmente devido a elevados erros individuais encontrados que tendem a projetar consideravelmente este indicador. Como foi já descrito no Capítulo 4, os consumidores que contribuíram para este DC sintetizado não coincidem com os consumidores alimentados por esta saída da subestação de distribuição de Ortigosa, o que compromete a precisão alcançada por este modelo. A opção ideal seria o recurso a um DC sintetizado composto por padrões de carga baseados em dados de consumo de uma amostra de consumidores locais.

As seguintes figuras (da Figura 6.34 à Figura 6.36) pretendem ilustrar alguns exemplos de previsão em diferentes épocas do ano e para períodos de treino e de teste.

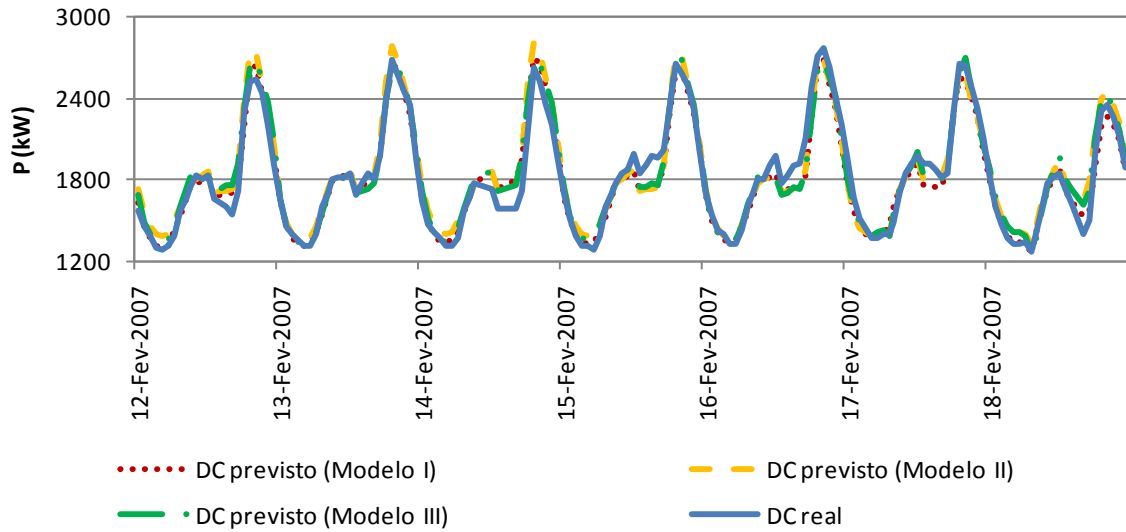


Figura 6.34 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir dos diferentes modelos de previsão (RN) numa semana típica de fevereiro de 2007 (período de treino) (2º EdC)

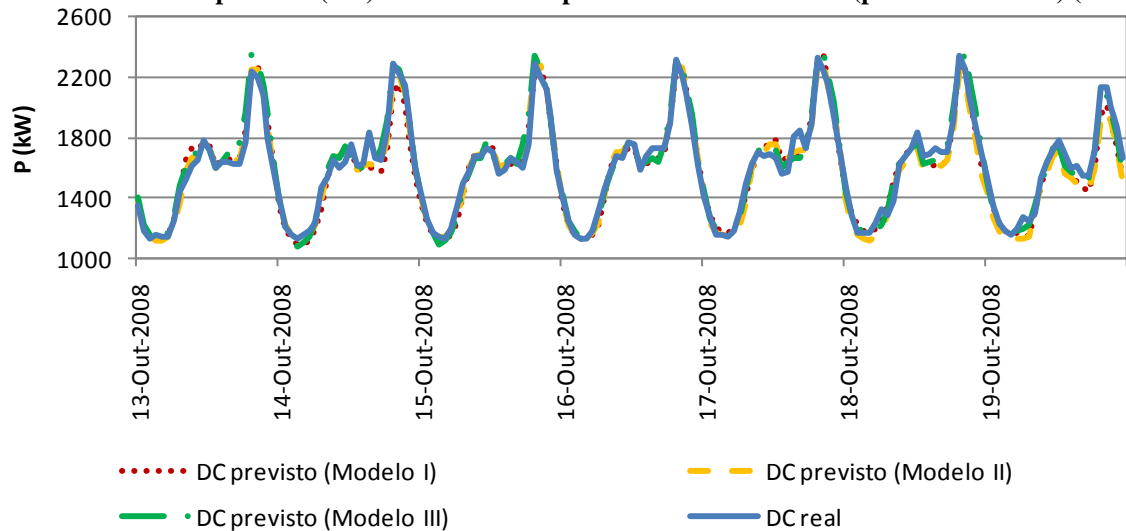


Figura 6.35 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir dos diferentes modelos de previsão (RN) numa semana típica de outubro de 2008 (período de treino) (2º EdC)

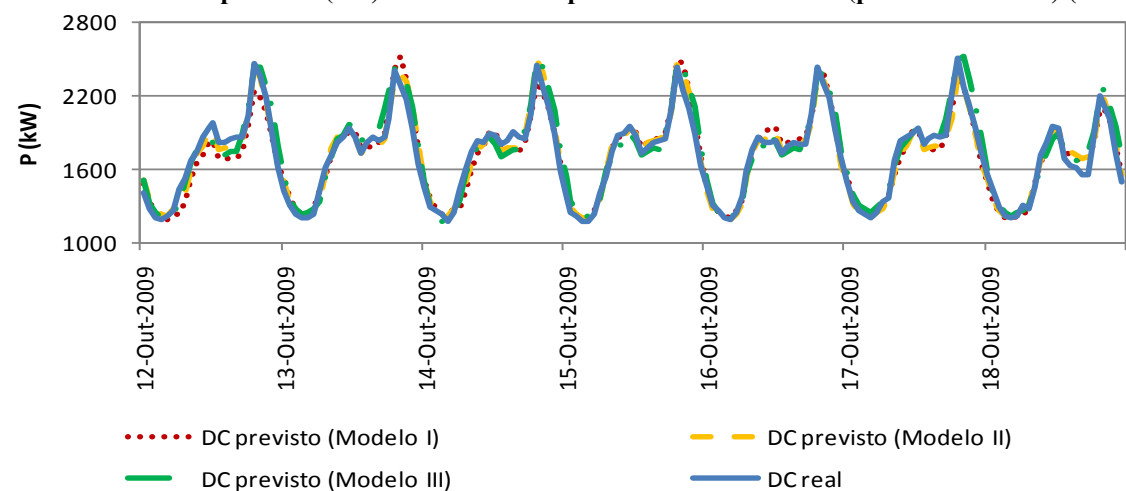
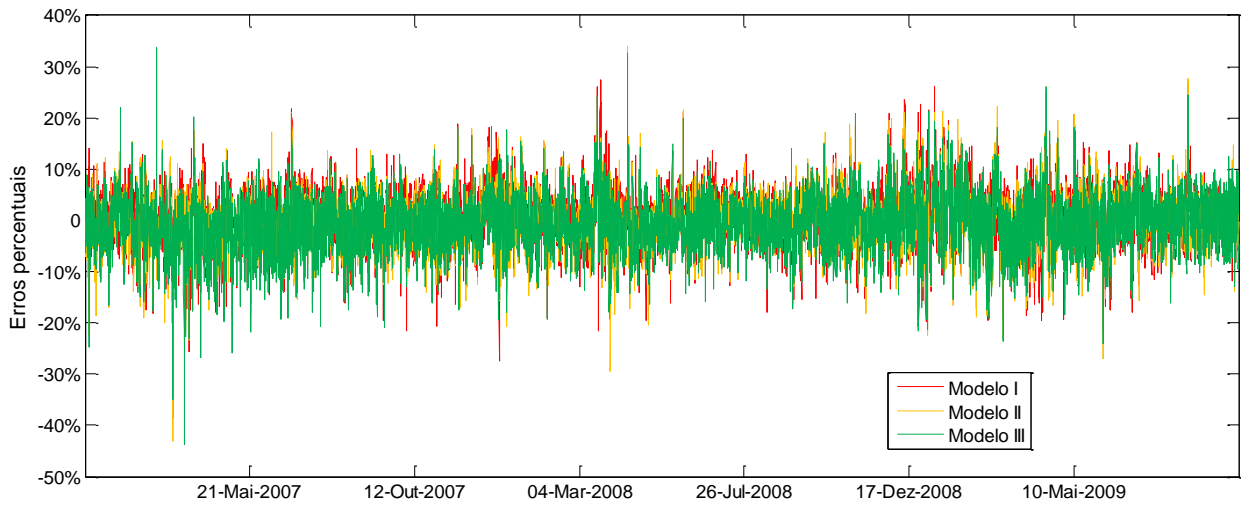
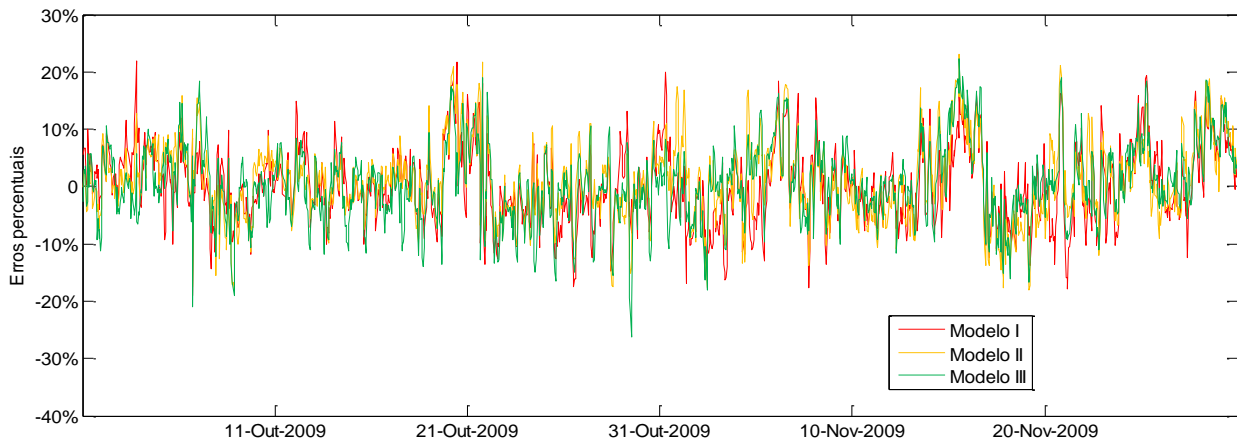


Figura 6.36 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir dos diferentes modelos de previsão (RN) numa semana típica de outubro de 2009 (período de teste) (2º EdC)

O erro percentual tem a distribuição segundo a Figura 6.37 para os dados de treino e segundo a Figura 6.38 para os dados de teste.

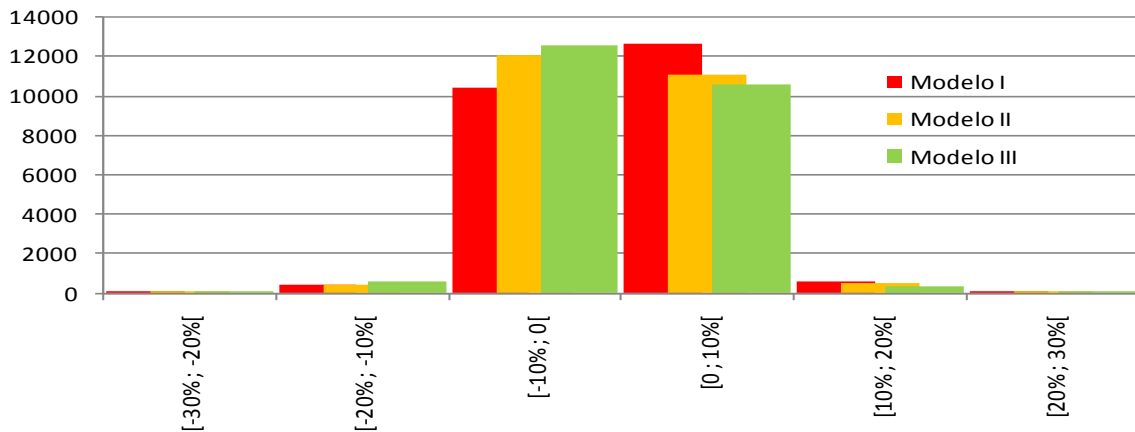


**Figura 6.37 – Distribuição dos erros percentuais por modelo (RN) durante o período de treino considerado (2º EdC)**

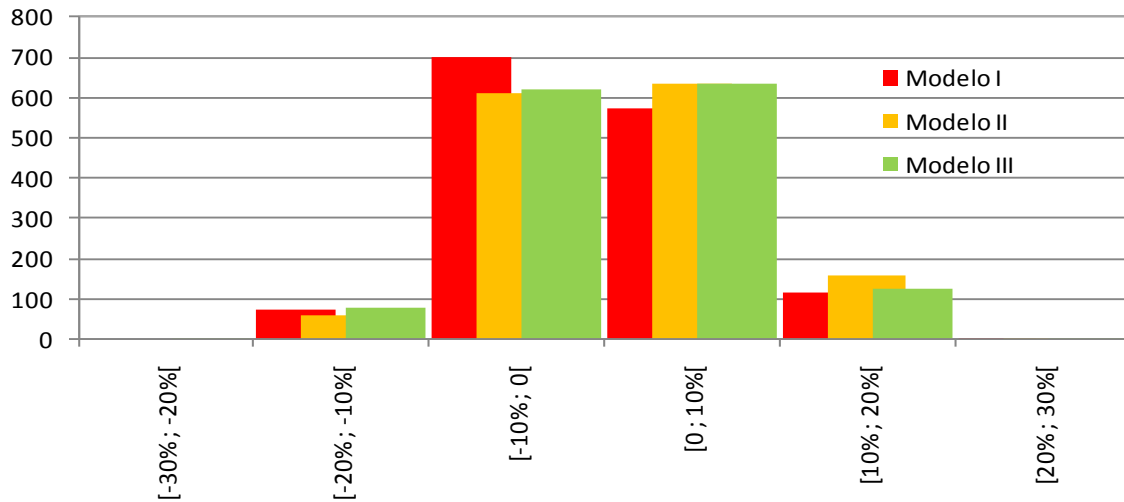


**Figura 6.38 – Distribuição dos erros percentuais por modelo (RN) durante o período de teste considerado (2º EdC)**

Nas seguintes figuras (Figura 6.39 e Figura 6.40) encontram-se histogramas dos erros para os dois subconjuntos de dados considerados. De uma forma sumária é possível verificar que para este estudo de caso não há já tanta discrepância entre os diferentes modelos considerados.

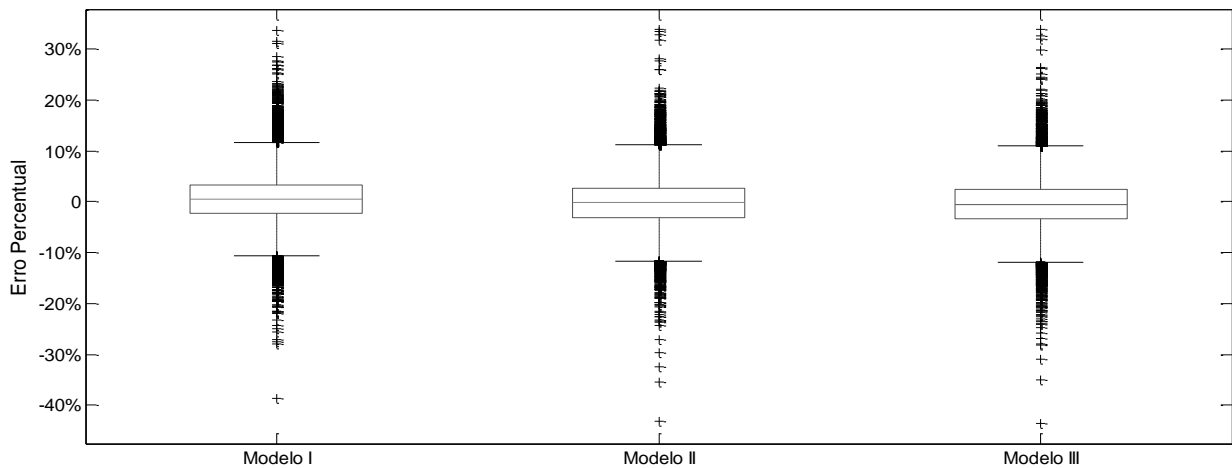


**Figura 6.39 – Histograma dos erros percentuais por modelo (RN) para o período de treino considerado (2º EdC)**

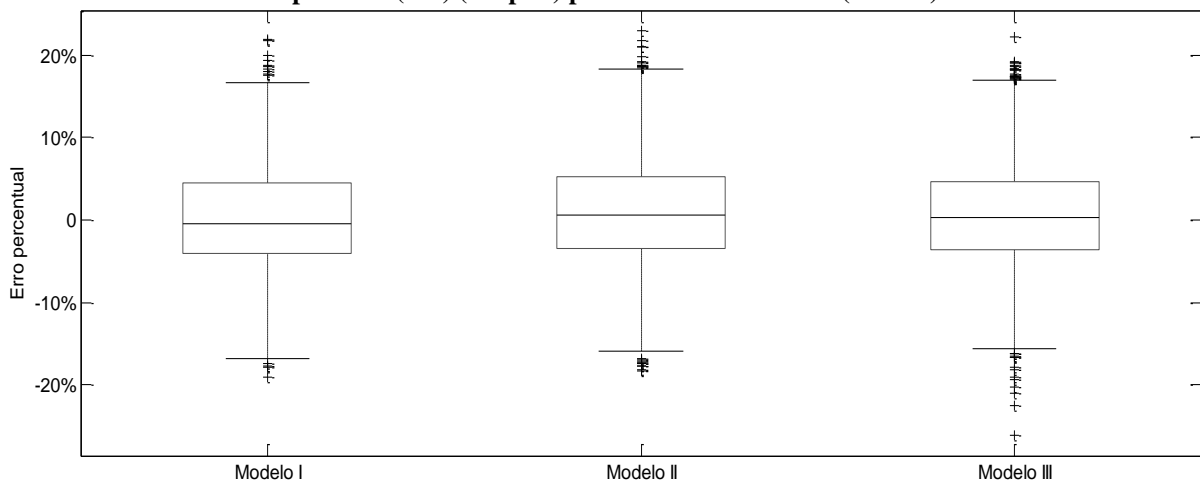


**Figura 6.40 – Histograma dos erros percentuais por modelo (RN) para o período de teste considerado (2º EdC)**

Na Figura 6.41 é apresentada uma distribuição do erro seguindo uma análise estatística centrada na mediana e nos 1º e 3º quartis para os dados de treino. As semelhanças são notórias, o que tornaria de extrema dificuldade a escolha entre os diferentes modelos a partir deste critério.



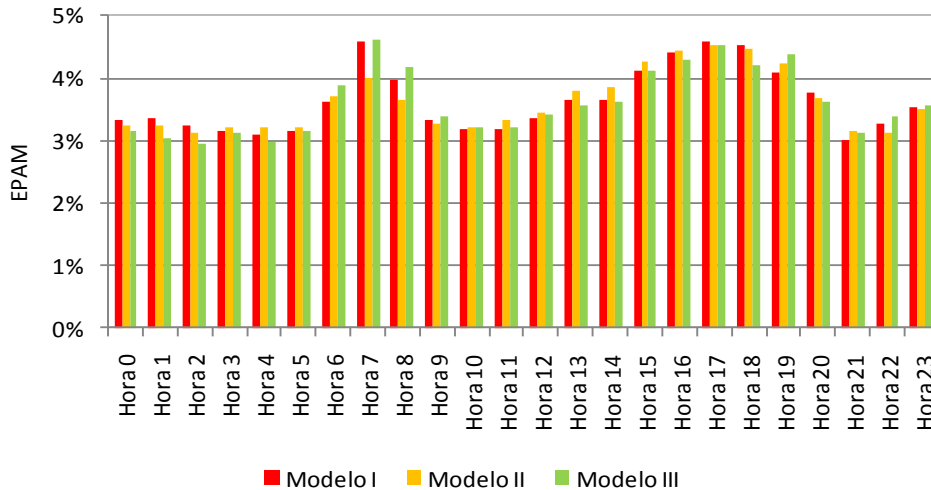
**Figura 6.41 – Comparação na distribuição de erros entre modelos de previsão (RN) (boxplot) para os dados de treino (2º EdC)**



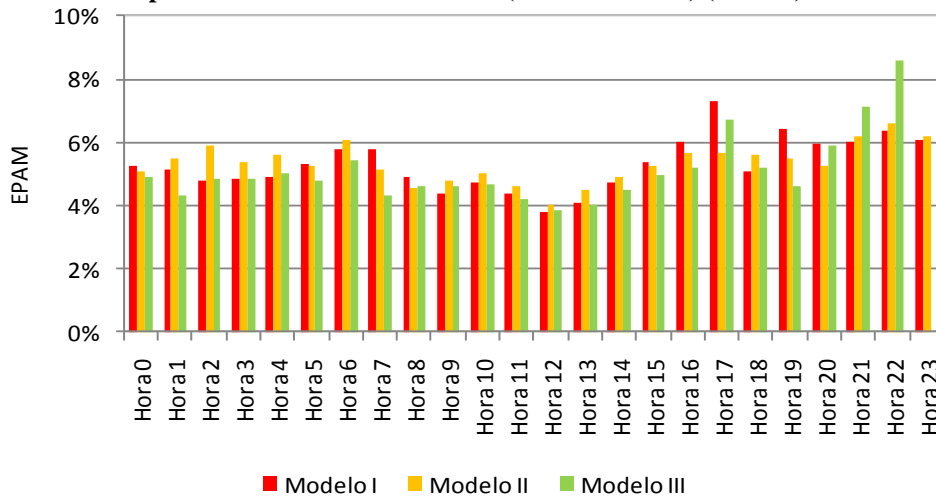
**Figura 6.42 – Comparação na distribuição de erros entre modelos de previsão (RN) (boxplot) para os dados de teste (2º EdC)**

Na análise dos dados de teste que constam na Figura 6.42, verifica-se que o valor da mediana dos erros no Modelo III é a que está mais próxima de zero (0,24%), comparativamente ao Modelo I (-0,51%) e ao Modelo II (0,63%). Em termos de amplitudes interquartis, o Modelo III apresenta uma amplitude menor (de 8,26%), enquanto o Modelo I apresenta uma amplitude de 8,64% e o Modelo II uma amplitude de 8,72%.

A distribuição do erro pelas diferentes horas do dia pode ser analisada na Figura 6.43 (para dados de treino) e na Figura 6.44 (para dados de teste).



**Figura 6.43 – Distribuição de erros em modelos de previsão (RN) para as diferentes horas do dia (dados de treino) (2º EdC)**



**Figura 6.44 – Distribuição de erros em modelos de previsão (RN) para as diferentes horas do dia (dados de teste) (2º EdC)**

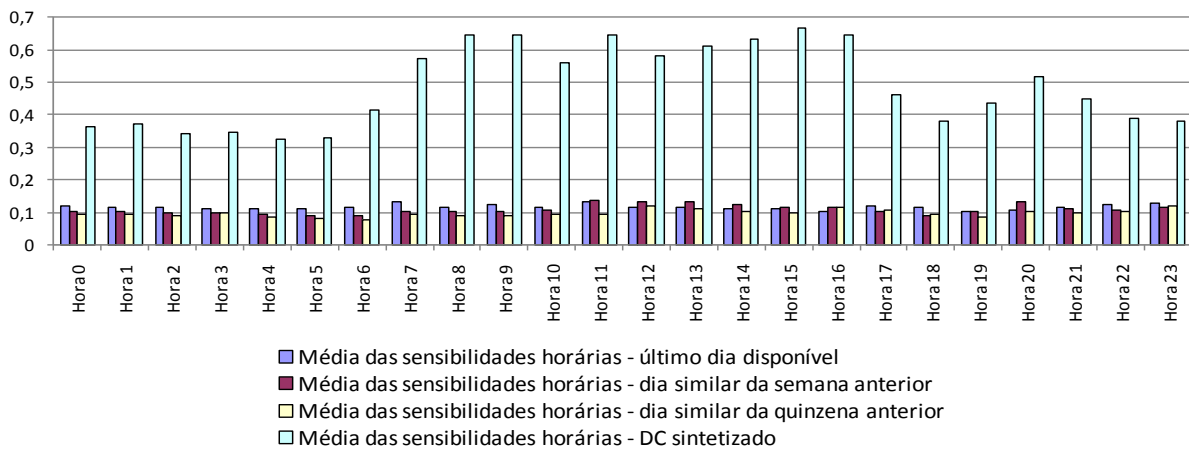
A forma como o erro se distribui pelas diferentes horas do dia é consideravelmente díspar do cenário apresentado no primeiro estudo de caso. As horas para as quais há uma maior probabilidade de se evidenciar um aumento do erro dizem respeito ao período do início da manhã (entre as 6 e as 8 horas) quando no setor residencial há um maior consumo nas habitações, bem como no final do dia (entre as 17 e as 20 horas) com o regresso dos ocupantes às respetivas habitações e conseqüente consumo de energia elétrica que provém de sistemas de

iluminação, cargas de aquecimento, ventilação e/ou ar-condicionado e de equipamento elétrico de cozinha.

