

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



Previsão de Consumos Elétricos, com Incertezas e Dependências, para Múltiplos Pontos de Carga

Jorge Emanuel Paulino e Pessoa

Mestrado Integrado em Engenharia Eletrotécnica e de Computadores

Orientador: Professor Doutor Cláudio Monteiro

22 de Julho de 2014

A Dissertação intitulada

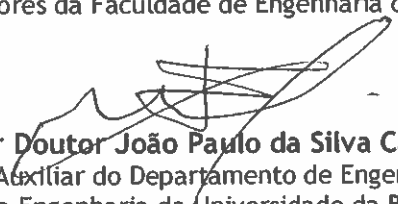
***“Previsão de Consumos Elétricos, com Incertezas e Dependências, para
Múltiplos Pontos de Carga”***

foi aprovada em provas realizadas em 17-07-2014

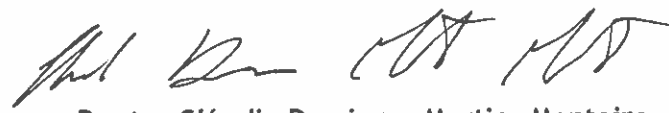
o júri



Presidente **Professor Doutor José Nuno Moura Marques Fidalgo**
Professor Associado do Departamento de Engenharia Eletrotécnica e de
Computadores da Faculdade de Engenharia da Universidade do Porto

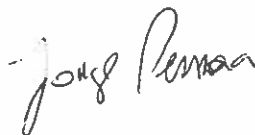


Professor Doutor João Paulo da Silva Catalão
Professor Auxiliar do Departamento de Engenharia Electrotécnica da Faculdade de
Ciências da Engenharia da Universidade da Beira Interior



Professor Doutor Cláudio Domingos Martins Monteiro
Professor Auxiliar do Departamento de Engenharia Eletrotécnica e de Computadores
da Faculdade de Engenharia da Universidade do Porto

O autor declara que a presente dissertação (ou relatório de projeto) é da sua exclusiva autoria e foi escrita sem qualquer apoio externo não explicitamente autorizado. Os resultados, ideias, parágrafos, ou outros extratos tomados de ou inspirados em trabalhos de outros autores, e demais referências bibliográficas usadas, são corretamente citados.



Autor - Jorge Emanuel Paulino e Pessoa

Faculdade de Engenharia da Universidade do Porto

Resumo

A premência em conseguir antever as necessidades energéticas da rede, assim como garantir a fiabilidade da mesma, de forma a que não ocorram sobrecargas nos transformadores nem nas linhas de transporte, evidenciam a necessidade do desenvolvimento de modelos de previsão que possam, com a melhor precisão possível, facilitar a tomada de decisões por parte das empresas e dos operadores do sistema.

É neste âmbito que se insere o tema desta dissertação, cuja finalidade é o desenvolvimento de modelos de previsão, a curto prazo, do consumo de eletricidade. Para isso, foram usadas ferramentas como o Microsoft Excel, fundamental no desenvolvimento e avaliação dos modelos, e um programa baseado no *Kernel Density Estimation* (KDE).

Numa primeira fase, foram estudadas as variáveis que, possivelmente, seriam mais pertinentes para a previsão de modo a serem incluídas nos modelos. Com o primeiro conjunto de variáveis determinadas, foram criados modelos, testados, posteriormente, em novos conjuntos de dados de forma a poder avaliar a sua performance em diferentes situações temporais. Em seguida, foram efetuadas dois novos modelos - inclusão de uma nova variável meteorológica e inclusão de histórico do consumo - de forma a perceber que influência teriam no resultado final da previsão. Foi ainda analisada a correlação entre os diferentes pontos de carga da rede para avaliar a importância que esta informação traria para o aumento da precisão dos modelos de previsão.

Constatou-se que os modelos criados com o primeiro conjunto de variáveis apresentaram uma performance satisfatória, visto que a sua utilização em dados referentes a outros anos provocou, na sua maioria, uma melhoria no valor do erro MAPE entre o consumo real e o consumo previsto. Já para a previsão da incerteza, o comportamento não foi tão evidente, levando ao aumento do erro MAPE em alguns casos, mas apresentando valores interessantes quando calculada para o Agregado das subestações.

Apesar de se tratar de uma situação nem sempre possível, em tempo real, pela dificuldade em se obter um histórico do consumo, nomeadamente referente ao dia anterior ao da previsão, a utilização desta variante nos modelos de previsão revelou-se fundamental para o decréscimo do erro entre o consumo real e o consumo previsto, pelo que obter este histórico será a melhor solução.

Palavras chave: previsão de consumo, incerteza da previsão, *Kernel Density Estimation*, MAPE.

Abstract

The urgency in being able to foresee the energy needs of the network, as well as ensuring its own reliability, in a way that overloads don't occur in transformers or transport lines, show the need of developing forecasting models which can facilitate, with the best possible accuracy, the process of making a decision by companies and system operators.

Is in this context that the dissertation theme fits, having as a goal the development of short-term forecasting models to the electricity consumption. To achieve that, tools like Microsoft Excel, absolutely essential in models development and evaluation, and a program based on Kernel Density Estimation (KDE) have been used.

In a first stage, the variables that possibly would have more interest to the forecast were studied in order to be included in models. With the first set of certain variables, models were created and, afterwards, they were tested in new sets of data in order to allow the evaluation of its performance in different time situations. Then, two new models were performed - inclusion of a new weather variable and inclusion of the consumption historic - in order to understand their influence on the final result of the forecast. It was also analyzed the correlation between the different network load points to evaluate the importance that this information would bring to the increase of accuracy in forecasting models.

It was found that created models with the first set of variables showed a satisfactory performance now that its using on data related to another years caused, in general, an improvement in the error value (MAPE) between the real consumption and the forecasted consumption. Otherwise, for the uncertainty forecast, the behavior was not so clear, causing an increase of the MAPE error in some cases, but showing interesting values when calculated for the Aggregate of the substations.

Despite being a situation not always possible, on real time, due to the difficulty of obtaining a consumption historic, particularly referring to the previous day of forecast, the using of this variant on forecasting models revealed itself as fundamental to the decrease of the error between real consumption and the forecasted one, thus obtaining this historic will be the best option.

Keywords: consumption forecasting, forecasting uncertainty, Kernel Density Estimation, MAPE.

Agradecimentos

Se sou Engenheiro, devo-o a várias pessoas, pelo que gostaria de agradecer,:

À Smartwatt, S.A. pelos dados que me possibilitaram concretizar esta dissertação.

Ao professor Dr. Cláudio Monteiro, pela oportunidade para desenvolver capacidades a nível académico e pessoal.

Ao Eng. Bruno Santos, pela flexibilidade que demonstrou, assim como apoio e empatia.

Aos meus companheiros desta reta final, Diana Lima, Marta Rolo e Vasco Seifert, pelos bons momentos de camaradagem, diversão e entreajuda que demonstraram.

Aos meus amigos, pois sem eles teria terminado o curso um ano mais cedo, mas me deram a oportunidade de consolidar conhecimentos para além de aproveitar momentos inesquecíveis. Obrigado Nuno Soares, Nuno Moreira, Hui Xia, entre outros de que agora não me lembro, mas que nunca serão esquecidos. #BB.

Ao Eng. Carlos Sampaio e sua equipa, Eng. José Oliveira, Eng. Pedro Araújo e Eng. Sérgio Carvalho, pela possibilidade que me facilitaram de, ao longo desta minha caminhada, nos meses de verão, contactar com o mundo empresarial da energia, assim como os momentos de descontração.

Ao meu pai por me ter transmitido a resiliência necessária para que não mudasse de curso e levasse esta etapa até ao fim.

À minha mãe, pelo carinho, compreensão e por ter sempre o gabinete de advocacia aberto para me defender.

À minha irmã, pela meiguice e capacidade de me fazer sorrir mesmo quando os momentos a isso não propiciavam.

Aos três, pela preocupação constante e apoio, desculpando a má disposição por noites mal dormidas.

Um bem-haja a todos,

Jorge Emanuel Paulino e Pessoa

“Life without regrets is not possible. We all have some options and choices in our lives (hence the possibility of error). We all have some perceived control over aspects of our lives (hence the possibility of feeling responsible for errors). Regret serve an important role in directing our movement in the world.”

Kenshin Himura

Conteúdo

1	Introdução	1
1.1	Enquadramento	1
1.2	Motivação e objetivos	2
1.3	Estrutura da Dissertação	3
2	Revisão Bibliográfica	5
2.1	Fatores influenciadores dos padrões de consumo	5
2.2	Classificação da previsão de consumo	6
2.3	Metodologias de previsão de consumo	7
2.3.1	Abordagem de dia semelhante	7
2.3.2	Métodos de Regressão	8
2.3.3	Séries Temporais	8
2.3.4	Redes Neurais	9
2.3.5	Lógica Difusa (<i>Fuzzy Logic</i>)	10
2.3.6	Ferramentas de Previsão Probabilística	10
2.4	Modelos numéricos de previsão meteorológica (NWP)	14
2.4.1	Modelos globais	15
2.4.2	Modelos regionais	16
2.5	Avaliação da precisão dos modelos de previsão	16
2.5.1	Coefficiente de determinação (R^2) e coeficiente de correlação de Pearson (R)	16
2.5.2	Índices de erro	17
3	Metodologia para a construção de modelos de previsão	19
3.1	Caso de estudo e dados utilizados	19
3.2	Obtenção e Tratamento de dados	25
3.2.1	Etapa 1 - Acerto da Hora	25
3.2.2	Etapa 2 - Caracterização dos dados	26
3.2.3	Etapa 3 - Normalização dos dados	29
3.2.4	Etapa 4 - Escolha das Variáveis	35
3.3	Construção dos modelos de previsão	44
3.3.1	Descrição do processo de construção	44
3.3.2	Modelo 1 - Modelo Base Simples	49
3.3.3	Modelo 2 - Modelo Simples	49
3.3.4	Modelo 3 - Modelo Simples com explicador Temperatura	49
3.3.5	Modelo 4 - Modelo Base com explicador Temperatura e Irradiância	49
3.3.6	Modelo 5 - Modelo Base com explicador Temperatura e Consumo do Dia Anterior	50

4	Análise de desempenho de modelos de previsão	51
4.1	Comparação dos Casos de Estudo	51
4.1.1	Caso de Estudo 1	51
4.1.2	Caso de Estudo 2	57
5	Previsão com interdependência geográfica	61
5.1	Divisão dos dados	61
5.2	Estudo da correlação	62
5.3	Cálculo da previsão	63
6	Conclusões e Trabalhos Futuros	67
6.1	Satisfação dos Objectivos	67
6.2	Trabalhos Futuros	68
A	Caso de Estudo 1	69
B	Caso de Estudo 2	77
C	Previsão com interdependência geográfica	95
C.1	Correlação	95
C.2	Previsão	104
	Referências	115

Lista de Figuras

1.1	Países que integram o projeto SiNGULAR. [1]	2
2.1	Histograma aplicado a uma função de densidade de probabilidade beta.	11
2.2	Função densidade de probabilidade da distribuição beta para diferentes valores dos coeficientes de forma α e β .	14
3.1	Consumo Anual e Ponta de Consumo em Creta até 2010. [33]	20
3.2	Consumo Anual e Ponta de Consumo em Creta até 2020. [33]	20
3.3	Produção Anual em Creta até 2010, por tipo de central. [33]	21
3.4	Produção Anual em Creta até 2010, por tipo de combustível usado. [33]	21
3.5	Divisão dos dados (in-sample e out-sample) para o caso de estudo 1.	23
3.6	Divisão dos dados (in-sample e out-sample) para o caso de estudo 2.	24
3.7	Etapas do processo de tratamento de dados.	25
3.8	Consumos por setor, em Creta, no ano de 2005. [33]	27
3.9	Consumo residencial por aplicação, em Creta, no ano de 2005. [33]	28
3.10	Etapas do processo de tratamento de dados.	29
3.11	Gráfico de Dispersão que relaciona os Padrões Diário (Adimensional) e Médio Diário (Adimensional)	31
3.12	Gráfico de Dispersão que relaciona os Padrões Diário (Adimensional) e Médio Diário (Adimensional), intersetado por retas de exclusão	31
3.13	Gráfico resultante da filtragem pelas retas de exclusão.	32
3.14	Gráfico de Dispersão que relaciona os Padrões Semanal (Adimensional) e Médio Semanal (Adimensional)	33
3.15	Gráfico de Dispersão que relaciona os Padrões Semanal (Adimensional) e Médio Semanal (Adimensional), intersetado por retas de exclusão	34
3.16	Gráfico resultante da filtragem pelas retas de exclusão.	34
3.17	Padrões de consumo em função das horas do dia (em Creta), separado por dias úteis, fins de semana e feriados.	36
3.18	Padrões de consumo em função das horas do dia (em Creta), em diferentes períodos do ano.	36
3.19	Padrões de consumo em função das horas do dia (em Creta), em diferentes períodos do ano.	37
3.20	Previsão probabilística de consumo, para um quantil de 10 e 90.	37
3.21	Gráfico de Dispersão que relaciona o Consumo com a Hora do Dia.	38
3.22	Gráfico de Dispersão que relaciona o Consumo com o Dia da Semana.	39
3.23	Gráfico de Dispersão que relaciona o Consumo com o Mês do Ano.	40
3.24	Gráfico de Dispersão que relaciona o Consumo com a Temperatura.	41

3.25 Diagrama de Carga do Agregado fornecido e da soma de consumo de todas as subestações.	42
3.26 Gráfico de Dispersão que relaciona o Consumo com a Irradiância.	43
3.27 Gráfico de Dispersão que relaciona o Consumo com o Consumo do Dia Anterior.	44
3.28 Etapas do Processo Iterativo de Criação dos Modelos de Previsão.	45
3.29 Comparação das probabilidades obtidas em cada banda de incerteza.	48
3.30 Comparação das probabilidades obtidas em cada banda de incerteza.	48
4.1 Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, usando o Modelo 1 - Modelo Base Simples.	53
4.2 Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, usando o Modelo 3.	54
4.3 Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, usando o Modelo 3.	54
4.4 Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, usando o Modelo 3.	55
4.5 Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, usando o Modelo 3.	55
A.1 Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Hrakleio2.	69
A.2 Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Hrakleio2.	70
A.3 Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Hrakleio2.	70
A.4 Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Hrakleio2.	71
A.5 Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Hrakleio3.	71
A.6 Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Hrakleio3.	72
A.7 Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Hrakleio3.	72
A.8 Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Hrakleio3.	73
A.9 Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Rethymno.	73
A.10 Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Rethymno.	74
A.11 Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Rethymno.	74
A.12 Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Rethymno.	75
B.1 Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para o Agregado.	77
B.2 Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para o Agregado.	78

B.3	Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para o Agregado.	78
B.4	Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para o Agregado.	79
B.5	Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Ag.Nikolaos.	79
B.6	Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Ag.Nikolaos.	80
B.7	Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Ag.Nikolaos.	80
B.8	Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Ag.Nikolaos.	81
B.9	Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Chania.	81
B.10	Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Chania.	82
B.11	Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Chania.	82
B.12	Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Chania.	83
B.13	Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Hrakleio1.	83
B.14	Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Hrakleio1.	84
B.15	Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Hrakleio1.	84
B.16	Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Hrakleio1.	85
B.17	Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Hrakleio2.	85
B.18	Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Hrakleio2.	86
B.19	Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Hrakleio2.	86
B.20	Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Hrakleio2.	87
B.21	Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Hrakleio3.	87
B.22	Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Hrakleio3.	88
B.23	Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Hrakleio3.	88
B.24	Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Hrakleio3.	89
B.25	Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Rethymno.	89
B.26	Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Rethymno.	90

B.27	Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Rethymno.	90
B.28	Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Rethymno.	91
B.29	Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Sitia.	91
B.30	Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Sitia.	92
B.31	Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Sitia.	92
B.32	Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Sitia.	93
C.1	Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Hrakleio2, para o conjunto 1 (dias úteis).	95
C.2	Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Hrakleio2, para o conjunto 2 (restantes dias).	96
C.3	Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Hrakleio3, para o conjunto 1 (dias úteis).	96
C.4	Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Hrakleio3, para o conjunto 2 (restantes dias).	97
C.5	Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Rethymno, para o conjunto 1 (dias úteis).	97
C.6	Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Rethymno, para o conjunto 2 (restantes dias).	98
C.7	Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio2 com o Consumo da subestação Hrakleio3, para o conjunto 1 (dias úteis).	98
C.8	Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio2 com o Consumo da subestação Hrakleio3, para o conjunto 2 (restantes dias).	99
C.9	Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio2 com o Consumo da subestação Rethymno, para o conjunto 1 (dias úteis).	99
C.10	Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio2 com o Consumo da subestação Rethymno, para o conjunto 2 (restantes dias).	100
C.11	Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio3 com o Consumo da subestação Hrakleio2, para o conjunto 1 (dias úteis).	100
C.12	Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio3 com o Consumo da subestação Hrakleio2, para o conjunto 2 (restantes dias).	101
C.13	Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio3 com o Consumo da subestação Rethymno, para o conjunto 1 (dias úteis).	101
C.14	Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio3 com o Consumo da subestação Rethymno, para o conjunto 2 (restantes dias).	102
C.15	Gráfico de Dispersão que relaciona o Consumo da subestação Rethymno com o Consumo da subestação Hrakleio2, para o conjunto 1 (dias úteis).	102
C.16	Gráfico de Dispersão que relaciona o Consumo da subestação Rethymno com o Consumo da subestação Hrakleio2, para o conjunto 2 (restantes dias).	103
C.17	Gráfico de Dispersão que relaciona o Consumo da subestação Rethymno com o Consumo da subestação Hrakleio3, para o conjunto 1 (dias úteis).	103
C.18	Gráfico de Dispersão que relaciona o Consumo da subestação Rethymno com o Consumo da subestação Hrakleio3, para o conjunto 2 (restantes dias).	104

C.19	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 1 (dias úteis), com base na correlação com o Agregado.	104
C.20	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 2 (restantes dias), com base na correlação com o Agregado.	105
C.21	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 1 (dias úteis), com base na correlação com o Agregado.	105
C.22	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 2 (restantes dias), com base na correlação com o Agregado.	106
C.23	Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 1 (dias úteis), com base na correlação com o Agregado.	106
C.24	Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 2 (restantes dias), com base na correlação com o Agregado.	107
C.25	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 1 (dias úteis), com base na correlação com a subestação Hrakleio2.	107
C.26	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 2 (restantes dias), com base na correlação com a subestação Hrakleio2.	108
C.27	Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 1 (dias úteis), com base na correlação com a subestação Hrakleio2.	108
C.28	Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 2 (restantes dias), com base na correlação com a subestação Hrakleio2.	109
C.29	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 1 (dias úteis), com base na correlação com a subestação Hrakleio3.	109
C.30	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 2 (restantes dias), com base na correlação com a subestação Hrakleio3.	110
C.31	Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 1 (dias úteis), com base na correlação com a subestação Hrakleio3.	110
C.32	Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 2 (restantes dias), com base na correlação com a subestação Hrakleio3.	111
C.33	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 1 (dias úteis), com base na correlação com a subestação Rethymno.	111
C.34	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 2 (restantes dias), com base na correlação com a subestação Rethymno.	112
C.35	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 1 (dias úteis), com base na correlação com a subestação Rethymno.	112
C.36	Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 2 (restantes dias), com base na correlação com a subestação Rethymno.	113

Lista de Tabelas

3.1	Capacidade Instalada nas centrais de Creta. [33]	21
3.2	Subestações existentes na ilha de Creta.	22
3.3	Modelos e subestações usados para o caso de estudo 1.	23
3.4	Modelos e subestações usados para o caso de estudo 2.	24
3.5	Mudanças de hora em Creta, nos anos de 2012, 2013 e 2014.	26
3.6	Comparação entre Hora Local e Hora UTC nos diferentes fusos horários.	26
3.7	Associação entre Dia da Semana e número auxiliar.	27
3.8	Feriados nos anos de 2012, 2013 e 2014.	28
3.9	Tabela demonstrativa do cálculo de EMD, PD e PMD.	30
3.10	Tabela demonstrativa do cálculo de EMS, PS e PMS.	33
3.11	Tabela demonstrativa do cálculo de EMA, PA e PMA.	35
3.12	Processo de seleção do melhor parâmetro h de cada variável.	46
4.1	Modelos e subestações usados para o caso de estudo 1.	52
4.2	Modelos e subestações usados para o caso de estudo 2.	57
4.3	Modelos e subestações usados para o caso de estudo 2.	58
4.4	Modelos e subestações usados para o caso de estudo 2.	59
5.1	Coefficiente R^2 e Coeficientes de Pearson para a relação entre o consumo do Agregado e o consumo das subestações.	62
5.2	Equações das linhas de tendência polinomial de 2ª ordem (do tipo $y = a \cdot x^2 + b \cdot x + c$) obtidas pela correlação entre o Agregado e cada uma das subestação, para o conjunto de dados especificado.	63
5.3	Valor do erro MAPE para a previsão com interdependência geográfica.	64
5.4	Comparação do erro MAPE entre a previsão com interdependência geográfica e a previsão usando o Modelo 3 - Modelo Base com explicador Temperatura.	65

Abreviaturas e Símbolos

FEUP	Faculdade de Engenharia da Universidade do Porto
SiNGULAR	<i>Smart and Sustainable Insular Electricity Grids Under Large-Scale Renewable Integration</i>
KDE	<i>Kernel Density Estimation</i>
RES	<i>Renewable Energy Sources</i>
PIB	Produto Interno Bruto
AR	<i>Autoregressive Model</i>
ARMA	<i>Autoregressive Moving-Average</i>
ARIMA	<i>Autoregressive Integrated Moving-Average</i>
ARIMAX	<i>Autoregressive Integrated Moving-Average with exogenous variables</i>
NN	Redes Neurais
ANN	Redes Neurais Artificiais
fdp	Funções de Densidade de Probabilidade
NWP	Modelos Numéricos de Previsão Meteorologia
GFS	<i>Global Forecasting Model</i>
NCEP	<i>National Centers for Environmental Prediction</i>
GEM	<i>Global Environmental Multiscale</i>
IFS	<i>Integrated Forecast System</i>
ECMWF	<i>European Centre for Medium-Range Weather Forecasts</i>
NAVGEN	<i>Fleet Numerical Meteorology and Oceanography Center</i>
WRF	<i>Weather Research and Forecasting Model</i>
NCAR	<i>National Center for Atmospheric Research</i>
NAM	<i>North American Mesoscale Model</i>
MM5	<i>Fifth Generation Penn State/NCAR Mesoscale Model</i>
HIRLAM	<i>High Resolution Limited Area Model</i>
ALADIN	<i>Aire Limitée Adaptation Dynamique Développement InterNational</i>
RMSE	<i>Root Mean Square Error</i>
nRMSE	<i>Normalized Root Mean Square Error</i>
MAE	<i>Mean Absolute Error</i>
MAPE	<i>Mean Absolute Percentage Error</i>
MSE	<i>Mean Square Error</i>
EMD	Energia Média Diária
PD	Padrão Diário
PMD	Padrão Médio Diário
EMS	Energia Média Semanal
PS	Padrão Semanal
PMS	Padrão Médio Semanal
EMA	Energia Média Anual
PA	Padrão Anual
PMA	Padrão Médio Anual
CDA	Consumo do Dia Anterior

Capítulo 1

Introdução

A presente dissertação foi realizada no âmbito do Mestrado Integrado em Engenharia Eletrotécnica e de Computadores, da Faculdade de Engenharia da Universidade do Porto (FEUP).

Quer os dados usados no decorrer do trabalho prático quer a ferramenta computacional utilizada foram fornecidos pela *Smartwatt, SA*, uma empresa de serviços de energia que fornece soluções para sistemas de energia. Esta empresa atua em três vetores distintos: energias renováveis, eficiência energética e Sistemas de Inteligência, com o objetivo de otimizar o uso dos recursos energéticos existentes.

Ao longo da dissertação foram utilizadas ferramentas como o Microsoft Excel, fundamental na modelização, desenvolvimento e avaliação dos modelos, criados a partir de um programa baseado no *Kernel Density Estimation* (KDE), que consiste numa ferramenta de inteligência computacional que produz as previsões.

1.1 Enquadramento

O tema para esta dissertação insere-se no âmbito do projeto europeu SiNGULAR [1] (Smart and Sustainable Insular Electricity Grids Under Large-Scale Renewable Integration). Trata-se de um projeto na área da sustentabilidade energética, tendo como principal foco a gestão de energia em ilhas de vários países através da construção de smart grids (redes inteligentes de gestão de energia).

Este projeto resulta de uma parceria entre onze empresas e cinco universidades de sete países (1.1) (Chipre, Espanha, Grécia, Itália, Portugal, Roménia e Suíça), tendo como objetivo comum a criação de soluções energéticas, no campo das energias renováveis, que garantam a autonomia dos recursos, assim como uma boa gestão dos mesmos. Esta gestão permitirá uma otimização da energia utilizada através da previsão das necessidades energéticas de forma a, por exemplo, compensar aumentos de consumo pela entrada em serviço de centrais controláveis ou possibilitar a compra atempada de combustível em centrais que necessitem dessa matéria prima, evitando, assim, falhas de abastecimento. Além disso, pode, ainda, propiciar que a compra seja feita em períodos em que se pratique preços mais reduzidos.



Figura 1.1: Países que integram o projeto SiNGULAR. [1]

Esta dissertação tem como foco rede elétrica da ilha de Creta, na Grécia, sendo objetivo principal a previsão dos consumos a curto prazo (até uma semana) e, assim, a permitir uma melhor e mais eficaz regulação da rede.

1.2 Motivação e objetivos

A eficiência energética é hoje, cada vez mais, um dos pilares da política energética sustentável, a par das energias renováveis. É, por isso, fundamental que se proceda a uma gestão da energia necessária em cada momento do dia, de forma a que as necessidades sejam colmatadas, mas também para que a utilização de centrais controláveis não ultrapasse o desejável. É desta intenção de gerir melhor os recursos que surge a importância da previsão dos consumos.

A previsão de consumo de energia elétrica baseia-se na tarefa de prever, com a melhor precisão possível, os valores de consumo de energia para diferentes horizontes temporais futuros. Trata-se de uma tarefa complexa pelo facto de o consumo ser influenciado por diversas variáveis, como é o caso do dia da semana, do mês do ano ou da temperatura, o que adiciona um fator de incerteza às previsões efetuadas.

Pelas incertezas inerentes ao processo há, por isso, a necessidade de que estas previsões de curto prazo sejam feitas com a maior precisão possível, de modo a que se possa tirar o melhor partido da informação proporcionada por estas previsões, quer ao nível da antecipação de situações de sobrecarga nos transformadores ou nas linhas de transporte quer pela sua importância na operação económica dos sistemas elétricos. Isto porque estes fatores são fundamentais para o balanceamento do mercado que opera com base nas propostas de compra e venda de energia elétrica efetuada com um dia de antecedência.

Há ainda que ter em consideração o risco inerente a estas previsões, já que, na eventualidade de uma previsão incorreta, a repercussão no mercado será significativa, pois se o consumo real ultrapassar a quantidade de energia adquirida, haverá a necessidade de se recorrer ao mercado para colmatar essa diferença o que causará sobrecustos desnecessários, pela emergência dessa necessidade.

Isto demonstra a importância de se proceder a um estudo cuidadoso das variáveis que integrem maior informação quanto ao comportamento dos padrões de consumo, assim como perceber qual a influência que os múltiplos pontos de carga da rede têm para a evolução dos diagramas de carga. Assim, previsões mais precisas permitem um maior lucro, por poupança na escolha mais eficaz da entrada em serviço de centrais controláveis e por redução da necessidade de recurso aos mercados para compensações inesperadas.

Tendo em conta a premência em efetuar boas previsões de forma a aumentar os lucros e a aumentar a fiabilidade do sistema, os objetivos a atingir nesta dissertação foram os seguintes:

- Desenvolver modelos de previsão a curto prazo (7 dias) de consumos de eletricidade, em múltiplos pontos de carga;
- Modelizar a incerteza nos vários pontos de carga, utilizando ferramentas de inteligência computacional.

1.3 Estrutura da Dissertação

Nesta dissertação, para além do capítulo introdutório, há outros quatro capítulos.

No capítulo 2, é feita a descrição do estado da arte, sendo apresentadas abordagens já utilizadas na resolução de problemas relacionados com a temática em análise.

No capítulo 3, é explicada a génese dos diferentes modelos de previsão, nomeadamente o programa utilizado, as variáveis consideradas e a metodologia seguida, entre outros.

No capítulo 4, são apresentados os resultados obtidos através dos modelos criados.

No capítulo 5, é criado e testado um modelo de previsão por interdependência geográfica.

No capítulo 6, são tiradas conclusões com base nos resultados obtidos e feitas considerações sobre eventuais trabalhos futuros.