### 6.3.1. Análise de sensibilidades da rede neuronal adotada no segundo estudo de caso considerado

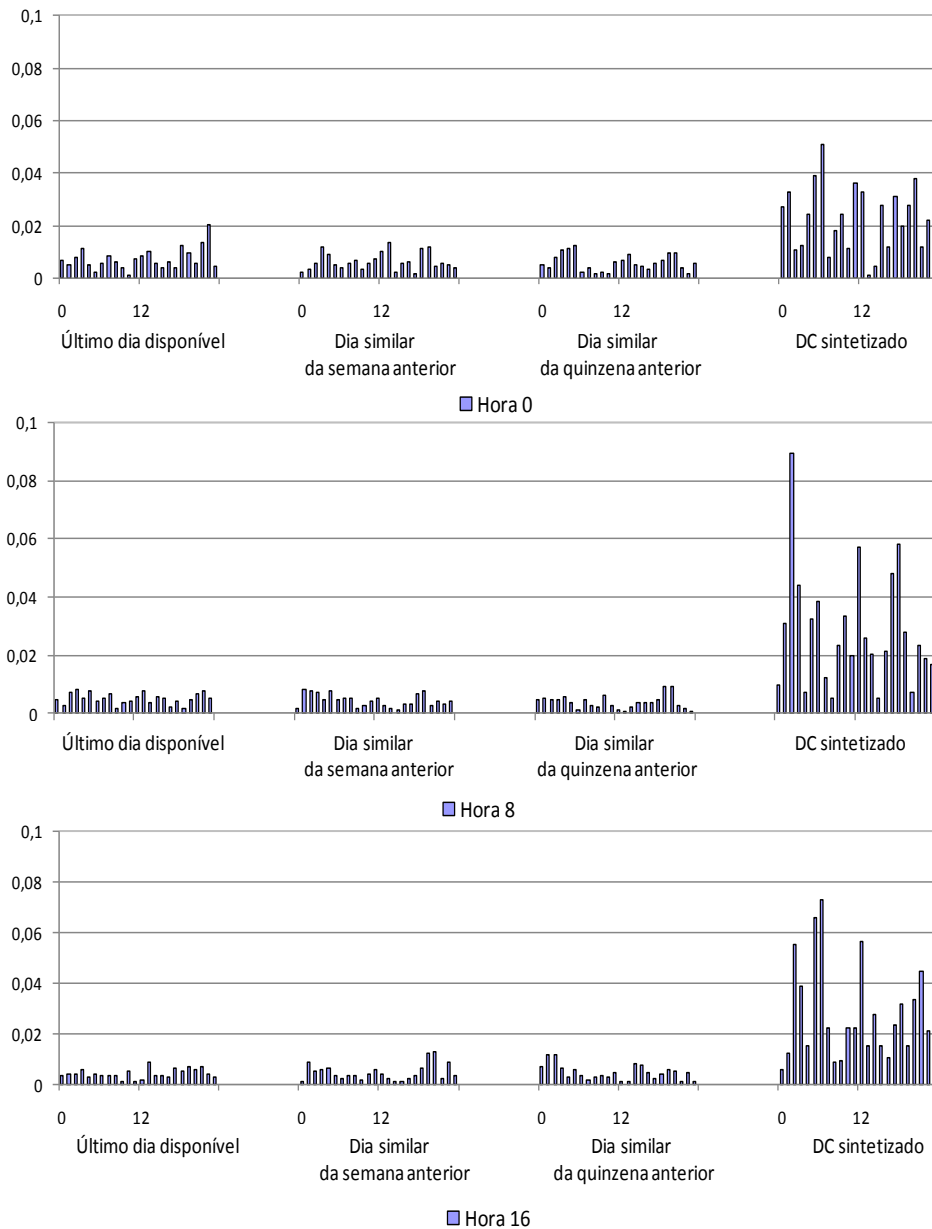
Esta subsecção torna a pretender avaliar a importância relativa que o diagrama de carga sintetizado toma relativamente a outros vetores de entrada usados (último dia disponível, dia similar da semana anterior e dia similar da quinzena anterior), agora com referência ao segundo estudo de caso. É novamente analisado em maior detalhe o Modelo III, já que, é aquele que inclui o diagrama de carga sintetizado na informação das entradas e também por ter resultado em erros de previsão mais reduzidos. O modelo que recorre a 96 entradas distintas para prever os 24 registos horários de potência ativa do dia seguinte merece um estudo de sensibilidades que será abordado de forma análoga ao que foi apresentado na Secção 6.1.1 para o primeiro estudo de caso.

A Figura 6.45 ilustra os resultados obtidos, onde se evidencia uma relevância do vetor associado ao diagrama de carga sintetizado face às entradas que dependem do passado recente, agora verificável para todas as horas do dia.



**Figura 6.45 – Médias das sensibilidades horárias de cada saída do Modelo RN relativamente às 24 h de cada vetor de entrada (2º EdC)**

É também novamente demonstrada a importância das diferentes entradas individuais para cada hora que se pretende prever, sendo apresentados alguns exemplos na Figura 6.46. As sensibilidades apresentadas nesta figura encontram-se novamente normalizadas, com a soma das diferentes sensibilidades para cada saída a ser unitária.



**Figura 6.46 – Sensibilidades manifestadas pelas saídas do modelo RN referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (2º EdC)**

Com a apresentação destes exemplos, é reforçada a ideia que o diagrama de carga sintetizado é genericamente uma contribuição relevante para prever os registos horários de potência ativa do dia seguinte. De uma forma sucinta pode-se avançar que o Modelo III ficará consideravelmente sensível à informação proveniente do DC sintetizado, o que torna expectável assumir-se que se esta informação dependesse de perfis de consumidores locais, o impacto que este facto teria na redução dos erros de previsão seria notável.

Mais uma vez torna-se importante concluir que o modelo multivariável já treinado não dependerá única e exclusivamente de horas congéneres dos dias anteriores e do DC sintetizado quando se pretende prever uma hora específica do dia seguinte, mas também de um conjunto de

outras horas do dia. Uma abordagem mais detalhada acerca da aplicação da análise de sensibilidades ao modelo adotado apresenta-se no Anexo B.

#### 6.4. Aplicação da metodologia baseada em máquinas de suporte vetorial para a previsão de consumos relativos ao segundo estudo de caso considerado

Uma vez mais foram usados como entradas do processo de previsão os registos horários de potência ativa referentes ao último dia disponível, do dia similar da semana anterior, do dia similar da quinzena anterior e do DC sintetizado.

Os modelos MSV criados para este propósito foram previamente sujeitos a um processo de otimização, implicando necessariamente uma busca de parâmetros MSV ajustados a este novo caso em estudo. Os critérios de otimização associados a *Simulated Annealing* foram os mesmos seguidos aquando do 1º estudo de caso (com os valores máximos referentes aos diferentes parâmetros expressos na Tabela 5.1 e parâmetros e critérios associados ao processo de *Simulated Annealing* revelados na Tabela 5.2).

##### 6.4.1. Descrição do processo de criação de diferentes modelos horários otimizados

Após o processo de otimização é possível apresentar alguns dados úteis, nomeadamente os parâmetros MSV obtidos e que resultaram no menor erro de previsão no subconjunto de validação, o número de vetores de suporte que influenciam diretamente cada modelo, bem como o valor do erro percentual absoluto médio para cada um dos subconjuntos de dados considerado. Na Tabela 6.9 são apresentados os referidos indicadores para alguns dos modelos horários.

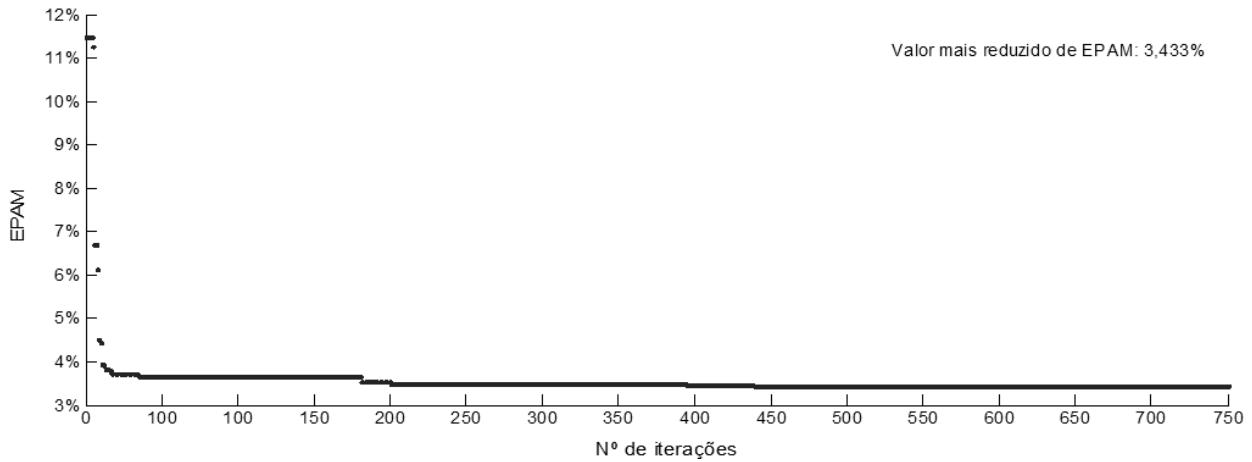
**Tabela 6.9 – Descrição de alguns dos modelos MSV obtidos (para horas específicas) (2º EdC)**

Modelo	Parâmetros MSV obtidos		Número de vetores de suporte	Dados de treino	Dados de validação	Dados de teste
				EPAM	EPAM	EPAM
Hora 0	$\gamma$	0,00119	518 <sup>(*)</sup>	2,50%	3,43%	4,08%
	C	14,057				
	$\varepsilon$	0,0992				
Hora 7	$\gamma$	0,00192	363 <sup>(*)</sup>	3,07%	3,79%	4,54%
	C	8,063				
	$\varepsilon$	0,201				
Hora 18	$\gamma$	0,00236	769 <sup>(*)</sup>	3,41%	4,70%	4,47%
	C	2,793				
	$\varepsilon$	0,0115				

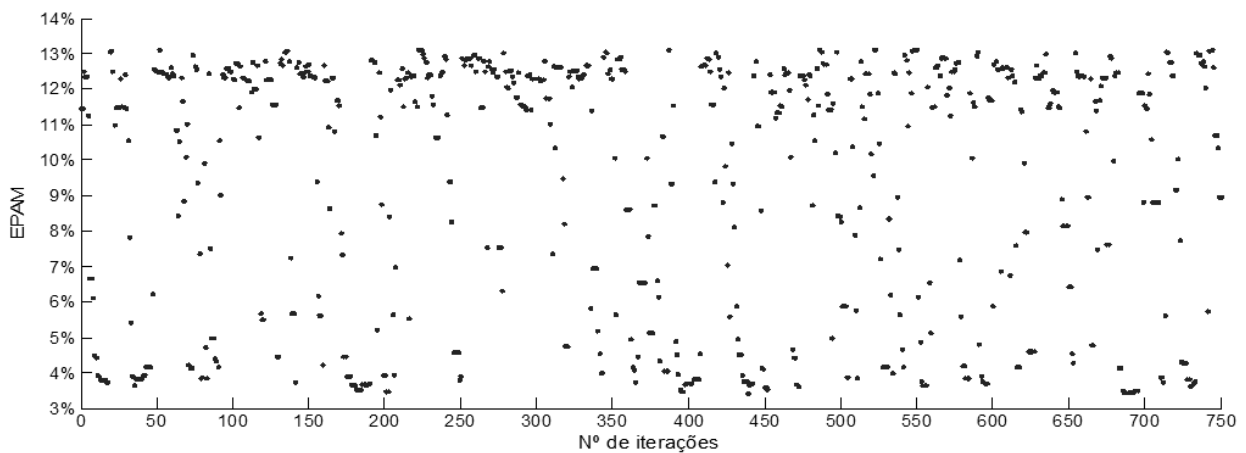
<sup>(\*)</sup> – Número de vetores de suporte reconhecidos de um total de 806 vetores considerados no período de treino



O processo de otimização para o modelo horário referente à hora 0 pretende ser mais explícito com o recurso às seguintes figuras. A Figura 6.47 caracteriza a evolução do valor mínimo do EPAM verificado nos dados de validação e a Figura 6.48 ilustra com maior detalhe os resultados obtidos para diferentes combinações exploradas ao longo do processo de otimização.



**Figura 6.47 – Evolução do mínimo resultado de EPAM para o conjunto de validação – modelo referente à hora 0 (2º EdC)**



**Figura 6.48 – Evolução do resultado EPAM verificado no conjunto de validação durante o processo de otimização – modelo referente à hora 0 (2º EdC)**

Para este modelo horário pode verificar-se que o indicador de previsão se situa entre os 3,43% e sensivelmente 13%, reforçando-se a necessidade de garantir alguma acuidade na determinação dos parâmetros MSV. A partir de outros modelos analisados durante o processo de otimização (alguns exemplos são apresentados no Anexo C) identificam-se erros de previsão na ordem dos 22,4%, especialmente em horas para as quais o consumo é superior.

6.4.2. Apresentação de resultados de previsão com modelos baseados em máquinas de suporte vetorial

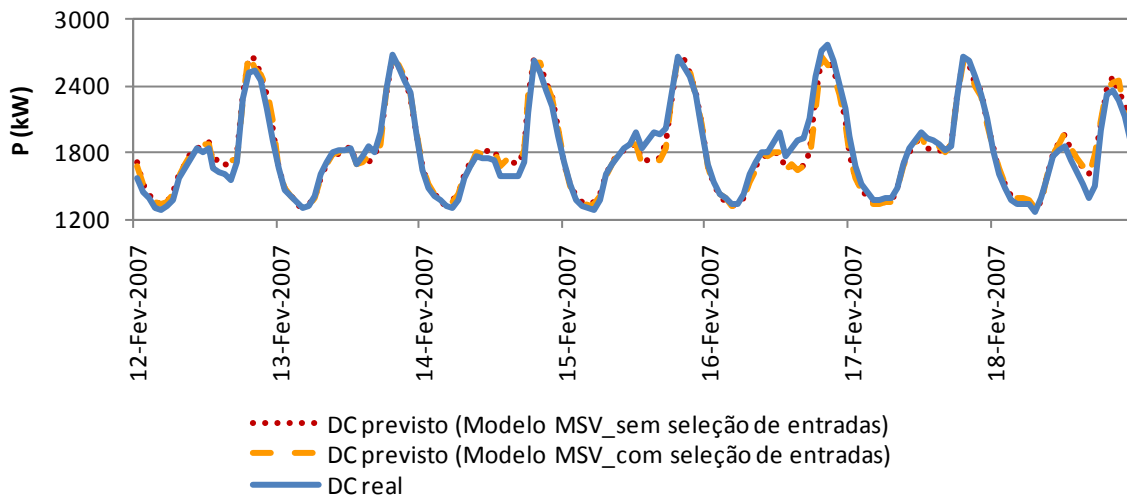
A Tabela 6.10 apresenta os resultados obtidos com um total de 96 entradas (semelhante ao Modelo III de redes neuronais), bem como resultados obtidos com o recurso à seleção de parâmetros explicativos conforme foi apresentado na Secção 5.2.3.

**Tabela 6.10 – Resultados de previsão baseada em máquinas de suporte vetorial para o segundo estudo de caso**

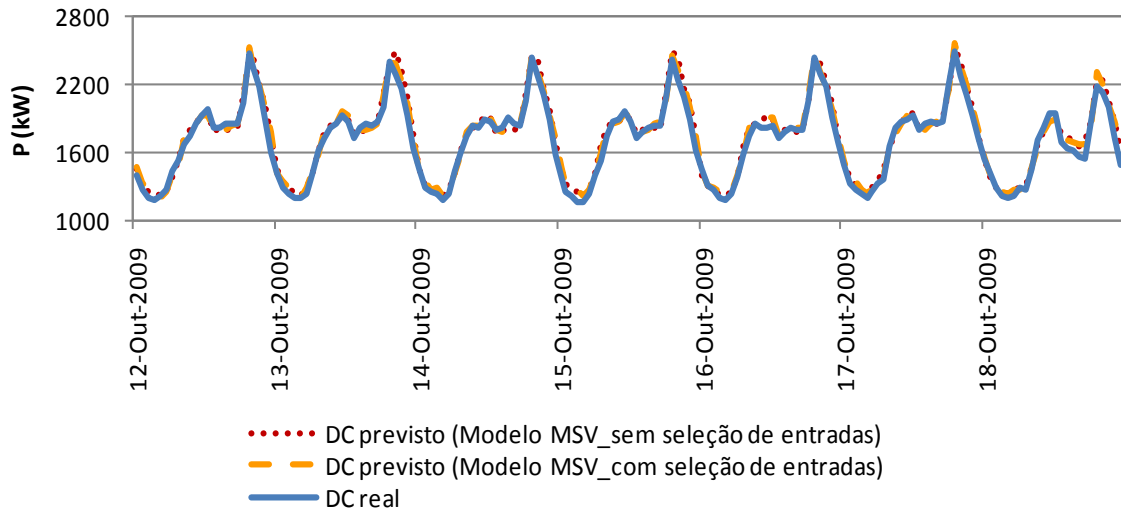
Descrição do modelo adotado	Medidas de Erro	EPAM	REPQM
total de 96 entradas	Dados de treino	3,00%	4,07%
	Dados de teste	4,65%	6,17%
com seleção de parâmetros explicativos	Dados de treino	2,98%	4,13%
	Dados de teste	4,77%	6,27%

É novamente observável que os modelos apresentam erros no conjunto de treino consideravelmente reduzidos, enquanto os erros no subconjunto de teste tornam a ser mais condizentes com os erros encontrados pelos modelos RN. Os dois modelos MSV apresentam indicadores de previsão muito semelhantes, pretendendo-se mais uma vez demonstrar com esta comparação que, uma eventual redução do número de entradas pode não condicionar a capacidade de generalização para simulação de modelos de previsão em novos dados apresentados.

Alguns resultados comparativos numa base semanal são apresentados na Figura 6.49 e na Figura 6.50.

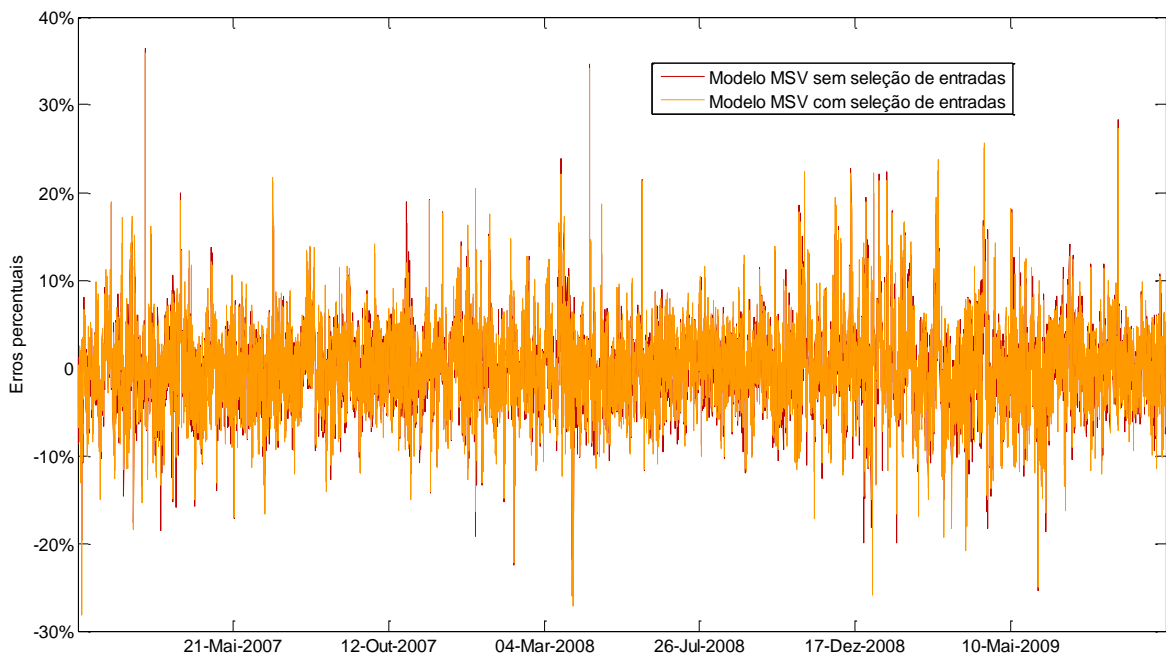


**Figura 6.49 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana típica de fevereiro de 2007 (período de treino) (2º EdC)**



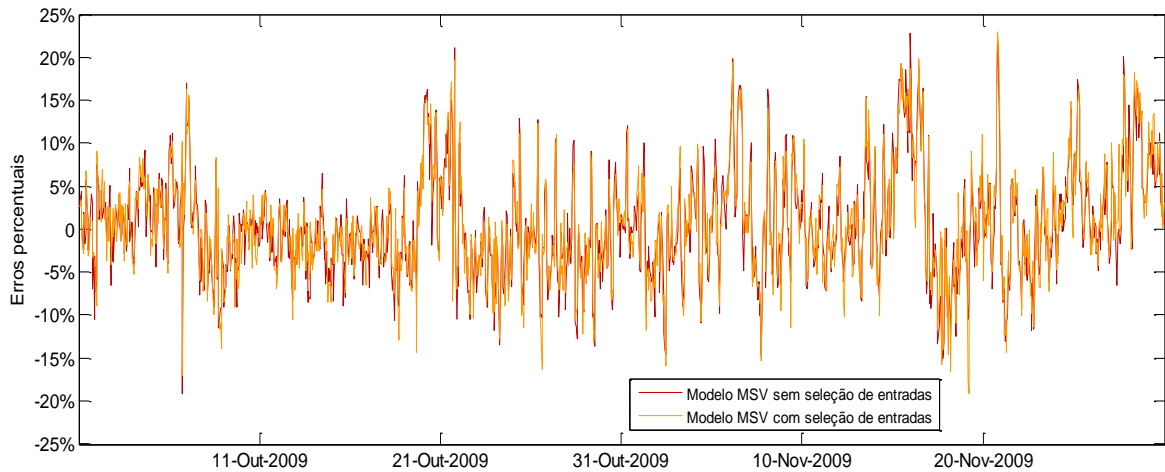
**Figura 6.50 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de modelos de previsão (MSV) numa semana típica de outubro de 2009 (período de teste) (2º EdC)**

A distribuição do erro para os dois subconjuntos de dados em análise encontra-se na Figura 6.51 (para os dados de treino) e na Figura 6.52 (para os dados de teste). Na distribuição de erros de treino existem registos de algum modo atípicos que são comuns às duas variantes de modelos MSV.



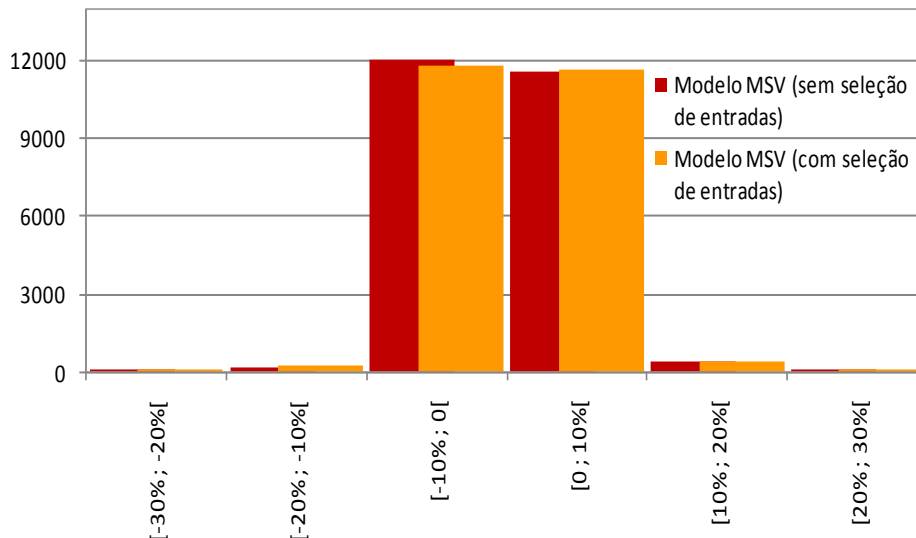
**Figura 6.51 – Distribuição dos erros percentuais por modelo (MSV) durante o período de treino considerado (2º EdC)**

Os erros de teste demonstram também um equilíbrio entre as duas variantes com os registos atípicos (embora numa gama menor do que o caso anterior) a serem concordantes entre os dois modelos.



**Figura 6.52 – Distribuição dos erros percentuais por modelo (MSV) durante o período de teste considerado (2º EdC)**

Na Figura 6.53 pode analisar-se um histograma de distribuição dos erros no período de treino e na Figura 6.54 uma análise semelhante para dados de teste. Os resultados apontam novamente para distribuições semelhantes nos diferentes modelos seguidos. Apesar de existirem erros percentuais no período de treino que em módulo são superiores a 30%, como estes registos são em número residual tal efeito não teria expressão, pelo que esta gama foi filtrada no histograma apresentado na Figura 6.53.



**Figura 6.53 – Histograma dos erros de previsão com modelos MSV para os períodos de treino (2º EdC)**

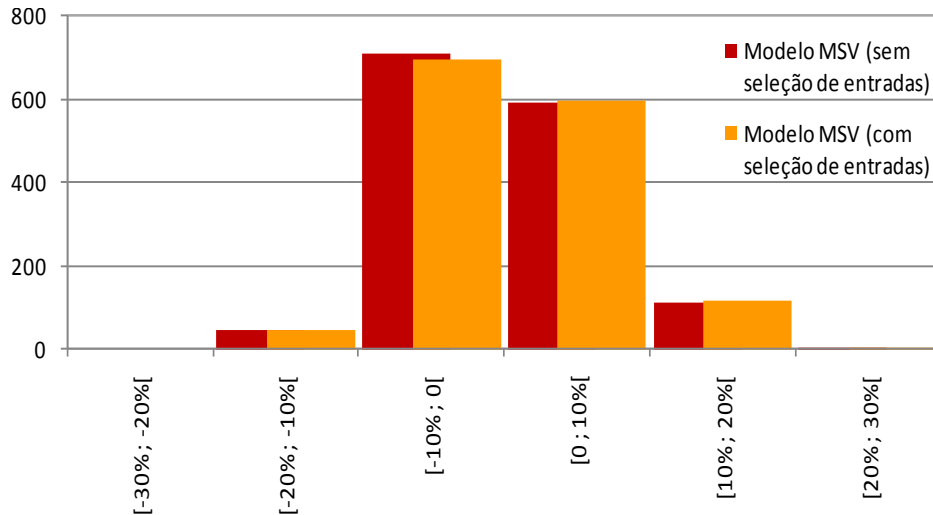


Figura 6.54 – Histograma dos erros de previsão com modelos MSV para os períodos de teste (2º EdC)

Como foi verificada a existência de registos atípicos, apresenta-se também uma representação *boxplot* dos erros de treino e de teste na Figura 6.55 e Figura 6.56, respetivamente, por forma a destacar estes *outliers*. Novamente podem ser confirmados registos reduzidos de mediana dos erros, bem como dos dois quartis. Nos dados de treino a similaridade é notória entre os dois modelos seguidos, enquanto nos dados de teste o modelo com seleção de entradas apresenta uma amplitude inter-quartil ligeiramente superior (passando-se de uma amplitude interquartil de 7,02% para o caso de inclusão de todos os parâmetros de entrada, para uma amplitude interquartil de 7,48% quando se opta por um determinado conjunto de entre o conjunto total de parâmetros de entrada).

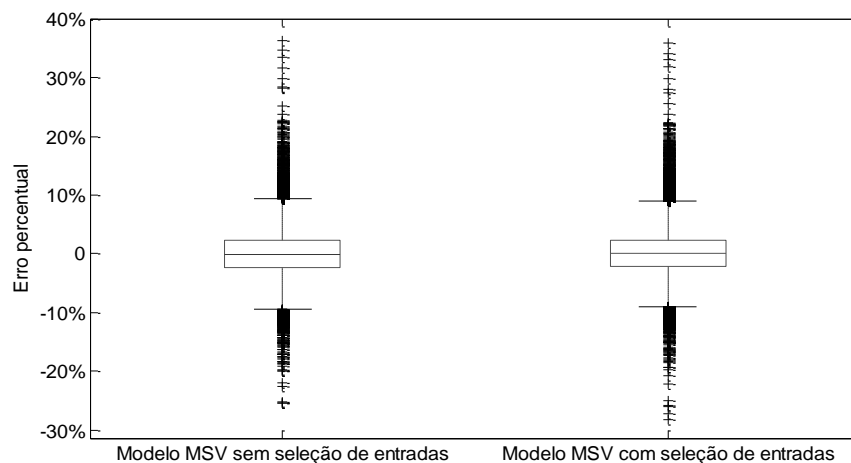
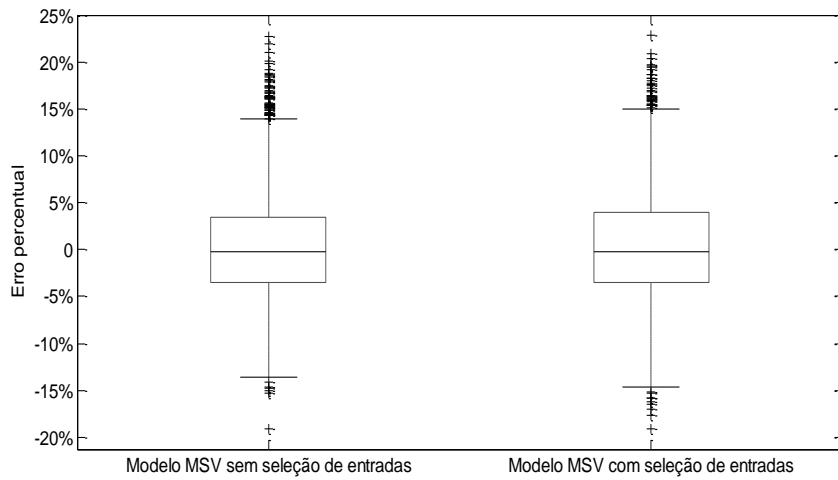
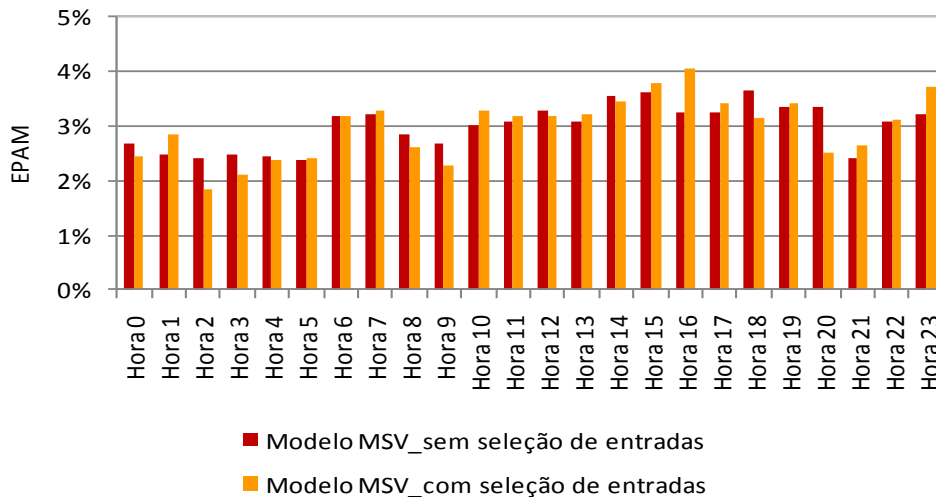


Figura 6.55 – Comparação na distribuição de erros entre modelos de previsão (MSV) (boxplot) para os dados de treino (2º EdC)

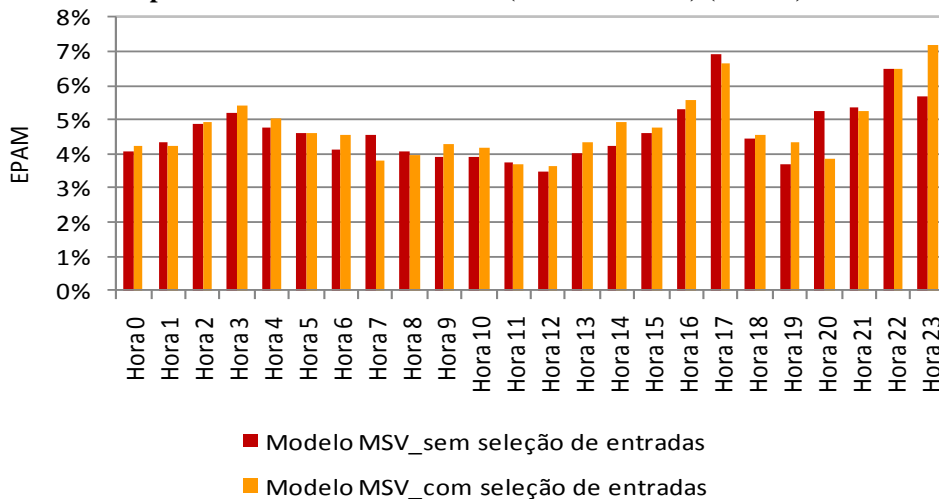


**Figura 6.56 – Comparação na distribuição de erros entre modelos de previsão (MSV) (boxplot) para os dados de teste (2º EdC)**

A distribuição do erro pelas diferentes horas do dia apresenta-se na Figura 6.57 (para os dados de treino) e na Figura 6.58 (para os dados de teste).



**Figura 6.57 – Distribuição de erros em modelos de previsão (MSV) para as diferentes horas do dia (dados de treino) (2º EdC)**



**Figura 6.58 – Distribuição de erros em modelos de previsão (MSV) para as diferentes horas do dia (dados de teste) (2º EdC)**

A forma como o erro se distribui para distintas horas do dia não é já tão comparável entre dados de treino e dados de teste, sendo que estes últimos diferem também das distribuições de erros verificadas com os modelos RN já apresentados nas Figura 6.43 e Figura 6.44. Como os modelos de suporte vetorial se caracterizam por ter um processo de otimização dedicado a cada hora em específico, fica patente uma certa dependência da escolha de parâmetros MSV que podem tender a afetar a capacidade de generalização para dados de teste. Com esta constatação pode ser concluído que a perenidade desta metodologia é efetivamente posta em causa quando não se verificar uma seleção criteriosa dos parâmetros provenientes da otimização que no estudo em causa é baseada em *Simulated Annealing*.

#### 6.4.3. Análise de importância das entradas mais relevantes no modelo de previsão baseado em máquinas de suporte vetorial

O procedimento já seguido para o primeiro estudo de caso foi estendido para a análise do segundo estudo de caso. A perspetiva passa por procurar verificar o impacto que cada vetor de entrada em particular apresenta para o indicador de erro de previsão. Nesse sentido, a Tabela 6.11 apresenta o efeito de remoção de cada um dos vetores de entrada individualmente.

**Tabela 6.11 – Efeito de remoção de diferentes vetores de entrada usando o modelo MSV sem seleção de entradas (no segundo estudo de caso)**

Descrição	Descrição do conjunto de dados	Medidas de Erro	
		EPAM	REPQM
4 vetores de entrada com 24 registos horários de potência incluídos (96 entradas)	Dados de treino	3,00%	4,07%
	Dados de teste	4,65%	6,17%
remoção do vetor de entrada associado ao último dia disponível (96-24= 72 entradas)	Dados de treino	3,85%	5,19%
	Dados de teste	5,50%	7,09%
remoção do vetor de entrada associado ao dia similar da semana anterior (96-24= 72 entradas)	Dados de treino	3,16%	4,28%
	Dados de teste	4,70%	6,21%
remoção do vetor de entrada associado ao dia similar da quinzena anterior (96-24= 72 entradas)	Dados de treino	3,14%	4,23%
	Dados de teste	4,85%	6,43%
remoção do vetor de entrada associado ao DC sintetizado (96-24= 72 entradas)	Dados de treino	3,27%	4,46%
	Dados de teste	4,84%	6,26%

A consulta desta tabela não resulta em conclusões explícitas, tal como sucedeu para o primeiro estudo de caso. Por um lado, o efeito de remoção de cada um dos vetores de entrada nunca é tão

nefasto para a previsão como se verificou no primeiro estudo de caso acima apresentado. Por outro lado, aparenta existir um certo equilíbrio entre os diferentes efeitos verificados, como se pode verificar muito facilmente com a remoção da informação da semana anterior e com a remoção da informação da quinzena anterior. Uma justificação reside na própria natureza da série temporal em questão, que sendo muito homogénea (sem grandes variações entre diferentes dias de semana e entre épocas do ano), torna a previsão mais tolerante à falta de um dos vetores de entrada considerados. Outra razão que deve ser apontada é que o teste do efeito de remoção de cada um dos vetores de entrada implicou sempre um novo treino com os novos parâmetros de entrada considerados, mantendo apenas os parâmetros MSV que foram ajustados ao cenário base. Uma justificação não menos importante reside no facto dos dados que definem o DC sintetizado não serem dedicados à localização exata dos dados reais de consumo monitorizado na saída de Monte Real, o que pode atenuar a importância que este vetor de entrada aparenta ter na previsão. Havendo novo processo de treino, e com resultados de previsão que se revelam tão próximos, revela-se uma tarefa de considerável complexidade despistar se o efeito de alteração da acuidade na previsão se deve ao conjunto de entradas usadas em cada modelo, ou mais propriamente ao processo de treino decorrente.

A mesma constatação foi verificada para o modelo MSV que inclui no processo de otimização uma fase de escolha de entradas a usar. A estratégia desta análise resulta novamente de uma remoção individual de cada bloco de entradas (relacionadas com as horas contributivas de cada vetor de entrada cujo impacto é avaliado) e no processo são mantidos os parâmetros MSV ( $\gamma$ ,  $C$  e  $\varepsilon$ ) identificados no processo de otimização durante o cenário base. Os resultados obtidos apresentam-se na Tabela 6.12.

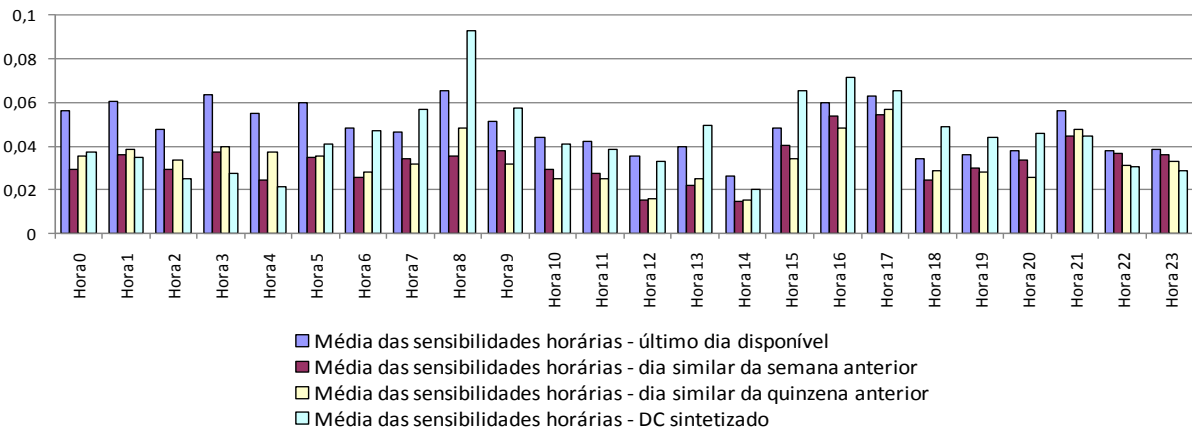
Algumas das considerações tecidas aquando da análise da Tabela 6.11 permanecem válidas, tornando-se de alguma forma complexo e, por ventura, pouco rigoroso tomar considerações em termos de *ranking* de importâncias dos diferentes vetores de entrada considerados.



**Tabela 6.12 – Efeito de remoção de diferentes vetores de entrada usando o modelo MSV com seleção de entradas (no segundo estudo de caso)**

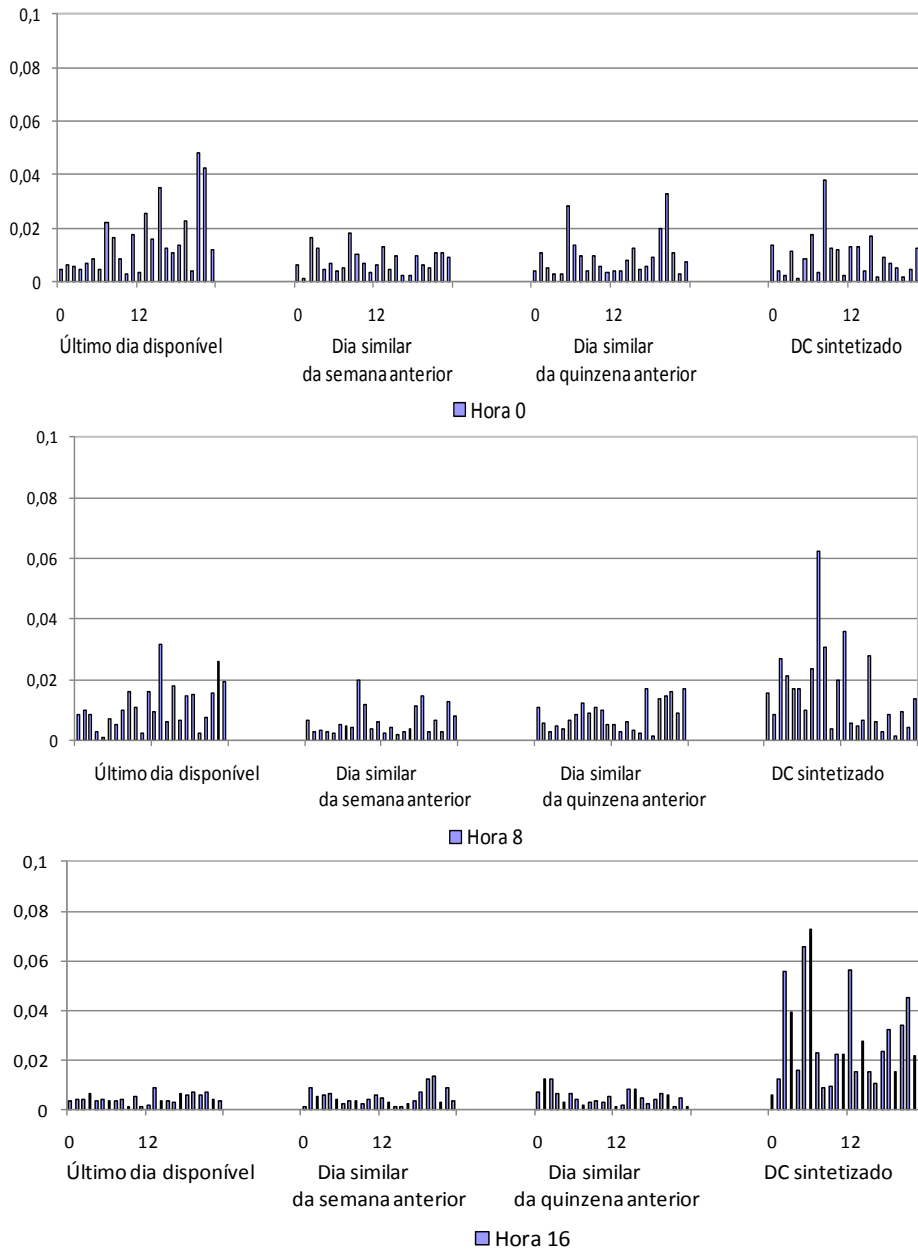
Descrição	Descrição do conjunto de dados	Medidas de Erro	
		EPAM	REPQM
4 vetores de entrada parcialmente incluídos	Dados de treino	2,98%	4,13%
	Dados de teste	4,77%	6,27%
remoção do vetor de entrada associado ao último dia disponível	Dados de treino	3,97%	5,42%
	Dados de teste	5,68%	7,40%
remoção do vetor de entrada associado ao dia similar da semana anterior	Dados de treino	3,20%	4,38%
	Dados de teste	4,90%	6,46%
remoção do vetor de entrada associado ao dia similar da quinzena anterior	Dados de treino	3,13%	4,30%
	Dados de teste	4,94%	4,84%
remoção do vetor de entrada associado ao DC sintetizado	Dados de treino	3,34%	4,63%
	Dados de teste	4,84%	6,28%

O critério que irá substanciar o apuramento da importância relativa dos diferentes vetores de entrada será através de análise de sensibilidades por meio da derivada parcial. Para os diferentes modelos horários que incorporam todas as horas de cada vetor de entrada, obtém-se a média das sensibilidades manifestadas por cada hora em particular, de acordo com a Figura 6.59.



**Figura 6.59 – Médias das sensibilidades horárias de cada saída de Modelo MSV (sem seleção de entradas) relativamente às 24 h de cada vetor de entrada (2º EdC)**

Particularizando algumas das horas podem ser obtidas distribuições das sensibilidades normalizadas, tal como se pretende ilustrar na Figura 6.60.

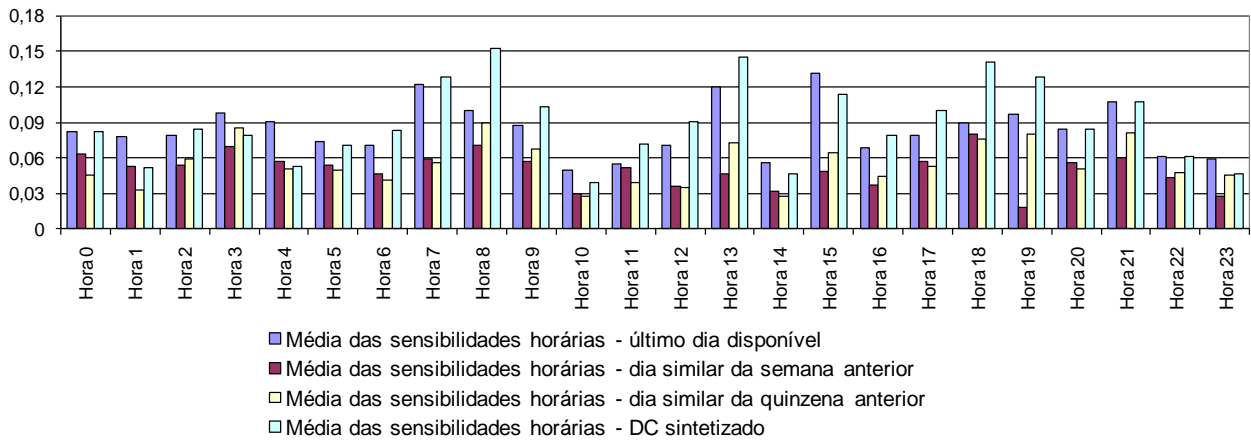


**Figura 6.60 – Sensibilidades manifestadas pelas saídas do modelo MSV (sem seleção de entradas) referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (2º EdC)**

As figuras anteriores permitem concluir que, apesar da informação que consta no DC sintetizado poder ser pouco discriminatória para a série temporal que se pretende prever (por não se tratar exatamente dos mesmos consumidores), os modelos MSV adotados são sensíveis à informação que resulta deste vetor de entrada. Com esta franca dependência dos modelos a estes registos de entrada, pode ser adiantado que os erros de previsão teriam margem para serem consideravelmente diminuídos, caso se utilizasse um DC sintetizado resultante da informação local dos consumidores que compõem a amostra em questão.

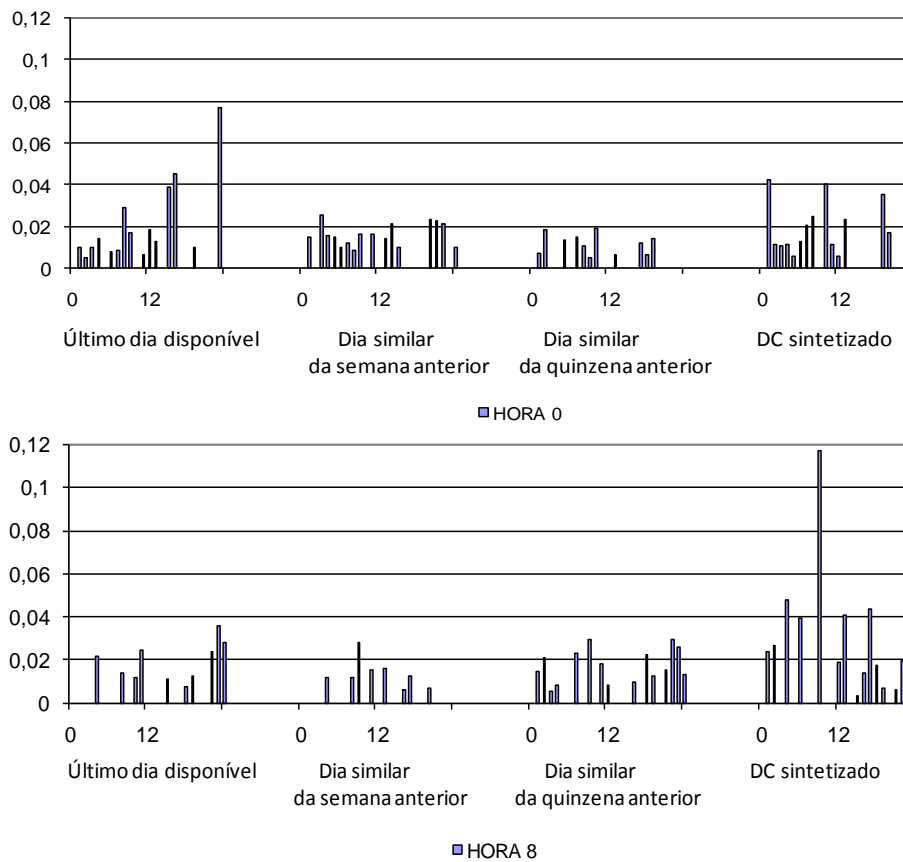
Estendendo a mesma análise para o cenário que contempla uma seleção criteriosa das entradas a usar, podem ser validadas as conclusões já apresentadas. Na Figura 6.61 apresentam-se as

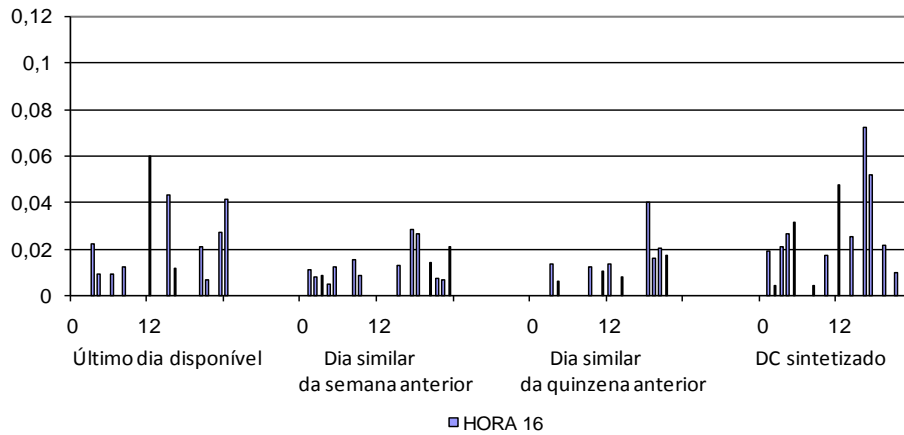
médias das sensibilidades horárias de cada saída de Modelo MSV relativamente aos registos horários de entrada que efetivamente contribuem para cada modelo.



**Figura 6.61 – Médias das sensibilidades horárias de cada saída de Modelo MSV (com seleção de entradas) relativamente às horas contributivas de cada vetor de entrada (2º EdC)**

A Figura 6.62 exemplifica a distribuição das sensibilidades normalizadas para alguns modelos horários considerados.





**Figura 6.62 – Sensibilidades manifestadas pelas saídas do modelo MSV (com seleção de entradas) referentes às horas 0, 8 e 16 do dia a prever relativamente ao conjunto das entradas individuais aplicadas (2º EdC)**

Numa análise global aos resultados apresentados é possível destacar que, apesar das limitações associadas ao DC sintetizado que foi utilizado, estes dados demonstraram ter uma pertinência considerável, já que os modelos ficam indubitavelmente sensíveis a estes registos.

### **6.5. Avaliação de modelo apenas baseado em informação de consumo resultante de perfis e em variáveis meteorológicas para previsão aplicada ao segundo estudo de caso**

Esta subsecção visa dar uma perspetiva do rigor de previsão de consumos do dia seguinte alcançável quando não existem disponíveis registos de consumo em determinadas áreas geográficas. Neste âmbito, no caso de não haver dados disponíveis ou de se pretender evitar os custos associados à necessidade de instalação de equipamentos de medição horária de consumos energéticos e consequente investimento financeiro e humano para gerir uma base com os dados de consumo registados, poder-se-á adotar um modelo de previsão que beneficia apenas de informação que resulta de perfis de carga dos consumidores envolvidos. Neste sentido, e para o segundo estudo de caso considerado, foram desenvolvidos modelos de previsão que apenas dependem de informação resultante do DC sintetizado e de variáveis meteorológicas disponíveis para o efeito (o cruzamento entre as variáveis meteorológicas e seu efeito no consumo energético foi já escamoteado na Secção 4.2.4 deste documento). A descrição dos diferentes modelos esquematiza-se da seguinte forma:

- o primeiro modelo resulta apenas da informação proveniente dos perfis de consumo, resultando assim num modelo que se baseia no designado diagrama de carga sintetizado;

- o segundo modelo recorre à mesma informação do primeiro modelo, incluindo ainda informação meteorológica, como as temperaturas máximas e mínimas dos dias a prever. O modelo incorpora um total de 26 entradas;

– uma vez que, em termos operacionais não são conhecidas com a antecedência de um dia as temperaturas máximas e mínimas que efetivamente se irão registar, o cenário ideal seria dispor de informação acerca da previsão local das variáveis climatéricas relativas ao dia a prever. Como esta informação não se encontrava disponível no âmbito deste estudo, optou-se por considerar um terceiro modelo que adota registos máximos e mínimos de temperatura no dia imediatamente anterior ao dia a prever. Tratando-se de valores extremos de temperatura e não de registos horários, pressupõe-se que estes registos sejam já conhecidos no momento em que se pretende estabelecer a previsão para o dia seguinte.

Os melhores resultados de previsão que foram obtidos com modelos únicos multivariáveis de redes neuronais apresentam-se em seguida.