Capítulo 2

Revisão Bibliográfica

A previsão de consumo é um processo fulcral no planeamento e funcionamento dos sistemas elétricos e, por isso, inúmeras companhias e centros de investigação investem, continuamente, na investigação e desenvolvimento de modelos mais precisos e eficazes. Todo este processo em busca da melhoria resultou na existência de diversas metodologias de previsão que divergem em horizonte temporal, erro de previsão e complexidade computacional.

Este capítulo não tem como objetivo a apresentação detalhada de todos estes modelos, mas explorar os fatores que influenciam os padrões de carga e demonstrar a existência de diferentes métodos para a previsão de consumo. Além disso, proceder-se-á à abordagem dos modelos de previsão meteorológica existentes e descrever-se-ão parâmetros usados para a avaliação da precisão dos modelos de previsão.

2.1 Fatores influenciadores dos padrões de consumo

As redes energéticas são compostas por uma agregação de diferentes pontos de consumo, o que provoca, pelas variações que estes originam, diferenças nos padrões de carga consoante as circunstâncias criadas. É, por isso, crucial um estudo pormenorizado dos fatores que vão influenciar o consumo, de forma a se apresentarem soluções que permitam minimizar a diferença entre o consumo real e o consumo previsto.

Existem diversos fatores que influenciam os padrões de consumo, diferindo, por exemplo, no momento de inicialização e na sua duração.

De acordo com [2], os fatores podem ser divididos em dois grupos: eventos especiais como, por exemplo, greves ou eventos desportivos e fatores “sempre presentes”, nomeadamente condições meteorológicas e conduta humana.

O grupo de eventos especiais é caracterizado pelo facto de, normalmente, o momento em que é inicializado e a sua duração ser conhecido, como é o caso dos feriados (período da Páscoa, do Natal e do Ano Novo), eventos desportivos, greves ou dias de eleições. No entanto, casos como blackout's, desastres naturais ou eventos com grande influência na rede como, por exemplo, o

encerramento de um grande consumidor de energia, também integram o grupo de eventos especiais apesar de ser mais complicado definir o seu momento de início ou a sua duração [3][4][5].

De entre os fatores “sempre presentes” mencionados, as condições meteorológicas são as que apresentam maior preponderância [3][6][7]. sendo de especial importância elementos como temperatura real, temperatura aparente e humidade. Ainda assim, os comportamentos humanos não são de excluir, pois a procura de conforto térmico leva à utilização de reguladores de temperatura, tais como ares-condicionados e aquecedores, que afetam o consumo. Esse consumo é também afetado por comportamentos menos esporádicos, como a utilização de equipamentos eletrónicos, cada vez mais indispensáveis no dia-a-dia, como os computadores, televisões, fornos, etc.

Há ainda a considerar outros fatores externos como, por exemplo, o nível de consumo na produção industrial e regulações no consumo derivados do estado económico do país e do seu PIB [2] [3] [8] [9].

2.2 Classificação da previsão de consumo

A previsão de carga pode ser dividida consoante o seu horizonte temporal, constituindo assim três tipos de previsão de consumo [2][10][11]:

- **previsões a curto prazo** - cobrem espaços temporais desde uma hora até uma semana. São usadas para estimar fluxos de carga e para tomar decisões que possam prevenir sobrecargas. A implementação oportuna de tais decisões leva à melhoria da fiabilidade da rede e à redução de ocorrências de falhas e blackouts nos equipamentos. Incluem métodos de aproximação por dia semelhante (*similar day approach*), de regressão, de séries temporais, redes neuronais, algoritmos de aprendizagem estatística, de lógica difusa (*fuzzy logic*) [10][12];
- **previsões a médio prazo** - estabelecidos para prazos de uma semana a um ano, são usadas para estudos de contingência, manutenção de equipamentos e planeamento financeiro. Recorre-se a métodos de uso final (*end-use*) e abordagens econométricas;
- **previsões a longo prazo** - para prazos superiores a um ano e cuja precisão depende fortemente da previsão meteorológica. Normalmente, os resultados são apresentados sob a forma de distribuição de probabilidades. Esta previsão utiliza os mesmos métodos da previsão a médio prazo.

No âmbito desta dissertação será tido em consideração apenas a previsão a curto prazo que, de entre as muitas técnicas desenvolvidas para este modelo, estas podem ser classificadas em técnicas estatísticas (abordagem de dia semelhante (*similar day approach*), modelos de regressão e séries temporais) e de inteligência artificial (redes neuronais, sistemas periciais, lógica difusa (*fuzzy logic*), e algoritmos de aprendizagem estatística).

Abordagens estatísticas requerem uma formulação de modelo matemático da carga como função de diferentes fatores, incluindo tempo, condições meteorológicas e tipo de utilizadores. Dentro dos modelos matemáticos, as principais categorias são:

- **modelos aditivos** - previsão de carga feita pela adição de um número de componentes [13], como apresentado na equação 2.2:

$$L = L_n + L_w + L_s + L_r \quad (2.1)$$

Em que:

- L - corresponde à carga total;
 - L_n - corresponde a um conjunto de configurações estandardizadas da carga para cada tipo de dia;
 - L_w - corresponde à componente sensível às alterações climáticas;
 - L_s - corresponde a eventos especiais que criam desvios nos padrões normais de consumo;
 - L_r - corresponde a um termo aleatório, que introduz ruído à equação.
- **modelos multiplicativos** - previsão de carga feita pelo produto de um número de componentes [14], como apresentado na equação 2.2:

$$L = L_n * F_w * F_s * F_r \quad (2.2)$$

Nesta equação, os fatores de correção F_w (condições meteorológicas), F_s (ocorrências especiais) e F_r (flutuações especiais) são definidos como números positivos que podem aumentar ou diminuir a carga total. Além destes fatores, podem, ainda ser tidos em consideração outros como preço de eletricidade e crescimento do consumo.

2.3 Metodologias de previsão de consumo

2.3.1 Abordagem de dia semelhante

Este método baseia-se na procura, nos dados históricos de consumo dos últimos anos, de características semelhantes ao dia a ser previsto, tais como condições meteorológicas, dia da semana e data. Uma vez encontrado o dia semelhante, a carga é considerada como uma previsão.

Ao invés de se restringir a similaridade a apenas um dia, a previsão pode resultar de uma combinação linear ou de um procedimento de regressão que pode incluir vários dias semelhantes. Os coeficientes de tendência podem ser usados para dias semelhantes aos anos anteriores.

2.3.2 Métodos de Regressão

O método de regressão é uma das técnicas estatísticas mais amplamente usada e é facilmente implementada. Permite modelação de relações entre o consumo e outros factores, como condições meteorológicas, utilização final e tipo de dias considerados.

Este método, como se pode verificar equação 2.3, assume a divisão do consumo em padrão de consumo *standard* e padrão linearmente dependente de alguns fatores que influenciam a carga [15].

$$L(t) = L(t) = L_n(t) + \sum a_i x_i(t) + e(t) \quad (2.3)$$

Em que:

- $L_n(t)$ - corresponde ao padrão *standard* do consumo no instante t ;
- $x_i(t)$ - corresponde aos fatores influenciadores independentes;
- a_i - corresponde aos coeficientes de influência definidos por $x_i(t)$;
- $e(t)$ - corresponde a uma componente de ruído (*white noise*).

A simplicidade do modelo reside no facto de serem necessários poucos parâmetros, que podem ser facilmente calculados a partir de dados históricos, aplicando a técnica da validação cruzada.

Mudança de carga sazonal, crescimento/aumento carga anual e a última mudança de carga diária também devem ser tidos em consideração.

2.3.3 Séries Temporais

Podem ser definidas como um conjunto sequencial de dados medidos durante um certo período de tempo, com um horizonte temporal definido.

Para o desenvolvimento deste modelo, é necessário, primeiramente, construir um padrão de correspondência para os dados disponíveis o mais preciso possível, passando, depois, por obter o valor previsto respeitando o modelo estabelecido.

As séries temporais assumem que os dados possuem entre si uma autocorrelação, tendência, ou variação sazonal e efetuam a deteção e exploração dessa estrutura.

As séries temporais mais usualmente utilizadas são:

- **Autoregressive Model (AR)** - modeliza o perfil de carga, assumindo uma combinação linear de cargas anteriores, como apresentado na equação 2.4 para um modelo autoregressivo de ordem m [16]:

$$L_k = \sum_{i=1}^m \alpha_{ik} * L_{k-i} + e_k \quad (2.4)$$

Em que:

- L_k - corresponde à carga prevista no instante k ;
 - e_k - corresponde a uma distribuição de carga aleatória;
 - α_i (com $i=1\dots m$) - corresponde a coeficientes desconhecidos, determinados pelo uso do algoritmos de mínimos quadrados.
- **Autoregressive Moving-Average (ARMA)** - baseado numa equação (2.5) para uma ordem (p,q) . O valor atual da série temporal $y(t)$ é representado linearmente em termos dos seus valores prévios e em termos de valores prévios de ruído (*white noise*). Os parâmetros c_i e k_i são determinados por esquemas recursivos.

$$y(t) = c_1 * 1 * y(t-1) + \dots + c_p * p * y(t-p) + a(t) - k_1 * 1 * a(t-1) - \dots - k_q * q * a(t-q) \quad (2.5)$$

- **Autoregressive Integrated Moving-Average (ARIMA)** - ao contrário do modelo ARMA, são usados processos não estacionários, representados pela introdução do operador ∇ aos modelos [17].
- **Autoregressive Integrated Moving-Average with exogenous variables (ARIMAX)** - pelo facto de o consumo depender de conjuntos de variáveis exógenas, tal como a temperatura e hora do dia, este tipo de modelo tem sido um dos mais utilizados [18].

2.3.4 Redes Neurais

De entre as ferramentas usadas para a previsão de consumos, as redes neuronais (NN) ou redes neuronais artificiais (ANN) apresentam-se no lote das mais utilizadas, suplantando muitas técnicas que foram usadas durante vários anos [19]. Isto deve-se à capacidade que apresentam para resolver uma ampla variedade de sistemas, incluindo processo lineares e não lineares. Além disso, o seu sucesso advém da sua capacidade de adaptação e da tolerância à introdução de dados considerados como ruído [20]. Os padrões escondidos nos dados são aprendidos por neurónios, fazendo generalizações destes padrões mesmo na presença de barulho ou informação em falta.

A aplicação de ANNs para previsões de carga eléctrica requiere a definição de parâmetros diferentes como o número de arquiteturas, o número e a conectividade de camadas e elementos, uso de ligações uni e bidireccionais, e o formato dos números (binário ou contínuo) a ser usado por inputs e outputs, e internamente.

De entre as várias arquiteturas existentes (exemplo Máquina de Boltzman, rede discreta de Hopfield, redes neuronais bidireccionais) a retro-propagação tornou-se a arquitetura de redes neuronais artificiais para previsões de carga eléctrica mais popular.

Este algoritmo surge do facto de os pesos atuais, atribuídos a elementos de entrada, serem determinados a partir de dados históricos de temperatura, por exemplo, para um dado output desejado, como é o caso de históricos de consumo.

2.3.5 Lógica Difusa (*Fuzzy Logic*)

Este método consiste na generalização da lógica Booleana usada para os circuitos digitais. Um dado input, com um certo alcance qualitativo associado, pela lógica booleana, toma valores discretos de “0” ou “1”, o que torna possível a dedução de outputs a partir desses inputs.

A previsão baseada na lógica difusa trabalha em duas etapas sequenciais. Numa fase inicial, é feito o treino usando dados históricos para gerar uma base de dados de padrões. Após concluída esta etapa, é conectado a um controlador que efetua a previsão das variações de carga. Caso seja encontrado um padrão com probabilidade de correspondência maior, é usado um processo de “defuzificação” que consiste na produção de outputs mais precisos, por identificação e aproximação de sistemas dinâmicos desconhecidos e, assim, obter um novo padrão de output [11][13].

2.3.6 Ferramentas de Previsão Probabilística

Tem havido um crescente interesse por este tipo de previsões, então, pelo facto de permitir o desenvolvimento de ferramentas de despacho e de agendamento avançadas, assim como a otimização de reservas flutuantes baseados na incerteza existente nas previsões [21].

2.3.6.1 Técnica KDE (*Kernel Density Estimation*)

Trata-se de uma forma não paramétrica de estimação de funções de densidade de probabilidade (fdp) para uma variável aleatória. É uma ferramenta utilizada para a suavização de dados de amostras finitas onde os dados atuais são obtidos pela observação de dados anteriores.

A estimação é caracterizada pelo *kernel* [22] [23] [24] que determina a forma da função de peso (*weighting function*), isto é, o peso a ser atribuído à vizinhança dos dados, assim como a largura de banda, que representa a distância para a qual é efetuada a procura entre dados vizinhos na função de peso. A essa largura está associada a quantidade de suavização (*smoothing*) que será necessário fazer.

Assim, numa fase inicial, para um conjunto de dados em que se desconheça a distribuição que lhes deu origem, o problema passa por:

- **Assumir uma fdp** - distribuição normal ou beta, por exemplo;
- **estimação dos parâmetros** - a estimação é feita como base no tipo de kernel e na largura de banda, de forma a otimizar as propriedades de $\hat{f}(x)$;
- **teste de precisão** - por se tratar de uma estimação, este processo obriga a um teste de precisão do método em que, caso a largura de banda não tenha sido a mais adequada, o processo seja reiniciado pela escolha de uma nova distribuição.

Obtidos os parâmetros, é construído um histograma de frequências relativas com base nos dados. A esse histograma é posteriormente adicionado uma linha de densidade estimada. O número de barras do histograma (bins) corresponde à largura de banda como se pode ver na figura 2.1 [25]. Deste modo, a função fdp faz a suavização do histograma sendo “capturadas” as principais características dos dados, que irão ser importantes para a previsão que se vai efetuar.

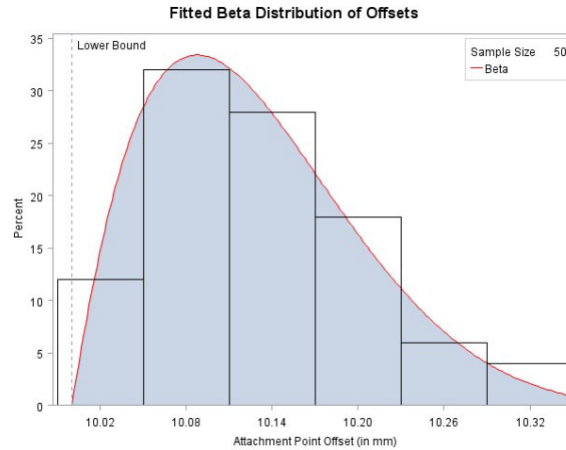


Figura 2.1: Histograma aplicado a uma função de densidade de probabilidade beta.

Nos problemas de previsão, a estimação da função de densidade condicional futura possui um papel de grande relevo pelo facto de descrever a relação existente entre as variáveis explicativas e as variáveis alvo.

Esta estimativa pode ser vista como uma generalização da regressão pelo facto de a estimativa de densidade condicional ter como objetivo a obtenção de uma função densidade de probabilidade completa $f_{y|x}(y|x)$ (2.6), enquanto a regressão tem como objetivo a estimação da média condicional $E(y|x)$ [26].

$$f_{y|x}(y|x) = \frac{f_{y,x}(y|x)}{f_x(x)} \quad (2.6)$$

Em que:

- $f_{y|x}(y,x)$ - corresponde à função de densidade conjunta;
- $f_x(x)$ - corresponde à função de densidade marginal.

A fdp obtida, nos problemas de previsão probabilística, pode ser utilizada para representar a incerteza.

Para os problemas de previsão [27][28] expressos em 2.7, a estimação da fdp condicional futura é feita para uma dada variável aleatória, num horizonte temporal $t+k$, dado um conjunto de aprendizagem com N amostras constituídas por pares de (X_n, P_N) sintetizando toda a informação histórica disponível até ao instante t .

Cada par consiste num conjunto de variáveis explicativas X_n e no valor medido correspondente da variável a prever.

Adicionalmente, é assumido que para o instante t , é conhecido um conjunto de variáveis explicativas $X_{t+k|t}$ para o horizonte temporal onde se quer prever.

$$\hat{f}_p(p_{t+k}|x_{t+k|t}) = \frac{f_{p,X}(p_{t+k}|x_{t+k|t})}{f_x(x_{t+k|t})} \quad (2.7)$$

Em que:

- p_{t+k} - corresponde à produção prevista para o horizonte temporal $t+k$;

Uma vez que as densidades conjuntas e marginais não são conhecidas, estas expressões são estimadas empregando o procedimento do KDE, para uma variável, usando as amostras históricas disponíveis, dado por (2.8):

$$\hat{f}_h(x) = \frac{1}{N \cdot h} \sum_{i=1}^N K \frac{x - X_i}{h} \quad (2.8)$$

Em que:

- h - corresponde à largura de banda;
- N - corresponde ao numero de amostras;
- K - corresponde à função de kernel;
- X_i - correspondente às amostras num dado instante i .

Caso se trate de um problema para múltiplas variáveis, o procedimento é dado por (2.9):

$$\hat{f}(x_1, \dots, x_d) = \frac{1}{N} \sum_{i=1}^N \left[\prod_{j=1}^d \frac{1}{h_j} K \left(\frac{x_j - X_{ij}}{h_j} \right) \right] \quad (2.9)$$

Em que:

- j - correspondente à variável a ser calcula num dado instante;
- d - número total de variáveis;

Existe um vasto conjunto de tipos de *kernel* entre os quais se pode seleccionar [28], sendo que esta seleção tem um impacto mínimo na qualidade da previsão quando comparado com a largura de banda escolhida. De entre os diferentes tipos de *kernel*, o mais utilizado é o gaussiano, representado na equação 2.10

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \quad (2.10)$$

Função de densidade de probabilidade da distribuição beta

Aplicado o KDE, o resultado obtido é aplicado a uma distribuição beta, definida por 2.11, por se tratar de uma boa aproximação para a modelização de variáveis, com um limite máximo (A) e com um limite mínimo (B).

$$f(x, \alpha, \beta) = \left(\frac{x^{\alpha-1} \cdot (1-x)^{\beta-1}}{B[\alpha, \beta]} \right) \quad (2.11)$$

Em que:

- x - está definido entre $[0,1]$, que garante o domínio entre $[0,1]$;
- B - corresponde à função beta;
- α - coeficiente de forma da função beta, maior que 0;
- β - coeficiente de forma da função beta, maior que 0.

Como o nome indica, os coeficiente de forma α e β definem a forma da distribuição beta e podem ser determinados por 2.12 e 2.13:

$$\alpha = \frac{(1-\mu) \cdot \mu^2}{\vartheta} - \mu \quad (2.12)$$

$$\beta = \frac{(1-\mu)}{\mu} \cdot \alpha \quad (2.13)$$

Em que:

- μ - corresponde ao valor da média original;
- ϑ - corresponde ao valor da variância original;

Na figura 2.2 é apresentada uma fdp da distribuição beta, para diferentes valores dos coeficientes de forma α e β [29].

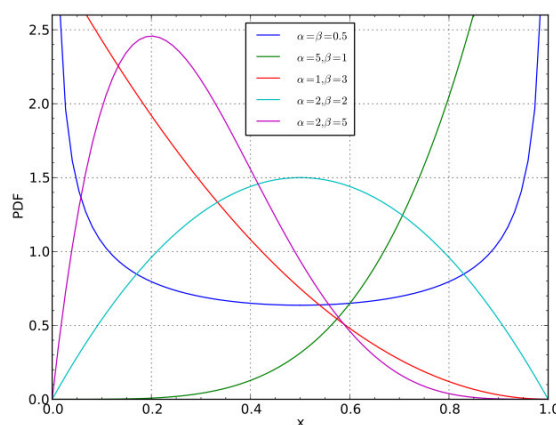


Figura 2.2: Função densidade de probabilidade da distribuição beta para diferentes valores dos coeficientes de forma α e β .

Funcionamento do KDE

O KDE faz uma seleção dos pontos vizinhos e cria uma gaussiana artificial, dando mais peso aos vizinhos mais próximos. O intervalo das variáveis não pode ser muito pequeno, pois isso pode originar um sobreajuste do modelo aos dados de treino (*overfitting*) havendo a necessidade de se determinar a largura ótima para cada variável. No entanto, esta largura apenas permite um ajustamento para o conjunto de dados atuais, pois será necessário voltar a determinar para um novo conjunto de variáveis de forma a minimizar o erro associado a esses novos dados.

A execução do programa é feita a partir dos valores de desvio padrão de cada uma das variáveis escolhidas e de um parâmetro associado à incerteza, ambos introduzidos pelo utilizador. É feito executado um processo iterativo de forma a diminuir o erro do modelo e da incerteza, por alteração dos parâmetros das diferentes variáveis.

Da execução do programa obtém-se os seguintes valores, que serão usados para estudos da precisão do modelo e da incerteza:

- “C” - consumo previsto;
- α - fator de forma da distribuição beta;
- β - fator de forma da distribuição beta;
- “A” - valor mínimo de um intervalo de confiança para a produção prevista;
- “B” - valor máximo de um intervalo de confiança para a produção prevista.

2.4 Modelos numéricos de previsão meteorológica (NWP)

Tal como constatado anteriormente, a variável temperatura contém informação que explica os padrões de consumo e, por isso, é importante a sua inclusão nos modelos de previsão. Este

interesse leva à necessidade de a previsão desta variável ser feita com o menor erro possível, de modo que a previsão de consumo se aproxime do consumo real. Esta previsão das variáveis meteorológicas pode ser feita através de métodos numéricos, como o modelo numérico de previsão meteorológica (NWP), com diferentes horizontes temporais e espaciais.

Esta metodologia é uma das mais utilizadas para previsões de curto prazo, usando para isso programas sofisticados que contêm modelos físicos e matemáticos que descrevem o comportamento da atmosfera e do oceano e que permitem, com base na informação do estado atual da meteorologia, obter uma previsão para o estado meteorológico no futuro.

Os modelos são classificados consoante a sua escala espacial e temporal com que as previsões são efetuadas, podendo ser divididos em [30] :

- **Modelos Globais** - grande área espacial, mas validade temporal reduzida;
- **Modelos Regionais (ou em mesoescala)** - curta área espacial, mas validade temporal alargada.

A construção dos modelos globais ou regionais é precedida de uma análise e observação da atmosfera, da superfície terrestre e dos oceanos o mais precisa possível, pois a sua qualidade está diretamente relacionada com a qualidade das previsões. Assim, quanto mais precisa for a análise e observação, mais preciso serão os dados das variáveis meteorológicas e, por conseguinte, mais preciso será o modelo de previsão.

2.4.1 Modelos globais

Este tipo de modelos estudam uma grande porção horizontal da atmosfera terrestre (podendo atingir os 100km) e tem um horizonte temporal até 15dias, havendo um decréscimo da sua performance caso haja um alargamento do horizonte temporal [31].

De entre os parâmetros meteorológicos recolhidos, destacam-se a temperatura, a precipitação e a cobertura das nuvens.

Dos modelos NWP globais existentes destacam-se:

- **Global Forecasting Model (GFS)** - criado por National Centers for Environmental Prediction (NCEP);
- **Global Environmental Multiscale (GEM)** - criado por Environment Canada;
- **Integrated Forecast System (IFS)** - European Centre for Medium-Range Weather Forecasts (ECMWF);
- **Navy Global Environmental Model (NAVGEN)** - Fleet Numerical Meteorology and Oceanography Center;

2.4.2 Modelos regionais

Os modelos regionais são usados para uma área mais limitada (entre 2 a 50km), produzindo, por isso, resultados mais detalhados e permitindo a obtenção de dados como a radiação que atinge a superfície terrestre [31].

De entre os modelos NWP globais existentes destacam-se:

- **Weather Research and Forecasting model (WRF)** - Cooperação entre National Centers for Environmental Prediction (NCEP), National Center for Atmospheric Research (NCAR) e a comunidade de pesquisa meteorológica;
- **North American Mesoscale model (NAM)** - criado por NCEP;
- **Fifth Generation Penn State/NCAR Mesoscale Model (MM5)** - criado por Penn State University e NCAR;
- **High Resolution Limited Area Model (HIRLAM)** - Cooperação entre diversos institutos meteorológicos Europeus;
- **Aire Limitée Adaptation dynamique Développement InterNational (ALADIN)** - Cooperação entre Météo-France e National Meteorological Services of Central and Eastern Europe

2.5 Avaliação da precisão dos modelos de previsão

Para que os modelos sejam eficazes e precisos, há a necessidade de se proceder à escolha correta das variáveis a serem utilizadas, pelo que terá que se perceber, numa fase inicial, se existe ou não correlação entre os dados.

Criados os modelos de previsão, a avaliação da precisão é feita com base na proximidade entre os valores resultantes da previsão e os valores reais de consumo, calculando-se, para cada modelo, o valor do erro entre esses valores.

2.5.1 Coeficiente de determinação (R^2) e coeficiente de correlação de Pearson (R)

O coeficiente de determinação (R^2) mede a percentagem da variância de uma das variáveis que pode ser explicada pela variância da outra. Varia entre “0” e “1”, sendo que “0” equivale à inexistência de relação e “1” corresponde à existência de uma relação linear perfeita entre as variáveis.

Este coeficiente é calculado por (2.14) [32] :

$$R^2 = 1 - \frac{\frac{1}{m} \sum_{t=1}^m (P_t - \hat{P}_t)^2}{\sum_{t=1}^m (P_t - P)^2} \quad (2.14)$$

Em que:

$$P = \frac{1}{m} \sum_{t=1}^m P_t \quad (2.15)$$

Pode também ser determinado pelo quadrado do coeficiente de correlação de Pearson (R) que revela a influência que o valor de uma variável aleatória tem sobre os seus vizinhos, sendo, por isso, uma ferramenta importante na determinação de padrões.

O coeficiente varia entre “-1” e “1”, sendo que:

- **R=1** - indica a presença de uma relação perfeita positiva entre as duas variáveis, ou seja, as variáveis são diretamente proporcionais;
- **R=0** - indica que não existe uma relação linear entre as variáveis, embora possam ter outro tipo de relação;
- **R=-1** - indica a presença de uma relação perfeita negativa entre as duas variáveis, ou seja, as variáveis são inversamente proporcionais.

A correlação é dada por (2.16) :

$$R(k) = \frac{E[(Y_t - \mu)(Y_{t+k} - \mu)]}{\sigma^2} \quad (2.16)$$

Em que:

- Y_t - variável aleatória, discreta e estacionária;
- k - distância com que se pretende medir a variável Y ;
- μ - valor da média;
- E - valor esperado;
- σ^2 - variância ou quadrado do desvio padrão.

Nota:

- Se $k=0$, o valor do coeficiente de correlação é “1”, pois significa que a variável está correlacionada com ela própria;

2.5.2 Índices de erro

Como foi descrito, a avaliação dos modelos é feita pelo cálculo do erro. De entre os índices de erro existentes, estes são os mais utilizados para avaliar a precisão dos modelos de previsão:

- Root Mean Square Error (RMSE) [32]

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (C_t - \hat{C}_t)^2} \quad (2.17)$$

Possui uma versão normalizada intitulada de erro médio quadrático normalizado (*normalized Root Mean Square Error*) [32], calculado por:

$$nRMSE = \sqrt{\frac{\frac{1}{n} \sum_{t=1}^n (C_t - \hat{C}_t)^2}{\sum_{t=1}^n C_t^2}} \quad (2.18)$$

- Mean Absolute Error (MAE) [32]

$$MAE = \frac{1}{n} \sum_{t=1}^n |C_t - \hat{C}_t| \quad (2.19)$$

- Mean Absolute Percentage Error (MAPE) [24]

$$MAPE = \frac{\frac{1}{n} \sum_{t=1}^n |C_t - \hat{C}_t|}{\frac{1}{n} \sum_{t=1}^n |C_t|} \quad (2.20)$$

- Mean Square Error (MSE) [32]

$$MSE = \frac{1}{n} \sum_{t=1}^n (\hat{C}_t - C_i)^2 \quad (2.21)$$

Em que:

- n - tamanho da amostra a ser avaliada;
- C_t - consumo real no instante t;
- $\hat{C}_t$ - consumo previsto no instante t;

Capítulo 3

Metodologia para a construção de modelos de previsão

No âmbito desta dissertação, a previsão efetuada corresponde a uma previsão de curto prazo, focalizada na previsão para o dia seguinte, podendo, no entanto, ser feita para horizontes temporais mais alargados como semanas, meses ou anos.

A importância da previsão para o dia seguinte prende-se não só com o facto deste tipo de previsão ser um dos requisitos para o controlo e planeamento dos sistemas elétricos, mas também pela existência de vários mercados energéticos, como é o caso do *MIBEL*. Neste modelo de mercado, o uso destas ferramentas de previsão revela-se fundamental para o balanceamento do mercado que opera com base nas propostas de compra e venda de energia elétrica efetuadas com um dia de antecedência.