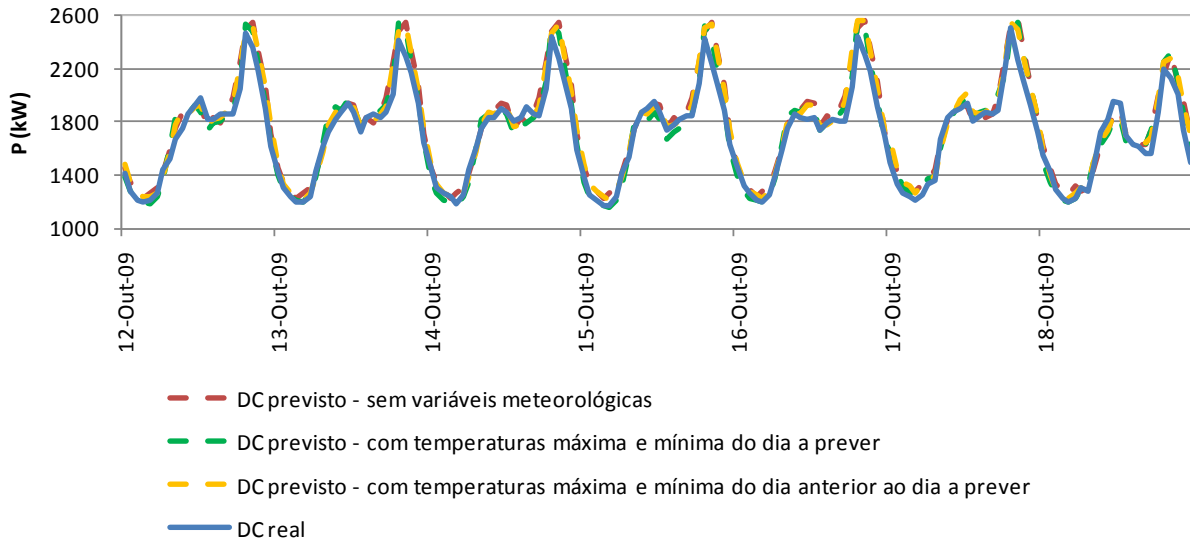
**Tabela 6.13 – Resultados de previsão com redes neuronais baseadas apenas em informação de perfis de consumo e em variáveis climatéricas**

Modelo de previsão adotado	Arquitetura da rede com melhores resultados (*)	Medidas de Erro	EPAM	REPQM
sem variáveis meteorológicas	24:13:24	Dados de treino	5,37%	7,02%
		Dados de teste	6,83%	8,91%
com variáveis meteorológicas relativas ao dia a prever	26:13:24	Dados de treino	3,78%	5,18%
		Dados de teste	5,66%	7,50%
com variáveis meteorológicas do dia anterior ao dia a prever	26:13:24	Dados de treino	4,15%	5,82%
		Dados de teste	6,09%	8,09%

(\*) Número de entradas: Número de neurónios na camada intermédia: Número de saídas

Os resultados da Tabela 6.13 permitem verificar que esta metodologia, não recorrendo a dados de consumo histórico recente, apresenta indicadores de previsão relativamente reduzidos, apesar de se terem usado perfis de consumo não afetos à zona geográfica em questão. Para além desse facto, a distribuição do consumo total desta saída de subestação por diferentes classes de consumidores também não era conhecida, tendo sido por isso estimada a contribuição das diferentes classes, o que naturalmente será outro fenómeno agravante na aproximação que se pretende ao diagrama de carga real. Ainda assim, obtêm-se resultados que parecem promissores para que o estudo seja extensível a zonas remotas em que se tenha dificuldade na obtenção e/ou no tratamento dos dados monitorizados. O efeito das variáveis meteorológicas é digno de registo, já que, como o padrão de consumo tem uma grande contribuição do setor residencial, é

expectável que haja influência significativa da temperatura no consumo. O terceiro modelo seguido, por não fazer coincidir diretamente o efeito da temperatura no dia em questão que se pretende prever, não apresenta um tão forte contributo, mas ainda assim melhor do que seria se simplesmente se ignorasse a influência climatérica. Na Figura 6.63 é possível verificar a previsão obtida pelos três modelos estudados, durante uma semana do período de teste.



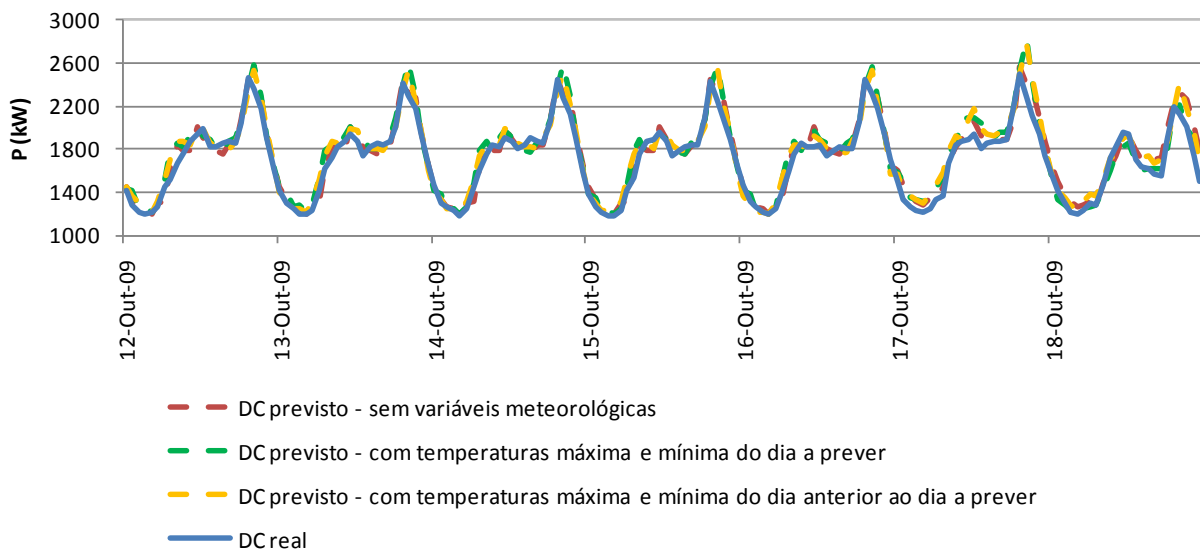
**Figura 6.63 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (RN) com recurso ao DC sintetizado e a variáveis meteorológicas (semana durante o período de teste)**

A previsão foi também testada com máquinas de suporte vetorial, sendo concebidos 24 modelos horários distintos para cada dia que se pretende prever. A escolha dos parâmetros MSV regeu-se pela mesma estratégia já apresentada anteriormente, recorrendo-se a um algoritmo de otimização e aos mesmos critérios já adotados anteriormente. Os resultados de previsão obtidos para as diferentes variantes consideradas apresentam-se na Tabela 6.14.

**Tabela 6.14 – Resultados de previsão com máquinas de suporte vetorial baseadas apenas em informação de perfis de consumo e em variáveis climatéricas**

Modelo de previsão adotado	Medidas de Erro	EPAM	REPQM
sem variáveis meteorológicas	Dados de treino	4,47%	6,07%
	Dados de teste	6,56%	8,79%
com variáveis meteorológicas relativas ao dia a prever	Dados de treino	3,57%	5,01%
	Dados de teste	6,46%	8,48%
com variáveis meteorológicas do dia anterior ao dia a prever	Dados de treino	3,70%	5,19%
	Dados de teste	6,54%	8,66%

Como se pretende realçar com a análise da Tabela 6.14, o uso das variáveis meteorológicas continua a ter um efeito benéfico na previsão, muito embora essa diferença não seja já tão considerável, sobretudo nos dados de teste. Uma vez que são otimizadas máquinas de suporte vetorial individuais por hora, torna-se mais facilitada a obtenção de modelos com maior rigor recorrendo-se apenas ao uso de informação provenientes dos perfis. A perspetiva de modelos únicos por hora torna a metodologia de previsão menos dependente das variáveis meteorológicas ou, de outro modo, a previsão é mais imune à falta de informação meteorológica ou, eventualmente, ao pouco rigor associado aos dados climatéricos utilizados. A Figura 6.64 ilustra também uma comparação entre o diagrama de carga real e os diagramas de carga previstos a partir dos diferentes modelos, incidindo numa semana específica do período de teste.



**Figura 6.64 – Comparação entre o diagrama de carga real e os diagramas de carga obtidos a partir de diferentes modelos de previsão (MSV) com recurso ao DC sintetizado e a variáveis meteorológicas (semana durante o período de teste)**





## *7. Conclusões e pistas de trabalho futuro*

Este estudo pretende avaliar o potencial associado à estratégia de incluir informação disponível em perfis de consumo para favorecer processos de previsão de consumos de curto prazo (neste caso, para prever registos de potência elétrica ativa das vinte e quatro horas do dia seguinte). Os perfis adotados encontram-se já validados e aprovados para permitirem o adequado funcionamento do mercado liberalizado em Portugal, tendo esta abordagem de previsão passado pela elaboração de um diagrama de carga sintetizado que resulta das contribuições aditivas dos diferentes perfis, salvaguardando-se naturalmente o peso que cada classe toma no consumo total verificado.

Este objetivo principal do trabalho ficou validado no Capítulo 6 para os dois estudos de caso apresentados e com recurso a duas estratégias distintas de previsão, uma baseada em redes neuronais e outra recorrendo a máquinas de suporte vetorial. O recurso a máquinas de suporte vetorial é deste modo determinante para atender a dois propósitos distintos já que, por um lado permite validar a gama de incerteza identificada com modelos de redes neuronais, tornando-se simultaneamente um método complementar de previsão que permite aferir a importância do diagrama de carga sintetizado.

De ressaltar que as duas variantes não devem ser alvo de uma comparação direta, já que partem de pressupostos distintos. As redes neuronais implementadas seguiram um modelo único multivariável, acomodando-se numa única rede as vinte e quatro saídas com informação para o dia seguinte. Por outro lado, as máquinas de suporte vetorial são dedicadas a cada hora em específico, ou seja, são criados diferentes modelos horários. Um modelo único multivariável de rede neuronal caracteriza-se por ter um menor tempo de computação associado ao período de treino, tornando-se também posteriormente mais fácil de interpretar e de ser usado para simular novos dados apresentados. Sendo que as redes neuronais resultam de um processo empírico por forma a identificar-se uma apropriada inicialização dos seus parâmetros, este processo pode revelar-se consideravelmente mais complexo quando se opta por recorrer a diferentes modelos específicos de cada hora. As máquinas de suporte vetorial adotam necessariamente modelos por hora, mas têm uma séria vantagem já que são por definição um método mais estruturado que apresenta uma solução única, impondo, contudo, a necessidade de se recorrer a um processo de

otimização para a escolha criteriosa de parâmetros a considerar. Adicionalmente, este processo de otimização sofreu uma alteração que lhe permitiu incluir a escolha das entradas a considerar nos diferentes modelos, tendo ficado provado com os exemplos apresentados que algumas entradas podem-se revelar redundantes ou pouco expressivas. Neste caso, estas entradas devem ser descartadas de modo a permitir a obtenção de modelos mais simples, sem que tal decisão afete a qualidade da previsão efetuada.

O recurso a análises de sensibilidades manifestadas pela(s) saída(s) às diferentes entradas aplicadas revelou-se de extrema utilidade quer para as redes neuronais, quer para as máquinas de suporte vetorial. Esta análise permite, por um lado, despistar de forma praticamente inequívoca a presença de entradas em excesso no processo (sempre que a derivada parcial indica uma débil relação entre uma saída e uma entrada específica), possibilitando também o apuramento do grau de significância que as entradas dependentes do diagrama de carga sintetizado apresentam, tipicamente sobrepondo-se à importância relativa de outros registos históricos ocorridos no passado recente. Uma vantagem prática reconhecida é que com a inclusão do diagrama de carga sintetizado os modelos ficam menos dependentes de registos anteriores de carga elétrica, permitindo-lhes uma melhor adaptação a eventuais alterações nas configurações da rede.

O estudo permitiu também avaliar o impacto que as variáveis meteorológicas poderiam ter na previsão de consumos. Deste modo, pode concluir-se que nos casos em que os consumos têm uma forte contribuição do setor residencial, a influência meteorológica deve ser tida em conta, já que esta classe reage de forma inequívoca quando existem variações significativas nas variáveis climáticas. Esta influência é mais notória em modelos de redes neuronais, uma vez que o recurso a modelos horários de máquinas de suporte vetorial torna o resultado de previsão menos dependente das variáveis climáticas.

A qualidade da previsão baseada em classificação de consumidores deve ser entendida como fortemente dependente de alguns pressupostos que são essenciais para a viabilidade da técnica. Uma adequada segmentação de consumidores, o rigor na fase de elaboração dos perfis de consumo que representam cada classe, uma seleção criteriosa da arquitetura de rede neuronal ou a escolha fundamentada de parâmetros em máquinas de suporte vetorial, uma escolha das entradas mais influentes no processo de previsão e a disponibilidade de informação comercial completa e atualizada acerca dos consumidores individuais que contribuem para um determinado consumo agregado revelam-se fatores preponderantes com um impacto direto na acuidade da previsão.

No que respeita a trabalho futuro haverá naturalmente margem para um melhor desenvolvimento do tema proposto, nomeadamente através dos seguintes pontos a destacar:

- o conceito de previsão de consumos a partir de informação proveniente de caracterização de consumos deve pressupor a determinação de perfis de consumo adequados, pelo que o recurso a métodos *clustering* poderá manifestar-se de sério interesse para o propósito em causa. Para além disso, os perfis de consumo utilizados neste estudo são os que servem o acerto de contas entre distribuidor e comercializadores, viabilizando o funcionamento do mercado de eletricidade em Portugal. Todavia, denota-se que os mesmos carecem de alguma representatividade em determinadas circunstâncias. Uma solução poderia passar por contemplar perfis com distinções entre diferentes dias de semana e atendendo às especificidades de feriados fixos e de feriados móveis;

- como ficou evidenciado na apresentação do segundo estudo de caso, o uso de informação local acerca dos consumidores efetivamente abastecidos por determinada subestação e os padrões de consumo destes consumidores pode ser determinante para o sucesso desta metodologia, pelo que é sugerível estender o estudo a uma subestação para a qual se disponha desta informação. A este respeito, é expectável que o conceito das *Smart Grids* possa tender a providenciar mais e melhor informação para atender a este requisito, já que pressupõe a instalação de um contador inteligente por instalação consumidora, com os dados de consumo a poderem ficar registados com resoluções pré-definidas.

- o recurso a técnicas de decomposição do sinal (como o *Wavelet Transform*) poderá permitir que se extraiam componentes de baixa frequência (próprias de variações suaves típicas em diagramas regulares) e componentes de alta frequência (por exemplo, o efeito de variáveis meteorológicas em consumidores residenciais, ou a influência de cargas intermitentes de frio alimentar pelo seu carácter altamente imprevisível). A sujeição de uma série reconstruída com base em diferentes componentes de frequência poderá ser um meio para favorecer processos de previsão, podendo optar-se por filtrar algumas componentes de maior frequência para despistar algum ruído identificado no sinal.

- o processo de otimização subjacente à previsão baseada em máquinas de suporte vetorial merece também um conjunto de testes adicionais que procurem robustecer a estratégia seguida. Por exemplo, em vez de se considerar um número máximo de iterações para se interromper o processo, poder-se-ia adotar um critério baseado na avaliação do decréscimo sentido na função objetivo, que poderia ser distinto para cada hora a prever (para cada modelo

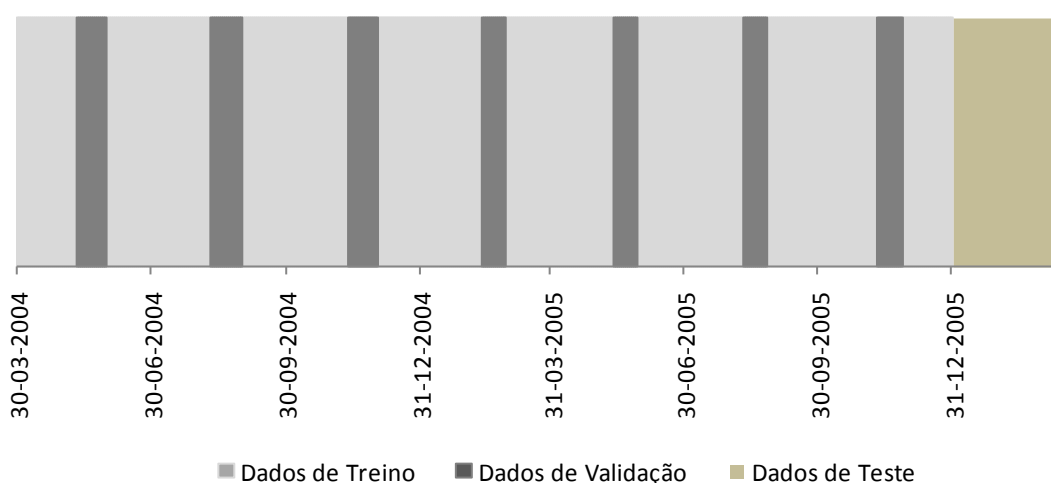
individual). O mesmo poderia ser considerado para a determinação de um critério que definisse o processo de *reannealing*. Em vez de se fixar um determinado número de iterações entre operações de *reannealing*, a escolha poderia passar por uma análise independente a cada caso horário, para que o *reannealing* fosse executado apenas quando apropriado (por exemplo, quando a função objetivo não revelasse alterações muito significativas após um número pré-definido de iterações). Uma futura melhoria poderia contemplar a exploração de uma lista das melhores soluções encontradas durante a fase de validação, para que posteriormente se cruzassem estas soluções e se optasse pela solução mais apropriada quando aplicada a dados de teste.

## ANEXOS

**Anexo A. – Segmentação dos dados em distintos períodos de treino, de validação e de teste**

Outros cenários alternativos de validação cruzada aplicados ao primeiro estudo de caso são apresentados de seguida.

O primeiro cenário alternativo à validação apresentada na Figura 5.1 é disponibilizado na Figura Anexo A.1.



**Figura Anexo A.1 – 1º Cenário alternativo de validação cruzada para o primeiro estudo de caso**

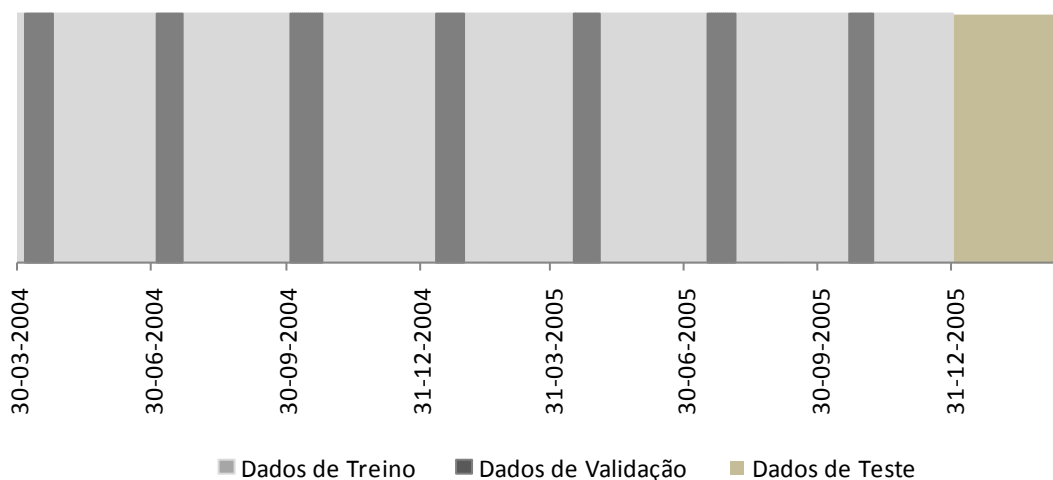
Os resultados mais favoráveis obtidos para este panorama de validação cruzada encontram-se na Tabela Anexo A.1, podendo ser confrontados com os resultados já apresentados na Tabela 6.1.

**Tabela Anexo A.1 – Resultados de erros de previsão com modelos RN para o primeiro cenário alternativo de validação cruzada aplicado ao primeiro estudo de caso**

Modelo de previsão adotado	Arquitetura da rede com melhores resultados <sup>(*)</sup>	Medidas de Erro	EPAM	REPQM
Modelo I	72:7:24	Dados de treino	6,06%	8,98%
		Dados de teste	6,93%	10,67%
Modelo II	75:9:24	Dados de treino	4,54%	6,24%
		Dados de teste	5,71%	7,61%
Modelo III	96:3:24	Dados de treino	4,36%	6,29%
		Dados de teste	5,18%	7,64%

<sup>(\*)</sup> Número de entradas: Número de neurónios na camada intermédia: Número de saídas

O segundo cenário alternativo à validação apresentada na Figura 5.1 é disponibilizado na Figura Anexo A.2.



**Figura Anexo A.2 – 2º Cenário alternativo de validação cruzada para o primeiro estudo de caso**

Os resultados mais favoráveis obtidos para este panorama de validação cruzada encontram-se na Tabela Anexo A.2, podendo ser confrontados estes resultados com os apresentados na Tabela 6.1.

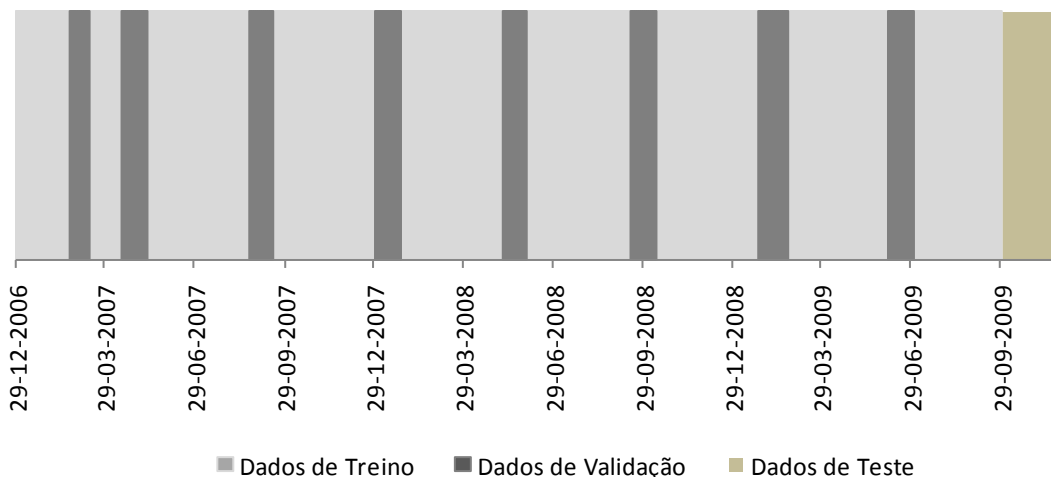
**Tabela Anexo A.2 – Resultados de erros de previsão com modelos RN para o segundo cenário alternativo de validação cruzada aplicado ao primeiro estudo de caso**

Modelo de previsão adotado	Arquitetura da rede com melhores resultados <sup>(*)</sup>	Medidas de Erro	EPAM	REPQM
Modelo I	72:7:24	Dados de treino	6,28%	9,94%
		Dados de teste	6,82%	12,38%
Modelo II	75:9:24	Dados de treino	4,13%	5,70%
		Dados de teste	5,68%	7,44%
Modelo III	96:3:24	Dados de treino	4,38%	5,97%
		Dados de teste	4,91%	6,46%

<sup>(\*)</sup> Número de entradas: Número de neurónios na camada intermédia: Número de saídas

Os diferentes cenários de validação cruzada aplicados ao segundo estudo de caso são apresentados de seguida.

O primeiro cenário alternativo à validação apresentada na Figura 5.2 é disponibilizado na Figura Anexo A.3.



**Figura Anexo A.3 – 1º Cenário alternativo de validação cruzada para o segundo estudo de caso**

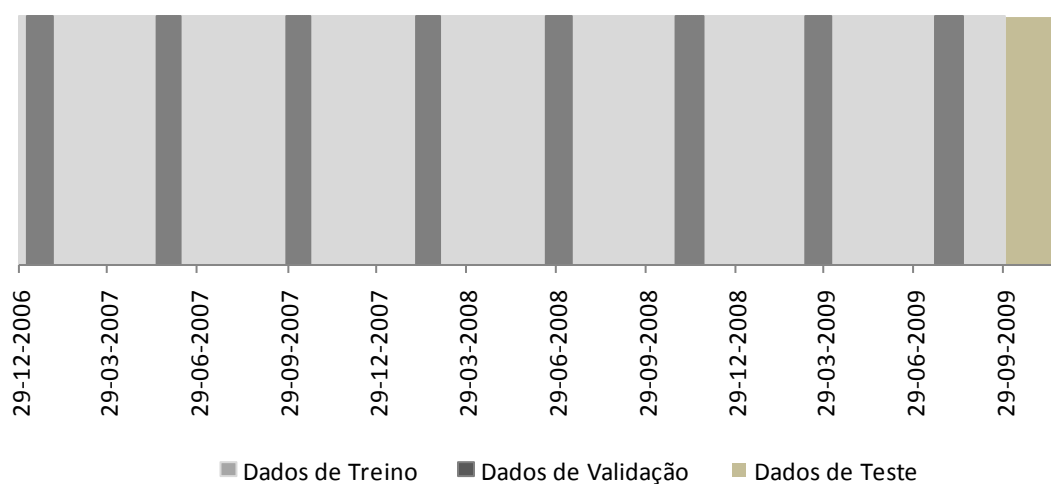
Os resultados mais favoráveis obtidos para este panorama de validação cruzada encontram-se na Tabela Anexo A.3, podendo ser confrontados com os resultados já apresentados na Tabela 6.8.

**Tabela Anexo A.3 – Resultados de erros de previsão com modelos RN para o primeiro cenário alternativo de validação cruzada aplicado ao segundo estudo de caso**

Modelo de previsão adotado	Arquitetura da rede com melhores resultados <sup>(*)</sup>	Medidas de Erro	EPAM	REPQM
Modelo I	72:7:24	Dados de treino	3,75%	4,95%
		Dados de teste	5,71%	7,33%
Modelo II	75:6:24	Dados de treino	3,75%	4,91%
		Dados de teste	5,89%	7,53%
Modelo III	96:7:24	Dados de treino	3,69%	4,90%
		Dados de teste	5,54%	7,15%

<sup>(\*)</sup> Número de entradas: Número de neurónios na camada intermédia: Número de saídas

O segundo cenário alternativo à validação apresentada na Figura 5.2 é disponibilizado na Figura Anexo A.4.



**Figura Anexo A.4 – 2º Cenário alternativo de validação cruzada para o segundo estudo de caso**

Os resultados mais favoráveis obtidos para este panorama de validação cruzada encontram-se na Tabela Anexo A.4, podendo ser confrontados com os resultados já apresentados na Tabela 6.8.

**Tabela Anexo A.4 – Resultados de erros de previsão com modelos RN para o segundo cenário alternativo de validação cruzada aplicado ao segundo estudo de caso**

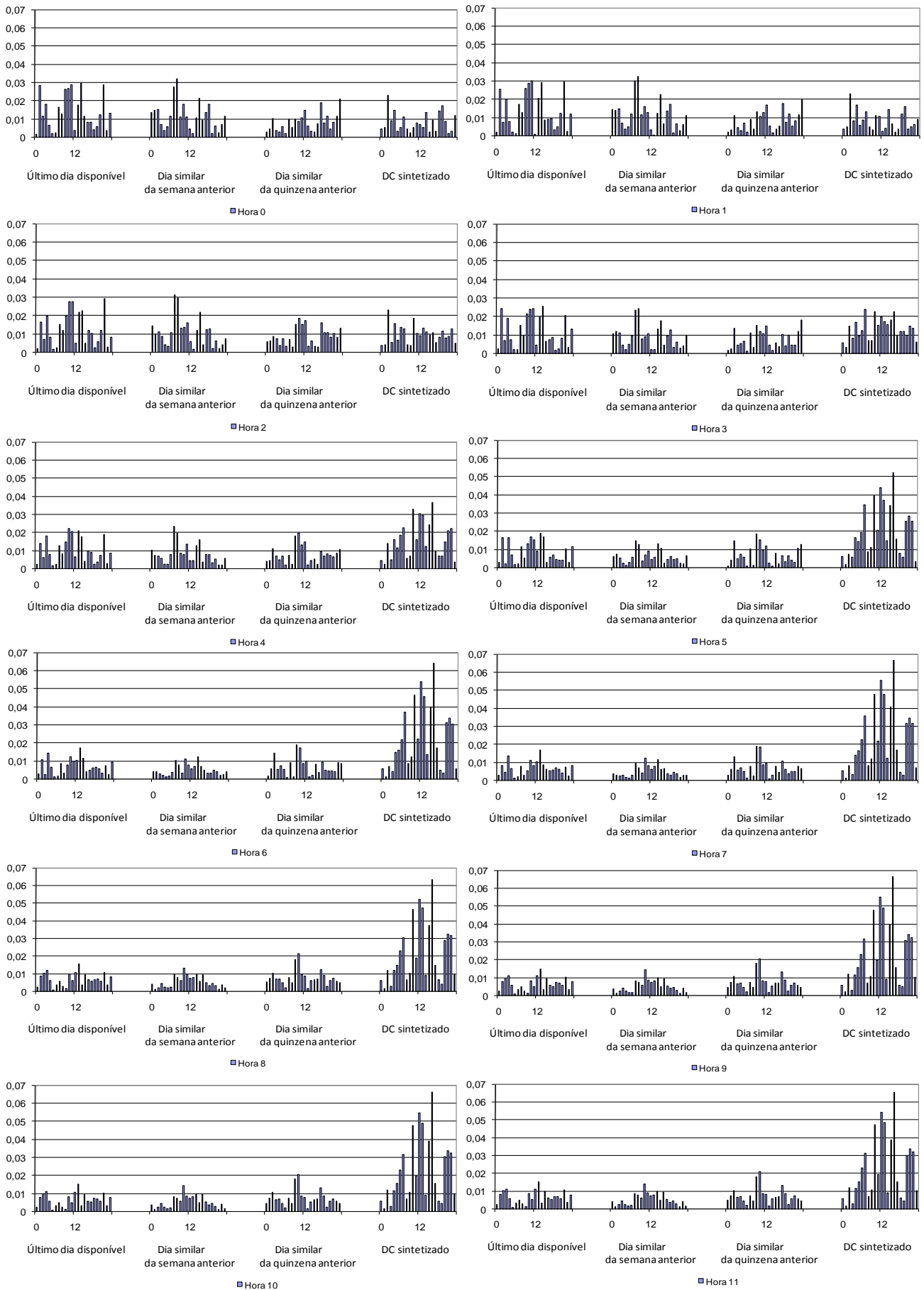
Modelo de previsão adotado	Arquitetura da rede com melhores resultados <sup>(*)</sup>	Medidas de Erro	EPAM	REPQM
Modelo I	72:7:24	Dados de treino	4,13%	5,50%
		Dados de teste	6,05%	7,86%
Modelo II	75:6:24	Dados de treino	3,98%	5,20%
		Dados de teste	5,97%	7,61%
Modelo III	96:7:24	Dados de treino	3,77%	4,97%
		Dados de teste	5,78%	7,72%

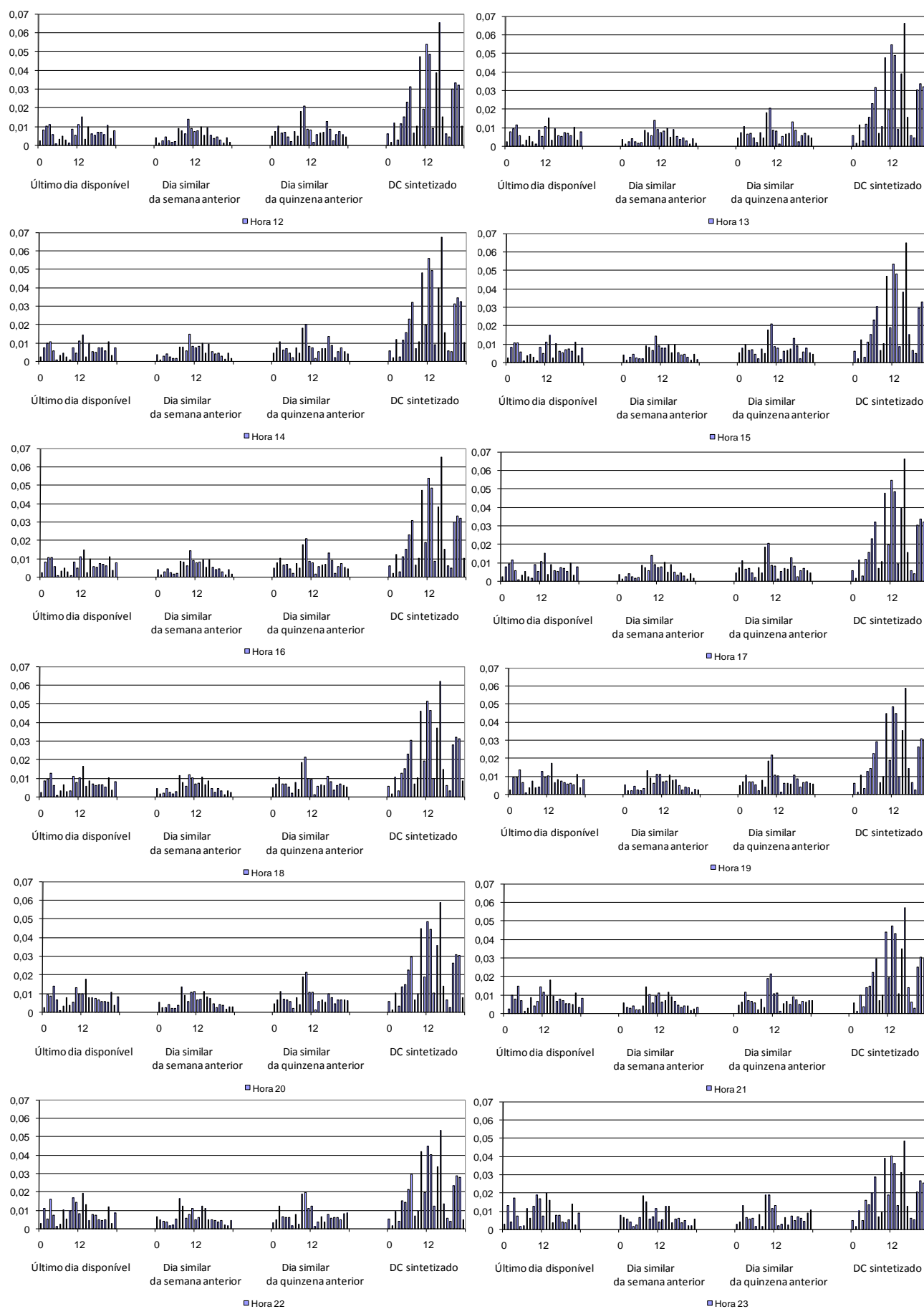
<sup>(\*)</sup> Número de entradas: Número de neurónios na camada intermédia: Número de saídas



***Anexo B. – Detalhe da análise de sensibilidades às diferentes horas do dia seguinte em modelos RN com DC sintetizado***

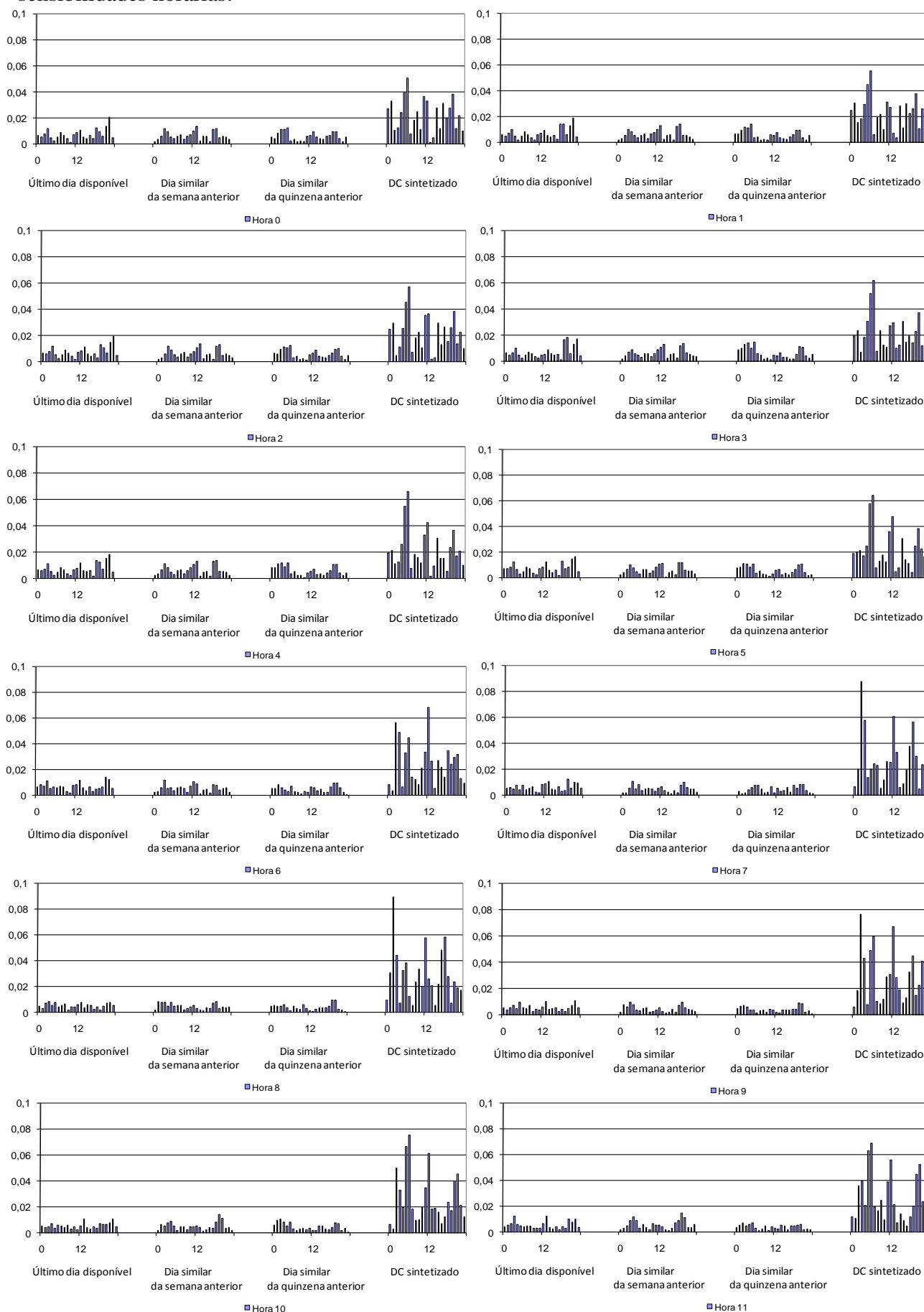
As diferentes sensibilidades apresentadas nas páginas reportam à dependência de cada saída individual às diferentes entradas usadas, estando devidamente normalizadas para que a soma seja unitária para cada caso. Nas páginas B.2 e B.3 serão apresentadas as sensibilidades verificadas no primeiro estudo de caso, enquanto nas páginas B.4 e B.5 serão apresentadas as sensibilidades manifestadas quando se incide no segundo estudo de caso.

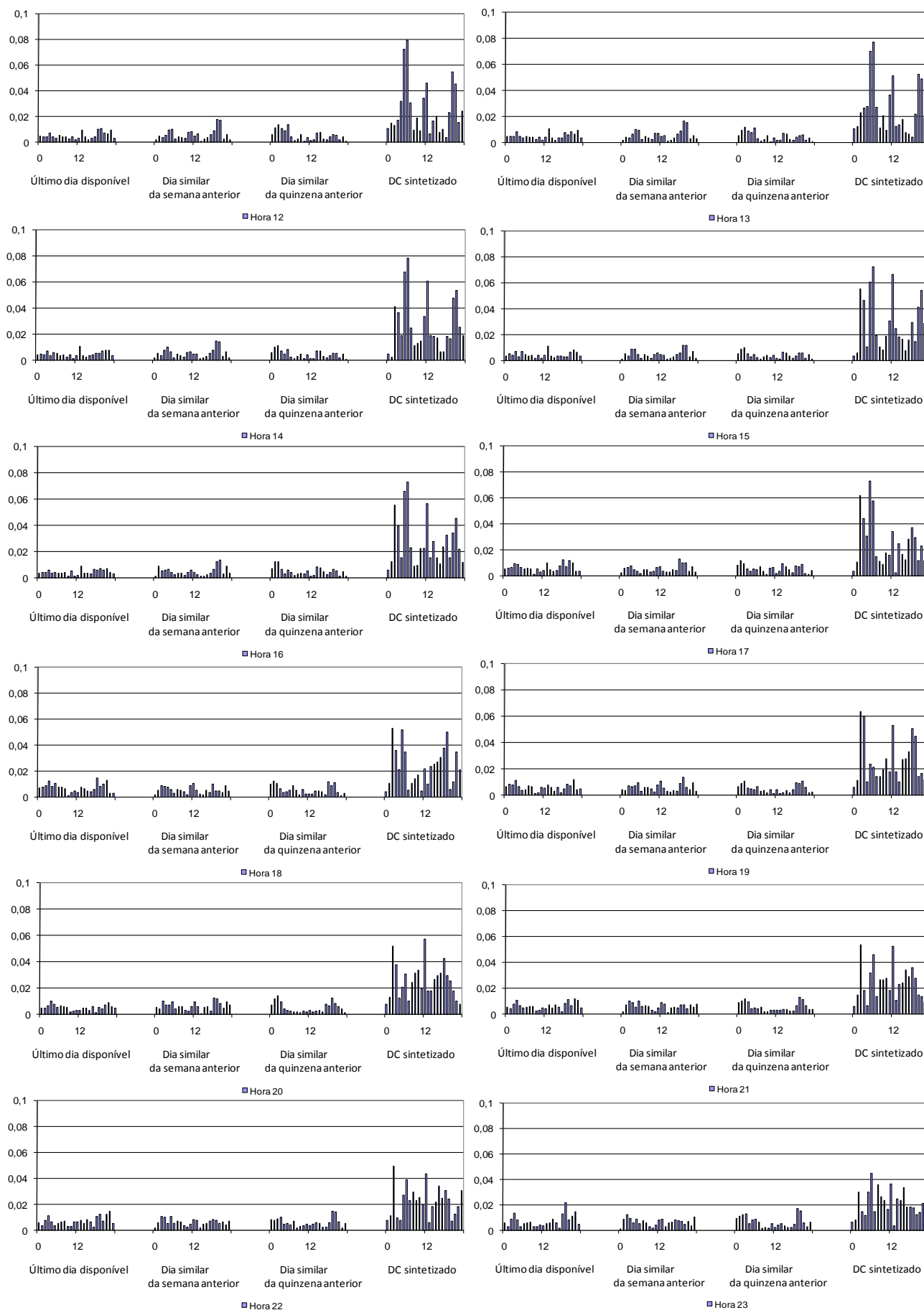




**Figura Anexo B.1 – Análise de sensibilidades horárias verificadas em modelos RN para o primeiro estudo de caso considerado**

Relativamente ao segundo estudo de caso, apresentam-se os seguintes resultados de sensibilidades horárias:





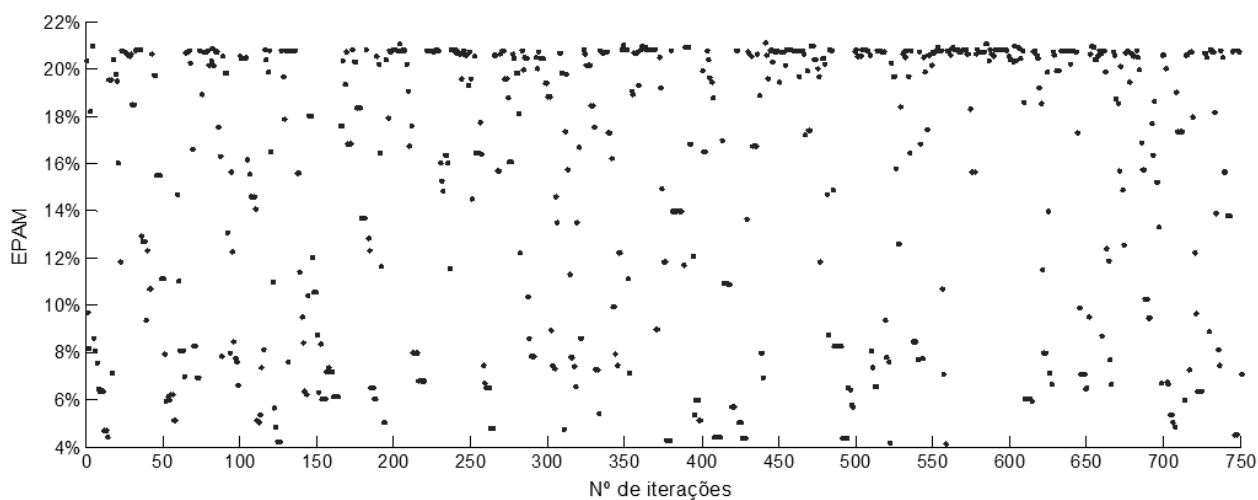
**Figura Anexo B.2– Análise de sensibilidades horárias verificadas em modelos RN para o segundo estudo de caso considerado**



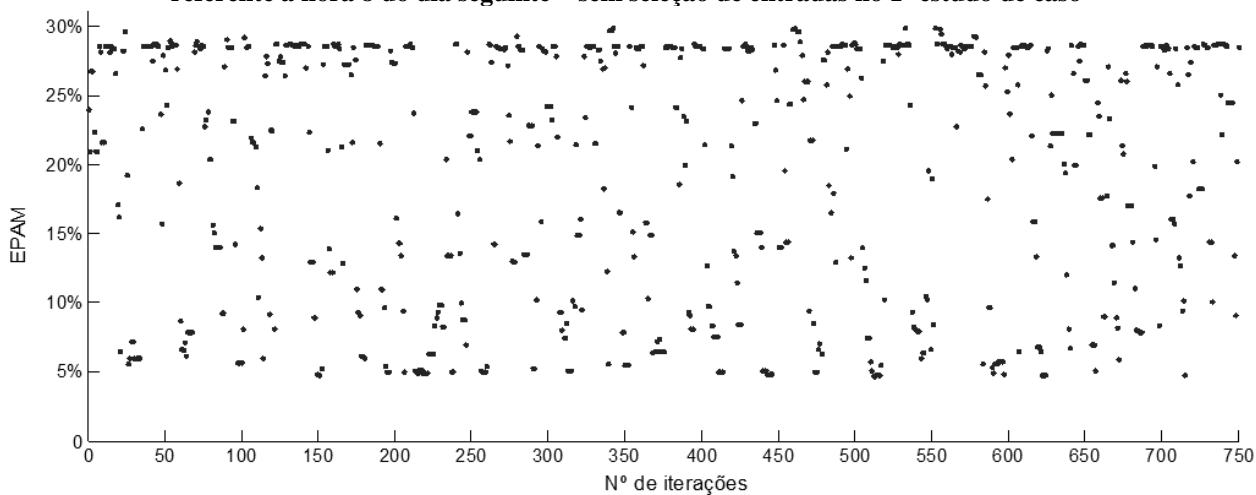
### **Anexo C. - Descrição do processo de otimização associado aos modelos MSV adotados**

Neste anexo são apresentados alguns exemplos de otimização resultantes de modelos horários baseados em máquinas de suporte vetorial.

Relativamente ao primeiro estudo de caso e recorrendo à estratégia de previsão com máquinas de suporte vetorial sem seleção de entradas obtêm-se os seguintes exemplos de otimização:



**Figura Anexo C.1 – Otimização do processo de previsão associado ao modelo horário referente à hora 8 do dia seguinte – sem seleção de entradas no 1º estudo de caso**



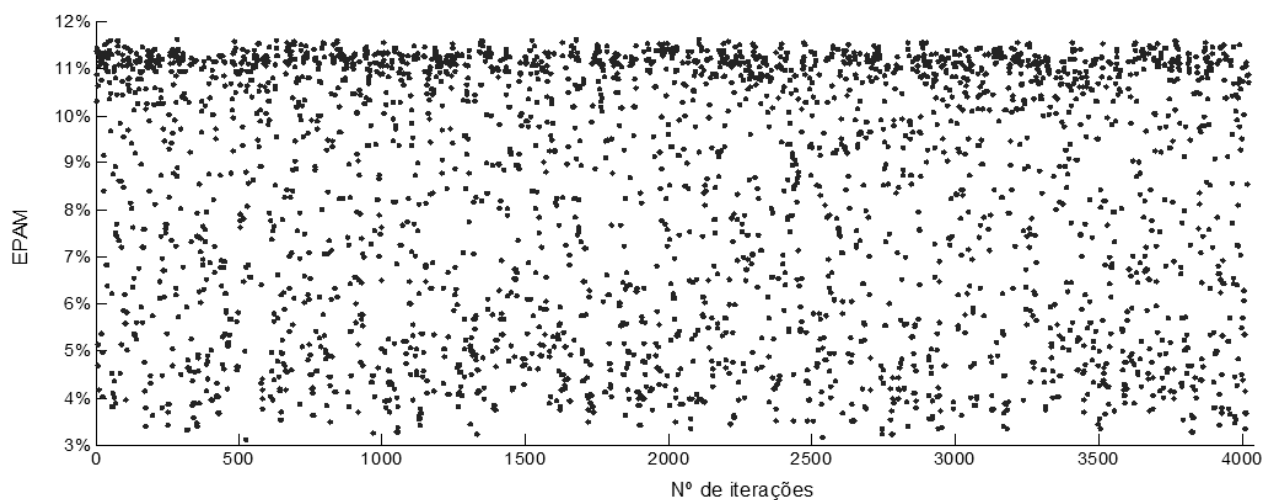
**Figura Anexo C.2 – Otimização do processo de previsão associado ao modelo horário referente à hora 16 do dia seguinte – sem seleção de entradas no 1º estudo de caso**

Tabela Anexo C.1 – Parâmetros MSV identificados e número de vetores de suporte resultantes após otimização – sem seleção de entradas no 1º estudo de caso

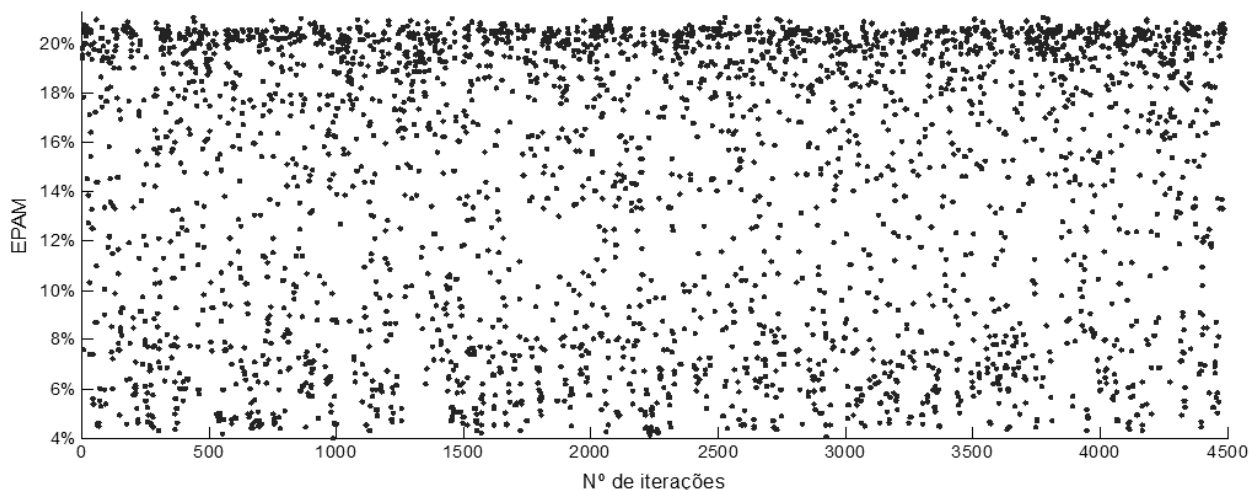
Modelos Testados	Parâmetros de entrada		Nº de Vetores de Suporte	Modelos Testados	Parâmetros de entrada		Nº de Vetores de Suporte
HORA 0	$\gamma$	0,001378	246/514	HORA 12	$\gamma$	0,0004018	506/514
	C	14,9719			C	14,6845	
	$\epsilon$	0,2063			$\epsilon$	0,005020	
HORA 1	$\gamma$	0,001459	502/514	HORA 13	$\gamma$	0,001635	514/514
	C	14,9885			C	14,9861	
	$\epsilon$	0,00918			$\epsilon$	0,000243	
HORA 2	$\gamma$	0,005539	448/514	HORA 14	$\gamma$	0,001815	503/514
	C	3,4938			C	14,1680	
	$\epsilon$	0,04779			$\epsilon$	0,003913	
HORA 3	$\gamma$	0,002204	495/514	HORA 15	$\gamma$	0,001637	182/514
	C	14,9838			C	14,4757	
	$\epsilon$	0,01601			$\epsilon$	0,171770	
HORA 4	$\gamma$	0,004261	459/514	HORA 16	$\gamma$	0,0006073	408/514
	C	12,7629			C	13,6782	
	$\epsilon$	0,048045			$\epsilon$	0,044070	
HORA 5	$\gamma$	0,004713	307/514	HORA 17	$\gamma$	0,002128	189/514
	C	9,9407			C	12,8274	
	$\epsilon$	0,167035			$\epsilon$	0,158722	
HORA 6	$\gamma$	0,002118	509/514	HORA 18	$\gamma$	0,001975	213/514
	C	14,9971			C	4,9807	
	$\epsilon$	0,006068			$\epsilon$	0,157369	
HORA 7	$\gamma$	0,003642	273/514	HORA 19	$\gamma$	0,003173	486/514
	C	11,9279			C	14,4974	
	$\epsilon$	0,166386			$\epsilon$	0,017323	
HORA 8	$\gamma$	0,002762	434/514	HORA 20	$\gamma$	0,003692	205/514
	C	14,9947			C	4,5107	
	$\epsilon$	0,026833			$\epsilon$	0,229810	
HORA 9	$\gamma$	0,003124	293/514	HORA 21	$\gamma$	0,001351	161/514
	C	9,2193			C	6,0741	
	$\epsilon$	0,082038			$\epsilon$	0,316092	
HORA 10	$\gamma$	0,00063111156	514/514	HORA 22	$\gamma$	0,001271	439/514
	C	12,9086			C	12,6026	
	$\epsilon$	0,000185			$\epsilon$	0,046565	
HORA 11	$\gamma$	0,003091	506/514	HORA 23	$\gamma$	0,001626	290/514
	C	14,2668			C	14,7026	
	$\epsilon$	0,003169			$\epsilon$	0,168227	