3.1 Caso de estudo e dados utilizados

Como mencionado anteriormente, esta dissertação tem como foco o consumo energético na ilha de Creta. Dados de 2010 revelam um consumo anual de aproximadamente 3 TWh, com uma ponta de consumo de 600 MW (3.1).

As expectativas apontam para um aumento do consumo anual, em 2020, para os 4 TWh, com uma ponta de consumo de 900 MW (3.2) [33].

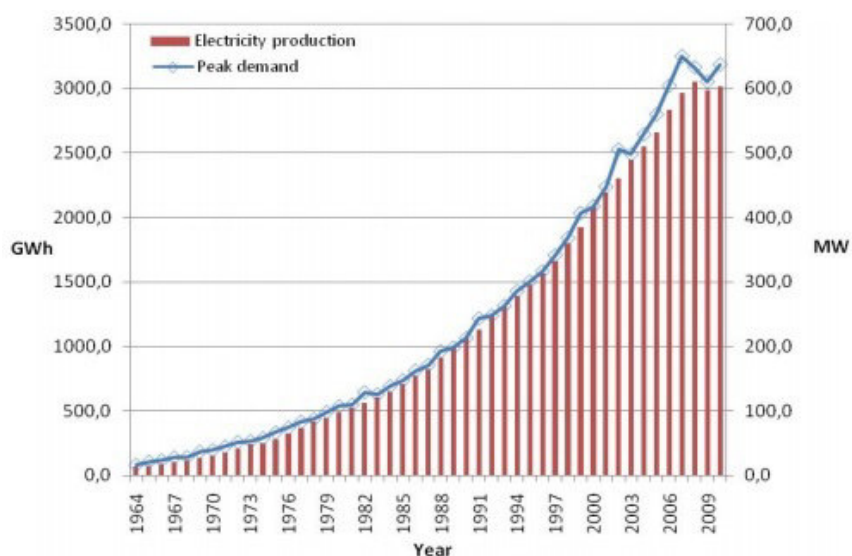


Figura 3.1: Consumo Anual e Ponta de Consumo em Creta até 2010. [33]

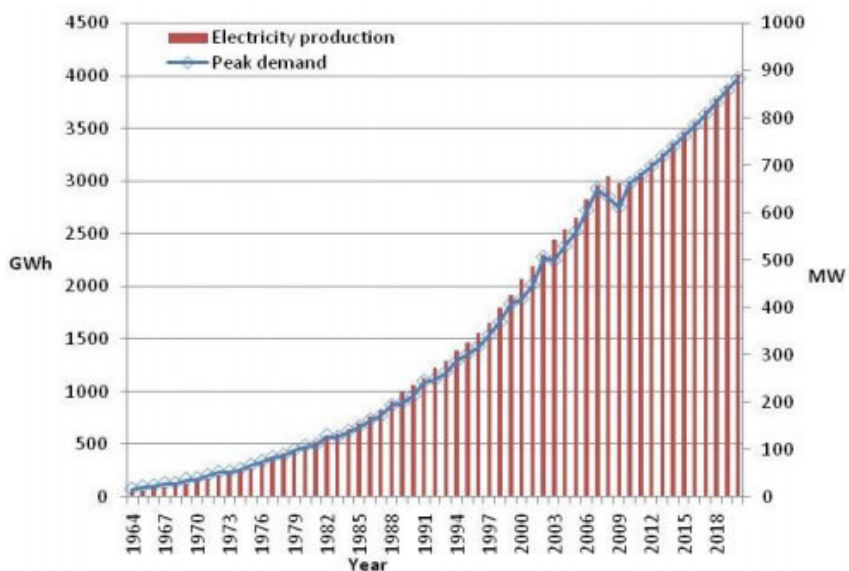


Figura 3.2: Consumo Anual e Ponta de Consumo em Creta até 2020. [33]

As necessidades energéticas são cobertas por centrais a vapor, a diesel, a gás, centrais de ciclo combinado, além de um conjunto de Recursos de Energia Renovável (RES) constituído por turbinas eólicas, centrais fotovoltaicas e centrais mini-hídricas.

Até 2010, essas centrais foram instaladas com a capacidade apresentada na tabela 3.1 :

Tabela 3.1: Capacidade Instalada nas centrais de Creta. [33]

Technology	Steam generators	Diesel generators	Gas turbines	Combined Cycle	Wind turbines	Hydro turbines	PV
Fuel type	Fuel oil	Fuel oil	Diesel	Diesel	-	-	-
Power[MW]	211.2	167.2	478.8	133.4	166.7	0.6	30.7

Produção energética em Creta, consoante o tipo de central (3.3) e consoante o tipo de combustível usado (3.4) :

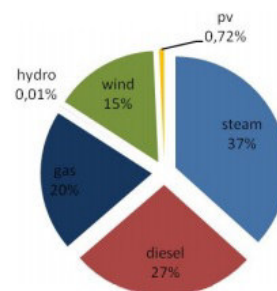


Figura 3.3: Produção Anual em Creta até 2010, por tipo de central. [33]

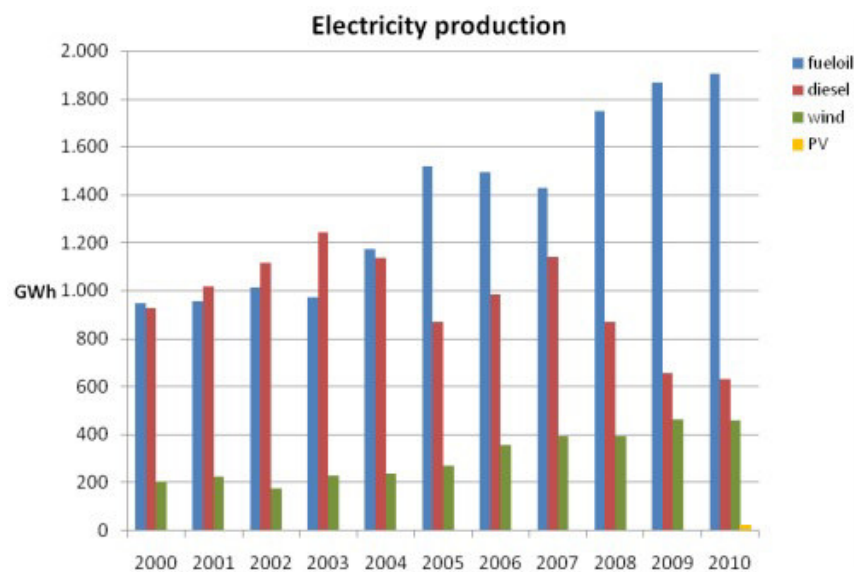


Figura 3.4: Produção Anual em Creta até 2010, por tipo de combustível usado. [33]

O sistema elétrico de energia na ilha de Creta engloba 18 subestações repartidas pela ilha (3.2)

Tabela 3.2: Subestações existentes na ilha de Creta.

Subestações	Potência Vazio (MW)	Potência Ponta (MW)	Potência Média (MW)	Capacidade PV instalada (MW)
Agyia	8.6	14.6	11.3	3.4
Agios Nikolaos	13.2	23.6	17.6	5.2
Chania	28.9	54.2	41.4	1.6
Ierapetra	11	19.1	14.2	7.1
Hrackleio1	15.5	34.3	25.6	1.1
Hrackleio2	32.2	58.6	46.2	7.7
Hrackleio3	16.7	28.9	23	3.4
Kasteli	8.6	15.3	11.3	3.4
Ljanoperamata	2.9	5.3	4	0.4
Moires	11.1	21.2	14.1	15.3
Praitoria	4.6	12.9	7.8	10.5
Rethymno	26.4	49.8	36.1	16.6
Sitia	5.9	12.2	7.7	9.3
Stalida	16.7	29.4	22.6	2.3
Vryses	5.9	10.3	7.6	3.4
Agregado	233.9	431.8	338.8	90.5

Nota: O *Agregado* corresponde ao conjunto de todas as subestações existentes na ilha de Creta.

O estudo nos modelos de previsão será feito, apenas, para algumas destas subestações, como se discriminará posteriormente.

Para a criação dos diferentes modelos de previsão foi necessário, dentro de uma vasta gama de variáveis, perceber quais as que teriam maior influência no processo de previsão do consumo e, posteriormente, efetuar um aperfeiçoamento dos parâmetros para os modelos pretendidos.

Foi um processo gradual, abordado sobre várias perspetivas, nomeadamente a alteração do conjunto das variáveis. Serão, assim, descritos e demonstrados os processos usados para a construção de 5 modelos:

- **Modelo 1 - Modelo Base Simples** - modelo criado com base em duas variáveis (Hora e Dia da Semana);
- **Modelo 2 - Modelo Base** - modelo criado com base em três variáveis (Hora, Dia da Semana e Mês);
- **Modelo 3 - Modelo Base com explicador Temperatura** - modelo criado com base em quatro variáveis (Hora, Dia da Semana, Mês e Temperatura);
- **Modelo 4 - Modelo Base com explicador Temperatura e Irradiância** - modelo criado com base em cinco variáveis (Hora, Dia da Semana, Mês, Temperatura e Irradiância);

- **Modelo 5 - Modelo Base com explicador Temperatura e Consumo do Dia Anterior** - modelo criado com base em cinco variáveis (Hora, Dia da Semana, Mês, Temperatura e Consumo do Dia Anterior);

Os dados a utilizar, quer de consumo quer de outras variáveis meteorológicas, pertencem aos anos 2012 e 2013. No entanto, verificou-se que, contrariamente ao que ocorre nos dados das subestações de 2012, os dados de consumo das subestações de 2013 (que é dado pela soma de toda a produção enegética associada a essa subestação, nesse ano) não comportam a produção fotovoltaica, pela dificuldade inerente à sua divisão pelas respetivas subestações. Embora isto aconteça para as subestações, o mesmo não acontece para o *Agregado* da ilha fornecido e que já possui a produção fotovoltaica.

Os casos de estudo efetuados na dissertação foram divididos da seguinte forma:

- Caso de Estudo 1:

Datas usadas:

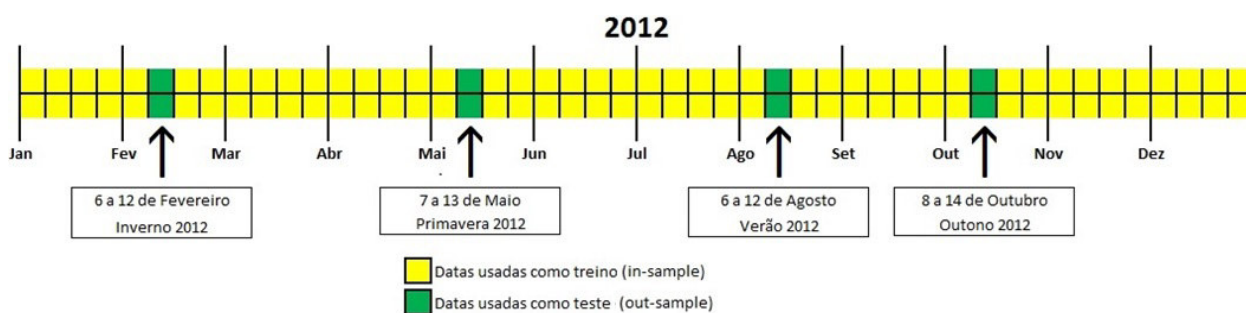


Figura 3.5: Divisão dos dados (in-sample e out-sample) para o caso de estudo 1.

Os modelos vão ser treinados usando os datas marcadas como “out-sample” (3.5) e, posteriormente, serão testados com as datas marcadas como “in-sample” (3.5) para serem comparados com outros modelos com as mesmas características (datas e subestações).

Modelos e Subestações utilizadas (As subestações têm produção fotovoltaica):

Tabela 3.3: Modelos e subestações usados para o caso de estudo 1.

Modelo	Subestações
Modelo 1 - Modelo Base Simples	Agregado
Modelo 2 - Modelo Base	Agregado
Modelo 3 - Modelo Base com explicador Temperatura	Agregado
	Hracleio 2
	Hracleio 3
	Rethymno

- Caso de Estudo 2:

Datas usadas:

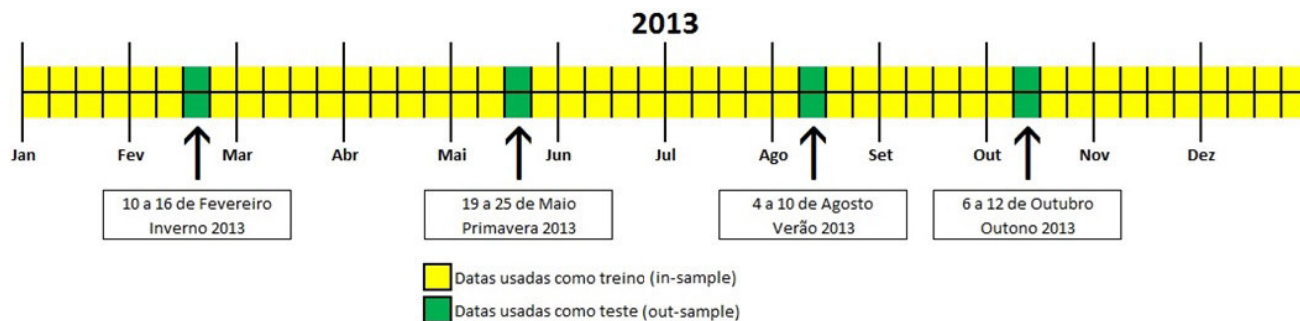


Figura 3.6: Divisão dos dados (in-sample e out-sample) para o caso de estudo 2.

Modelos e Subestações utilizadas (As subestações não têm incluída a produção fotovoltaica):

Tabela 3.4: Modelos e subestações usados para o caso de estudo 2.

Modelo	Subestações
Modelo 3 - Modelo Base com explicador Temperatura	Agregado
	Ag. Nikolaos
	Chania
	Hrakteio 1
	Hrakteio 2
	Hrakteio3
	Rethymno
Modelo 4 - Modelo Base com explicador Temperatura e Irradiância	Sitia
	Agregado
	Ag. Nikolaos
	Chania
	Hrakteio 1
	Hrakteio 2
	Hrakteio3
Modelo 5 - Modelo Base com explicador Temperatura e Consumo do Dia Anterior	Rethymno
	Sitia
	Agregado
	Ag. Nikolaos
	Chania
	Hrakteio 1
	Hrakteio 2
	Hrakteio3
	Rethymno
	Sitia

3.2 Obtenção e Tratamento de dados

Para todos os modelos referidos anteriormente foi necessário, numa fase inicial, efetuar o processamento dos dados para que pudessem ser avaliados de uma forma similar, assim como extrair características especiais que poderiam revelar-se importantes para a caracterização dos modelos. Assim, é fundamental, por exemplo, distinguir se os dados de consumos se referem a um dia de semana, fim de semana ou até mesmo um feriado.

Foram então seguidas as etapas representadas na figura 3.7.:

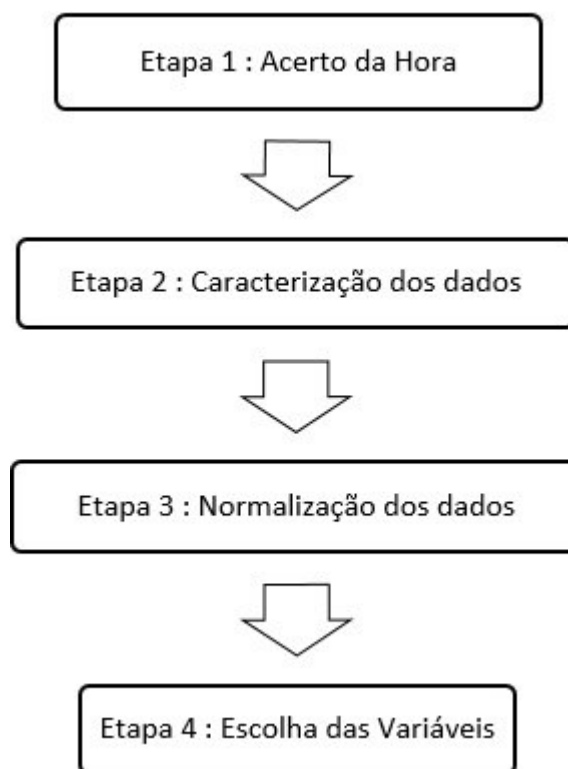


Figura 3.7: Etapas do processo de tratamento de dados.

3.2.1 Etapa 1 - Acerto da Hora

Nos dados de consumo verificou-se que a mudança da horário de verão (UTC+2) para horário de inverno (UTC+3) não se encontrava considerada sendo que, por esse motivo, foi necessário proceder a essas alterações manualmente, consoante as datas e horas expressas na tabela 3.5 [34][35][36].

Tabela 3.5: Mudanças de hora em Creta, nos anos de 2012, 2013 e 2014.

Antes		Depois	
Data	Hora	Data	Hora
25/03/2012	3h00	25/03/2012	4h00
28/10/2012	4h00	28/10/2012	3h00
31/03/2013	3h00	31/03/2013	4h00
27/10/2013	4h00	27/10/2013	3h00
30/03/2014	3h00	30/03/2014	4h00
26/10/2014	4h00	26/10/2014	3h00

Nota:

- **Horário de Verão** – os dados de consumos referentes às 3h00 do dia de mudança de horário foram eliminados, passando imediatamente das 2h00 para as 4h00;
- **Horário de Inverno** – os dados de consumos referentes às 3h00 do dia de mudança de horário foram repetidos, passando a haver dois registos para as 3h00 com o mesmo consumo.

Como a origem dos dados de consumos e dos dados meteorológicos não foi a mesma (2.4), houve a necessidade de fazer a interligação entre ambos, de modo a que, para um dado valor de consumo, referente a uma certa hora no fuso horário de Creta UTC+2 no caso de horário de verão ou UTC+3 no caso de horário de inverno), fosse associado o valor correto dos dados meteorológicos que se encontram no fuso horário UTC.

Sendo assim, foi necessário sincronizar os dados de consumo com os dados meteorológicos, fazendo associar a um dado consumo o dado meteorológico “atrasado” duas ou três horas, consoante se trate de horário de verão ou horário de inverno, como se pode constatar na tabela 3.6.

Tabela 3.6: Comparação entre Hora Local e Hora UTC nos diferentes fusos horários.

Data	Hora Local	Hora UTC	Fuso Horário
25/03/2012	5h00	3h00	Inverno (UTC+2)
25/03/2012	5h00	2h00	Verão (UTC+3)
28/10/2012	5h00	2h00	Verão (UTC+3)
28/10/2012	5h00	3h00	Inverno (UTC+2)

3.2.2 Etapa 2 - Caracterização dos dados

Um dos principais fatores que explicam a variação dos padrões de consumo é o Dia da Semana, pois há uma grande influência caso se trate de um dia útil, em que tudo estará a contribuir em pleno para o consumo de energia elétrica, nomeadamente os transportes (que representam a maior fonte de consumo energético em Creta (3.8) [33] ou se trate de um fim de semana ou de um feriado,

visto que, por exemplo, haverá uma menor afluência à utilização dos transportes, o que provocará um decréscimo no consumo.

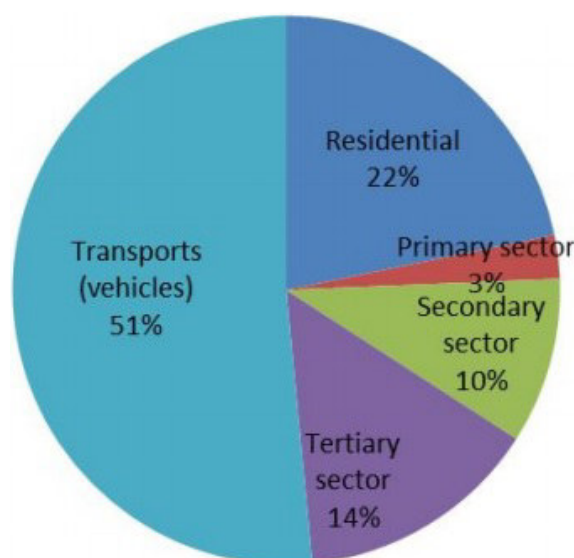


Figura 3.8: Consumos por setor, em Creta, no ano de 2005. [33]

Devido à pertinência desta nuance, a distinção dos diferentes dias da semana apresenta-se como uma variável bastante relevante para efeitos de diferenciação, o que melhorará substancialmente os resultados das previsões.

Com o auxílio do Microsoft Excel, procedeu-se à associação entre a data e o dia da semana correspondente, sendo que a cada dia da semana foi associado um número de forma a auxiliar em cálculos futuros (3.7).

Tabela 3.7: Associação entre Dia da Semana e número auxiliar.

Dia da Semana	Número
domingo	1
segunda-feira	2
terça-feira	3
quarta-feira	4
quinta-feira	5
sexta-feira	6
sábado	7
feriado	8

De notar a inserção de “feriados” na tabela semanal, pois trata-se de um dia especial cujo padrão de consumo se aproxima de “domingo”, independentemente do dia da semana em que esse dia de descanso ocorra.

Salvo raras exceções de feriados locais, a tabela 3.8 apresenta os feriados considerados na ilha de Creta para o período em estudo [37][38][39].

Tabela 3.8: Feriados nos anos de 2012, 2013 e 2014.

2012	2013	2014
01/01/2012	01/01/2013	01/01/2014
06/01/2012	06/01/2013	06/01/2014
30/01/2012	30/01/2013	30/01/2014
27/02/2012	18/03/2013	03/03/2014
25/03/2012	25/03/2013	25/03/2014
13/04/2012	01/05/2013	18/04/2014
15/04/2012	03/05/2013	20/04/2014
16/04/2012	05/05/2013	21/04/2014
01/05/2012	06/05/2013	
04/06/2012	24/06/2013	
24/07/2012	24/07/2013	
15/08/2012	15/08/2013	
28/10/2012	28/10/2013	
17/11/2012	17/11/2013	
25/12/2012	25/12/2013	
26/12/2012	26/12/2013	

Analisando a figura 3.9, constata-se que a utilização de dispositivos, aquecedores ou ar-condicionados, para maior conforto térmico, representa uma grande fatia (52%) do consumo residencial, o que demonstra a grande importância das variações da temperatura para o consumo, pois estas provocam uma alteração comportamental na população que tem necessidade de regular a temperatura, utilizando esses aparelhos.

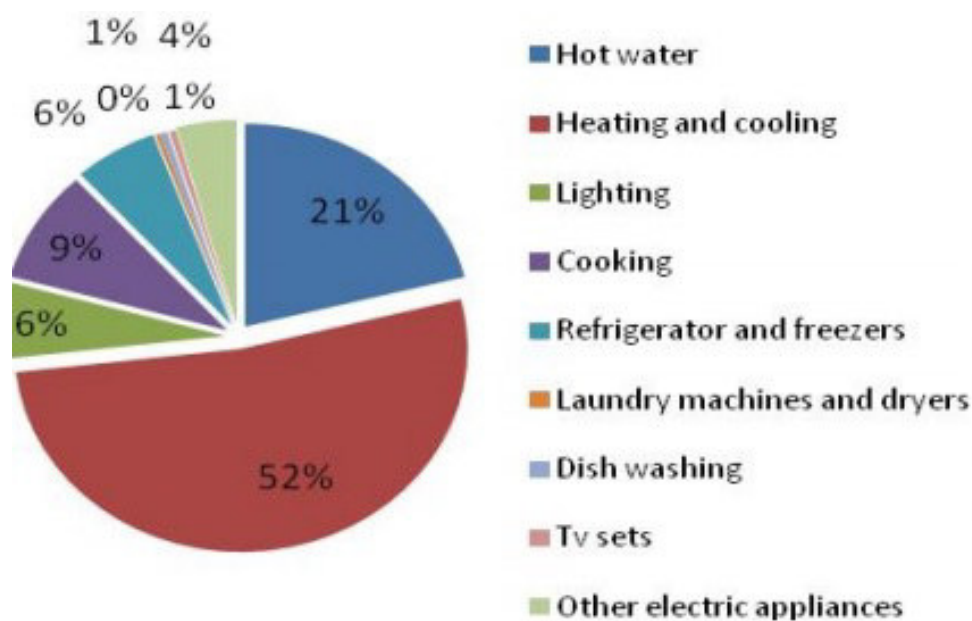


Figura 3.9: Consumo residencial por aplicação, em Creta, no ano de 2005. [33]

Excluindo casos esporádicos de aumentos de temperatura anormais no inverno ou decréscimos acentuados no verão em que o comportamento da população pode contrariar o padrão para essa época do ano, verifica-se que as estações do ano terão, também, influência nos padrões de carga.

Assim, para uma melhor caracterização dos dados e tendo em conta que as alterações comportamentais estão relacionadas com as estações do ano e, por conseguinte, com o mês do ano, a criação da variável diferenciadora Mês do Ano poderá vir a trazer melhorias na previsão do consumo, algo que será estudado, posteriormente, de forma mais detalhada.

3.2.3 Etapa 3 - Normalização dos dados

Como mencionado anteriormente, os dados de consumos obtidos são referentes a diferentes anos civis, podendo originar casos de variação considerável dos níveis de carga, devido a possíveis modificações no sistema como, por exemplo, aumento ou decréscimo da potência de uma ou várias subestações.

Para evitar discrepâncias expressivas entre os valores, houve a necessidade de se proceder à sua normalização de forma a conseguir obter padrões de similaridade entre os dados e para facilitar a determinação e posterior remoção de casos em que o consumo se desviasse significativamente dos valores padrão, nomeadamente em situações em que os dados não tenham sido corretamente registados ou mesmo em casos de falhas na rede.

A normalização foi feita através do cálculo dos padrões Diário, Semanal e Anual, sendo que este último não influenciou a identificação de novas anomalias, pois os restantes padrões revelaram-se suficientemente eficazes na sua remoção.

3.2.3.1 Padrão Diário

A normalização dos dados pelo padrão Diário pressupõe as etapas descritas na figura 3.10:

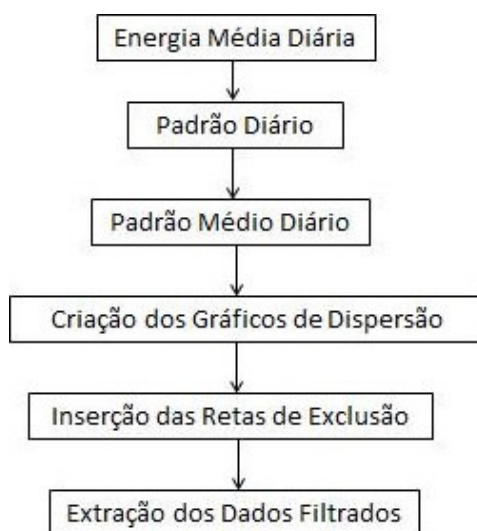


Figura 3.10: Etapas do processo de tratamento de dados.

$$EMD(d) = \frac{\sum_{i=0}^{dda} Consumo(dda)}{\sum_{i=0}^{dda} Consumo} [MW] \quad (3.1)$$

Nota: A Energia Média Diária (EMD) (3.1) é o valor médio, para cada dia do ano (dda), para a gama de valores tidos em conta, ou seja, caso se considere o período de um ano, este cálculo irá resultar em 365 valores, sendo que a cada um corresponderá a média de todos os valores de consumo que estiveram associados ao dia em questão, por exemplo, a energia média de todos os consumos do primeiro dia desse ano.

$$PD(d, h) = \frac{Consumo(d, h)}{EMD(d)} [Adimensional] \quad (3.2)$$

O Padrão Diário (PD) (3.2) é dado pelo quociente entre o valor de consumo real, de um dado dia (d), a uma dada hora (h) e pelo valor correspondente a esse dia do ano de EMD.

$$PMD(h) = \frac{\sum_{i=0}^h PD(h)}{\sum_{i=0}^t PD} [Adimensional] \quad (3.3)$$

Nota: O cálculo do Padrão Médio Diário (PMD) (3.3) associa a cada valor de consumo, de uma dada hora (h), a média do Padrão Diário para essa hora, ou seja, agrega, por exemplo, a todos os valores de consumo às 0 horas a média de todos os PD correspondentes a essa hora.

Para uma melhor compreensão do processo, apresenta-se, na tabela 3.9, um exemplo de dados obtidos para dois dias consecutivos dos cálculos descritos anteriormente.

Tabela 3.9: Tabela demonstrativa do cálculo de EMD, PD e PMD.