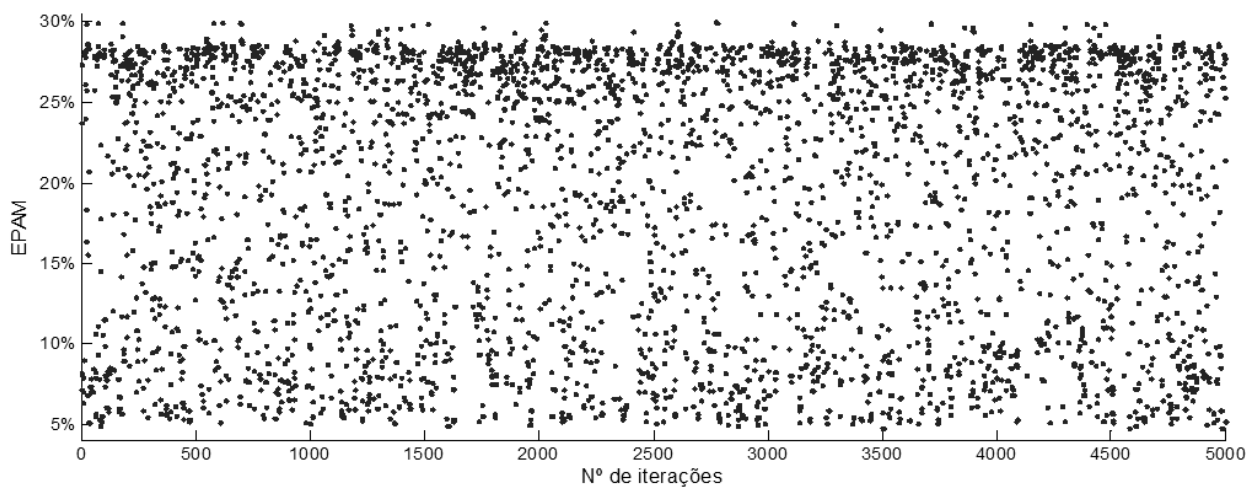
Relativamente ao primeiro estudo de caso e recorrendo à estratégia de previsão com máquinas de suporte vetorial com seleção de entradas:



**Figura Anexo C.3 – Otimização do processo de previsão associado ao modelo horário referente à hora 0 do dia seguinte – com seleção de entradas no 1º estudo de caso**



**Figura Anexo C.4 – Otimização do processo de previsão associado ao modelo horário referente à hora 8 do dia seguinte – com seleção de entradas no 1º estudo de caso**

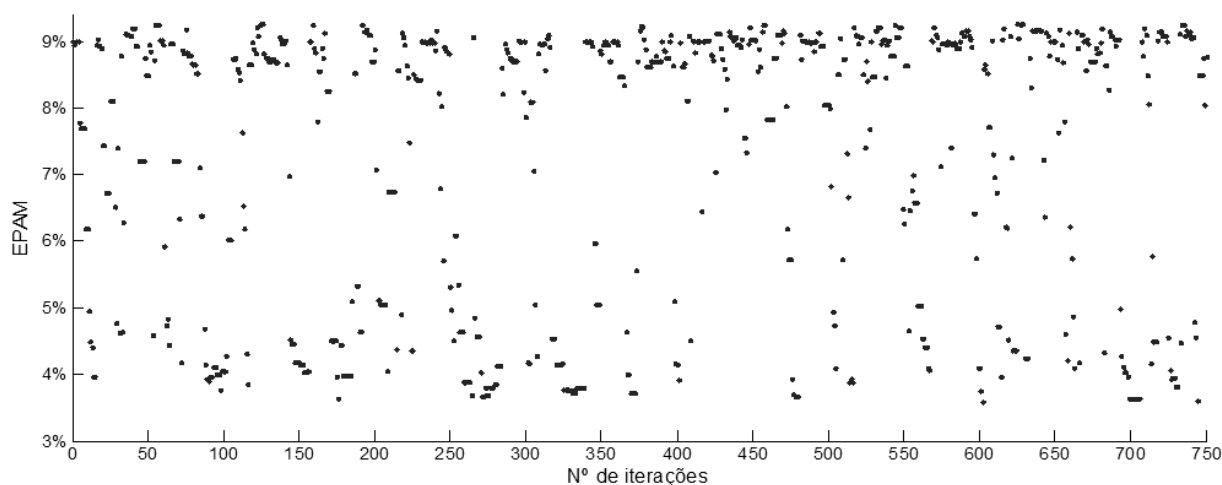


**Figura Anexo C.5 – Otimização do processo de previsão associado ao modelo horário referente à hora 16 do dia seguinte – com seleção de entradas no 1º estudo de caso**

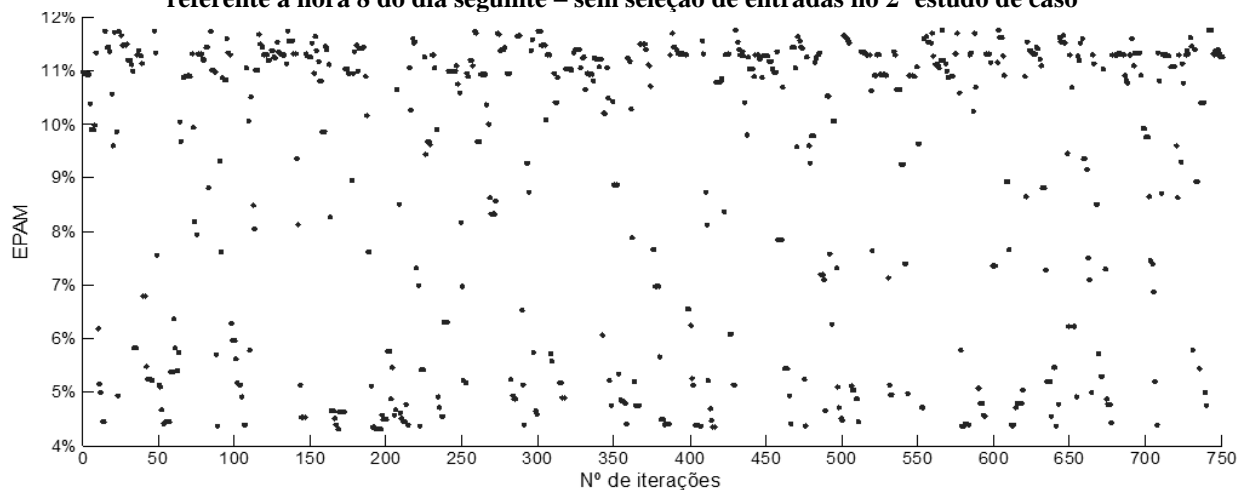
Tabela Anexo C.2 – Parâmetros MSV identificados e número de vetores de suporte resultantes após otimização – com seleção de entradas no 1º estudo de caso

Modelos Testados	Parâmetros de entrada		Nº de Vetores de Suporte	Modelos Testados	Parâmetros de entrada		Nº de Vetores de Suporte
	$\gamma$				$\gamma$		
HORA 0	$\gamma$	0,004816	455/514	HORA 12	$\gamma$	0,001006	489/514
	C	13,6985			C	13,2183	
	$\varepsilon$	0,045924			$\varepsilon$	0,010159	
HORA 1	$\gamma$	0,007761	339/514	HORA 13	$\gamma$	0,002385	462/514
	C	14,8299			C	14,1323	
	$\varepsilon$	0,142058			$\varepsilon$	0,020987	
HORA 2	$\gamma$	0,005406	381/514	HORA 14	$\gamma$	0,001088	514/514
	C	9,0277			C	14,9823	
	$\varepsilon$	0,102108			$\varepsilon$	0,0000054	
HORA 3	$\gamma$	0,006764	512/514	HORA 15	$\gamma$	0,001917	385/514
	C	13,4487			C	13,3016	
	$\varepsilon$	0,001678			$\varepsilon$	0,053115	
HORA 4	$\gamma$	0,008143	511/514	HORA 16	$\gamma$	0,0009237	511/514
	C	14,9280			C	14,3016	
	$\varepsilon$	0,002343			$\varepsilon$	0,000630	
HORA 5	$\gamma$	0,01431	509/514	HORA 17	$\gamma$	0,002702	204/514
	C	4,1244			C	14,1003	
	$\varepsilon$	0,004149			$\varepsilon$	0,151973	
HORA 6	$\gamma$	0,006509	494/514	HORA 18	$\gamma$	0,01015	291/514
	C	9,7899			C	14,9262	
	$\varepsilon$	0,012534			$\varepsilon$	0,112375	
HORA 7	$\gamma$	0,002884	502/514	HORA 19	$\gamma$	0,004488	319/514
	C	13,1864			C	5,631	
	$\varepsilon$	0,005414			$\varepsilon$	0,1083	
HORA 8	$\gamma$	0,009060	417/514	HORA 20	$\gamma$	0,003505	330/514
	C	14,9099			C	12,1880	
	$\varepsilon$	0,037823			$\varepsilon$	0,122919	
HORA 9	$\gamma$	0,008652	504/514	HORA 21	$\gamma$	0,002363	464/514
	C	14,9603			C	14,9262	
	$\varepsilon$	0,003441			$\varepsilon$	0,036627	
HORA 10	$\gamma$	0,001312	379/514	HORA 22	$\gamma$	0,004328	236/514
	C	14,2103			C	10,0316	
	$\varepsilon$	0,053303			$\varepsilon$	0,220843	
HORA 11	$\gamma$	0,003740	329/514	HORA 23	$\gamma$	0,003802	131/514
	C	12,7598			C	14,9872	
	$\varepsilon$	0,072102			$\varepsilon$	0,342237	

No que diz respeito ao segundo estudo de caso considerado e recorrendo a modelos MSV sem serem alvo de processo de seleção de entradas, podem analisar-se alguns exemplos de otimização para algumas horas específicas.



**Figura Anexo C.6 – Otimização do processo de previsão associado ao modelo horário referente à hora 8 do dia seguinte – sem seleção de entradas no 2º estudo de caso**

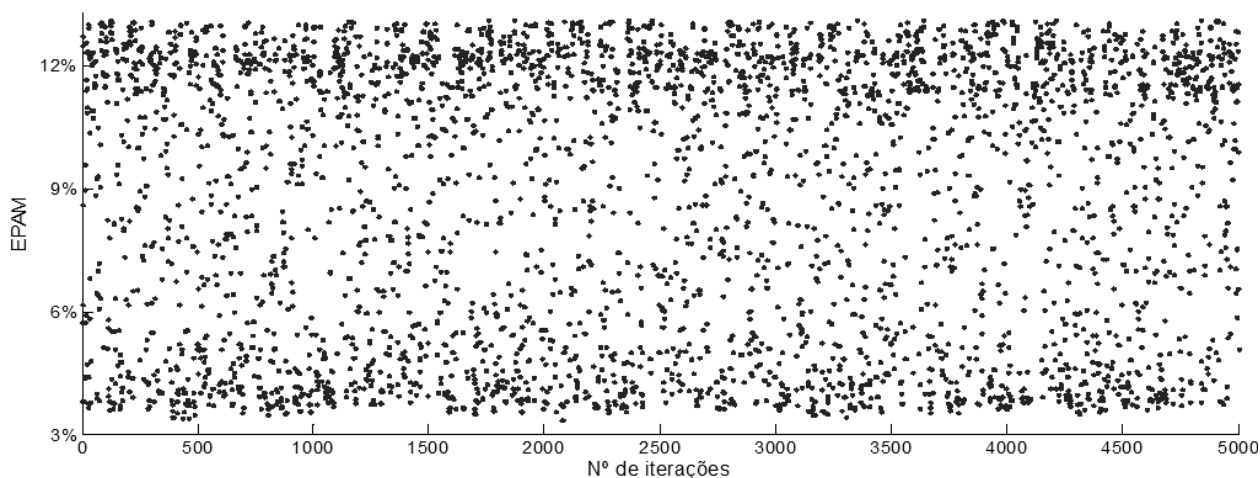


**Figura Anexo C.7 – Otimização do processo de previsão associado ao modelo horário referente à hora 16 do dia seguinte – sem seleção de entradas no 2º estudo de caso**

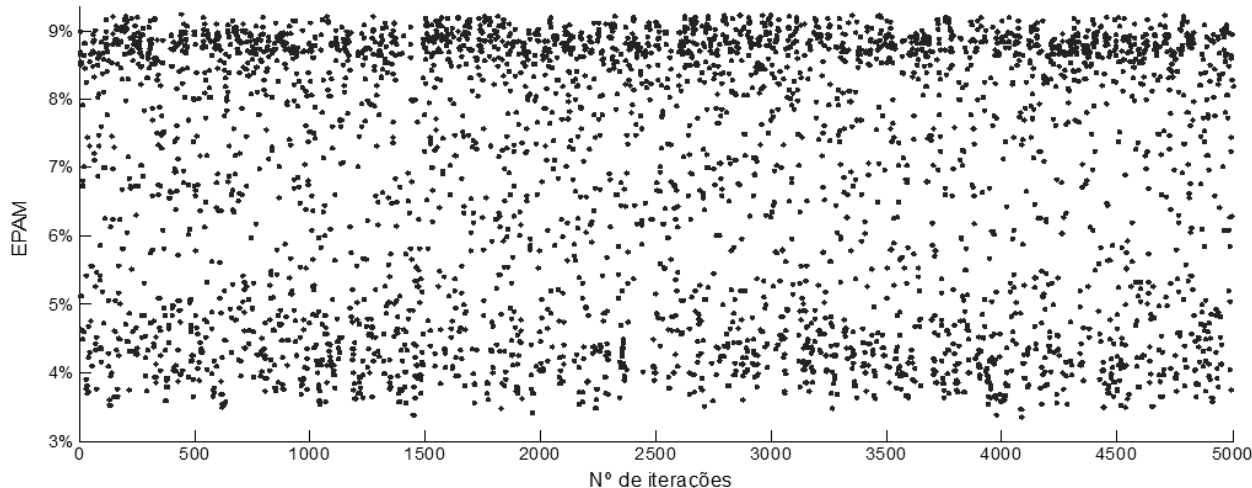
Tabela Anexo C.3 – Parâmetros MSV identificados e número de vetores de suporte resultantes após otimização – sem seleção de entradas no 2º estudo de caso

Modelos Testados	Parâmetros de entrada		Nº de Vetores de Suporte	Modelos Testados	Parâmetros de entrada		Nº de Vetores de Suporte
HORA 0	$\gamma$	0,001195	518/806	HORA 12	$\gamma$	0,0003419	651/806
	C	14,0571			C	6,9866	
	$\epsilon$	0,099187			$\epsilon$	0,081198	
HORA 1	$\gamma$	0,003109	419/806	HORA 13	$\gamma$	0,002993	676/806
	C	13,9003			C	3,0722	
	$\epsilon$	0,140004			$\epsilon$	0,059497	
HORA 2	$\gamma$	0,005332	343/806	HORA 14	$\gamma$	0,004085	737/806
	C	5,1796			C	0,423357	
	$\epsilon$	0,173917			$\epsilon$	0,037053	
HORA 3	$\gamma$	0,005119	290/806	HORA 15	$\gamma$	0,0009279	757/806
	C	11,4104			C	14,9935	
	$\epsilon$	0,224161			$\epsilon$	0,02874	
HORA 4	$\gamma$	0,004012	437/806	HORA 16	$\gamma$	0,003797	638/806
	C	5,7805			C	9,2299	
	$\epsilon$	0,143006			$\epsilon$	0,088229	
HORA 5	$\gamma$	0,001364	805/806	HORA 17	$\gamma$	0,004578	475/806
	C	14,0296			C	12,1743	
	$\epsilon$	0,000183			$\epsilon$	0,141028	
HORA 6	$\gamma$	0,001524	262/806	HORA 18	$\gamma$	0,002362	769/806
	C	13,8786			C	2,7932	
	$\epsilon$	0,265645			$\epsilon$	0,011452	
HORA 7	$\gamma$	0,001918	363/806	HORA 19	$\gamma$	0,003072	405/806
	C	8,0632			C	4,2405	
	$\epsilon$	0,201024			$\epsilon$	0,140730	
HORA 8	$\gamma$	0,001355	635/806	HORA 20	$\gamma$	0,003174	125/806
	C	14,0908			C	10,5923	
	$\epsilon$	0,089577			$\epsilon$	0,391048	
HORA 9	$\gamma$	0,003741	806/806	HORA 21	$\gamma$	0,009613	284/806
	C	3,1127			C	14,5373	
	$\epsilon$	0,000334			$\epsilon$	0,253802	
HORA 10	$\gamma$	0,001099	217/806	HORA 22	$\gamma$	0,003126	103/806
	C	12,6129			C	14,8415	
	$\epsilon$	0,421029			$\epsilon$	0,494734	
HORA 11	$\gamma$	0,001072	227/806	HORA 23	$\gamma$	0,007315	154/806
	C	13,7119			C	5,5095	
	$\epsilon$	0,380905			$\epsilon$	0,415711	

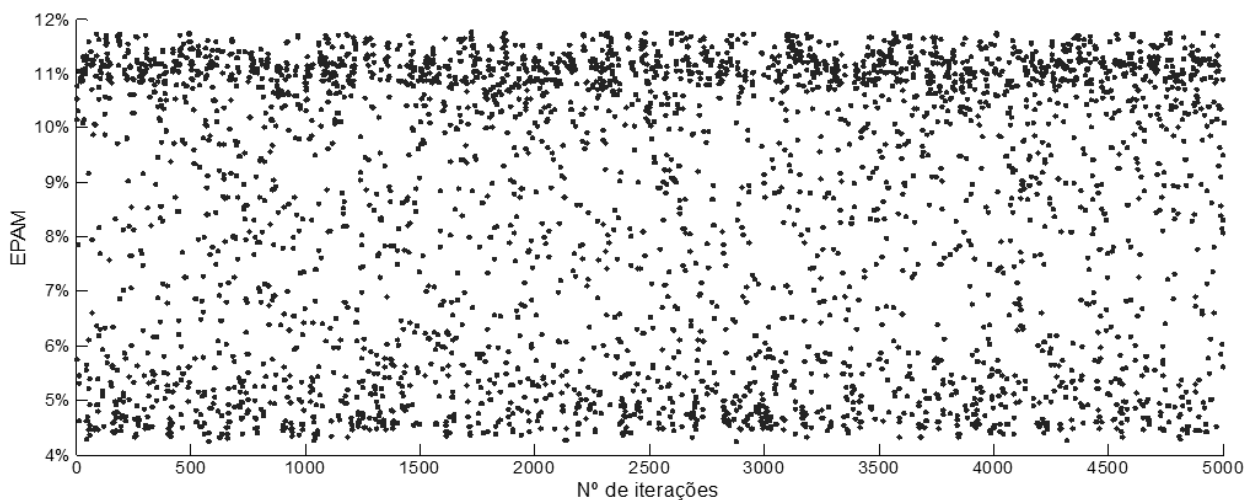
Relativamente ao segundo estudo de caso considerado e recorrendo a modelos MSV que incorporam uma estratégia de seleção de entradas, podem analisar-se alguns exemplos de otimização para algumas horas específicas.



**Figura Anexo C.8 - Otimização do processo de previsão associado ao modelo horário referente à hora 0 do dia seguinte – com seleção de entradas no 2º estudo de caso**



**Figura Anexo C.9 – Otimização do processo de previsão associado ao modelo horário referente à hora 8 do dia seguinte – com seleção de entradas no 2º estudo de caso**



**Figura Anexo C.10 – Otimização do processo de previsão associado ao modelo horário referente à hora 16 do dia seguinte – com seleção de entradas no 2º estudo de caso**

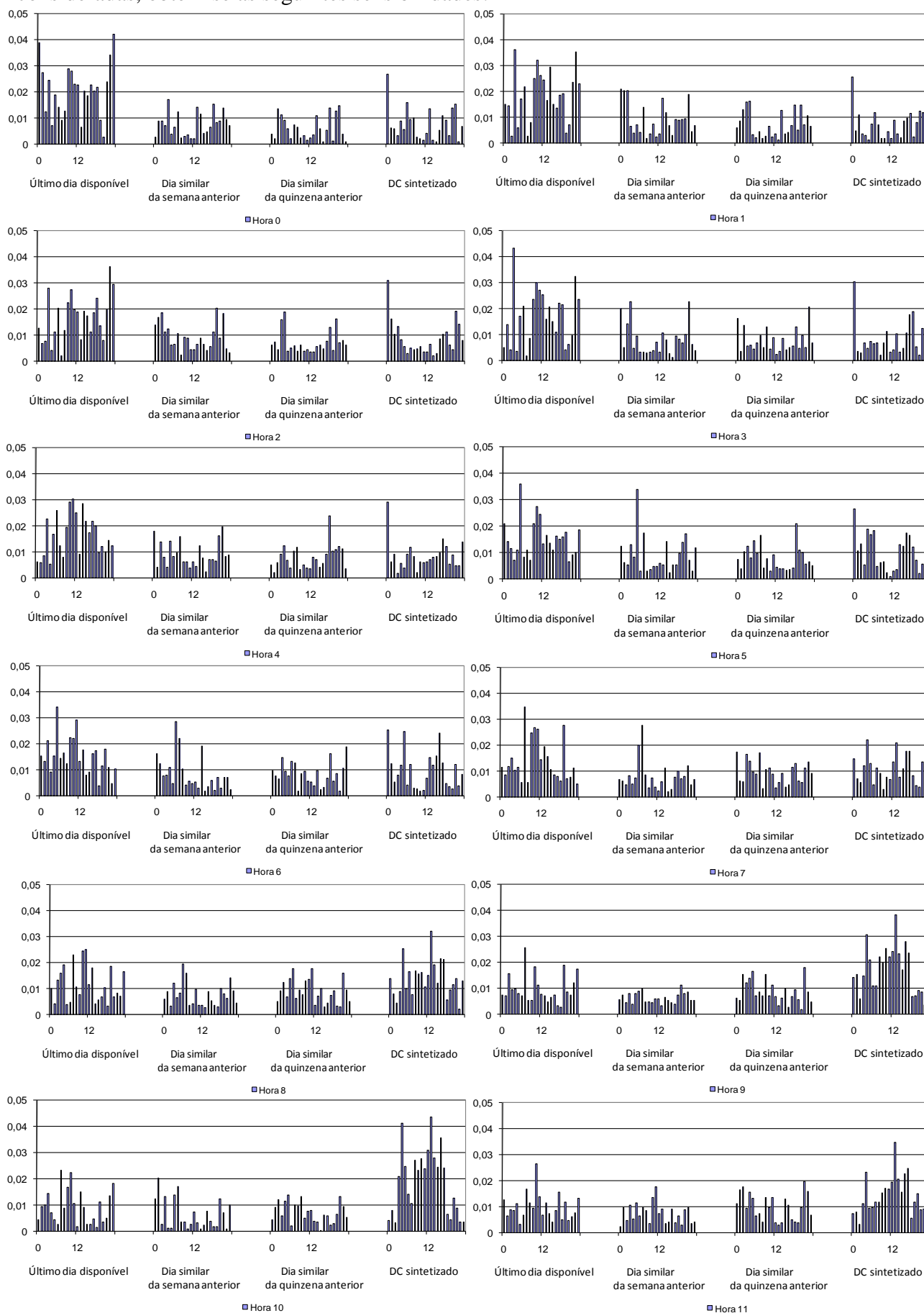
Tabela Anexo C.4 – Parâmetros MSV identificados e número de vetores de suporte resultantes após otimização – com seleção de entradas no 2º estudo de caso

Modelos Testados	Parâmetros de entrada		Nº de Vetores de Suporte	Modelos Testados	Parâmetros de entrada		Nº de Vetores de Suporte
HORA 0	$\gamma$	0,005061	500/806	HORA 12	$\gamma$	0,001987	407/806
	C	14,7805			C	6,8610	
	$\epsilon$	0,101283			$\epsilon$	0,228459	
HORA 1	$\gamma$	0,006545	253/806	HORA 13	$\gamma$	0,009692	788/806
	C	6,4851			C	2,2668	
	$\epsilon$	0,240573			$\epsilon$	0,006921	
HORA 2	$\gamma$	0,01430	806/806	HORA 14	$\gamma$	0,01153	374/806
	C	4,5112			C	0,8434	
	$\epsilon$	0,000019			$\epsilon$	0,255548	
HORA 3	$\gamma$	0,006495	790/806	HORA 15	$\gamma$	0,006406	280/806
	C	14,8501			C	8,3668	
	$\epsilon$	0,005847			$\epsilon$	0,369277	
HORA 4	$\gamma$	0,006437	578/806	HORA 16	$\gamma$	0,004035	415/806
	C	11,0333			C	2,9233	
	$\epsilon$	0,086698			$\epsilon$	0,221772	
HORA 5	$\gamma$	0,003424	686/806	HORA 17	$\gamma$	0,005806	801/806
	C	14,2952			C	6,6393	
	$\epsilon$	0,044900			$\epsilon$	0,000749	
HORA 6	$\gamma$	0,002381	320/806	HORA 18	$\gamma$	0,007701	805/806
	C	14,9998			C	14,9361	
	$\epsilon$	0,229298			$\epsilon$	0,000338	
HORA 7	$\gamma$	0,008439	520/806	HORA 19	$\gamma$	0,009178	442/806
	C	4,1107			C	6,8332	
	$\epsilon$	0,137029			$\epsilon$	0,127766	
HORA 8	$\gamma$	0,006636	738/806	HORA 20	$\gamma$	0,02096	622/806
	C	8,0973			C	1,9367	
	$\epsilon$	0,034126			$\epsilon$	0,061041	
HORA 9	$\gamma$	0,02711	644/806	HORA 21	$\gamma$	0,01077	284/806
	C	1,8519			C	8,03305	
	$\epsilon$	0,080203			$\epsilon$	0,265695	
HORA 10	$\gamma$	0,0004878	177/806	HORA 22	$\gamma$	0,0004443	668/806
	C	14,9977			C	11,0362	
	$\epsilon$	0,497897			$\epsilon$	0,058273	
HORA 11	$\gamma$	0,002073	155/806	HORA 23	$\gamma$	0,0009121	138/806
	C	14,8632			C	12,6577	
	$\epsilon$	0,494849			$\epsilon$	0,4949	

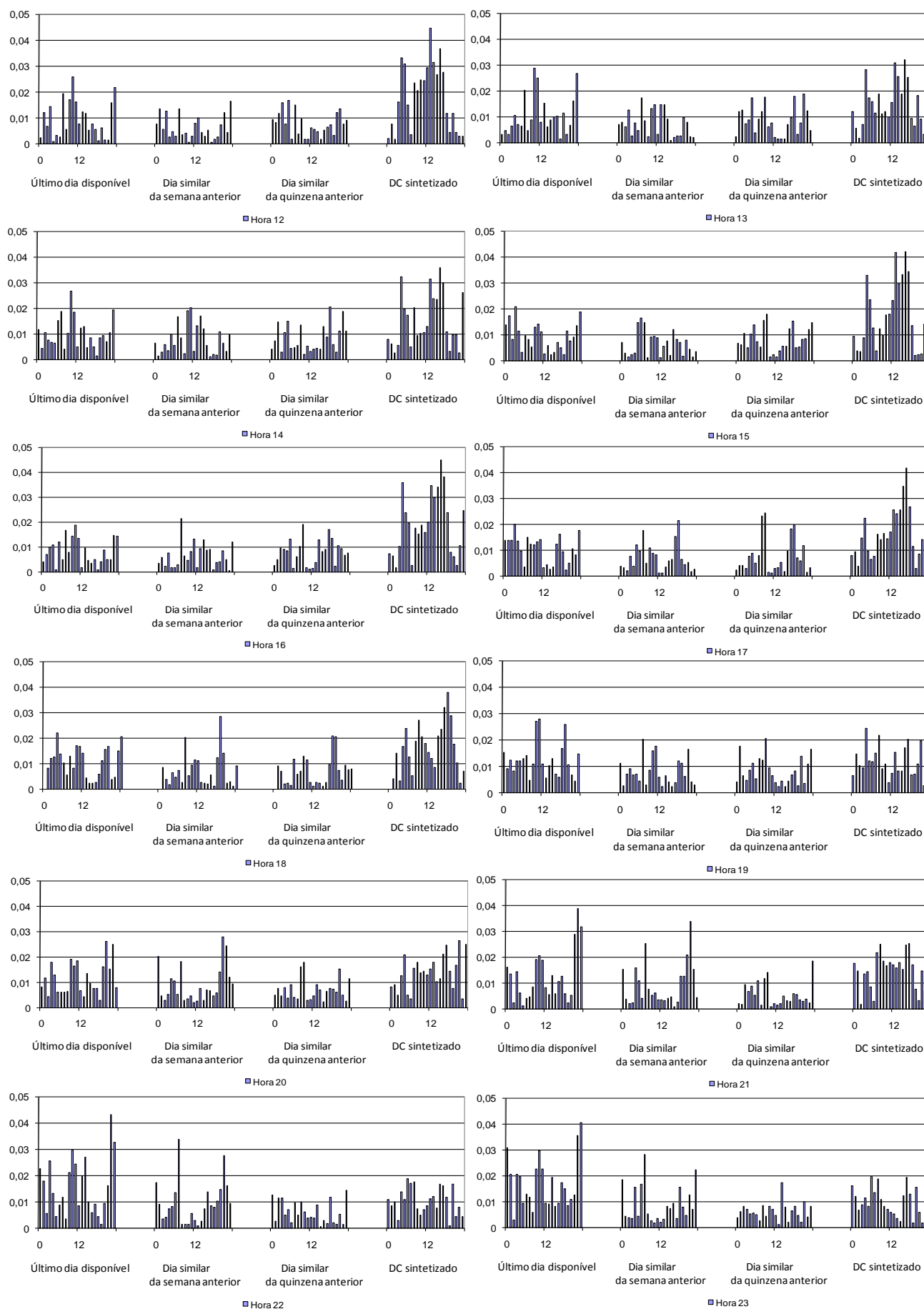
***Anexo D. - Detalhe da análise de sensibilidades às diferentes horas do dia seguinte em modelos MSV com DC sintetizado***

As diferentes sensibilidades apresentadas neste anexo reportam à dependência da saída de cada modelo às diferentes entradas usadas, estando devidamente normalizadas para que a soma seja unitária para cada modelo.

Para o primeiro estudo de caso e usando modelos MSV que integram todas as entradas consideradas, obtêm-se as seguintes sensibilidades:

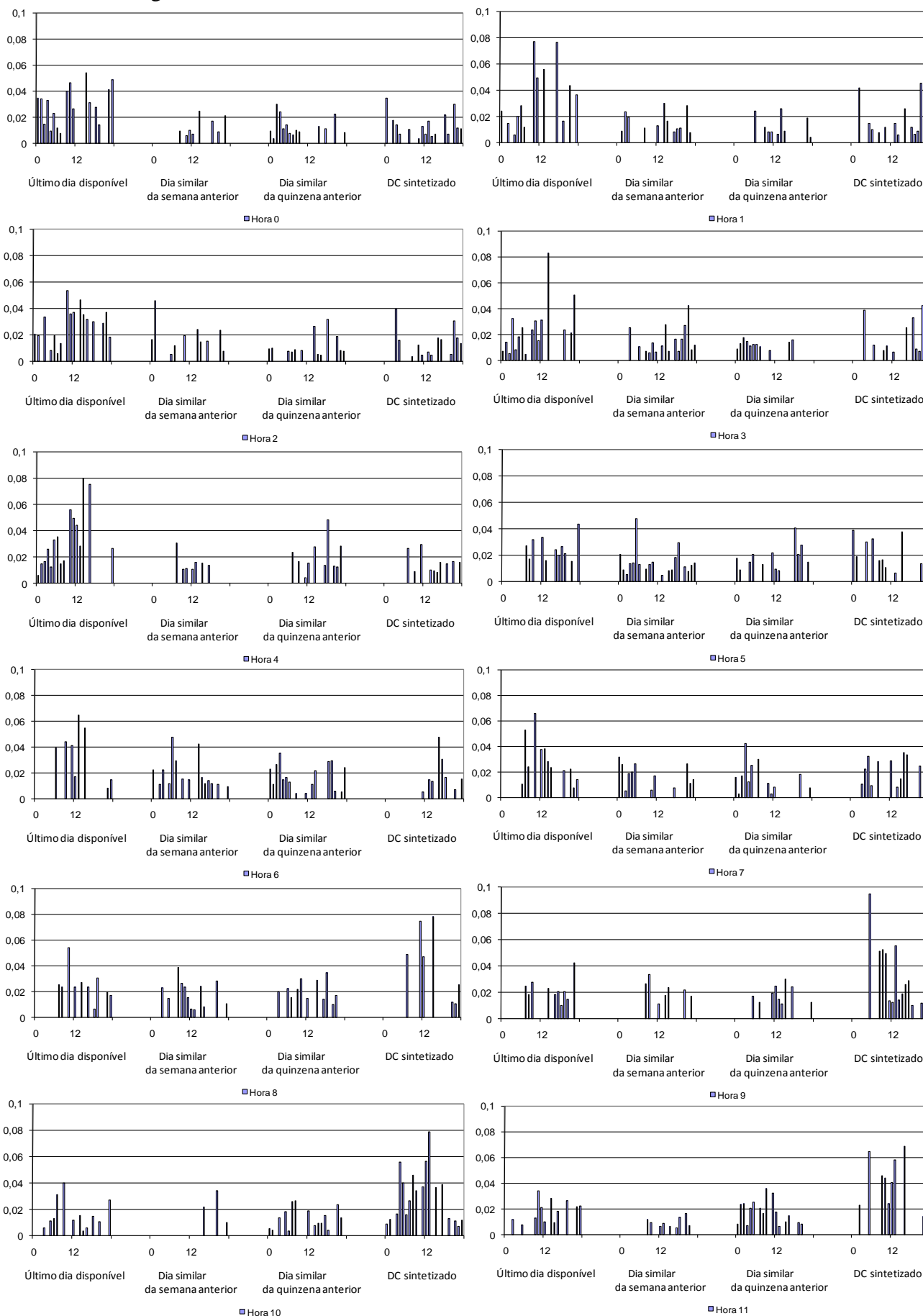


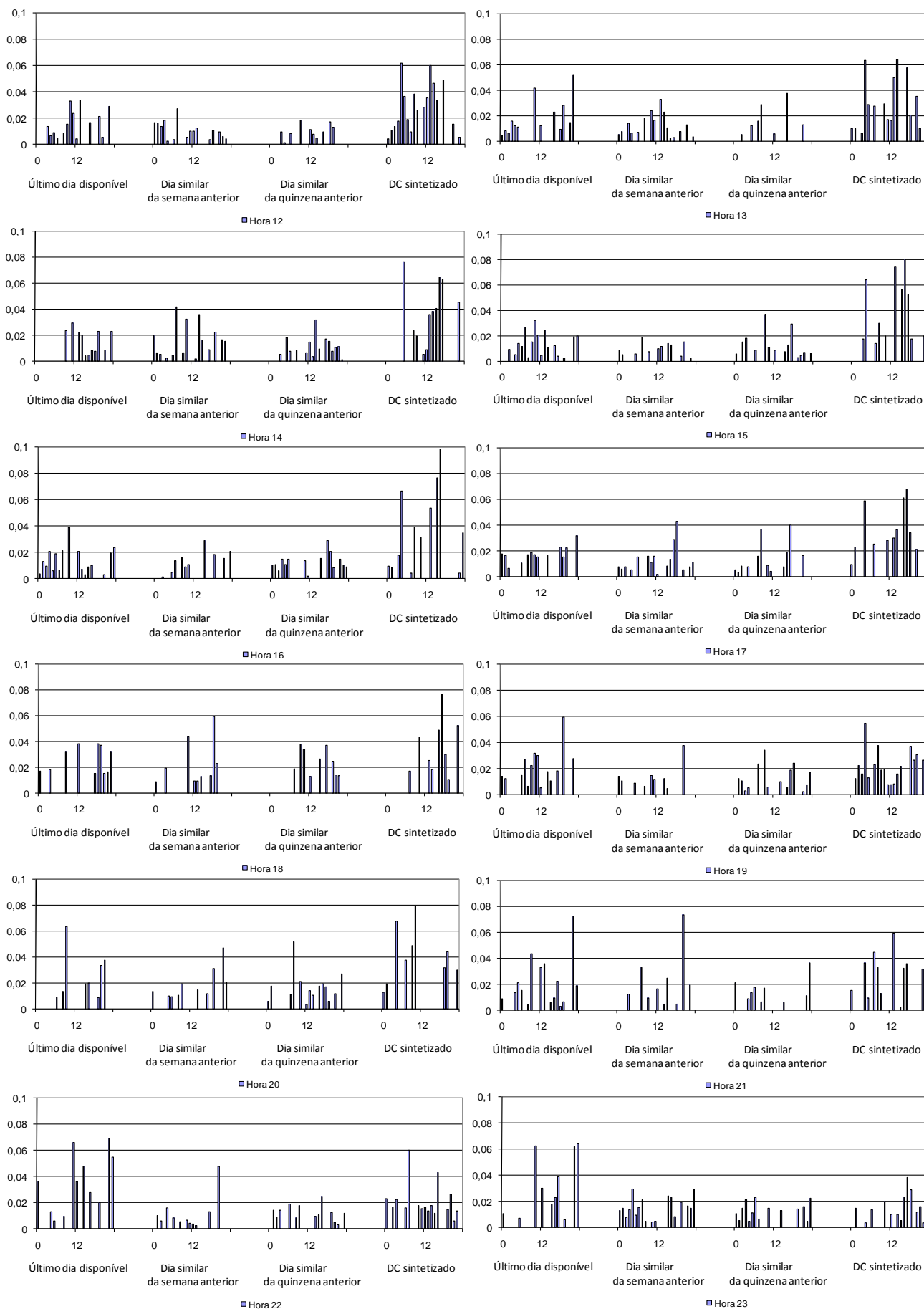




**Figura Anexo D.1 – Análise de sensibilidades horárias verificadas em modelos MSV sem seleção de entradas para o primeiro estudo de caso considerado**

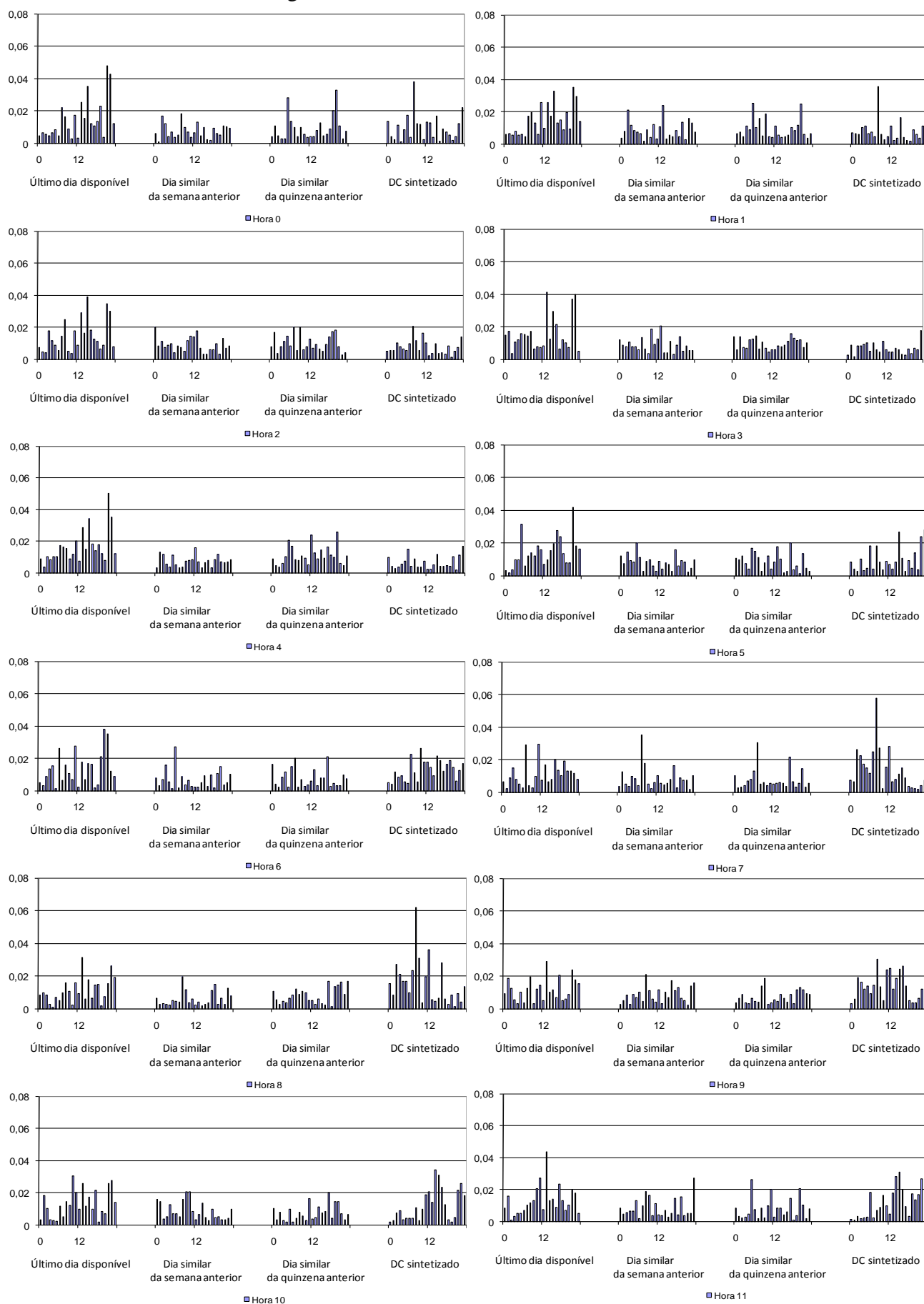
Relativamente ainda ao primeiro estudo de caso e usando modelos com seleção de entradas, obtêm-se os seguintes resultados:

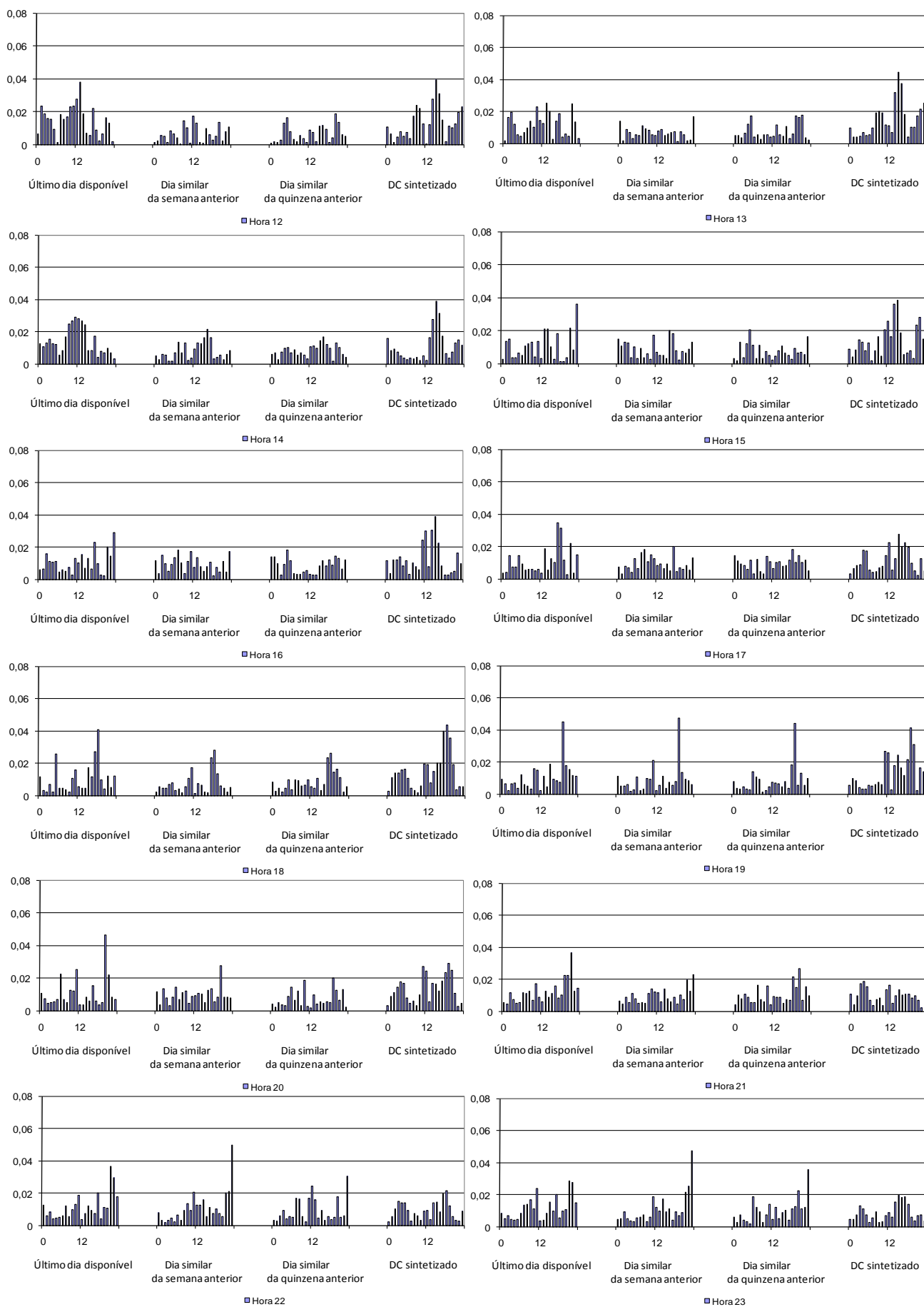




**Figura Anexo D.2 – Análise de sensibilidades horárias verificadas em modelos MSV com seleção de entradas para o primeiro estudo de caso considerado**

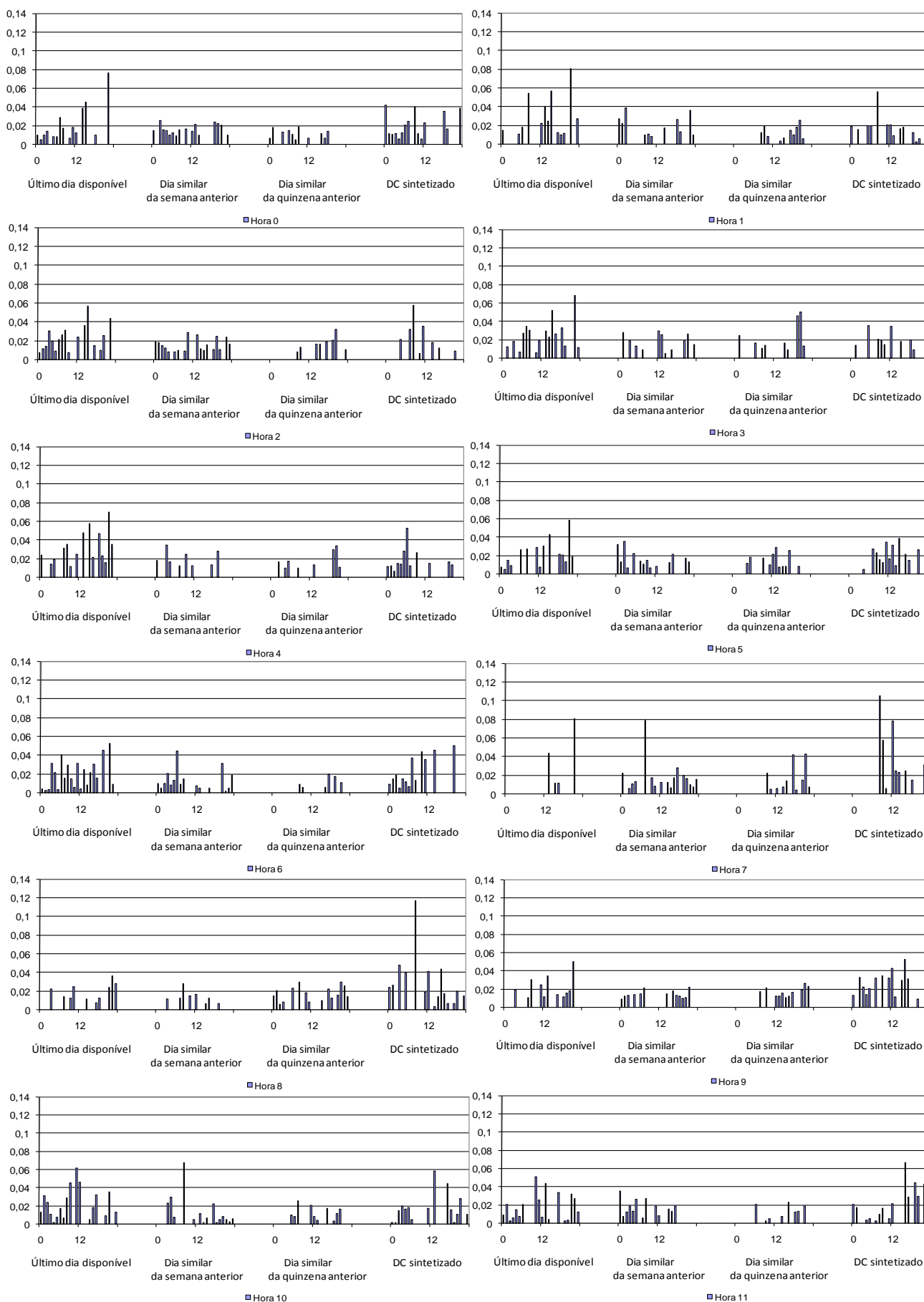
Para o segundo estudo de caso e usando modelos MSV que integram todas as entradas consideradas, obtêm-se as seguintes sensibilidades:

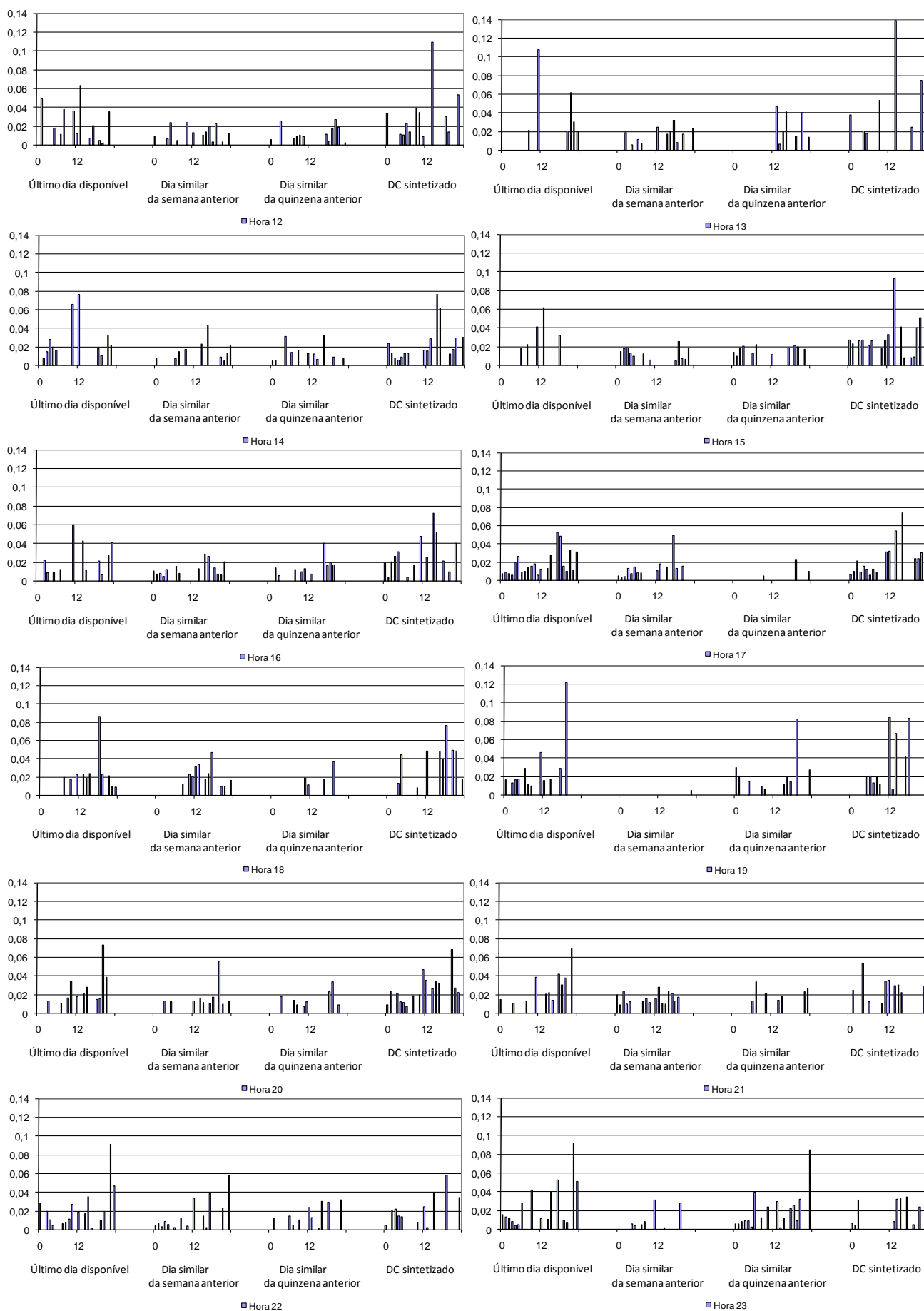




**Figura Anexo D.3 – Análise de sensibilidades horárias verificadas em modelos MSV sem seleção de entradas para o segundo estudo de caso considerado**

Para o segundo estudo de caso e usando modelos MSV com seleção de entradas, obtêm-se as seguintes sensibilidades:





**Figura Anexo D.4 – Análise de sensibilidades horárias verificadas em modelos MSV com seleção de entradas para o segundo estudo de caso considerado**





## ***Referências bibliográficas***

Abbas S.R., Arif M. (2006). «Electric Load Forecasting Using Support Vector Machines Optimized by Genetic Algorithm», in Proceedings of the INMIC'06 IEEE Multitopic Conference, 23<sup>th</sup> - 24<sup>th</sup> December 2006; p. 395-399, doi: 10.1109/INMIC.2006.358199.