Data	Hora	Consumo (MW)	DDA	EMD (MW)	PD (Adim)	PMD (Adim)
01/01/2013	0	289.49	0	286.83	1.01	0.86
01/01/2013	1	259.21	0	286.83	0.9	0.77
01/01/2013	2	245.17	0	286.83	0.85	0.72
01/01/2013	3	228.91	0	286.83	0.8	0.69
01/01/2013	4	216.07	0	286.83	0.75	0.68
01/01/2013	5	209.04	0	286.83	0.73	0.68
⋮	⋮	⋮	⋮	⋮	⋮	⋮
02/01/2013	0	257.39	1	300.73	0.86	0.86
02/01/2013	1	231.63	1	300.73	0.77	0.77
02/01/2013	2	214.74	1	300.73	0.71	0.72
02/01/2013	3	202.79	1	300.73	0.67	0.69
02/01/2013	4	197.69	1	300.73	0.66	0.68
02/01/2013	5	198.25	1	300.73	0.66	0.68
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Calculados os valores de PD e PMD, criou-se um gráfico de dispersão (3.11) de modo a observar os padrões criados e perceber quais os valores que estariam fora desse padrão.



Figura 3.11: Gráfico de Dispersão que relaciona os Padrões Diário (Adimensional) e Médio Diário (Adimensional)

Como se pode observar na figura 3.11, alguns valores encontram-se bastante desviados do padrão, pelo que, de modo a evitar que estas anomalias afetem as previsões, se procedeu à sua omissão. Esta exclusão é feita com base em duas retas que filtram os dados de Padrão Médio para uma gama de valores entre um valor máximo (estabelecido para cada ponto pela reta superior) e mínimo (estabelecido para cada ponto pela reta inferior).

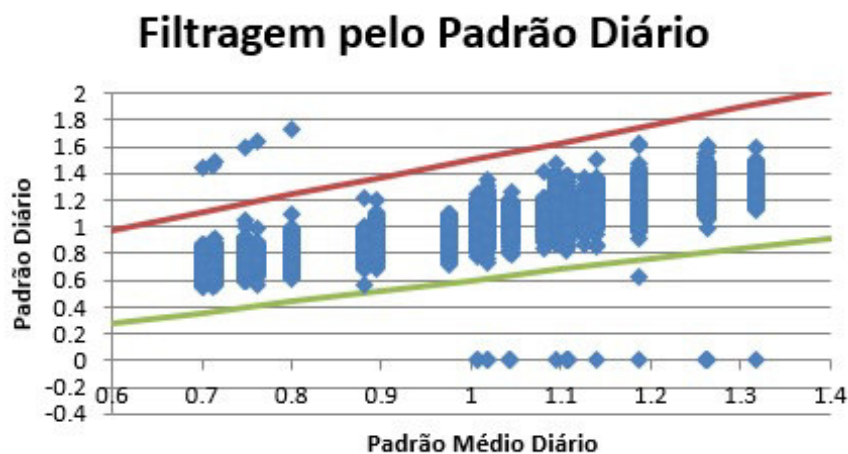


Figura 3.12: Gráfico de Dispersão que relaciona os Padrões Diário (Adimensional) e Médio Diário (Adimensional), intersetado por retas de exclusão

Esta filtragem permite, então, separar os dados a serem utilizados para a previsão e os dados a serem excluídos (3.13).

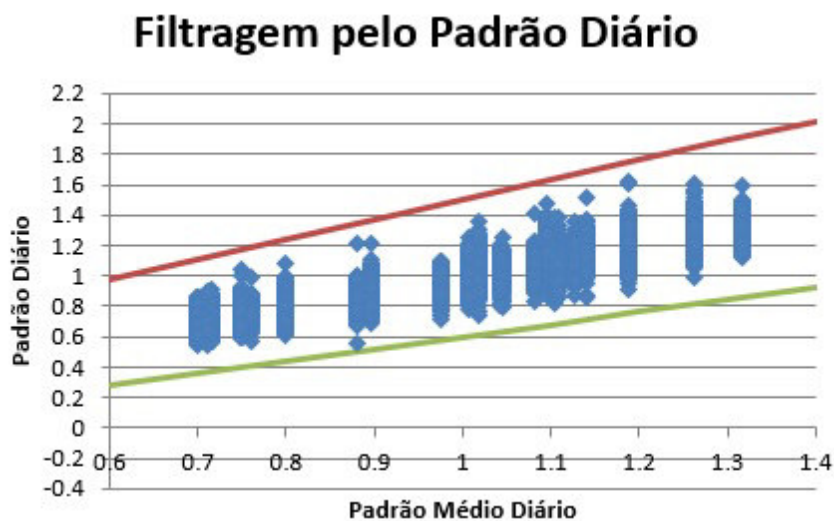


Figura 3.13: Gráfico resultante da filtragem pelas retas de exclusão.

3.2.3.2 Padrão Semanal

Processo semelhante ao descrito em 3.2.3.1 com a diferença de que os valores calculados de Energia Média Semanal (EMS) (3.4), Padrão Semanal (PS) (3.5) e Padrão Médio Semanal (PMS) (3.6) passam a fazer referência à semana do ano (sda).

$$EMS(d) = \frac{\sum_{i=0}^{sda} Consumo(sda)}{\sum_{i=0}^{sda} Consumo} \quad [MW] \quad (3.4)$$

$$PS(d, h) = \frac{Consumo(d, h)}{EMS(d)} \quad [Adimensional] \quad (3.5)$$

$$PMS(h) = \frac{\sum_{i=0}^h PS(h)}{\sum_{i=0}^t PS} \quad [Adimensional] \quad (3.6)$$

Exemplo dos valores obtidos para os mesmos dois dias considerados anteriormente (3.10).

Tabela 3.10: Tabela demonstrativa do cálculo de EMS, PS e PMS.

Data	Hora	Consumo (MW)	SDA	EMS (MW)	PS (Adim)	PMS (Adim)
01/01/2013	0	289.49	0	306.51	0.94	0.97
01/01/2013	1	259.21	0	306.51	0.94	0.97
01/01/2013	2	245.17	0	306.51	0.94	0.97
01/01/2013	3	228.91	0	306.51	0.94	0.97
01/01/2013	4	216.07	0	306.51	0.94	0.97
01/01/2013	5	209.04	0	306.51	0.94	0.97
⋮	⋮	⋮	⋮	⋮	⋮	⋮
02/01/2013	0	257.39	1	306.51	0.98	1.02
02/01/2013	1	231.63	1	306.51	0.98	1.02
02/01/2013	2	214.74	1	306.51	0.98	1.02
02/01/2013	3	202.79	1	306.51	0.98	1.02
02/01/2013	4	197.69	1	306.51	0.98	1.02
02/01/2013	5	198.25	1	306.51	0.98	1.02
⋮	⋮	⋮	⋮	⋮	⋮	⋮

O processo de filtragem é feito como descrito anteriormente, embora para o gráfico de dispersão sejam agora usados os valores de PS e PMS.



Figura 3.14: Gráfico de Dispersão que relaciona os Padrões Semanal (Adimensional) e Médio Semanal (Adimensional)

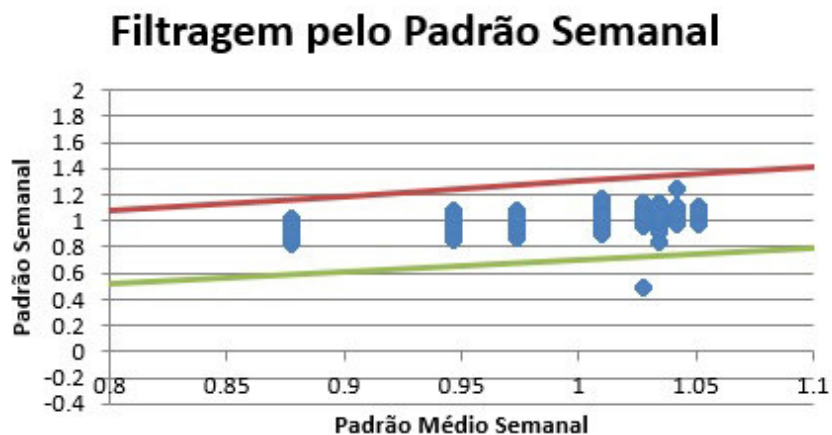


Figura 3.15: Gráfico de Dispersão que relaciona os Padrões Semanal (Adimensional) e Médio Semanal (Adimensional), intersestado por retas de exclusão

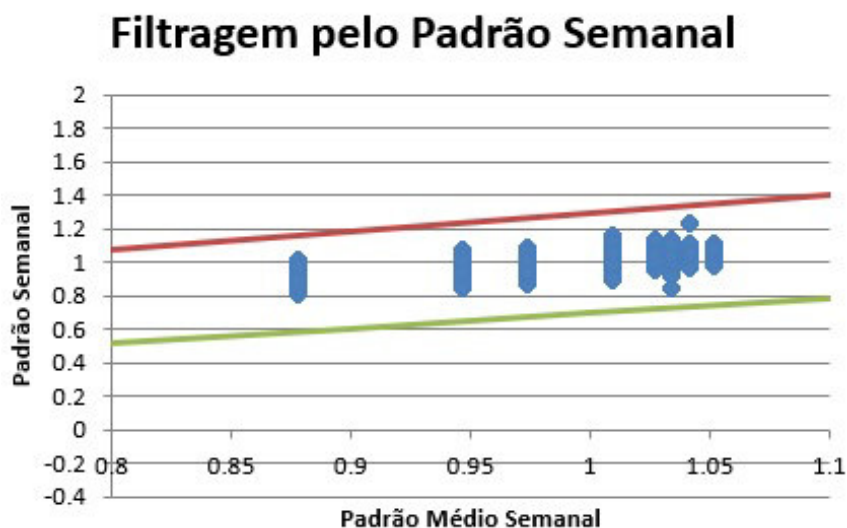


Figura 3.16: Gráfico resultante da filtragem pelas retas de exclusão.

3.2.3.3 Padrão Anual

O processo é novamente similar para este ponto em que os valores da Energia Média Anual (EMA) (3.7), do Padrão Anual (PA) (3.8) e do Padrão Médio Anual (PMA) (3.9), calculados para o mês do ano (mda), são dados por:

$$EMA(d) = \frac{\sum_{i=0}^{mda} \text{Consumo}(mda)}{\sum_{i=0}^{mda} \text{Consumo}} \quad [MW] \quad (3.7)$$

$$PA(d, h) = \frac{\text{Consumo}(d, h)}{EMA(d)} \quad [Adimensional] \quad (3.8)$$

$$PMA(h) = \frac{\sum_{i=0}^h PA(i)}{\sum_{i=0}^t PA} \quad [Adimensional] \quad (3.9)$$

Exemplificação dos valores obtidos, para os mesmos dois dias:

Tabela 3.11: Tabela demonstrativa do cálculo de EMA, PA e PMA.

Data	Hora	Consumo (MW)	MDA	EMA (MW)	PA (Adim)	PMA (Admin)
01/01/2013	0	289.49	0	313.69	0.96	0.96
01/01/2013	1	259.21	0	313.69	0.96	0.96
01/01/2013	2	245.17	0	313.69	0.96	0.96
01/01/2013	3	228.91	0	313.69	0.96	0.96
01/01/2013	4	216.07	0	313.69	0.96	0.96
01/01/2013	5	209.04	0	313.69	0.96	0.96
⋮	⋮	⋮	⋮	⋮	⋮	⋮
02/01/2013	0	257.39	1	313.69	0.96	0.96
02/01/2013	1	231.63	1	313.69	0.96	0.96
02/01/2013	2	214.74	1	313.69	0.96	0.96
02/01/2013	3	202.79	1	313.69	0.96	0.96
02/01/2013	4	197.69	1	313.69	0.96	0.96
02/01/2013	5	198.25	1	313.69	0.96	0.96
⋮	⋮	⋮	⋮	⋮	⋮	⋮

A filtragem destes valores não se mostrou necessária, visto que os resultados produzidos pelos pontos anteriores cobriam todas as anomalias surgidas inicialmente. Assim, apesar de os valores terem sido calculados, não houve necessidade da sua utilização, contrariamente ao que inicialmente era expectável.

3.2.4 Etapa 4 - Escolha das Variáveis

A necessidade de se desenvolverem ferramentas de previsão para realizar previsões de consumo a curto prazo é amplamente reconhecida por ser uma tarefa essencial para a otimização de atividades operacionais como o agendamento de entrada e saída de serviço de centrais controláveis, despacho de carga, avaliação de segurança e atribuição de capacidade de reserva [40].

De entre uma vasta gama de variáveis, aquelas que melhor explicam os padrões de consumo são usados como input no KDE de forma a incluir a incerteza nas previsões e, assim, facilitar e melhorar as tomadas de decisão.

Como a janela temporal das previsões de curto prazo é de uma hora até uma semana, fatores como horas do dia (3.17) (que permitem distinguir períodos de cheia e vazio), dias da semana (3.18) (dias úteis, fim-de-semana, feriados) e semana do ano apresentam-se como variáveis explicativas para este tipo de previsão, passíveis de serem usados como input no KDE.

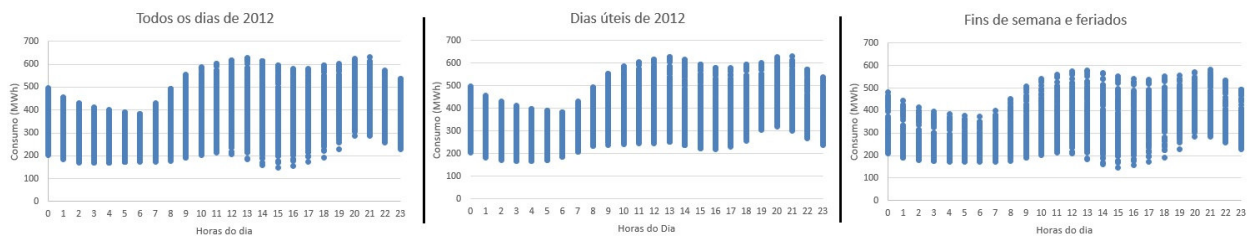


Figura 3.17: Padrões de consumo em função das horas do dia (em Creta), separado por dias úteis, fins de semana e feriados.

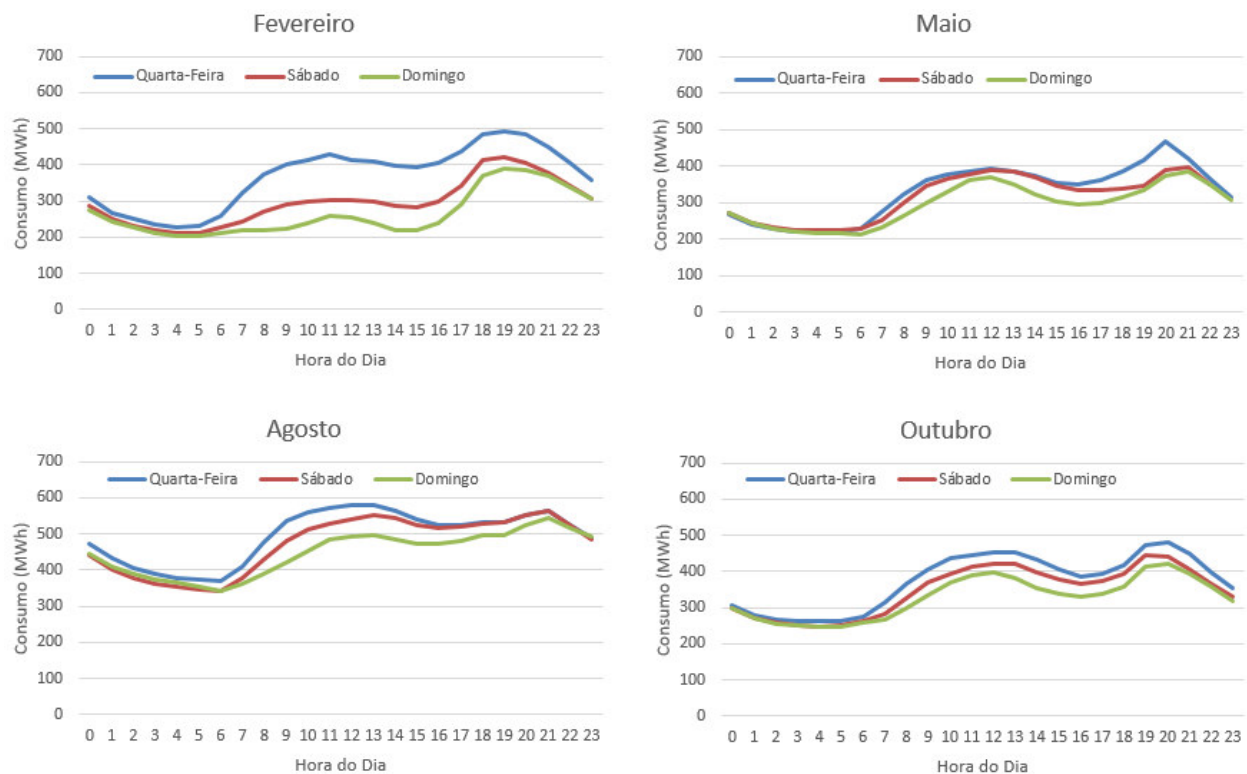


Figura 3.18: Padrões de consumo em função das horas do dia (em Creta), em diferentes períodos do ano.

Pela observação das figuras 3.17 e 3.18 pode-se constatar que os padrões de consumo na ilha de Creta variam:

- consoante os diferentes dias da semana, pois apresentam valores superiores nos dias úteis;
- consoante a altura do ano, pois ao estudar o comportamento dos mesmo dias (quarta-feira, sábado e domingo), em diferentes meses do ano (fevereiro, maio, agosto e outubro), correspondentes às diferentes estações (inverno, primavera, verão, outono), verifica-se que o comportamento do diagrama de carga sofre variações.

Como variável explicativa pode ainda ser considerado o fator da temperatura (3.19), pois o conforto térmico é muito importante para as populações e, em situações de temperaturas acima ou abaixo dos valores normais, há a possibilidade de recorrer a aparelhos reguladores térmicos que vão introduzir aumentos de consumo e alteração dos diagramas de carga, o que não acontece quando esses aparelhos não são usadas, caso a temperatura seja mais amena.

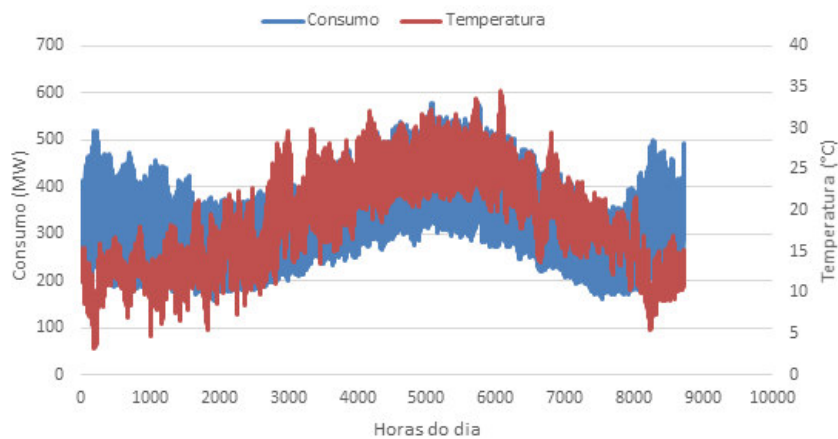


Figura 3.19: Padrões de consumo em função das horas do dia (em Creta), em diferentes períodos do ano.

A figura 3.19 é exemplificativa do que foi afirmado, pois constata-se que nos momentos de maior e menor temperatura o consumo é superior em relação ao verificado para outras temperaturas.

A introdução destas variáveis explicativas no KDE dão origem a previsões probabilísticas de consumo, com as apresentadas na figura 3.20 [24] :

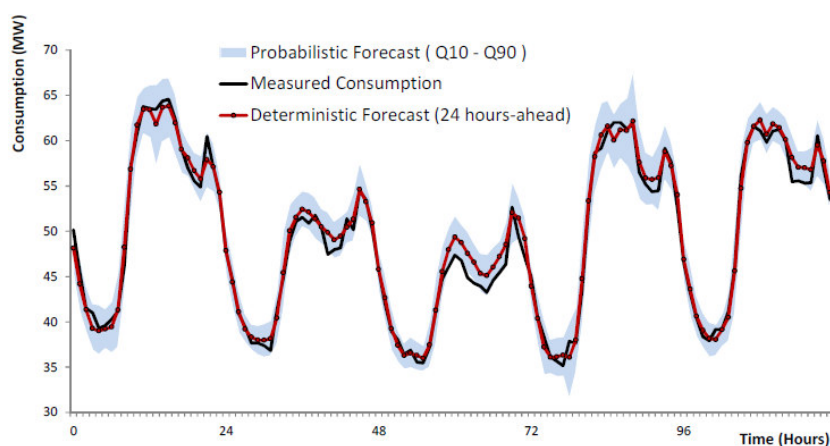


Figura 3.20: Previsão probabilística de consumo, para um quantil de 10 e 90.

Tendo em consideração este pressupostos e feita a normalização dos valores, o passo seguinte,

para a construção dos modelos de previsão, consiste na procura das variáveis que o vão compor. O objetivo passa por escolher as variáveis que melhor expliquem os padrões de consumo de modo a que a previsão seja o mais precisa possível.

Esta escolha terá por base uma análise crítica aos gráficos de dispersão que relacionam o consumo como as diferentes variáveis, de forma a fundamentar as opções tomadas.

Anteriormente, neste capítulo, já se evidenciou a importância de variáveis como Dia da Semana e Mês do Ano mas, agora, fundamentar-se-á essa decisão analiticamente, para além de se analisar a relevância de utilização de outro tipo de variáveis, nomeadamente, Hora do Dia e Temperatura.

Hora do Dia

Como é do conhecimento de todos, os consumos de eletricidade variam de acordo com as horas de vazio e de ponta (consumo em baixa tensão) e horas de vazio normal, horas de super vazio, horas de ponta e horas de cheias (consumos em média tensão). Assim, a variável Hora do Dia mostra-se indispensável para a criação de modelos de previsão.



Figura 3.21: Gráfico de Dispersão que relaciona o Consumo com a Hora do Dia.

Pela visualização da figura 3.21 conclui-se existir uma dependência do Consumo relativamente à Hora do Dia, o que é confirmado pelo facto de as linhas de tendência polinomial de 5ª ordem seguir os valores médios das horas do dia.

Dia da Semana

As alterações que ocorrem nos padrões de consumo da rede, consoante se trate de um dia útil, de um fim de semana ou feriado, revelam que a adição do Dia da Semana como uma variável para a construção dos modelos de previsão trará informação importante que ajudará a tornar mais precisos os resultados da previsão.

Com o auxílio do Microsoft Excel foi criado o gráfico de dispersão (3.22) que estabelece a comparação entre a variável dependente Consumo e a variável independente Dia da Semana, sendo também apresentado o valor médio dos dias da semana.

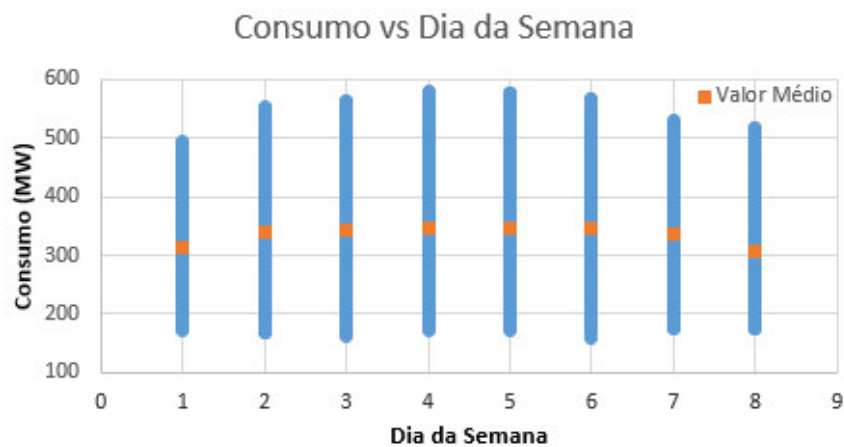


Figura 3.22: Gráfico de Dispersão que relaciona o Consumo com o Dia da Semana.

Através da observação desta figura, verifica-se que, tal como era expectável, os valores de consumo são diferentes caso se trate, de fim de semana (1 e 7), feriado (8) ou dia útil (2, 3, 4, 5 e 6), sendo superior nos dias uteis. Verifica-se, também, que os valores médios de cada dia da semana acompanham a forma do gráfico, o que corrobora o que foi afirmado.

Mês do Ano

A potencial importância desta variável já foi, anteriormente, descrita e será corroborada com o auxílio da ferramenta Microsoft Excel.

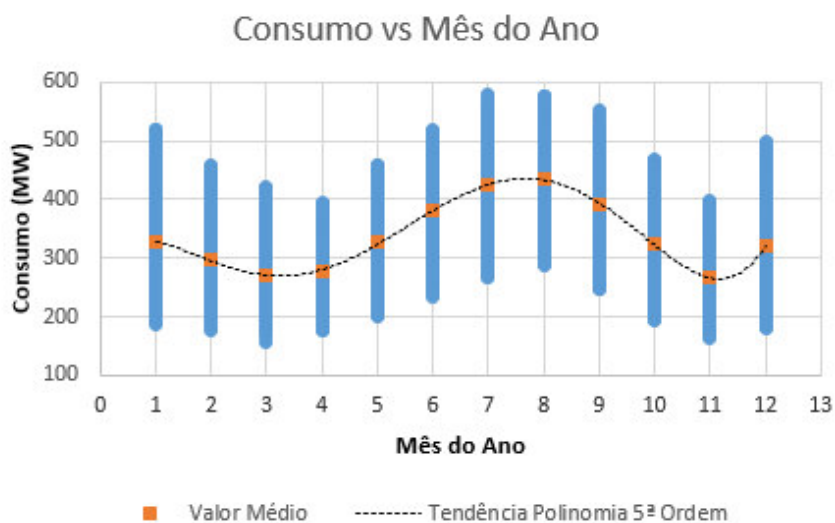


Figura 3.23: Gráfico de Dispersão que relaciona o Consumo com o Mês do Ano.

Uma rápida análise da figura 3.23 permite constatar que o consumo é menor em meses associados a temperaturas mais amenas e maior nos restantes casos, o que está relacionado com a utilização de aparelhos para conforto térmico: aquecedores ou ares-condicionados.

A linha de tendência polinomial de 5ª ordem apresenta a tendência do gráfico que, como se pode observar, acompanha os valores médios dos meses.

Temperatura

A decisão pela utilização desta variável está relacionada com o que foi descrito para a variável Mês do Ano, pois nessa variável constatou-se que em meses que a temperatura é amena os valores de consumo são inferiores aos dos restantes meses, ou seja, espera-se que o estudo da relação entre as variáveis Temperatura e Consumo seja revelador da correlação entre estas duas variáveis.

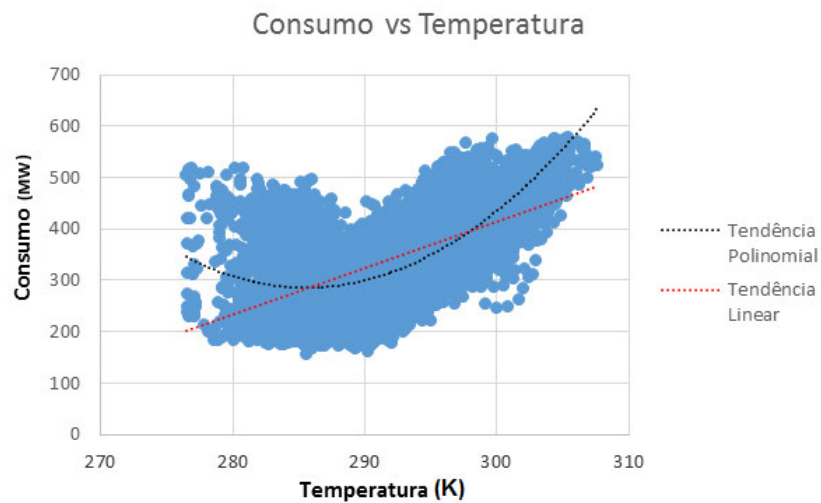


Figura 3.24: Gráfico de Dispersão que relaciona o Consumo com a Temperatura.

A nuvem de pontos mostra, de forma clara, o que foi descrito anteriormente, pois verifica-se que, para valores intermédios de temperatura (correspondente a temperaturas mais amenas), os níveis de consumo são inferiores aos registados para valores mais altos ou mais baixos de temperatura.

As linhas de tendência, nomeadamente a linha de tendência polinomial, seguem a forma da nuvem de pontos o que é um bom indicador da relação entre as variáveis.

Irradiância

Como mencionado aquando da caracterização dos dados (3.1), contrariamente ao que ocorre nos dados das subestações em 2012, os dados de consumo das subestações em 2013 (obtido pela soma de toda a produção de energia associada a essa subestação, nesse ano) não comportam a produção fotovoltaica, pela dificuldade inerente à sua divisão pelas respetivas subestações.

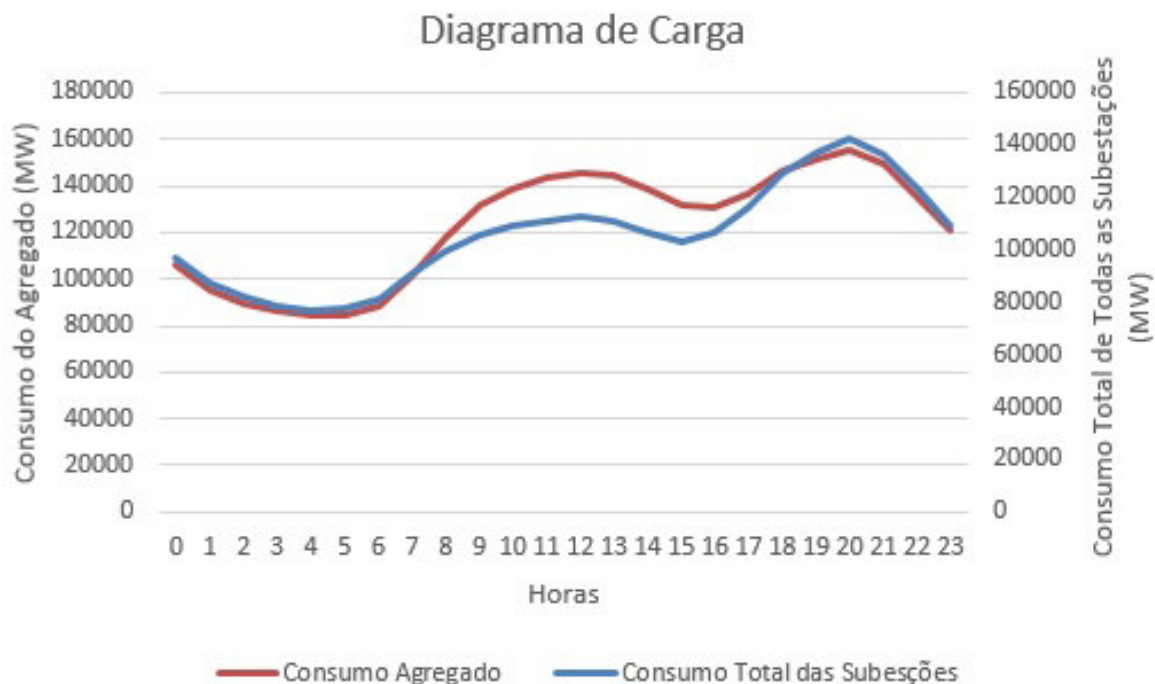


Figura 3.25: Diagrama de Carga do Agregado fornecido e da soma de consumo de todas as subestações.

Como se pode observar na figura 3.25, no período das 8 às 17 horas o comportamento do diagrama de carga do total de consumo das subestações é bastante diferente do que se verifica no *Agregado*. Esta discrepância deve-se ao facto de os valores de consumo total das subestações serem obtidos pela adição de todas as fontes de energia conectadas a cada uma das subestações (excetuando a produção fotovoltaica pela razão mencionada), visto não haver nem compra nem venda de energia, ou seja, o consumo será igual à energia que for produzida, sem incluir a produção fotovoltaica. Pelo contrário, para o caso do *Agregado*, como o próprio nome indica, é feita a agregação de todos os pontos de produção de energia da ilha, onde já se encontra incluída a produção fotovoltaica.

Determinada a razão dos desvios entre o diagrama de carga do *Agregado* e do total de consumo das subestações, considerou-se a utilização da variável Irradiância. A associação desta variável é feita de forma a avaliar se será suficiente para colmatar a falta dos valores da produção fotovoltaica nos consumos das subestações.

Apresenta-se, tal como em avaliações já descritas, o gráfico produzido em Microsoft Excel (3.26) que estabelece a comparação entre a variável dependente Consumo e a variável independente Irradiância, ao qual foram aplicadas linhas de tendência lineares e polinomiais.

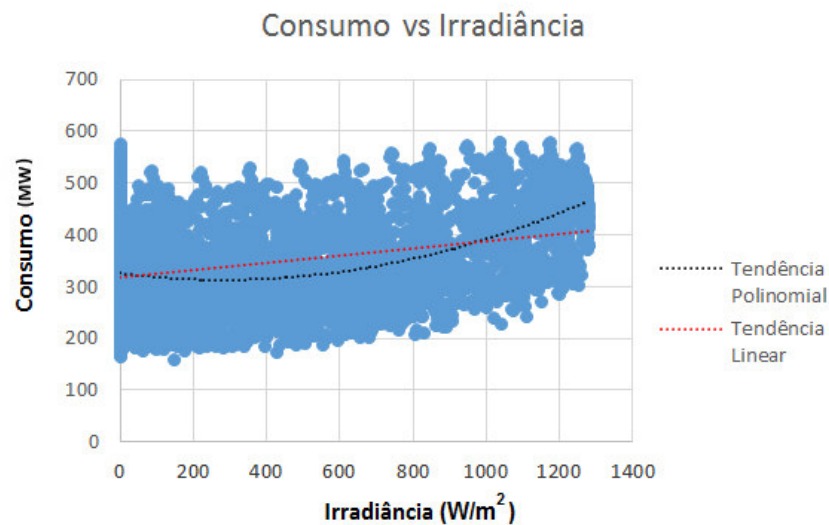


Figura 3.26: Gráfico de Dispersão que relaciona o Consumo com a Irradiância.

Constata-se, assim, que as linhas de tendência, nomeadamente a linha de tendência polinomial, seguem a forma da nuvem de pontos, sendo este facto um bom indicador da relação entre as variáveis.

Consumo do Dia Anterior

Um fator que se pode revelar de extrema importância para o aperfeiçoamento dos modelos de previsão é a inserção de uma variável que comporte os valores de consumo correspondentes ao dia anterior àquele cuja previsão se vai efetuar.

Esta informação extra poderá ajudar a tornar a previsão mais precisa, conduzindo, assim, a uma expectável redução do erro MAPE, pois facultará ao modelo uma capacidade extra que lhe permitirá estar “melhor preparado”.

Apesar de esta variável poder ser relevante para os modelos de previsão, a sua utilização é utópica numa situação em tempo real, pois nem sempre os criadores dos modelos de previsão têm acesso aos dados de consumo do dia anterior (por vezes nem ao consumo dos dois dias anteriores conseguem ter acesso) de modo a poder usá-los para prever o consumo do dia seguinte.

Ainda que seja difícil empregar esta solução, como se trata de um estudo procedeu-se à sua inserção para perceber, caso o fornecimento da informação de consumo fosse eficiente, qual seria a sua influência nas previsões.

Posto isto, foi estudada, por análise da figura e da linha de tendência (3.27), a correlação existente entre a variável dependente Consumo e a variável independente Consumo do Dia Anterior (CDA).

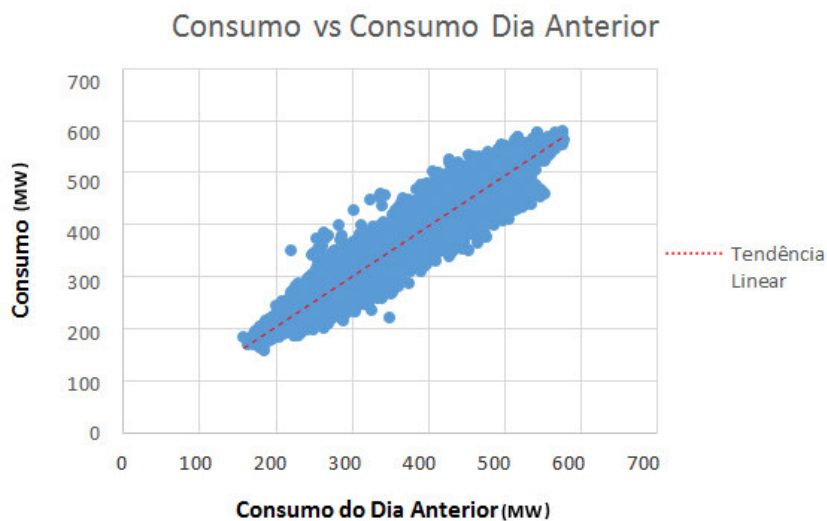


Figura 3.27: Gráfico de Dispersão que relaciona o Consumo com o Consumo do Dia Anterior.

A linha de tendência demonstra a enorme correlação entre estas duas variáveis, o que evidencia a grande importância que a inserção desta variável pode ter para os modelos.

3.3 Construção dos modelos de previsão

Concluído todo o processo de obtenção e tratamento de dados, base comum a todos os modelos explorados, passou-se, por fim, à criação dos modelos.

3.3.1 Descrição do processo de construção

Tendo por base o descrito quanto ao funcionamento do KDE (2.3.6.1), este processo decorreu da seguinte forma:

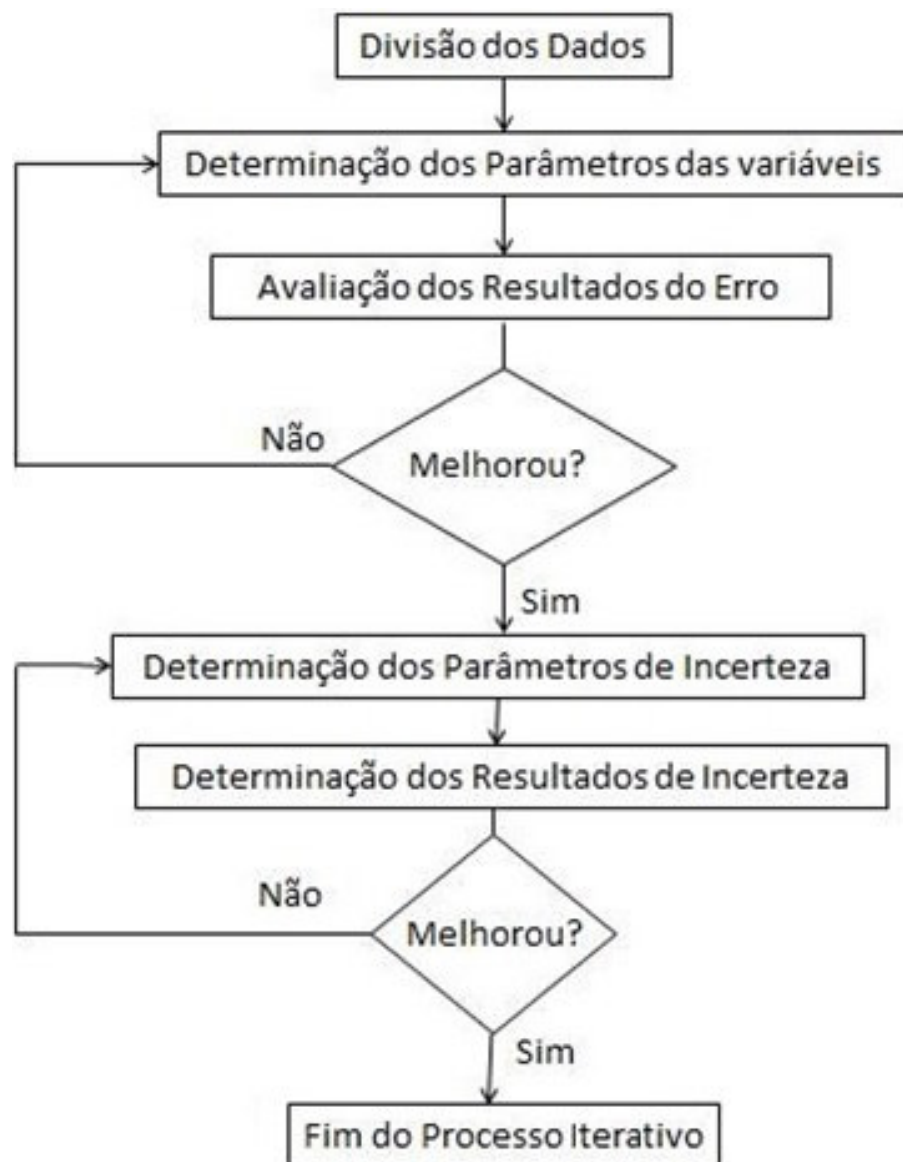


Figura 3.28: Etapas do Processo Iterativo de Criação dos Modelos de Previsão.

O processo iterativo de construção de modelos de previsão passa, numa primeira fase, pela divisão dos dados necessários, ou seja, escolha das variáveis a usar para o modelo que se for tratar, assim como o ano a que se refere (divisão dos anos apresentada em 3.1).

Escolhidos os dados e as variáveis a usar, o passo seguinte passa por determinar o parâmetro “h” a introduzir no KDE, que é apresentado na expressão usado no procedimento do KDE para múltiplas variáveis (2.9) representado em 2.3.6.1.

Esse parâmetro “h” é inicializado com 10% do valor de desvio padrão das variáveis escolhidas para o modelo que se for analisar. A otimização deste parâmetro de cada variável é feita através da alteração progressiva de cada um deles, de forma a que a previsão resultante do programa, para o período de teste considerado, tenha o menor erro possível (erro MAPE) quando comparada com

os valores reais nesse mesmo período (2.5.2).

O procedimento descrito encontra-se exemplificado na tabela 3.12 onde se apresentam, na primeira linha os valores iniciais de “h” de duas variáveis e o erro MAPE obtido com base nesses valores.

Na segunda linha o parâmetro “h” da primeira variável já é o valor final resultante do processo iterativo para essa variável, tendo resultado, juntamente com o parâmetro inicial “h” da outra variável, no erro MAPE da mesma linha.

Por fim, na última linha é apresentado o valor final do parâmetro “h” de ambas as variáveis, assim como o melhor erro MAPE obtido com esse conjunto.

Tabela 3.12: Processo de seleção do melhor parâmetro h de cada variável.

Hora do Dia	Dia da Semana	MAPE
0.2107	0.69045	20.1%
0.83442	0.69045	17.3%
0.83442	0.44071	14.9%

Determinado o melhor conjunto de variáveis e, com isso, o menor erro MAPE possível, a próxima etapa passa pela determinação do parâmetro que irá influenciar o valor da incerteza.

A última variável a ser introduzida no KDE, para todos os modelos, é o Consumo Real nesse período e o parâmetro “h” associado a esta variável irá permitir determinar o valor da incerteza. O parâmetro desta variável é inicializado com 1% do valor máximo de consumo registrado na subestação em estudo.

Para avaliar a evolução da incerteza são usados os valores extraídos do KDE (2.3.6.1), nomeadamente α , β , “A” (mínimo) e “B” (máximo), aos quais é aplicada a função “inv.beta” do Microsoft Excel, para uma dada probabilidade, obtendo-se o inverso da função densidade de probabilidade cumulativa beta. Concretizando, com os valores de α , de β , de “A”, de “B” e a partir de um valor de probabilidade, pertencente a um dado quantil, determina-se um valor do consumo.

A modelização da incerteza é feita para diferentes intervalos probabilísticos, com bandas de diferentes larguras, sendo “A” e “B” os delimitadores do intervalo de confiança.

As bandas consideradas têm larguras que variam de 0.1 a 0.9, o que significa que o ideal seria que existissem de 10% a 90%, respetivamente, de probabilidade de os valores reais de consumos se encontrarem dentro da banda probabilística considerada.

Para o cálculo da percentagem de valores que se encontram dentro da banda considerada é necessário começar por determinar esses mesmos valores.

$$Incerteza(t, b_{min}, b_{max}) = \begin{cases} 1, & inv.Beta(b_{min}, \alpha, \beta, A, B) < C_{real} < inv.Beta(b_{max}, \alpha, \beta, A, B) \\ 0, & \text{outros casos} \end{cases} \quad (3.10)$$

Em que:

- **Incerteza(t,bmin,bmax)** – valores que são validados pela condição, para um dado instante t ;
- **t** – instante de tempo considerado;
- **bmin** – valor mínimo da banda probabilística;
- **bmax** – valor máximo da banda probabilística;
- α – valor da variável α dado pelo KDE;
- β – valor da variável β dada pelo KDE;
- **A** – valor mínimo da banda de confiança para o valor de consumo previsto;
- **B** – valor máximo da banda de confiança para o valor de consumo previsto;
- **Creal** – valor de Consumo Real no instante de tempo t .

Por fim, o cálculo da probabilidade é dado pela tabela 3.11.

$$P(C_{real} \in [b_{min}, b_{max}]) = \frac{\sum_{i=0}^t Incerteza(t)}{t} \quad (3.11)$$

Como a intenção é obter a melhor incerteza possível para cada banda, ou seja, que a percentagem de valores dentro da banda seja o mais similar possível ao valor da largura da mesma, o erro da incerteza é calculado segundo a equação 3.12.

$$Erro_{inc} = \frac{|b_{max} - b_{min}| - P(C_{real} \in [b_{min}, b_{max}])}{|b_{max} - b_{min}|} \quad (3.12)$$

O valor da largura da banda e a percentagem associada à probabilidade de o consumo se encontrar dentro dessa mesma banda pode ser observado na figura 3.29, para um caso geral.

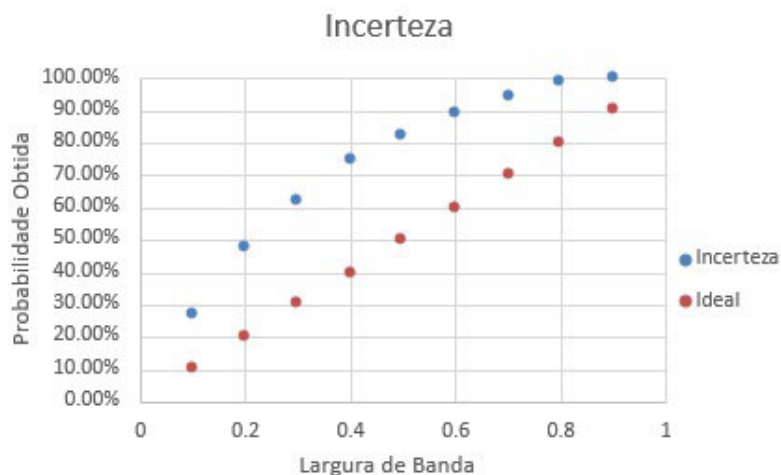


Figura 3.29: Comparação das probabilidades obtidas em cada banda de incerteza.

Com os valores dos erros de incerteza, calculados para as diferentes bandas consideradas, é feita uma média desses mesmos erros que, no processo iterativo de melhoria da incerteza, se pretende que seja o menor possível correspondendo a uma precisão cada vez maior da previsão do consumo.

$$Media_{erro} = \frac{\sum_{i=0}^9 Erro_{inc}(p)}{9}, \quad p \in [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9] \quad (3.13)$$

Atingido o mínimo possível da média dos erros das incertezas, a figura de comparação das probabilidades obtidas em cada banda de incerteza deve aparentar uma forma semelhante à apresentada em 3.30.

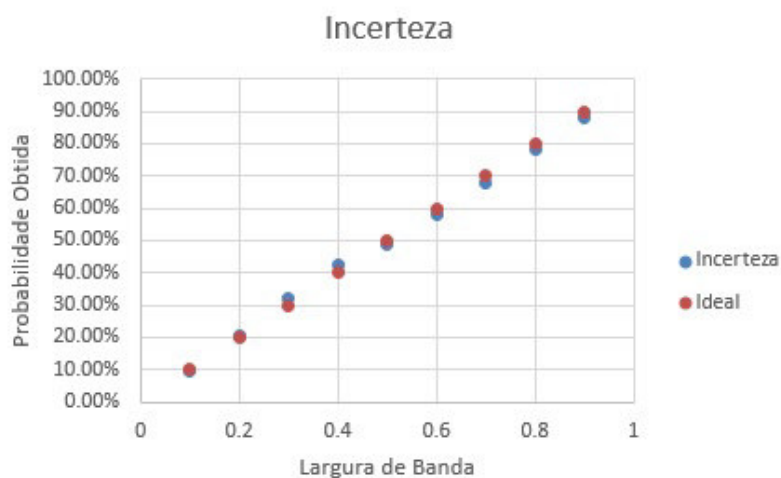


Figura 3.30: Comparação das probabilidades obtidas em cada banda de incerteza.

Após se obter o melhor parâmetro h para a variável associada à incerteza, o processo de criação do modelo considera-se concluído.

3.3.2 Modelo 1 - Modelo Base Simples

Neste modelo, as variáveis a serem usados são:

- Parâmetro 1 - Hora do Dia;
- Parâmetro 2 - Dia da Semana;

3.3.3 Modelo 2 - Modelo Simples

Neste modelo, recorre-se às variáveis:

- Parâmetro 1 - Hora do Dia;
- Parâmetro 2 - Dia da Semana;
- Parâmetro 3 - Mês do Ano;

3.3.4 Modelo 3 - Modelo Simples com explicador Temperatura

Neste modelo, foram aplicadas as variáveis:

- Parâmetro 1 - Hora do Dia;
- Parâmetro 2 - Dia da Semana;
- Parâmetro 3 - Mês do Ano;
- Parâmetro 4 - Temperatura;

3.3.5 Modelo 4 - Modelo Base com explicador Temperatura e Irradiância

Neste modelo, foram utilizadas as variáveis:

- Parâmetro 1 - Hora do Dia;
- Parâmetro 2 - Dia da Semana;
- Parâmetro 3 - Mês do Ano;
- Parâmetro 4 - Temperatura;
- Parâmetro 5 - Irradiância;

3.3.6 Modelo 5 - Modelo Base com explicador Temperatura e Consumo do Dia Anterior

Neste modelo, as variáveis a serem usados são:

- Parâmetro 1 - Hora do Dia;
- Parâmetro 2 - Dia da Semana;
- Parâmetro 3 - Mês do Ano;
- Parâmetro 4 - Temperatura;
- Parâmetro 5 - Consumo do Dia Anterior;

Capítulo 4

Análise de desempenho de modelos de previsão

Após a descrição da base de criação dos vários modelos previsionais (3.3), neste capítulo será feita uma comparação dos resultados obtidos, do erro MAPE e da incerteza, nos dois casos de estudo descritos em 3.1.

4.1 Comparação dos Casos de Estudo

4.1.1 Caso de Estudo 1

No final do processo iterativo (3.3.1), obtiveram-se os seguintes valores de erro MAPE (calculado pela equação 2.20) e de incerteza (calculado pela equação 3.13):

Tabela 4.1: Modelos e subestações usados para o caso de estudo 1.

Modelo	Subestações	Período	MAPE	Incerteza
Modelo 1 - Modelo Base Simples	Agregado	Inverno 2012	7.3%	8.0%
		Primavera 2012	9.6%	
		Verão 2012	31.8%	
		Outono 2012	5.3%	
		Global	14.9%	
Modelo 2 - Modelo Base	Agregado	Inverno 2012	5.6%	7.6%
		Primavera 2012	9.0%	
		Verão 2012	9.0%	
		Outono 2012	5.0%	
		Global	6.8%	
Modelo 3 - Modelo Base com explicador Temperatura	Agregado	Inverno 2012	4.4%	6.5%
		Primavera 2012	6.1%	
		Verão 2012	7.3%	
		Outono 2012	4.5%	
		Global	6.2%	
	Hrakteio 2	Inverno 2012	4.6%	7.6%
		Primavera 2012	5.6%	
		Verão 2012	7.3%	
		Outono 2012	3.1%	
		Global	5.4%	
	Hrakteio 3	Inverno 2012	6.8%	8.5%
		Primavera 2012	11.2%	
		Verão 2012	12.1%	
		Outono 2012	5.1%	
		Global	8.8%	
	Rethymno	Inverno 2012	9.8%	6.3%
		Primavera 2012	6.7%	
		Verão 2012	2.5%	
		Outono 2012	8.1%	
		Global	8.8%	

Como se pode atestar pela tabela 4.1, no *Modelo 1 - Modelo Base Simples* - os valores do erro MAPE, obtidos com a utilização das variáveis Hora do Dia e Dia da Semana, foram bastante elevados, nomeadamente no período do verão, pelo facto de o modelo não ter conseguido prever, com sucesso, o consumo real mais elevado neste período, como se pode comprovar com a figura 4.1:

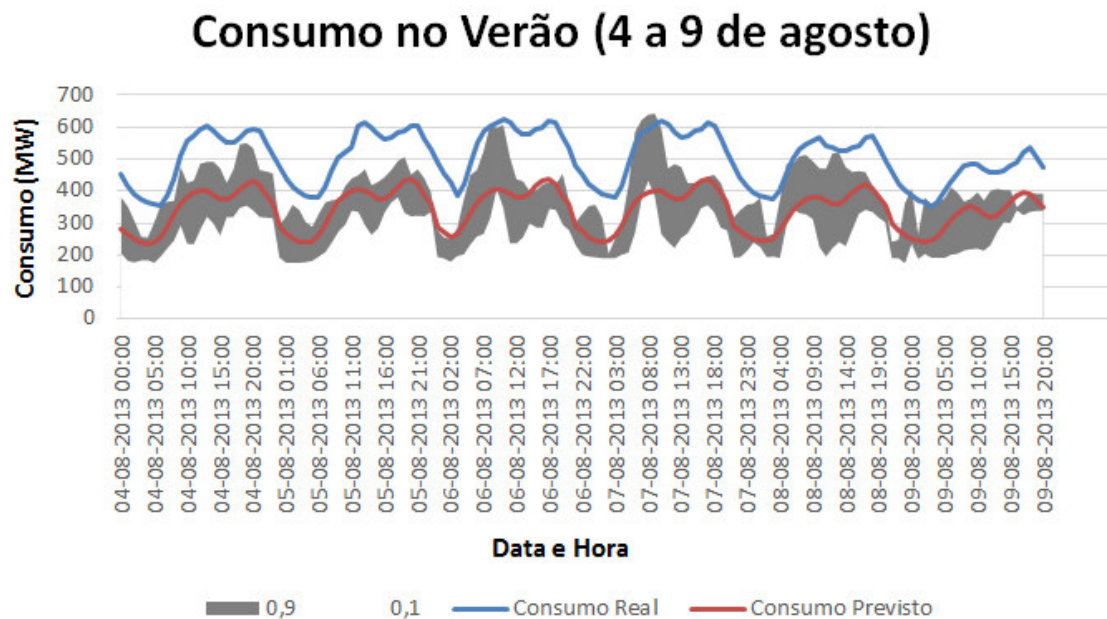


Figura 4.1: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, usando o Modelo 1 - Modelo Base Simples.

Ao criar o *Modelo 2 - Modelo Base* - com a introdução da variável Mês do Ano, verificou-se uma descida substancial nos valores de erro MAPE, o que facilitou a previsão mais correta do consumo no período de verão.

Por fim, a introdução da variável Temperatura no *Modelo 3 - Modelo Base com explicador Temperatura* voltou a revelar-se ligeiramente mais precisa, havendo uma redução generalizada do erro de MAPE.

Nas figuras (4.2 - 4.5) apresenta-se o comportamento, para cada um dos períodos, do Consumo Real e Consumo Previsto, obtido por utilização do *Modelo 3 - Modelo Base com explicador Temperatura* - para o Agregado.

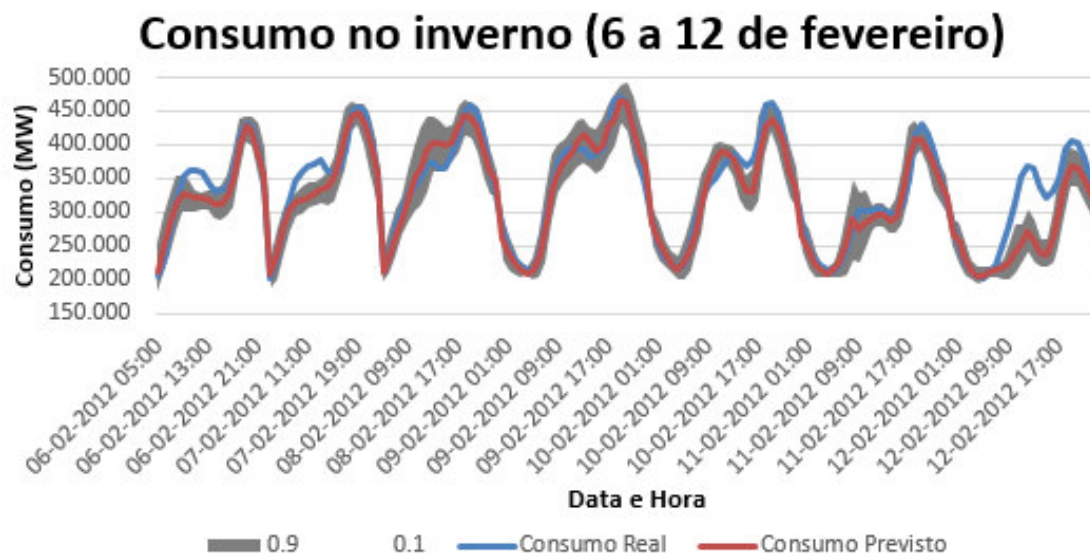


Figura 4.2: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, usando o Modelo 3.