Afshin M., Sadeghian A., Raahemifar K. (2007), «On Efficient Tuning of LS-SVM Hyper-Parameters in Short-Term Load Forecasting: A Comparative Study», in Proceedings of the IEEE Power Engineering Society General Meeting, 24<sup>th</sup> - 28<sup>th</sup> June 2007; p. 1-6, doi: 10.1109/PES.2007.385613.

AlFuhaid A., El-Sayed M., Mahmoud M. (1997), «Cascaded artificial neural networks for short-term load forecasting», *IEEE Transactions on Power Systems*, **12** (4), p. 1524-1529, doi: 10.1109/59.627852.

Bakirtzis A., Petridis V., Klartzis S., Alexiadis M. (1996), «A Neural Network Short Term Load Forecasting Model for the Greek Power System», *IEEE Transactions on Power Systems* **11** (2), p. 858-863.

Bunn D. W., Farmer E.D. (1985), *Comparative Models for Electrical Load Forecasting*, John Wiley & Sons Eds, ISBN-10: 0471906352; ISBN-13: 978-0471906353.

Campbell P., Adamson K. (2006), «Methodologies for load forecasting», in Proceedings of the IEEE 3<sup>rd</sup> International Conference on Intelligent Systems, London UK, 4<sup>th</sup>-6<sup>th</sup> September 2006; p. 800-806.

Cao L. J, Tay F. E. H. (2000), «Feature Selection for Support Vector Machines in Financial Time Series Forecasting», in Proceedings of the Second International Conference on Intelligent Data Engineering and Automated Learning – IDEAL 2000, Springer Lecture Notes in Computer Science, Data Mining, Financial Engineering, and Intelligent Agents, Hong Kong, China, December 13<sup>th</sup>-15<sup>th</sup>, 2000, Vol.1983, p. 268-273, doi: 10.1007/3-540-44491-2\_38.

Casazza J., Delea, F. (2003), *Understanding Electric Power Systems – An overview of the technology and the marketplace*, IEEE Press, A John Wiley & Sons Inc. Publication, 2003, Print ISBN: 9780471446521, doi: 10.1002/0471667854.

Cavallaro F. (2005), «Electric Load Analysis using an Artificial Neural Network», *International Journal of Energy Research*, **29** (5), p. 377-392, doi: 10.1002/er.1054.

- Chang P., Fan C., Lin J. (2011), «Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network», *Elsevier – International Journal of Electrical Power and Energy Systems*, 33 (1), p. 17-27, doi: 10.1016/j.ijepes.2010.08.008.
- Chang C., Chih-Jen Lin (2011), «LIBSVM: a library for support vector machines», *ACM Transactions on Intelligent Systems and Technology*, 2:27:1--27:27, 2011. Software disponível em: <http://www.csie.ntu.edu.tw/~cjlin/libsvm> [consultado a 08 de março de 2013].
- Chen B., Chang M., Lin C. (2004), «Load Forecasting Using Support Vector Machines: A Study on EUNITE Competition 2001», *IEEE Transactions on Power Systems*, **19** (4), p. 1821-1830, doi: 10.1109/TPWRS.2004.835679.
- Chen C. S., Kang M. S., Hwang J. C., Huang, C. W. (2001), «Temperature Effect to Distribution System Load Profiles and Feeder Losses», *IEEE Transactions on Power Systems*, **16** (4), p. 916-921, doi: 10.1109/59.962446.
- Chen H., Canizares C., Singh A (2001a), «ANN-based short-term load forecasting in electricity markets», in *Proceedings of the IEEE Power Engineering Society Winter Meeting*, vol. 2, p. 411-415, doi: 10.1109/PESW.2001.916876.
- Chicco G., Ilie, I. S. (2009), «Support Vector Clustering of Electrical Load Pattern Data», *IEEE Transactions on Power Systems*, **24** (3), p. 1619-1628, doi: 10.1109/TPWRS.2009.2023009.
- Chicco G., Napoli R., Piglione F. (2003), «Application of Clustering Algorithm and Self Organizing Maps to Classify Electricity Consumers», in *Proceedings of the IEEE PowerTech Conference, Bologna, Italy, 23<sup>rd</sup> – 26<sup>th</sup> June 2003*, vol. 1, doi: 10.1109/PTC.2003.1304160.
- Chicco G., Napoli R., Piglione F. (2001), «Load pattern clustering for short-term load forecasting of anomalous days», in *Proceedings of the IEEE PowerTech Conference, Porto 10<sup>th</sup>-13<sup>th</sup> September 2001* Vol. 2, Print ISBN: 0-7803-7139-9, doi: 10.1109/PTC.2001.964745.
- de Vicente J., Lanchares J., Hermida R. (2003), «Placement by thermodynamic simulated annealing», *Elsevier - Physics Letters A*, **317** (5)(6), p. 415-423, doi: 10.1016/j.physleta.2003.08.070.
- Demuth H., Beale M., Hagan M. (2007), *Neural Network Toolbox 5 – User’s Guide*, The Mathworks Inc., Revision History, September 2007.
- Drezga I., Rahman S. (1998), «Input variable selection for ANN-based short-term load forecasting», *IEEE Transactions on Power Systems*, **13** (4), p. 1238-1244.

Elattar E. E., Goulermas J., Wu Q. H. (2010), «Electric Load Forecasting Based on Locally Weighted Support Vector Regression», *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews*, **40** (4), p. 438-447, doi: 10.1109/TSMCC.2010.2040176.

ERSE (2012) – *Guia de medição, leitura e disponibilização de dados de energia elétrica em Portugal Continental*, Disponível no sítio da Entidade Reguladora dos Serviços Energéticos, janeiro de 2012.

ERSE (2011) – *Regulamento Tarifário do Sector Elétrico*, Disponível no sítio da Entidade Reguladora dos Serviços Energéticos, julho de 2011.

Eurelectric (2000) – *Metering, Load profiles and Settlement in Deregulated Systems*, System Tariff Issues Working Group, Union of the Electricity Industry – Eurelectric, March 2000, Ref. 2000-220-0004.

Espinoza M., Suykens J. A. K, Belmans R., DeMoor B. (2007), «Electric Load Forecasting using kernel-based modeling for nonlinear system identification», *IEEE Control Systems Magazine*, Special Issue on Applications of System Identification, **27** (5), p. 43 - 57, doi 10.1109/MCS.2007.904656.

Espinoza M., Joye C., Belmans R., DeMoor B. (2005), «Short-Term Load Forecasting, Profile Identification, and Customer Segmentation: A Methodology Based on Periodic Time Series», *IEEE Transactions on Power Systems*, **20** (3), p. 1622 - 1630, doi: 10.1109/TPWRS.2005.852123.

Everitt B. S., Landau S., Leese M. (2001), *Cluster Analysis*, Fourth Edition, Arnold Publishers.

Falvo M., Lamedica R., Prudenzi A. (2006), «Meteorological Parameters Influence for Medium Term Load Forecasting», in Proceedings of the IEEE Transmission and Distribution Conference and Exhibition, 1<sup>st</sup> - 24<sup>th</sup> May 2006, p. 1296-1301, doi: 10.1109/TDC.2006.1668698.

Fan S., Hyndman R. J. (2012), «Short-term Load Forecasting based on a Semi-Parametric Additive Model», *IEEE Transactions on Power Systems*, **27** (1), p. 134-141, doi: 10.1109/TPWRS.2011.2162082.

Feinberg E. A., Genethliou D. (2005), «Load Forecasting», in Chow J. H., Wu F.F., Momoh J.J. (eds.), in *Applied Mathematics for Restructured Electric Power Systems: Optimization, Control, and Computational Intelligence*, Springer US 2005, Chapter 12, p. 268-285.

Ferreira V. H., da Silva A.P.A. (2007), «Toward Estimating Autonomous Neural Network-Based Electric Load Forecasters», *IEEE Transactions on Power Systems*, **22** (4), p. 1554-1562, doi: 10.1109/TPWRS.2007.908438.

Fidalgo J. N., Matos M. A., Ribeiro L. (2012), «A new clustering algorithm for load profiling based on billing data», *Electric Power System Research*, **82** (1), p. 27-33, doi: 10.1016/j.epsr.2011.08.016.

Fidalgo J. N., Matos M. A. (2007), «Forecasting Portugal Global Load with Artificial Neural Networks», in Proceedings of the ICANN2007 - International Congress on Artificial Neural Networks 2007, Springer-Verlag, Part II, p. 728-737, ISBN: 3-540-74693-5 978-3-540-74693-5.

Fidalgo J. N., Peças Lopes J. A. (2005), «Load Forecasting Performance Enhancement when Facing Anomalous Events», *IEEE Transactions on Power Systems*, **20** (1), p. 408-415, doi: 10.1109/TPWRS.2004.840439.

Fidalgo J. N., Peças Lopes J. A. (2003), «Forecasting Active and Reactive Power at Substations' Transformers», in Proceedings of the IEEE Bologna PowerTech Conference, 23<sup>th</sup> – 26<sup>th</sup> June 2003, Vol. 1, ISBN: 0-7803-7967-5, doi: 10.1109/PTC.2003.1304157.

Fidalgo J.N. (2001), «Feature subset selection based on ANN sensitivity analysis – a practical study», in Proceedings of the WSES International Conference on Neural Networks and Applications, Tenerife, Spain, 11<sup>th</sup> –15<sup>th</sup> February; 2001. p. 206–211.

Gama J., Rodrigues P.P. (2008), «Electricity load forecast using data streams techniques», in *Knowledge discovery from sensor data*, Ganguly A.R., Gama J, Omiaomu O.A., Gaber M. M., Vatsavai R.R. (eds),. Florida: CRC Press; 2008. p. 131–47.

Gavrilas M., Gavrilas G., Sfintes, C. V. (2010), «Application of Honey Bee Mating Optimization Algorithm to Load Profile Clustering», in Proceedings of the IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA), Taranto, Italy, 6<sup>th</sup>-8<sup>th</sup> October 2010, doi: 10.1109/CIMSA.2010.5611759.

Gavrilas M. (2002), «Neural Network based forecasting for Electricity Markets», paper accepted for presentation at Telmark, London, UK, 2<sup>nd</sup> - 4<sup>th</sup> September 2002.

Gavrilas M., Ciutea I., Tanasa C. (2001), «Medium-Term Load Forecasting with Artificial Neural Network Models», Proceedings of the 16<sup>th</sup> International Conference and Exhibition on Electricity Distribution - CIRED, 18<sup>th</sup> -21<sup>th</sup> June 2001, Vol 482, doi: 10.1049/cp:20010930.

Gerbec D., Gasperic S., Smon I., Gubina F. (2005), «Allocation of the load profiles to consumers using probabilistic neural networks», *IEEE Transactions on Power Systems*, **20** (2), p. 548–555, doi: 10.1109/TPWRS.2005.846236.

Gerbec D., Gasperic S., Gubina F. (2004), «Determining the load profiles of consumers based on fuzzy logic and probability neural networks», *IEE Proceedings on Generation, Transmission and Distribution*, **151** (3), p. 395-400, doi: 10.1049/ip-gtd:20040472.

Gerse Á. (2007), «Temperature Sensitivity of the Electricity Demand in Hungary», in Proceedings of the First International Youth Conference on Energetics - IYCE 2007, Budapest, Hungary, 31<sup>st</sup> May – 2<sup>nd</sup> June 2007.

Guajardo J., Miranda J., Weber R.A. (2005), «Hybrid forecasting methodology using feature selection and support vector regression», in Proceedings of the HIS' 05 – fifth International Conference on Hybrid Intelligent Systems, Rio de Janeiro, Brasil, November 6<sup>th</sup>- 9<sup>th</sup> 2005, doi: 10.1109/ICHIS.2005.9

Guan C., Luh P. B., Coolbeth M. A., Zhao Y., Michel L. D., Chen Y., Manville C. J., Friedland P. B., Rourke S. J. (2009), «Very Short-term Load Forecasting: Multilevel Wavelet Neural Networks with Data Pre-Filtering», IEEE PES'09 – Power & Energy Society General Meeting, Calgary, Canada, 26<sup>th</sup>- 30<sup>th</sup> July 2009, p. 1-8, doi: 10.1109/PES.2009.5275296

Guo Y.C. (2009), «An Integrated PSO for Parameter Determination and Feature Selection of SVR and its Application in STLF», in Proceedings of the Eight International Conference on Machine Learning and Cybernetics, Baoding, China, 12<sup>th</sup> -15<sup>th</sup> July 2009, Vol. 1, p. 359-364, doi: 10.1109/ICMLC.2009.5212569.

Guo Y.C., Niu D., Chen Y. (2006), «Support Vector Machine Model in Electricity Load Forecasting», in Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, China, 13<sup>th</sup> -16<sup>th</sup> August 2006; p. 2982-2896, doi: 10.1109/ICMLC.2006.259076.

Guo X.C., Chen Z., Ge H., Liang Y. (2004), «Short-term load forecasting using neural network with principal component analysis», in Proceedings of the International Conference on Machine Learning and Cybernetics, Shanghai, China, 26<sup>th</sup> - 29<sup>th</sup> August 2004; Vol. 6, p. 3365-3369, doi: 10.1109/ICMLC.2004.1380362.

Hagan M. T., Demuth, H. B., Beale M. (1996), *Neural Network Design*, Boston MA, USA: PWS Publishing Company; 1996, ISBN: 0-9717321-0-8.

Hagan M.T., Menhaj, M.B. (1994), «Training feedforward networks with the Marquardt algorithm», *IEEE Transactions on Neural Networks*, **5** (6), p. 989–993.

Hamel L. (2009), *Knowledge Discovery with Support Vector Machines*, John Wiley and Sons, Inc., 2009, ISBN: 9780470371923, doi: 10.1002/9780470503065.

Hippert H.S., Taylor J.W. (2010), «An evaluation of Bayesian techniques for controlling model complexity and selecting inputs in a neural network for short-term load forecasting», *Elsevier - Neural Networks*, **23** (3), p.: 386-395, doi: 10.1016/j.neunet.2009.11.016.

Hippert H., Pedreira C. (2004), «Estimating temperature profiles for short-term load forecasting: neural networks compared to linear models», *IEE Proceedings on Generation, Transmission, Distribution*, **151** (4), p. 543-547, doi: 10.1049/ip-gtd:20040491.

Hippert H., Pedreira C., Souza R. C. (2001), «Neural Networks for Short-Term Load Forecasting: A Review and Evaluation», *IEEE Transactions on Power Systems*, **16** (1), p. 44-55, doi: 10.1109/59.910780.

IEA (2003), *The Power to choose – Enhancing demand response in liberalised markets*, Phil Harrington – International Energy Agency, Presented at Metering Europe; Billing Europe, CRM/CIS Europe, Paris 1<sup>st</sup> - 3<sup>rd</sup> October 2003.

Ivanov O., Gavrilas M. (2007), «Load Profiling by a Genetic Algorithm», in Proceedings of the First International Youth Conference on Energetics - IYCE 2007, Budapest, Hungary, 31<sup>st</sup> May – 2<sup>nd</sup> June 2007

Jardini J., Tahan C., Ahn S., Figueiredo F. M. (2000), «Daily Load Profiles for Residential, Commercial and Industrial Low Voltages Consumers», *IEEE Transactions on Power Delivery*, **15** (1), p. 375-380, doi: 10.1109/61.847276.

Jota P., Silva V., Jota F. (2011), «Building load management using cluster and statistical analyses», *Elsevier - International Journal of Electrical Power and Energy Systems*, **33** (8), p. 1498–1505, doi: 10.1016/j.ijepes.2011.06.034.

Kirschen D.S. (2003), «Demand-Side View of Electricity Markets», *IEEE Transactions on Power Systems – Invited Paper*, **18** (2), p. 520-527, doi: 10.1109/TPWRS.2003.810692.

Lamedica R., Prudenzi A., Sforza M., Caciotta M., Cencelli, V. (1996), «A neural network based technique for short-term forecasting of anomalous load periods», *IEEE Transactions on Power Systems*, **11** (4), p. 1749 – 1756, doi: 10.1109/59.544638.

- Lee M. (2009), «Using support vector machine with a hybrid feature selection method to the stock trend prediction», *Elsevier - Expert Systems with Applications*, **36** (8), p. 10896-10904, doi: 10.1016/j.eswa.2009.02.038.
- Mandal P., Senjyu T., Urasaki N., Funabashi T. (2006), «A neural network based several-hour-ahead electric load forecasting using similar days approach», *Elsevier - International Journal of Electrical Power and Energy Systems*; **28** (6), p. 367-373, doi: 10.1016/j.ijepes.2005.12.007.
- Mathworks Inc. (2009), *Genetic Algorithm and Direct Search Toolbox 2 User Guide*, Revised for Version 2.4.2 (Release 2009b), disponível em setembro de 2012.
- Matos M., Fidalgo J., Ribeiro L. (2005), «Deriving LV Load Diagrams for Market Purposes using Commercial Information», in Proceedings of the 13<sup>th</sup> International Conference on Intelligent Systems Application to Power Systems, 6<sup>th</sup> - 10<sup>th</sup> November 2005, p. 105-110, doi: 10.1109/ISAP.2005.1599248.
- Metaheuristics (2012), Documentação disponível em [www.metaheuristics.net](http://www.metaheuristics.net) e consultada em setembro de 2012.
- Miranda V., Monteiro C. (2000), «Fuzzy Inference in Spatial Load Forecasting», in Proceedings of the Power Engineering Society Winter Meeting, 23<sup>th</sup> - 27<sup>th</sup> January 2000, vol. 2, p. 1063-1068, doi: 10.1109/PESW.2000.850087.
- Miranda V., Monteiro C. (1999), «Fuzzy Inference Applied to Spatial Load Forecasting», in Proceedings of the International Conference on Electric Power Engineering - PowerTech, Budapest, Hungary, 29<sup>th</sup> August – 02<sup>nd</sup> September 1999, doi: 10.1109/PTC.1999.826435.
- Mohandes M. (2002), «Support Vector Machines for short-term electrical load forecasting», *Elsevier - International Journal of Energy Research*, **26** (4), p. 335-345, doi: 10.1002/er.787.
- Paatero J. V., Lund P. D. (2006), «A Model for Generating Household Electricity Load Profiles», *Elsevier - International Journal of Energy Research*, **30** (5), p. 273-290, doi: 10.1002/er.1136.
- Pai P., Hong W. (2005), «Support vector machines with simulated annealing algorithms in electricity load forecasting», *Elsevier - Energy Conversion and Management*, **46** (17), p. 2669-2688, doi: 10.1016/j.enconman.2005.02.004.

- Paiva J. P. Sucena (2005), *Redes de Energia Eléctrica: Uma análise sistemática*, IST Press, 2005, ISBN: 978-989-8481-06-1.
- Palit A.J., Popovic, D. (2005), *Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications*, Springer-Verlag London Limited 2005, ISBN-10: 1852339489.
- Ramezani M., Falaghi H., Haghifam M. and Shahryari G. A. (2005), «Short-Term Electric Load Forecasting Using Neural Networks», in Proceedings of the International Conference on Computer as a Tool - EUROCON, 21<sup>st</sup> – 24<sup>th</sup> November 2005, Vol. 2, p. 1525 – 1528, doi: 10.1109/EURCON.2005.1630255.
- Rudnick H. (1994), «Chile: Pioneer in Deregulation of the Electric Power Sector», *IEEE Power Engineering Review*, **14** (6), p. 28-30, doi: 10.1109/MPER.1994.286546.
- Santos P., Martins A., Pires A. (2007), «Designing the input vector to ANN-based models for Short-term load forecast in electricity distribution systems», *Elsevier - International Journal of Electrical Power and Energy Systems*, **29** (4), p. 338-347, doi: 10.1016/j.ijepes.2006.09.002.
- Santos P., Martins A., Martins F., Mendes R. (2006), «Short-term Load Forecast using Trend Information and Process Reconstruction», *International Journal of Energy Research*, **30** (10), p. 811-822, doi: 10.1002/er.1187.
- Santos P.C. (2005), *Previsão de Cargas em Horizonte de Curto-Prazo Aplicada a Subestações de Distribuição*, Tese de Doutoramento defendida na Universidade de Coimbra, Dezembro de 2005.
- Sarle W.S. (2000), «How to measure importance of inputs», part of *Neural Networks Frequently Asked Questions*, Cary, NC, USA: SAS Institute Inc., <ftp://ftp.sas.com/pub/neural/importance.html>, (revisto a 23.06.2000) (consultado a 05 de fevereiro de 2013).
- Schölkopf B., Smola A. J. (2002), *Learning with Kernels - Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2002, ISBN: 0262194759.
- Schölkopf B., Burges C. J., Smola A. J. (1998), *Advances in Kernel Methods: Support Vector Learning*, MIT Press, 1998, ISBN-10: 0262194163.



Soares I., Sarmiento P. (2010), *Does unbundling really matter? The Telecommunication and Electricity Cases*, Faculdade de Economia da Universidade do Porto, FEP Working Papers, n° 380, julho de 2010.

Sousa J. C., Jorge H. M., Neves L. P. (2013), «Short-term Load Forecasting based on Support Vector Regression and Load Profiling», aceite para publicação em *Wiley-Blackwell - International Journal of Energy Research*.

Sousa J. C., Neves L. P., Jorge H. M. (2012), «Assessing the relevance of load profiling information in electrical load forecasting based on neural networks», *Elsevier – International Journal of Electrical Power and Energy Systems*, **40** (1), p. 85-93, doi: 10.1016/j.ijepes.2012.02.008.

Srinivasan D., Tan S., Chang C., Chan E. (1999), «Parallel Neural Network-Fuzzy Expert System Strategy for Short-term Load Forecasting: System Implementation and Performance Evaluation», *IEEE Transactions on Power Systems*, **14** (3), p. 1100-1106, doi: 10.1109/59.780934.

Tao X., Renmu H., Peng W., Dongjie X. (2004), «Input Dimension Reduction for Load Forecasting Based on Support Vector Machines», in Proceedings of the IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies (DRPT2004), 5<sup>th</sup> -8<sup>th</sup> April 2004, Vol. 2, p. 510-514, doi: 10.1109/DRPT.2004.1338036.

Tikka J. (2009), «Simultaneous input variable and basis function selection for RBF networks», *Elsevier – Neurocomputing*, **72** (10-12), p. 2649-2658, doi: 10.1016/j.neucom.2008.10.003.

Üstün B., Melssen W., Buydens L. (2007), «Visualisation and interpretation of Support Vector Regression models», *Elsevier - Analytica Chimica Acta*, p. 299-309, doi: 10.1016/j.aca.2007.03.023.

Willis H.L. (2004), *Power Distribution Planning Reference Book*, Marcel Dekker, Inc. 2004, Second Edition, Revised and Expanded, ISBN: 0824700988.

Willis H.L., NorthCote-Green J.E.D (1983), «Spatial Electric Load Forecasting: A tutorial review», in Proceedings of the IEEE, **71** (2), p. 232-253, doi: 10.1109/PROC.1983.12562.

Wu H.C., Tsai C.S., Lu C.N. (1999), «Load forecast in a competitive retail market», in Proceedings of the IEEE Conference on Transmission and Distribution, New Orleans LA, 11<sup>th</sup> -16<sup>th</sup> Apr 1999, p. 160-164, doi: 10.1109/TDC.1999.755333.

Xia C., Wang J., McMenemy K. (2010), «Short, medium and long term load forecasting model and virtual load forecaster based on radial basis function neural networks», *International Journal of Electrical Power and Energy Systems*, **32** (7), p. 743–750, doi: 10.1016/j.ijepes.2010.01.009.

Yeung D.S., Cloete I., Shi D., Ng W.W.Y. (2010), *Sensitivity Analysis for Neural Networks*, Natural Computing Series, Springer-Verlag Berlin Heidelberg, 2010, ISBN: 978-3-642-02531-0, doi: 10.1007/978-3-642-02532-7.

Zhao Y., Luh P. B., Bomgardner C., Beerel G. H. (2009), «Short-term load forecasting: Multi-level wavelet neural networks with holiday corrections», IEEE PES'09 – Power & Energy Society General Meeting, Calgary, Canada, 26<sup>th</sup> - 30<sup>th</sup> July 2009, p. 1-7, doi: 10.1109/PES.2009.5275304.

Ziser C. J., Dong Z. Y., Saha T. (2005), «Investigation of Weather Dependency and Load Diversity on Queensland Electricity Demand», in Proceedings of the Australasian Universities Power Engineering Conference, Michael Negnevitsky (eds.), Hobart, Tasmania, 25<sup>th</sup>- 28<sup>th</sup> September 2005, p. 457-462, ISBN: 1-86295-277-9.