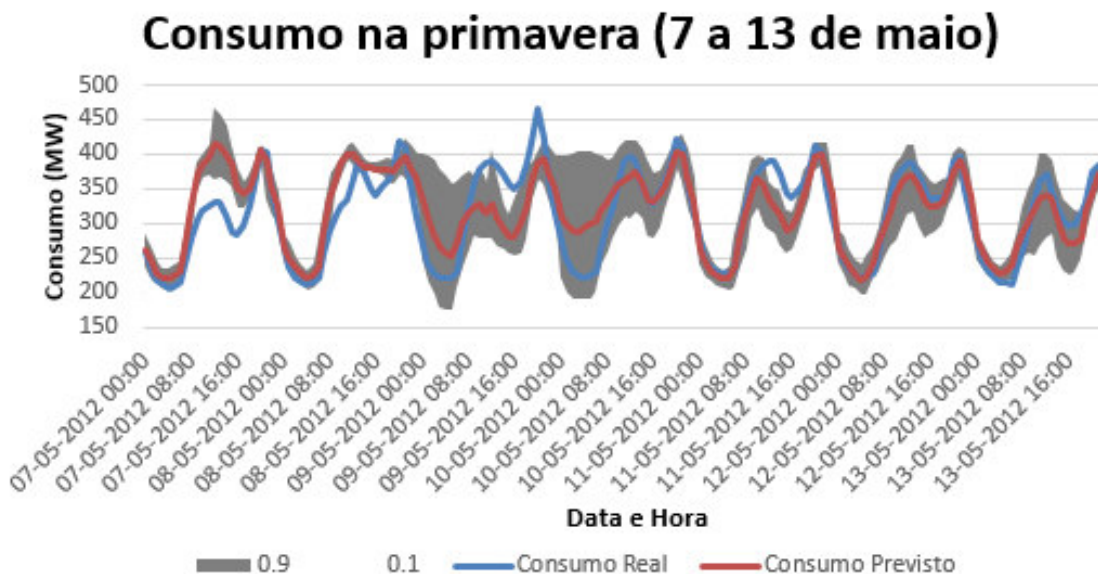


Figura 4.3: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, usando o Modelo 3.

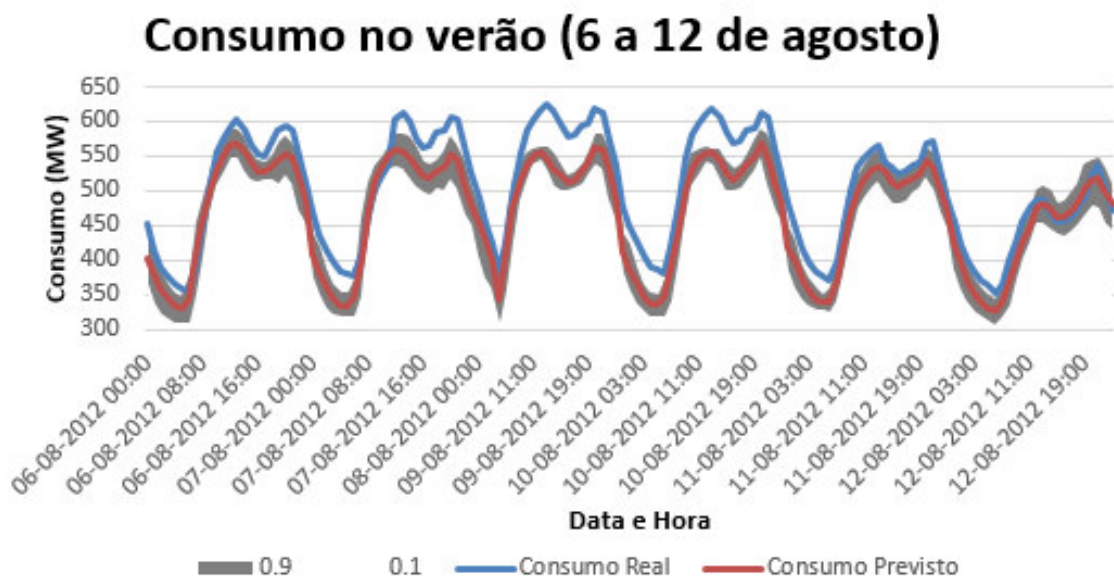


Figura 4.4: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, usando o Modelo 3.

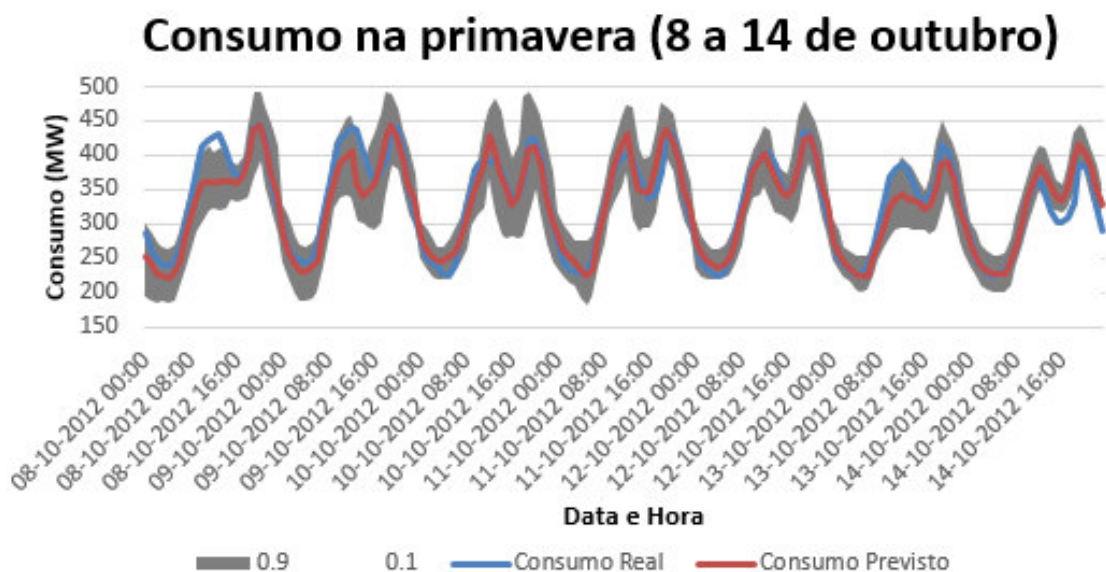


Figura 4.5: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, usando o Modelo 3.

Analogamente ao que aconteceu para o erro MAPE, o erro médio da incerteza também melhorou com a introdução de novas variáveis, embora não de forma tão significativa.

Em suma, notou-se uma clara evolução positiva dos valores do erro MAPE e da incerteza entre os modelos criados para o *Agregado* o que demonstra que, para os dados de consumo disponibilizados, referentes a 2012, e para o conjunto de variáveis consideradas, o *Modelo 3 - Modelo*

Base com explicador Temperatura é, sem dúvida, o modelo que originará melhores resultados de previsão.

Quanto às restantes subestações usadas no *Modelo 3 - Modelo Base com explicador Temperatura*, os valores do erro MAPE mantêm-se aproximadamente equivalentes entre os períodos considerados, embora, normalmente, esses valores sejam superiores na primavera e no verão. Este facto pode ser explicado por consumos inesperados (inferiores no caso da primavera e superiores no caso do verão) como se pode verificar nas figuras (A.1 - A.12). O erro médio da incerteza nestas subestações é, também, relativamente reduzido.

4.1.2 Caso de Estudo 2

Tabela 4.2: Modelos e subestações usados para o caso de estudo 2.

Modelo	Subestações	Período	MAPE	Incerteza
Modelo 3 - Modelo Base com explicador Temperatura	Agregado	Inverno 2013	5.3%	6.6%
		Primavera 2013	6.2%	
		Verão 2013	2.1%	
		Outono 2013	7.5%	
		Global	5.0%	
	Ag. Nikolaos	Inverno 2013	8.1%	2.4%
		Primavera 2013	7.6%	
		Verão 2013	3.2%	
		Outono 2013	15.2%	
		Global	7.9%	
	Chania	Inverno 2013	7.2%	7.4%
		Primavera 2013	5.0%	
		Verão 2013	2.9%	
		Outono 2013	5.0%	
		Global	5.1%	
	Hrackleio 1	Inverno 2013	2.4%	6.4%
		Primavera 2013	6.2%	
		Verão 2013	3.7%	
		Outono 2013	14.9%	
		Global	6.0%	
	Hrackleio 2	Inverno 2013	5.7%	2.1%
		Primavera 2013	8.4%	
		Verão 2013	5.4%	
		Outono 2013	4.9%	
		Global	6.1%	
	Hrackleio3	Inverno 2013	5.4%	7.8%
		Primavera 2013	6.6%	
		Verão 2013	3.4%	
		Outono 2013	3.6%	
		Global	4.7%	
	Rethymno	Inverno 2013	9.7%	2.5%
		Primavera 2013	8.8%	
		Verão 2013	3.2%	
		Outono 2013	7.3%	
		Global	6.9%	
	Sitia	Inverno 2013	12.5%	2.9%
		Primavera 2013	11.8%	
		Verão 2013	6.3%	
		Outono 2013	9.8%	
		Global	9.8%	

Tabela 4.3: Modelos e subestações usados para o caso de estudo 2.

Modelo	Subestações	Período	MAPE	Incerteza
Modelo 4 - Modelo Base com explicador Temperatura e Irradiância	Agregado	Inverno 2013	5.7%	0.3%
		Primavera 2013	5.1%	
		Verão 2013	7.2%	
		Outono 2013	2.1%	
		Global	5.7%	
	Ag. Nikolaos	Inverno 2013	6.9%	1.7%
		Primavera 2013	6.5%	
		Verão 2013	10.0%	
		Outono 2013	2.7%	
		Global	6.9%	
	Chania	Inverno 2013	5.8%	4.2%
		Primavera 2013	7.7%	
		Verão 2013	6.2%	
		Outono 2013	3.4%	
		Global	5.8%	
	Hrackleio 1	Inverno 2013	8.6%	4.1%
		Primavera 2013	7.1%	
		Verão 2013	6.6%	
		Outono 2013	3.2%	
		Global	8.6%	
	Hrackleio 2	Inverno 2013	6.5%	2.1%
		Primavera 2013	5.7%	
		Verão 2013	8.3%	
		Outono 2013	6.2%	
		Global	6.5%	
	Hrackleio3	Inverno 2013	5.3%	9.9%
		Primavera 2013	5.7%	
		Verão 2013	7.1%	
		Outono 2013	3.8%	
		Global	5.3%	
	Rethymno	Inverno 2013	8.0%	6.5%
		Primavera 2013	10.8%	
		Verão 2013	11.1%	
		Outono 2013	3.1%	
		Global	8.0%	
	Sitia	Inverno 2013	8.7%	6.3%
		Primavera 2013	11.1%	
		Verão 2013	10.5%	
		Outono 2013	5.8%	
		Global	8.7%	

Tabela 4.4: Modelos e subestações usados para o caso de estudo 2.

Modelo	Subestações	Período	MAPE	Incerteza
Modelo 5 - Modelo Base com explicador Temperatura e Consumo do Dia Anterior	Agregado	Inverno 2013	2.8%	2.6%
		Primavera 2013	4.3%	
		Verão 2013	2.3%	
		Outono 2013	1.7%	
		Global	2.8%	
	Ag. Nikolaos	Inverno 2013	4.3%	4.0%
		Primavera 2013	8.8%	
		Verão 2013	3.8%	
		Outono 2013	2.8%	
		Global	4.3%	
	Chania	Inverno 2013	3.6%	3.9%
		Primavera 2013	5.1%	
		Verão 2013	3.0%	
		Outono 2013	2.5%	
		Global	3.6%	
	Hrackleio 1	Inverno 2013	5.1%	8.6%
		Primavera 2013	5.1%	
		Verão 2013	4.9%	
		Outono 2013	2.6%	
		Global	5.1%	
	Hrackleio 2	Inverno 2013	3.9%	2.8%
		Primavera 2013	5.0%	
		Verão 2013	3.3%	
		Outono 2013	3.6%	
		Global	3.9%	
	Hrackleio3	Inverno 2013	3.7%	3.1%
		Primavera 2013	5.1%	
		Verão 2013	4.8%	
		Outono 2013	2.5%	
		Global	3.7%	
	Rethymno	Inverno 2013	6.4%	4.0%
		Primavera 2013	9.8%	
		Verão 2013	6.7%	
		Outono 2013	2.5%	
		Global	6.4%	
	Sitia	Inverno 2013	9.3%	3.8%
		Primavera 2013	12.0%	
		Verão 2013	12.4%	
		Outono 2013	5.8%	
		Global	9.3%	

Como referido anteriormente, o *Modelo 4 - Modelo Base com explicador Temperatura e Irradiância* - surgiu da necessidade de tentar colmatar a discrepância nos dados de consumo pela falta da inclusão dos dados relativos à produção fotovoltaica. Pela tabela 4.3, percebe-se que, num cômputo geral, a introdução da Irradiância não se apresentou propriamente frutífera, visto que houve um aumento do erro MAPE em relação ao *Modelo 3 - Modelo Base com explicador Temperatura*, tendência que se manteve também para a incerteza.

Aquando da caracterização dos dados (3.1) foi apresentada a capacidade de centrais fotovoltaicas instaladas por subestação (3.2) sendo que, contrariamente ao que se esperava, às centrais com maior capacidade não esteve associada uma melhoria no erro MAPE. Apenas houve uma melhoria muito ligeira do erro MAPE na 2ª e 3ª subestações com maior capacidade instalada, de entre as consideradas, Sitia e Hrakleio 2 respetivamente.

No *Modelo 5 - Modelo Base com explicador Temperatura e Consumo do Dia Anterior* - verifica-se uma melhoria generalizada do erro MAPE quando comparado com o *Modelo 3 - Modelo Base com explicador Temperatura*, o que evidencia a importância da inclusão do Consumo do Dia Anterior, quando possível. No que diz respeito à incerteza, a melhoria já não foi tão evidente, havendo no entanto um decréscimo significativo no valor da incerteza do *Agregado*

Nas figuras em anexo (B.1 - B.32) pode-se observar o comportamento, para cada um dos períodos, do Consumo Real e Consumo Previsto, obtido pela utilização do *Modelo 5 - Modelo Base com explicador Temperatura e Consumo do Dia Anterior* - para as subestações consideradas.

Capítulo 5

Previsão com interdependência geográfica

5.1 Divisão dos dados

A ilha de Creta pode ser considerada um sistema fechado onde as necessidades energéticas são completamente satisfeitas pela produção de energia na ilha. Isto significa que não existe ligação a outra rede externa não havendo, assim, trocas de energia.

A modelização da correlação entre variáveis é importante quando são necessários dados desagregados e sincronizados de consumo em diversos pontos geográficos da rede, como, por exemplo, para utilizar em trânsito de potências, pois permite obter o consumo numa subestação a partir da previsão do *Agregado* ou a partir de outras subestações

Nesta secção é apresentada uma proposta para recriar previsões sincronizadas e coerentes, num conjunto de pontos distribuídos geograficamente.

Para corroborar essa importância foi feito um estudo que tem por base as mesmas datas e subestações usadas como in-sample no *Modelo 3 - Modelo Base com explicador Temperatura* - para o Caso de Estudo 1 (3.1), procedendo-se, posteriormente, à sua comparação. Assim,

- Datas usadas: dias correspondentes às semanas de 6 a 12 de fevereiro, 7 a 13 de maio, 6 a 12 de agosto e 8 a 14 de outubro de 2012;
- Subestações consideradas: Hrakleio 2, Hrakleio 3 e Rethymno;
- Conjuntos: vão ser considerados dois conjuntos (dia úteis - 1; fins de semana - 2), por apresentarem características diferentes e para que seja feita uma previsão mais precisa.

Nota 1: Foi usado apenas o Caso de Estudo 1, pois, como referido anteriormente, nas diferentes subestações estão associados dados de consumo mais fiáveis que no Caso de Estudo 2, pelo facto de a potência fotovoltaica estar incluída.

5.2 Estudo da correlação

Pelo facto de se pretender efetuar uma comparação entre os valores obtidos neste tipo de previsão e os obtidos no *Modelo 3 - Modelo Base com explicador Temperatura* - (Caso de Estudo 1 (3.1)), o estudo da correlação é feito através da análise da correlação entre uma subestação base de correlação e a subestação a prever. Os dois estudos efetuados foram feitos:

- através da análise da correlação entre o valor de previsão do *Agregado* obtido no *Modelo 3 - Modelo Base com explicador Temperatura* - (Caso de Estudo 1 (3.1)) e o valor de consumo real das subestações descritas 5.1.
- através da análise da correlação entre o valor de previsão e o valor de consumo real das subestações descritas 5.1.

Tabela 5.1: Coeficiente R^2 e Coeficientes de Pearson para a relação entre o consumo do Agregado e o consumo das subestações.

Subestação base de correlação	Subestação a prever	Conjunto	R^2 Polinomial	Pearson Polinomial
Agregado	Hracleio2	1	0.6779	0.8234
		2	0.4446	0.6668
	Hracleio3	1	0.8447	0.9191
		2	0.7512	0.8667
	Rethymno	1	0.8732	0.9345
		2	0.8249	0.9082
Hracleio2	Hracleio3	1	0.5367	0.7326
		2	0.6414	0.8009
	Rethymno	1	0.8447	0.9191
		2	0.5262	0.7254
Hracleio3	Hracleio2	1	0.6254	0.7908
		2	0.4953	0.7038
	Rethymno	1	0.6848	0.8275
		2	0.6211	0.7881
Rethymno	Hracleio2	1	0.71	0.8426
		2	0.4927	0.7019
	Hracleio3	1	0.7397	0.8601
		2	0.7077	0.8413

Através das figuras em anexo (C.1 - C.18) e da tabela (5.1), constata-se que a correlação é bastante significativa, sendo que os valores de correlação R^2 e de Pearson apresentam valores, na sua maioria, bastante próximos de “1” o que representa uma correlação perfeita.

5.3 Cálculo da previsão

Com base nas equações das linhas de tendência polinomial de 2ª ordem (do tipo $y = a \cdot x^2 + b \cdot x + c$), apresentadas nas figuras em anexo (C.1 - C.18) e de forma mais clara na tabela 5.2, é feita a previsão do consumo para as subestações consideradas em 5.1, tendo por base o valor de consumo de diferentes subestações.

Tabela 5.2: Equações das linhas de tendência polinomial de 2ª ordem (do tipo $y = a \cdot x^2 + b \cdot x + c$) obtidas pela correlação entre o Agregado e cada uma das subestação, para o conjunto de dados especificado.

Subestação base de correlação	Subestação a prever	a	b	c	Conjunto
Agregado	Hracleio2	-0.00026	0.30715	-23.58896	1
		-0.00017	0.19061	0.86258	2
	Hracleio3	0.00001	0.05158	2.79588	1
		-0.00001	0.05535	3.50272	2
	Rethymno	0.00012	0.03265	13.10143	1
		0.00011	0.03891	13.47507	2
Hracleio2	Hracleio3	-0.00871	1.26847	-17.48061	1
		-0.00789	1.17697	-14.17582	2
	Rethymno	-0.0143	2.18606	-29.71874	1
		-0.01229	2.07942	-25.97258	2
Hracleio3	Hracleio2	-0.05214	3.78564	-11.0655	1
		-0.05214	3.78564	-11.0655	2
	Rethymno	0.10176	-1.94474	34.50895	1
		0.10176	-1.94474	34.50895	2
Rethymno	Hracleio2	-0.02633	3.28581	-36.42356	1
		-0.01706	2.04978	-7.44681	2
	Hracleio3	-0.00301	0.78981	-3.71469	1
		-0.00527	0.86689	-3.7898	2

Fazendo a previsão, por exemplo, para Hracleio 2 no conjunto 1, com base no *Agregado*, parte-se da respetiva equação e substitui-se o valor corresponde a “x” pela previsão real nessa subestação, para o mesmo período. Ao fazer-se essa substituição, o valor de “y” que se obtém é a previsão, para o período considerado, da subestação Hracleio 2, com base na correlação existente entre essa mesma subestação e o *Agregado*.

Para se estudar a performance destas previsões, é calculado o erro MAPE entre os valores obtidos por este processo (apresentados nas figuras C.19 - C.36) e os valores reais de consumo, tendo-se obtido os valores de erro MAPE apresentados na tabela 5.3:

Tabela 5.3: Valor do erro MAPE para a previsão com interdependência geográfica.

Subestação base de correlação	Subestação a prever	Conjunto	MAPE
Agregado	Hracleio2	1	5.3%
		2	6.2%
	Hracleio3	1	4.8%
		2	6.7%
	Rethymno	1	4.7%
		2	4.9%
Hracleio2	Hracleio3	1	14.1%
		2	4.9%
	Rethymno	1	17%
		2	6.8%
Hracleio3	Hracleio2	1	11.2%
		2	5.1%
	Rethymno	1	12.4%
		2	5.8%
Rethymno	Hracleio2	1	10%
		2	4.9%
	Hracleio3	1	10.1%
		2	4.2%

Por fim, foi feita a comparação entre o erro MAPE da tabela 5.3 e o erro MAPE do *Modelo 3 - Modelo Base com explicador Temperatura* - para o Caso de Estudo 1 (3.1), dividido pelos mesmos conjuntos.

Tabela 5.4: Comparação do erro MAPE entre a previsão com interdependência geográfica e a previsão usando o Modelo 3 - Modelo Base com explicador Temperatura.

Subestação base de correlação	Subestação a prever	Conjunto	MAPE Inter	MAPE M3
Agregado	Hracleio2	1	5.3%	3.5%
		2	6.2%	4.1%
	Hracleio3	1	4.8%	4.7%
		2	6.7%	5.2%
	Rethymno	1	4.7%	4.3%
		2	4.9%	4.6%
Hracleio2	Hracleio3	1	14.1%	4.7%
		2	4.9%	5.2%
	Rethymno	1	17%	4.3%
		2	6.8%	4.6%
Hracleio3	Hracleio2	1	11.2%	3.5%
		2	5.1%	4.1%
	Rethymno	1	12.4%	4.3%
		2	5.8%	4.6%
Rethymno	Hracleio2	1	10%	3.5%
		2	4.9%	4.1%
	Hracleio3	1	10.1%	4.7%
		2	4.2%	5.2%

Pela tabela 5.4 percebe-se, na maioria dos casos e para as mesmas condições, o valor do erro MAPE é ligeiramente superior neste modelo de previsão com interdependência geográfica.

Capítulo 6

Conclusões e Trabalhos Futuros

Neste capítulo, pretende-se perceber o cumprimento dos objetivos propostos e, com base no trabalho elaborado e nos conhecimentos adquiridos, são apresentadas propostas de eventuais trabalhos futuros que permitam a continuação do estudo realizado nesta dissertação.

6.1 Satisfação dos Objectivos

Os objetivos desta dissertação passavam pelo tratamento e análise dos dados de diferentes variáveis a inserir na previsão dos consumos na ilha de Creta, e construir modelos de previsão, recorrendo a ferramentas de inteligência computacional.

Feito o estudo das variáveis que se revelariam pertinentes na previsão, foram criados 5 modelos para diferentes períodos com o objetivo de se perceber o seu comportamento com dados referentes a diferentes anos.

O trabalho foi moroso, pois sempre que se pretendia inserir uma nova variável era necessário recorrer ao processo iterativo para obtenção dos parâmetros ótimos das variáveis. Contudo, as conclusões obtidas ajudarão, seguramente, na melhoria das previsões que auxiliarão em eventuais trabalhos feitos a partir desta dissertação.

Se a introdução das variáveis Hora do Dia e Dia da Semana (*Modelo 1 - Modelo Base Simples*) se traduziram num erro MAPE elevadíssimo, particularmente no verão, quando se acrescentou a variável Mês do Ano esse erro desceu em todos os valores (*Modelo 2 - Modelo Base*), o mesmo acontecendo quando, no *Modelo 3 - Modelo com explicador Temperatura*, se recorreu à variável Temperatura. Quanto ao erro da incerteza, verificou-se uma melhoria progressiva, nos três modelos, embora com pouco significado.

Uma vez que os dados a que se teve acesso, referentes ao ano de 2013, não indicavam a contribuição da produção fotovoltaica nos valores de cada subestação, introduziu-se a variável Irradiância (*Modelo 4 - Modelo com explicador Temperatura e Irradiância*), mas o erro MAPE aumentou, tal como a Incerteza, como era previsível.

“Uma dificuldade é uma oportunidade”, pelo que se deverá reforçar a constatação de que a introdução da variável Consumo do Dia Anterior (*Modelo 5 - Modelo com explicador Temperatura*

e *Consumo do Dia Anterior* -) se revelou frutífera, aumentando consideravelmente a precisão do modelo, para que a utilização destes dados sonogados nas habituais previsões de consumo energético seja possível. Se a utilização dos dados relacionados com a Hora do Dia, Dia da Semana, Mês do Ano e Temperatura acrescentam valor às previsões, o recurso à variável *Consumo do Dia Anterior* seria “ouro sobre azul”.

Após a criação dos modelos, passou-se a uma análise da dependência entre as cargas, nomeadamente pelo estudo da correlação existente entre as subestações e o *Agregado* ou mesmo entre diferentes subestações. Procedeu-se deste modo para se poder realizar uma previsão com interdependência geográfica, importante quando são necessários dados desagregados e sincronizados de consumo em diversos pontos geográficos da rede, como por exemplo, para utilizar em trânsito de potências.

6.2 Trabalhos Futuros

Neste trabalho, os dias feriado foram tidos em consideração e diferenciados dos restantes, para que os resultados fossem mais precisos. Do mesmo modo, futuramente, dever-se-ão, também, ter em conta dias de acontecimentos importantes, tais como eleições ou eventos desportivos, visto provocarem na população comportamentos diferentes dos que têm no quotidiano, havendo por isso variações nos padrões de carga diários.

No que diz respeito às variáveis a introduzir nos modelos, seria interessante, ainda, para além da variável temperatura utilizada, considerar a utilização duma variável associada à humidade, pois esta sensação térmica pode provocar a necessidade de recorrer a reguladores térmicos, apesar de os valores reais de temperatura não justificarem essa utilização.

O processo iterativo de otimização dos parâmetros das variáveis é, como afirmado anteriormente, um trabalho bastante moroso embora possa não garantir que sejam encontrados os melhores valores para o modelo em estudo. É, por isso, necessária a construção de uma ferramenta computacional que detete estes parâmetros de forma automática.

Para efeitos de melhor conhecimento da correlação entre os diferentes pontos de carga, além de uma previsão mais precisa por modelos de interdependência geográfica, é fundamental a obtenção de dados mais precisos dos consumos, nomeadamente para o estudo da correlação com o *Agregado* desses pontos e para que se possa efetuar a modelização da incerteza, tópico que não se conseguiu concretizar.

Para além deste possíveis trabalhos futuros, poder-se-ão desenvolver ainda geradores de cenários sincronizados de consumos previstos, assim como ser feita uma assimilação de dados do SCADA para ajuste dinâmico.

Anexo A

Caso de Estudo 1

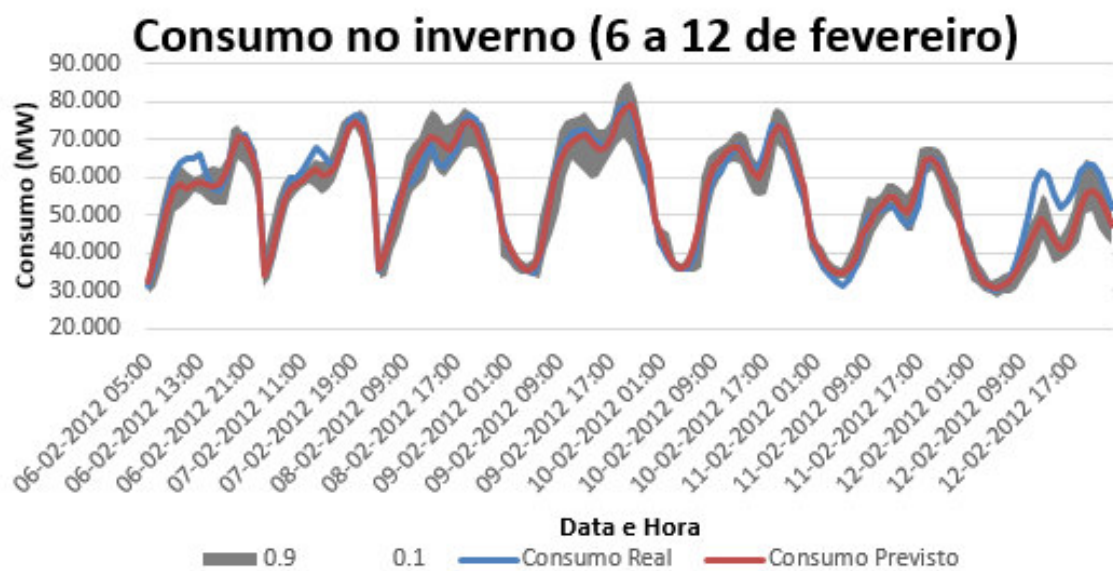


Figura A.1: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Hrakleio2.

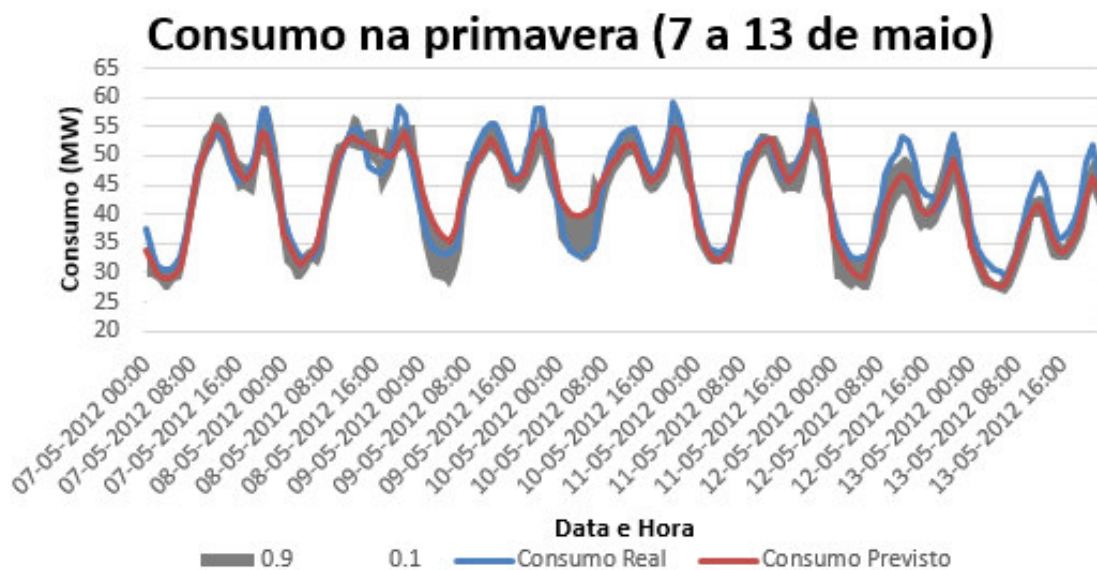


Figura A.2: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Hrakleio2.

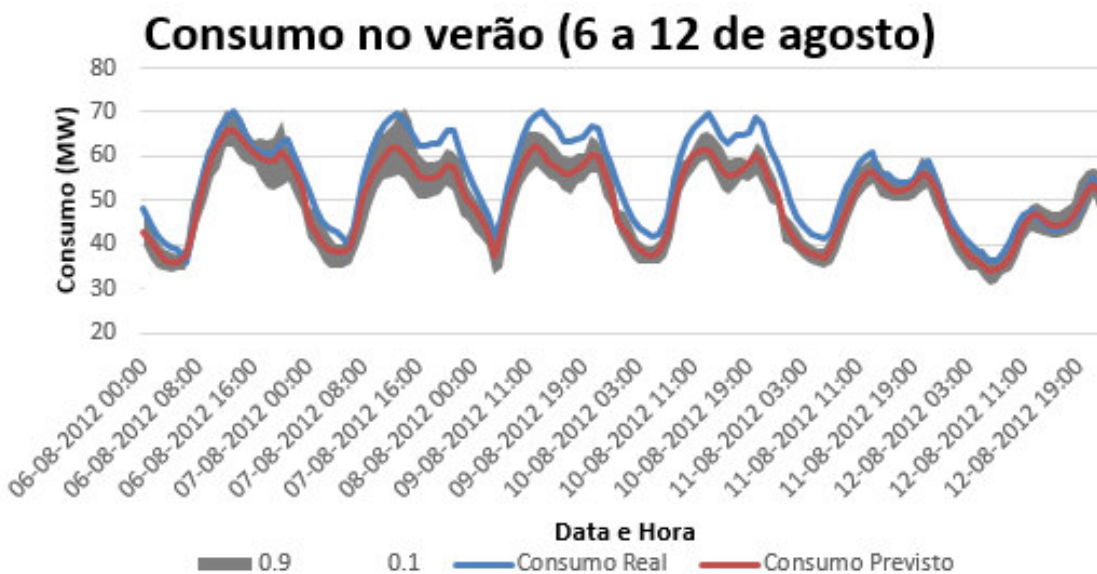


Figura A.3: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Hrakleio2.

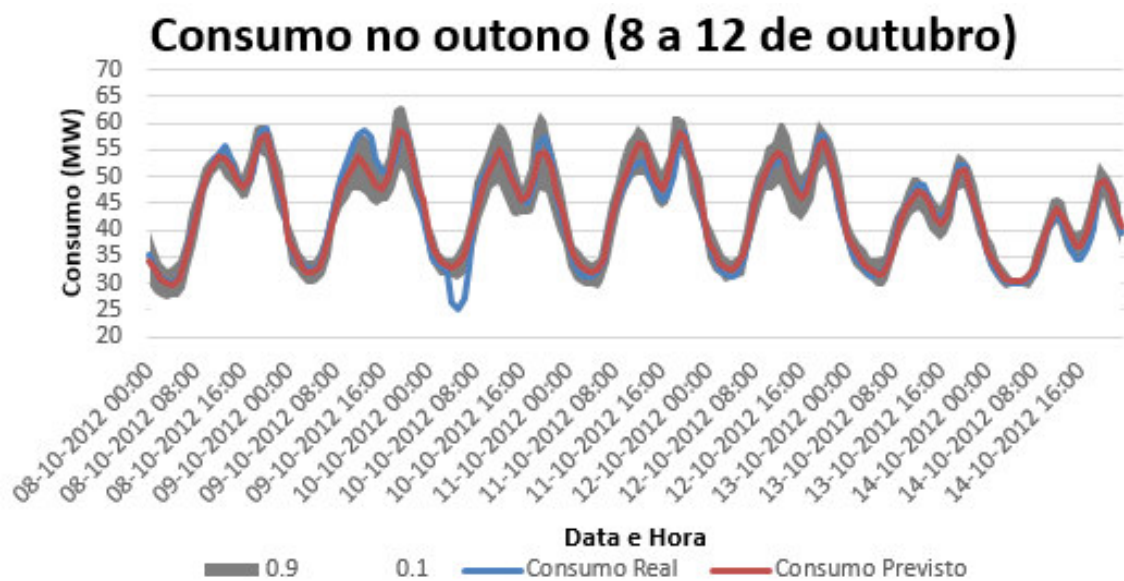


Figura A.4: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Hrakleio2.

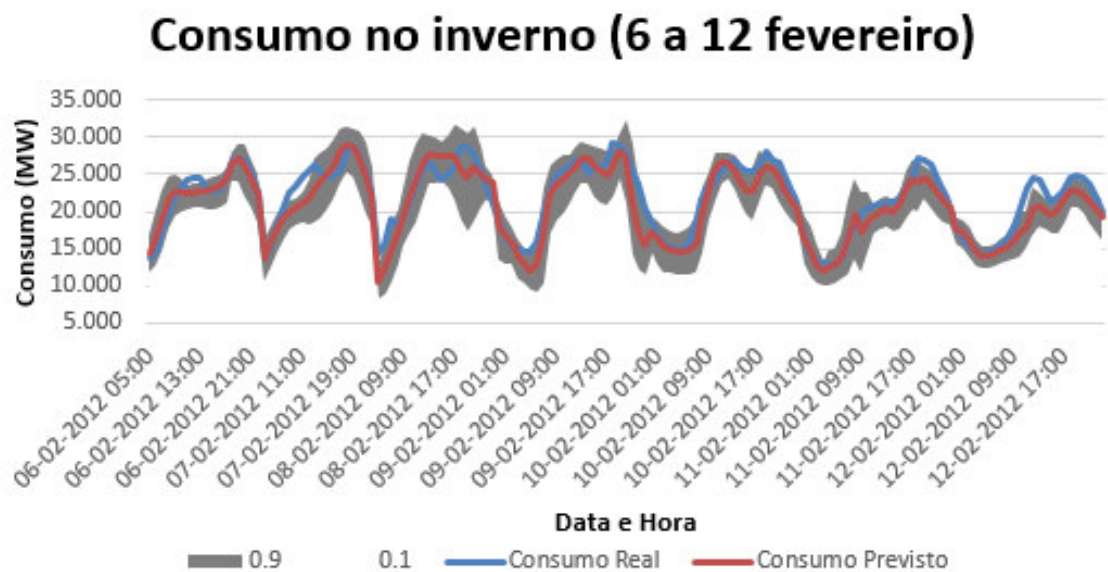


Figura A.5: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Hrakleio3.

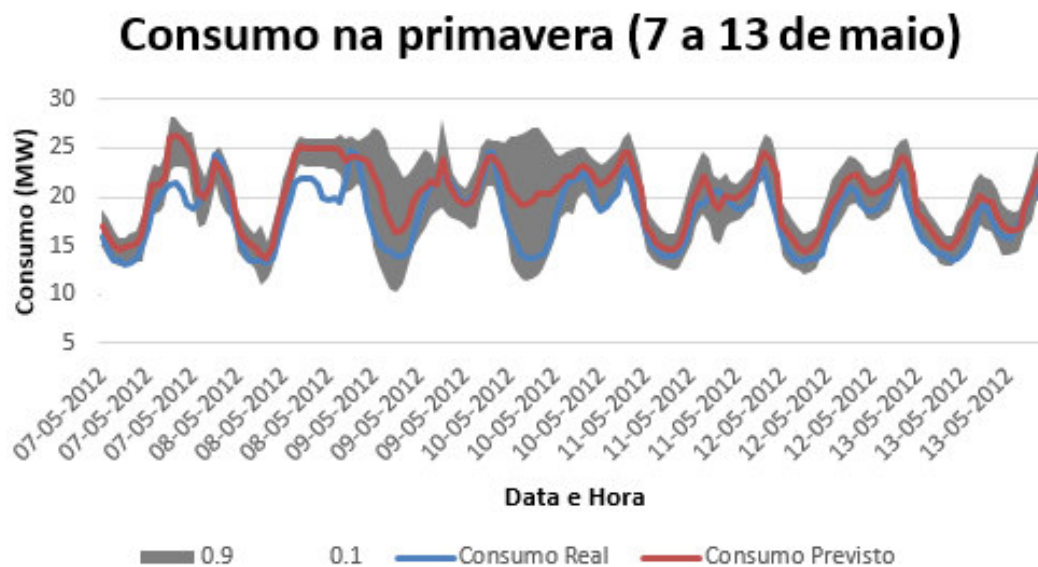


Figura A.6: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Hrakleio3.

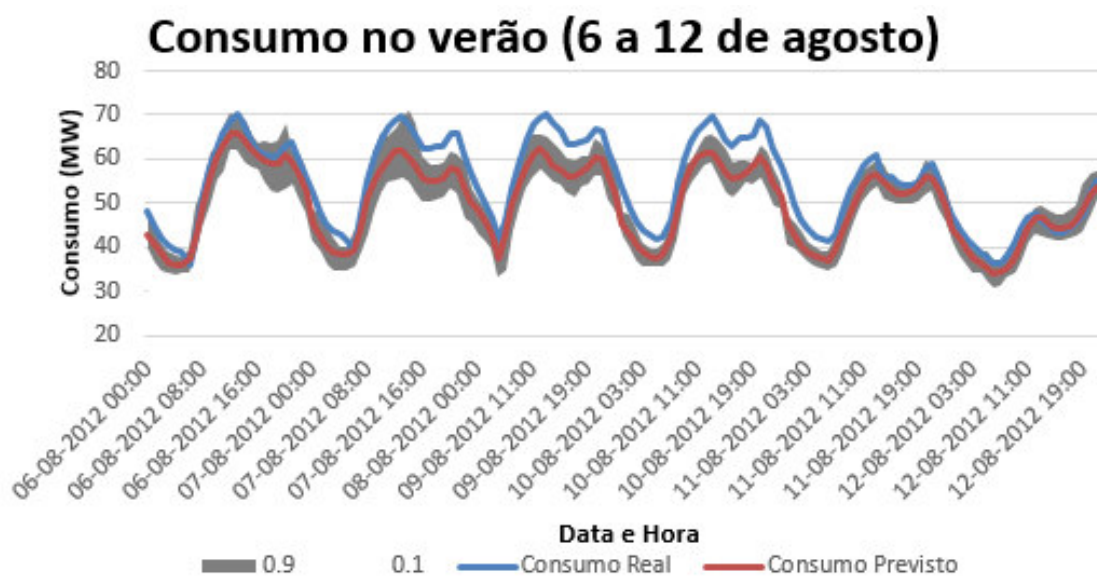


Figura A.7: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Hrakleio3.

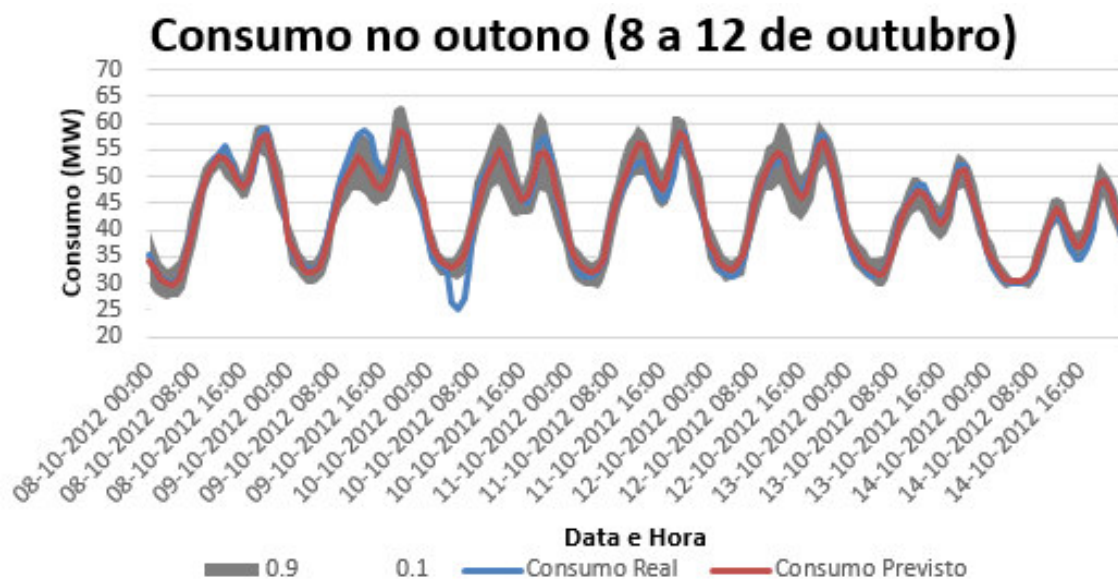


Figura A.8: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Hrakleio3.

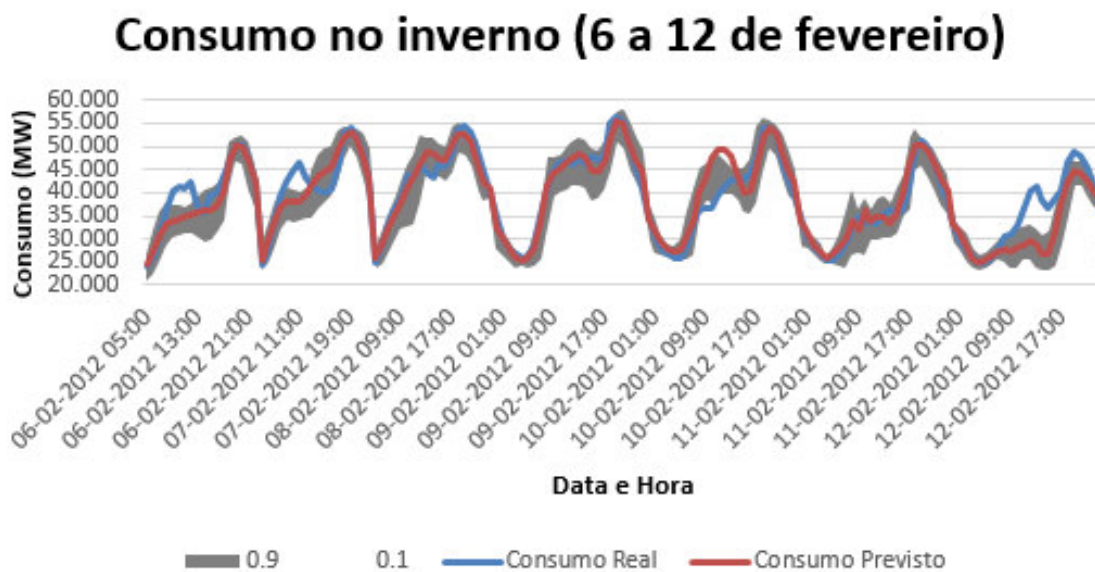


Figura A.9: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Rethymno.

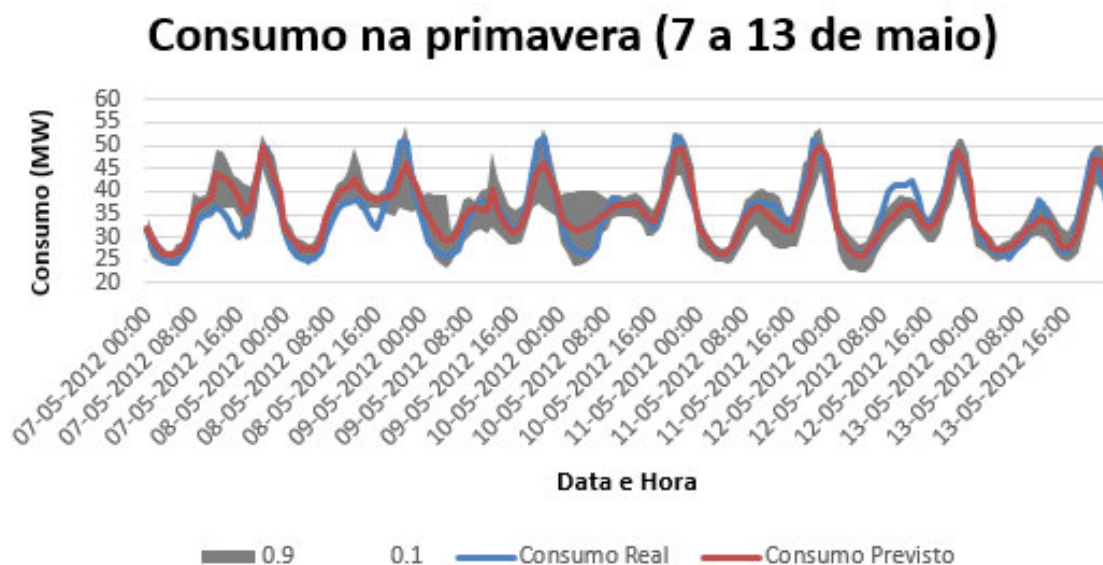


Figura A.10: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Rethymno.

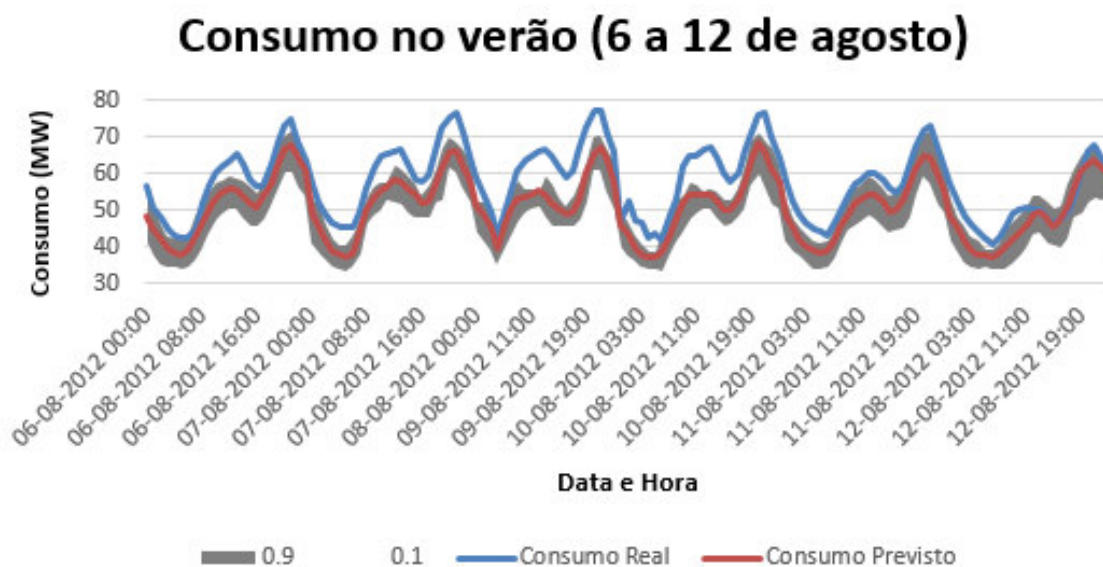


Figura A.11: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Rethymno.

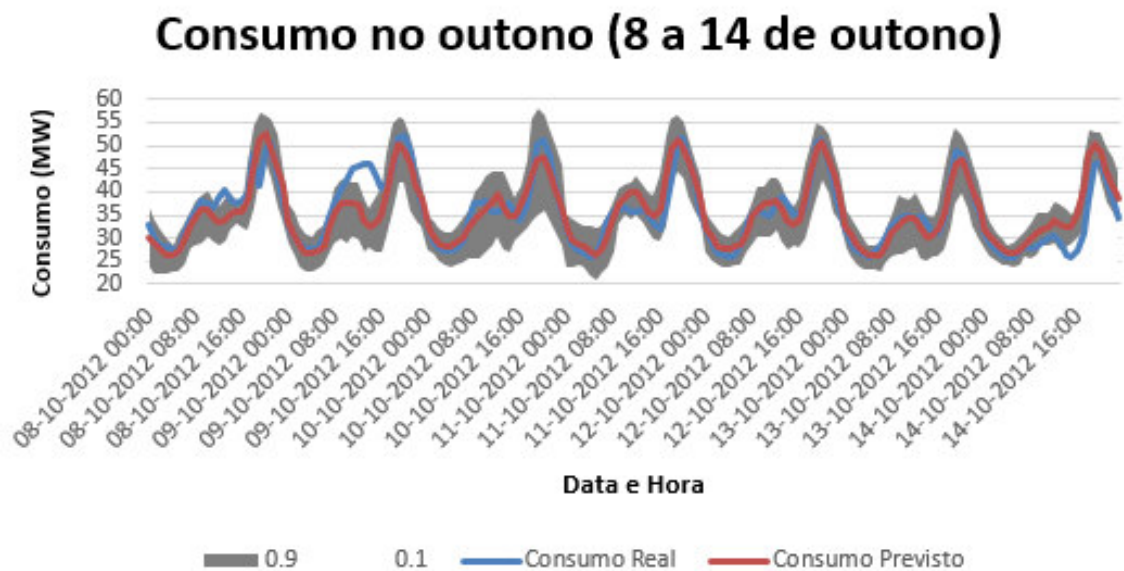


Figura A.12: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Rethymno.

Anexo B

Caso de Estudo 2

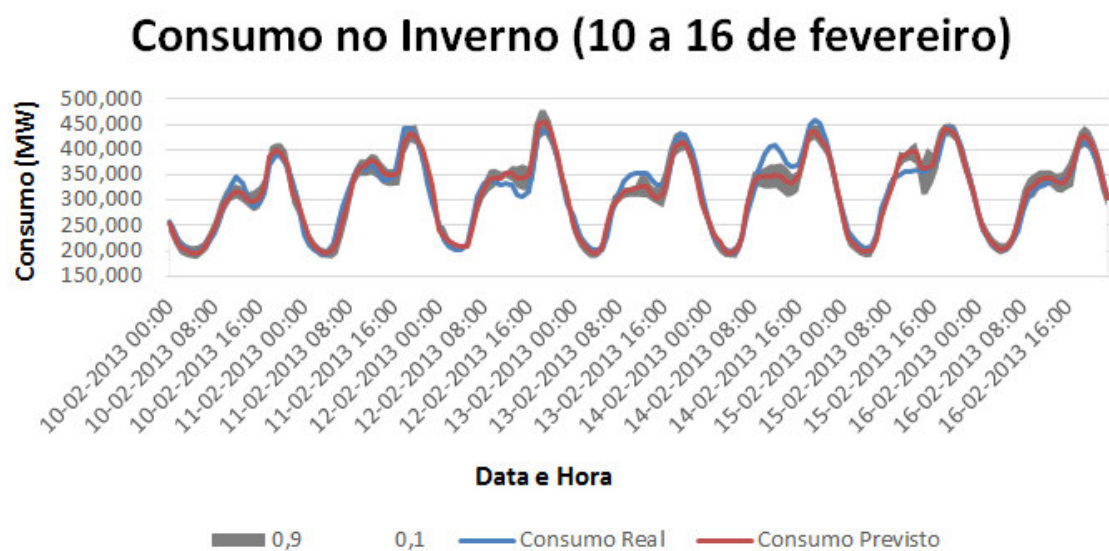


Figura B.1: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para o Agregado.

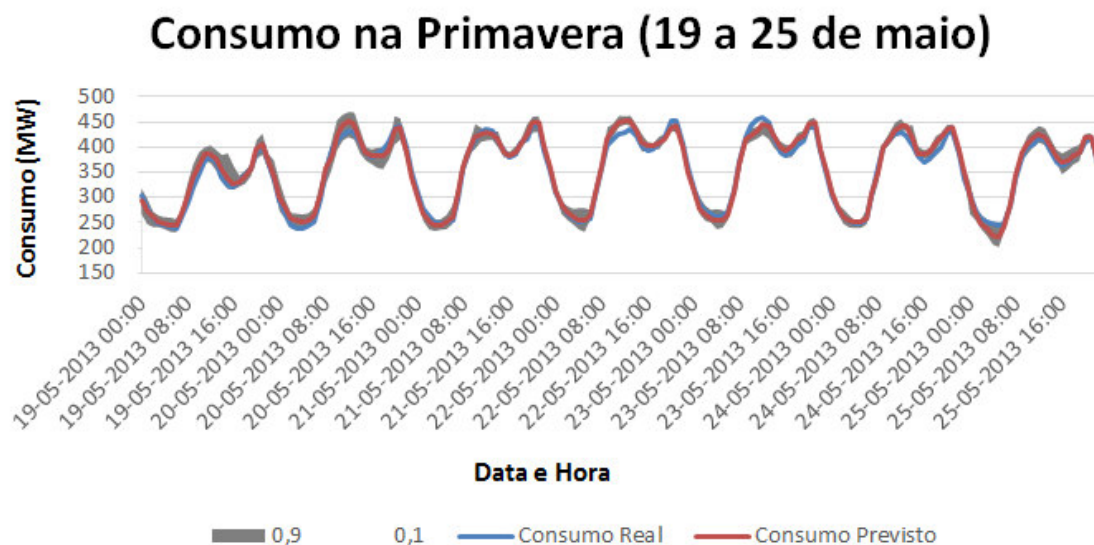


Figura B.2: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para o Agregado.

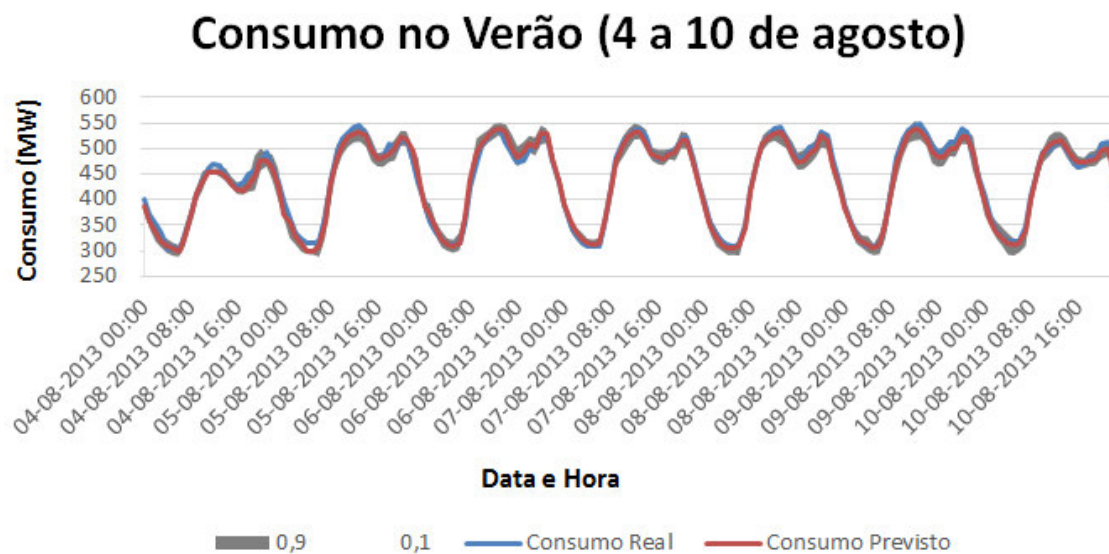


Figura B.3: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para o Agregado.

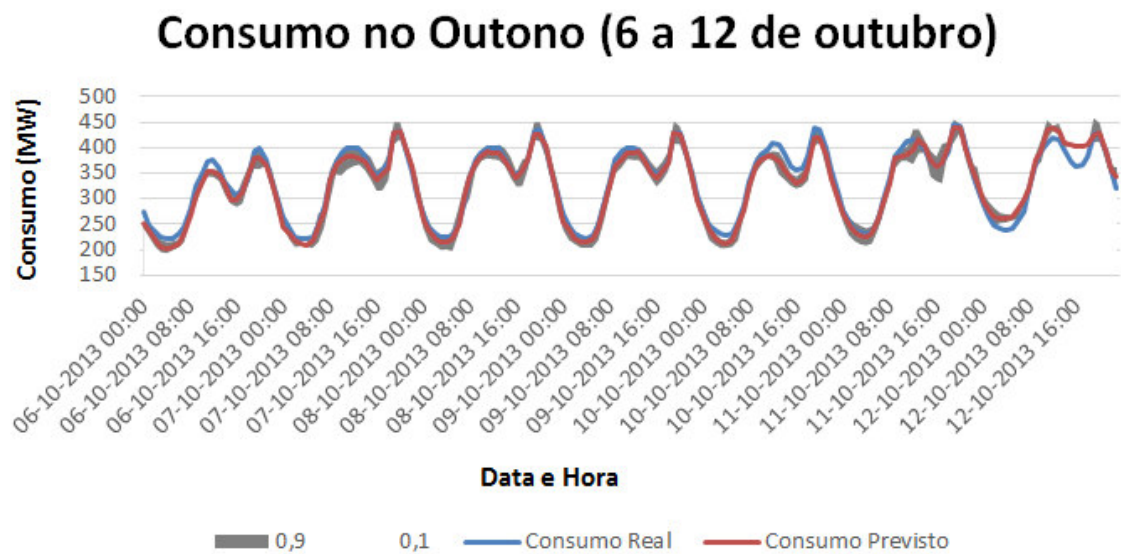


Figura B.4: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para o Agregado.

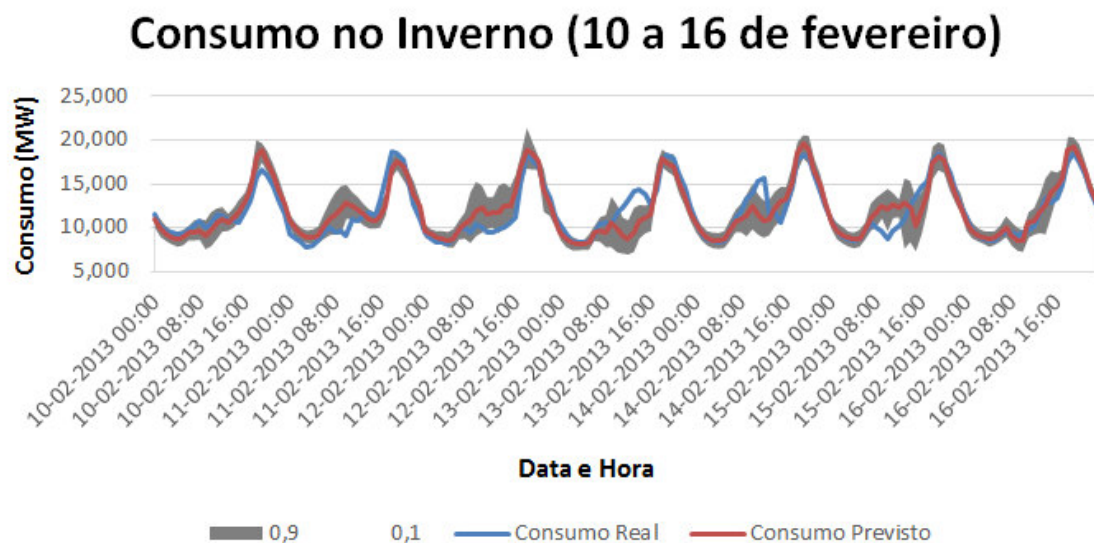


Figura B.5: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Ag.Nikolaos.

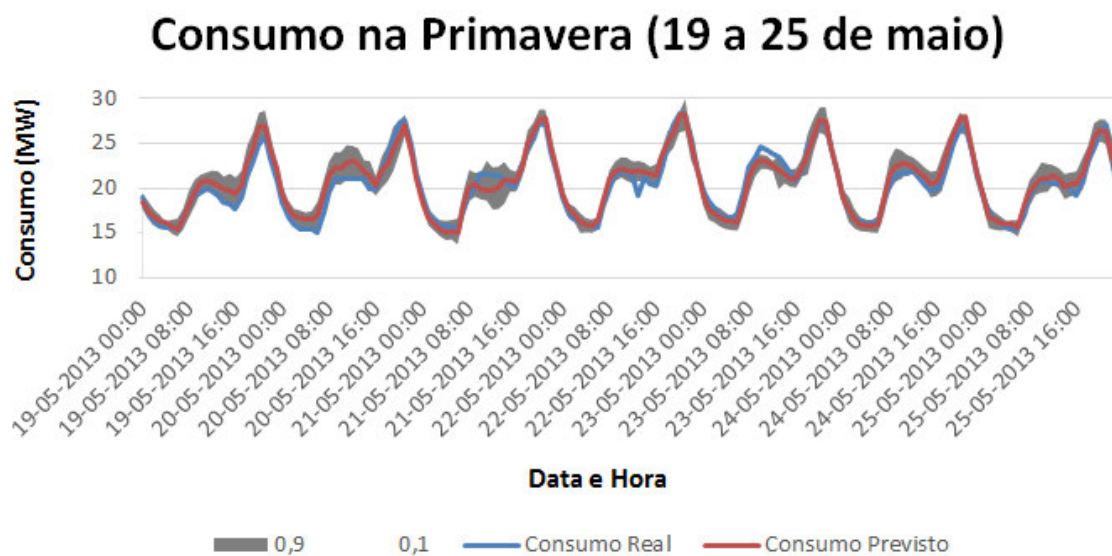


Figura B.6: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Ag.Nikolaos.

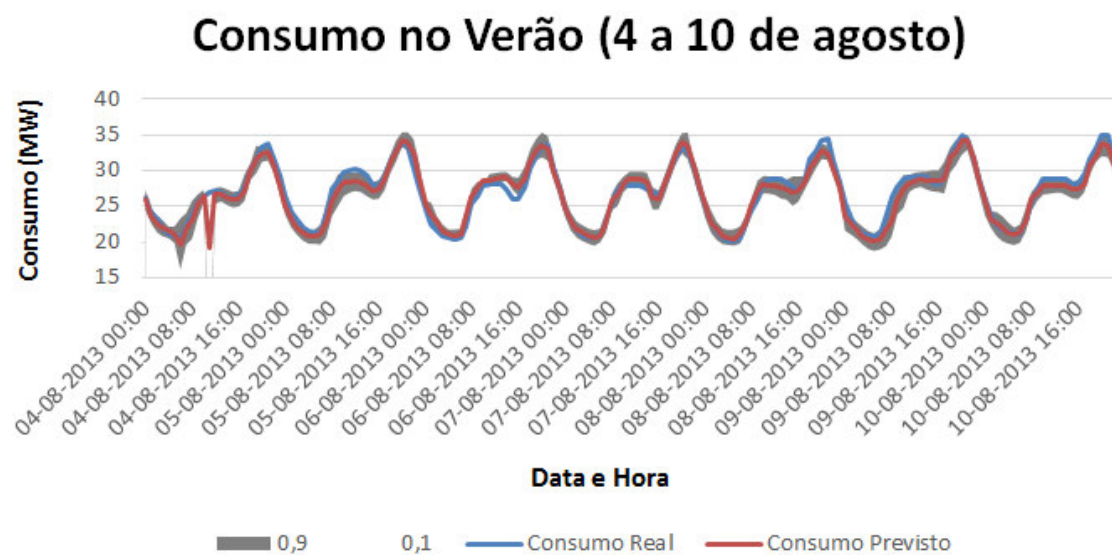


Figura B.7: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Ag.Nikolaos.

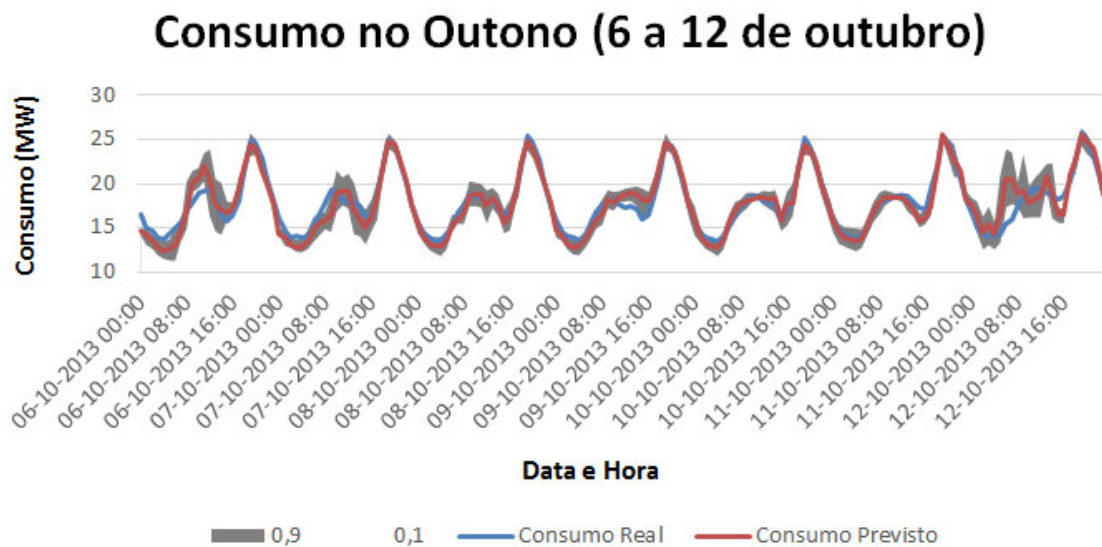


Figura B.8: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Ag.Nikolaos.

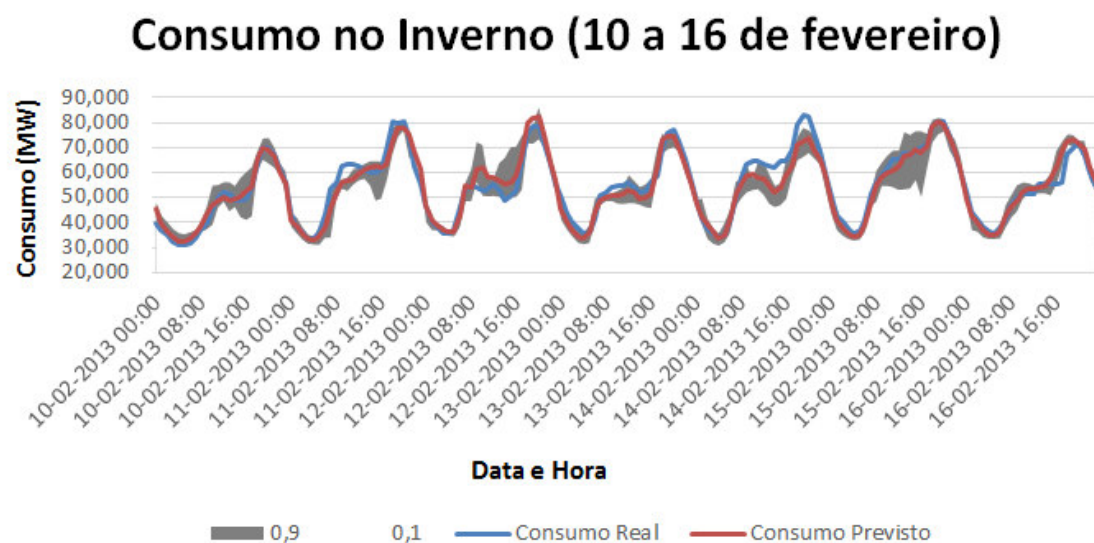


Figura B.9: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Chania.

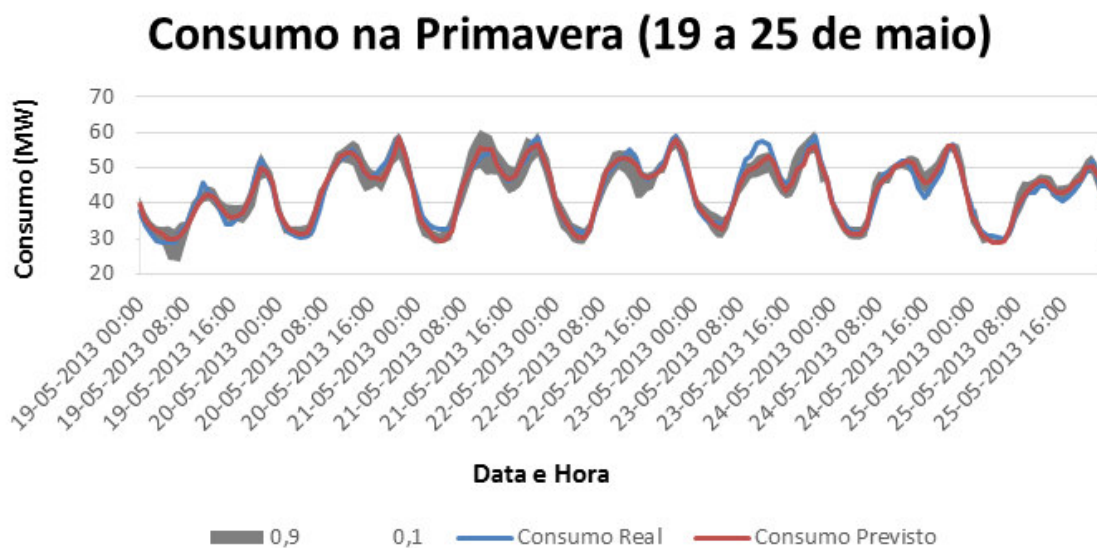


Figura B.10: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Chania.

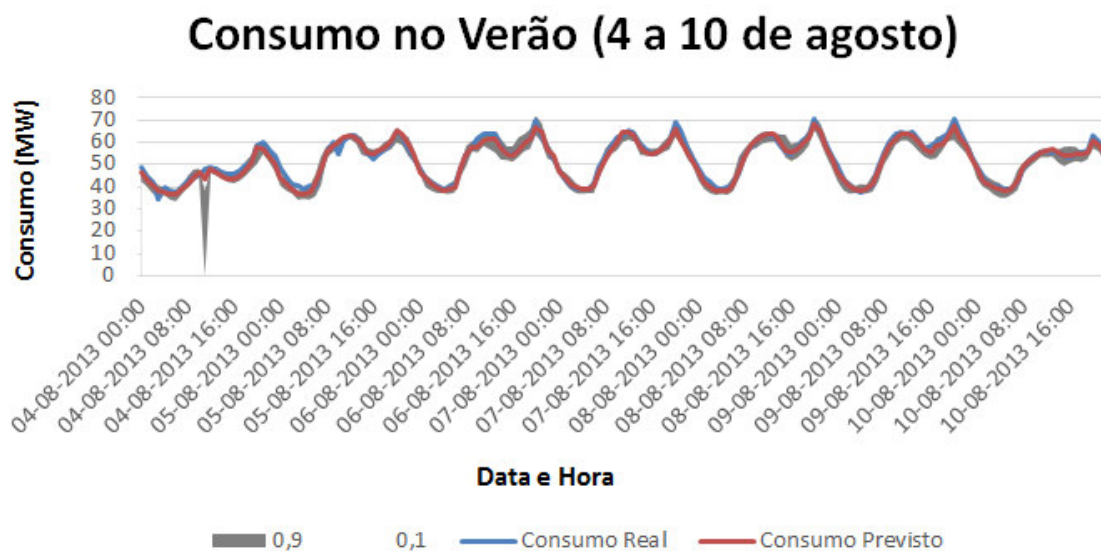


Figura B.11: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Chania.

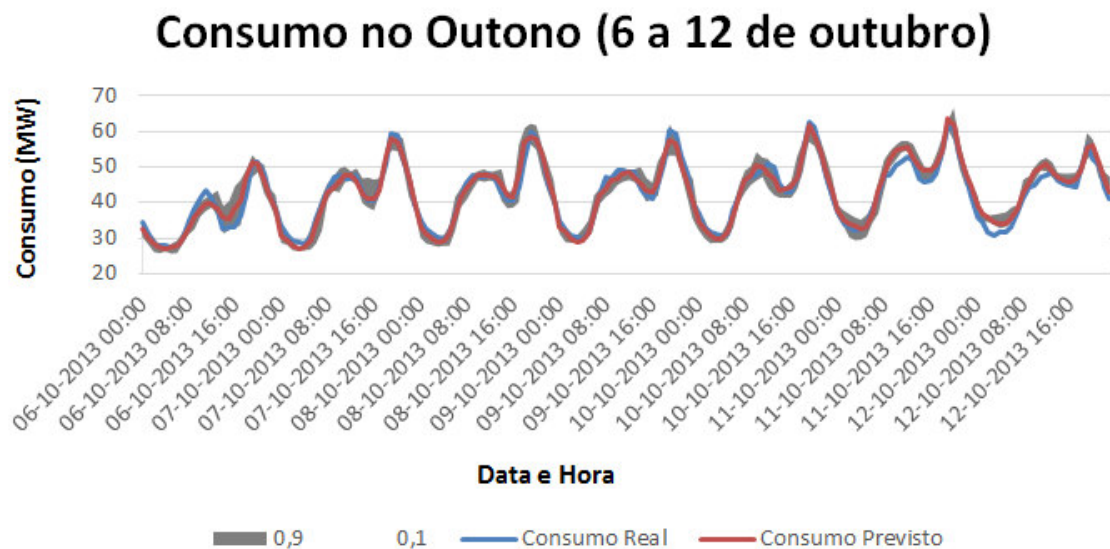


Figura B.12: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Chania.

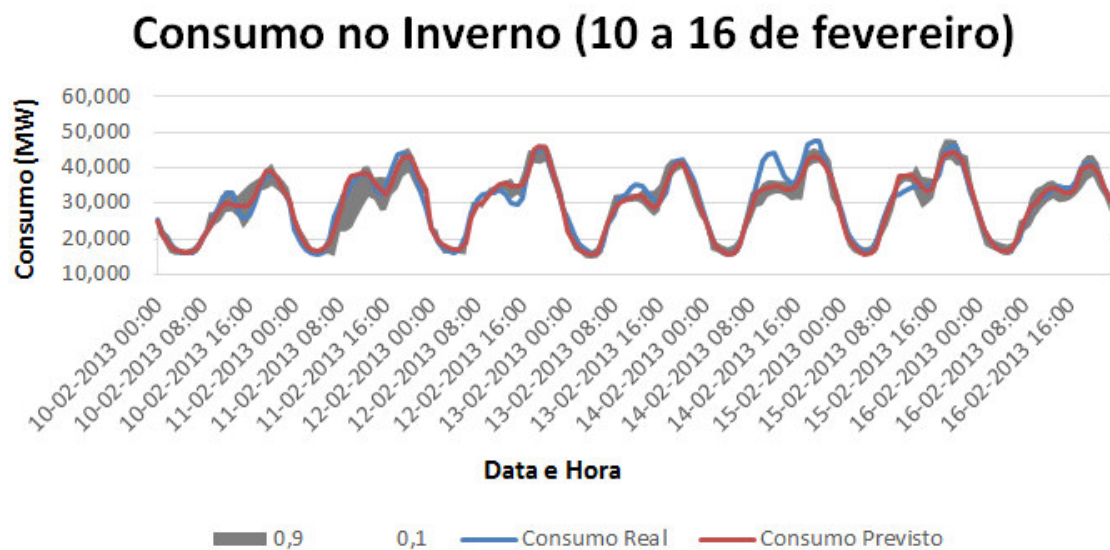


Figura B.13: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Hrakleio1.

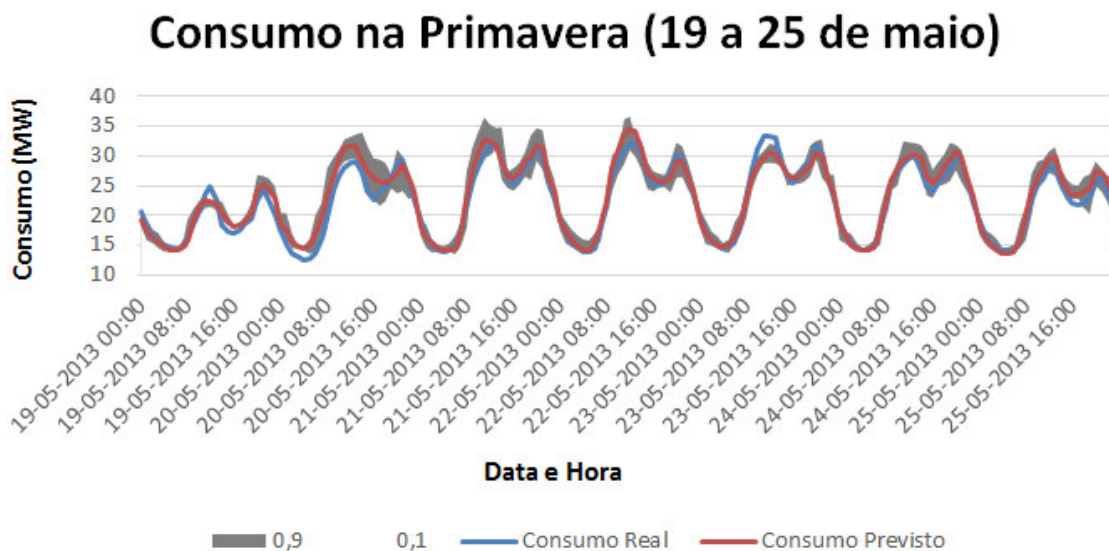


Figura B.14: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Hrakleio1.

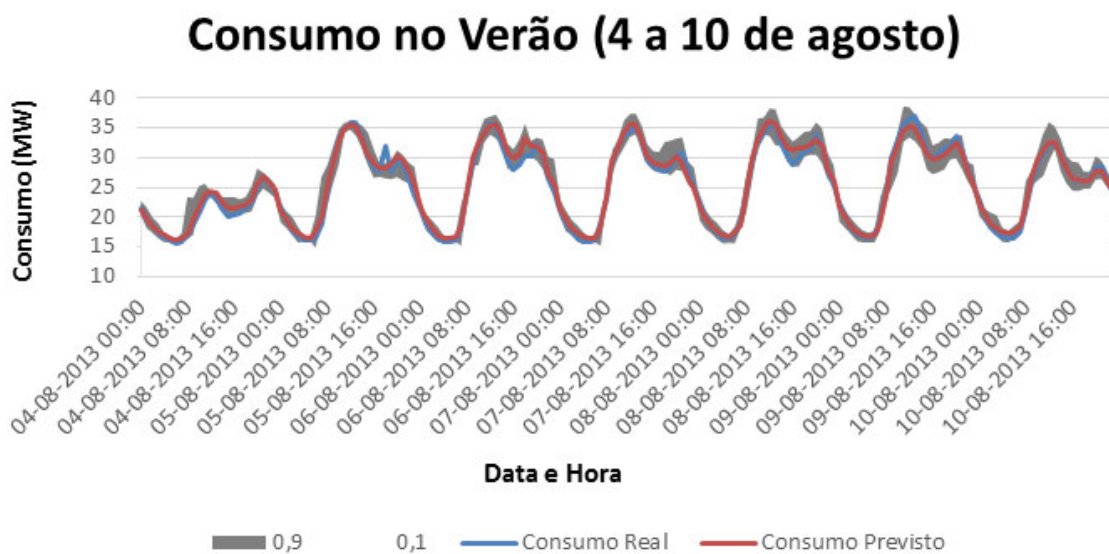


Figura B.15: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Hrakleio1.

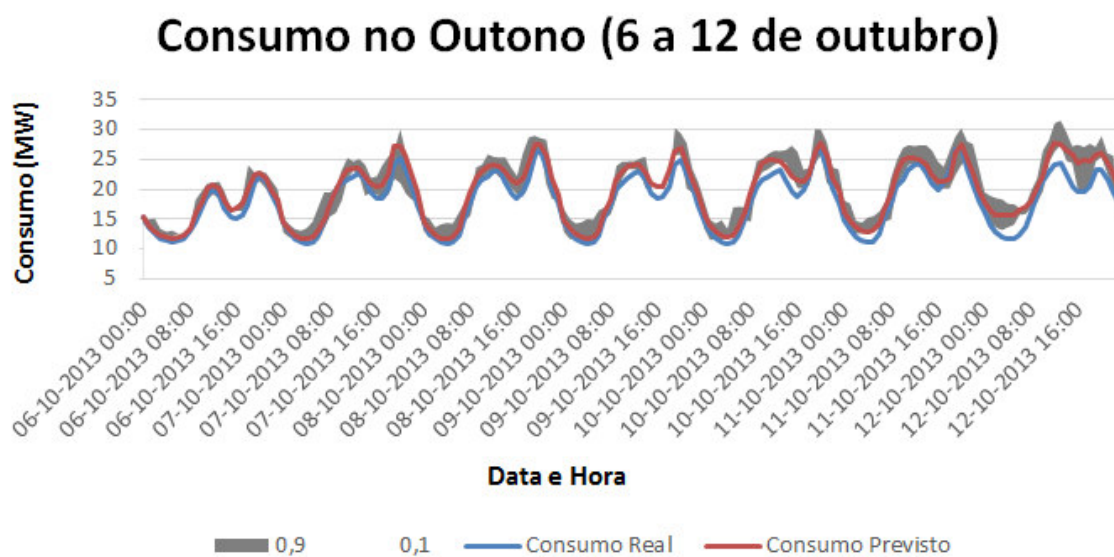


Figura B.16: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Hrakleio1.

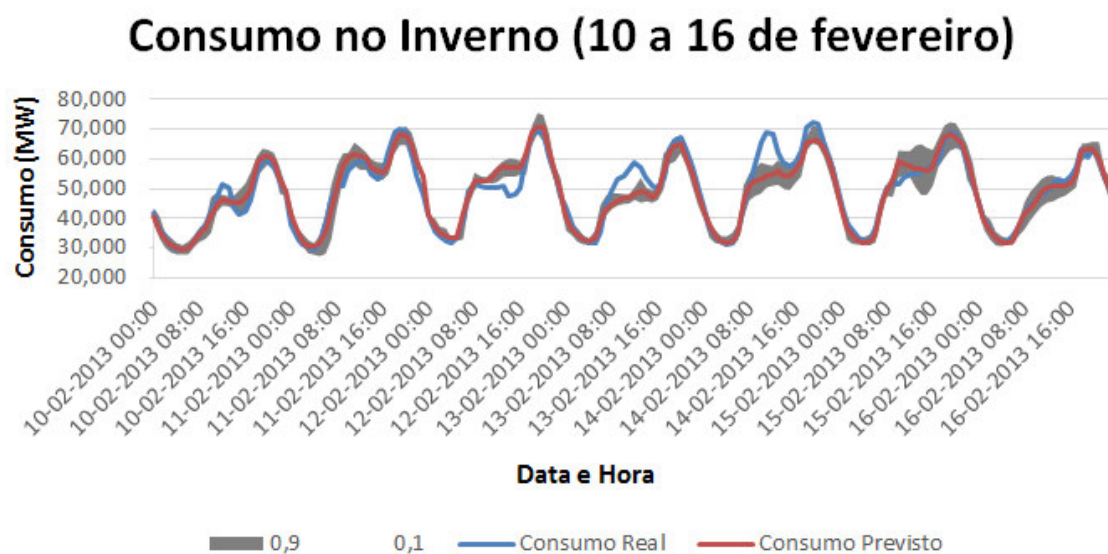


Figura B.17: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Hrakleio2.

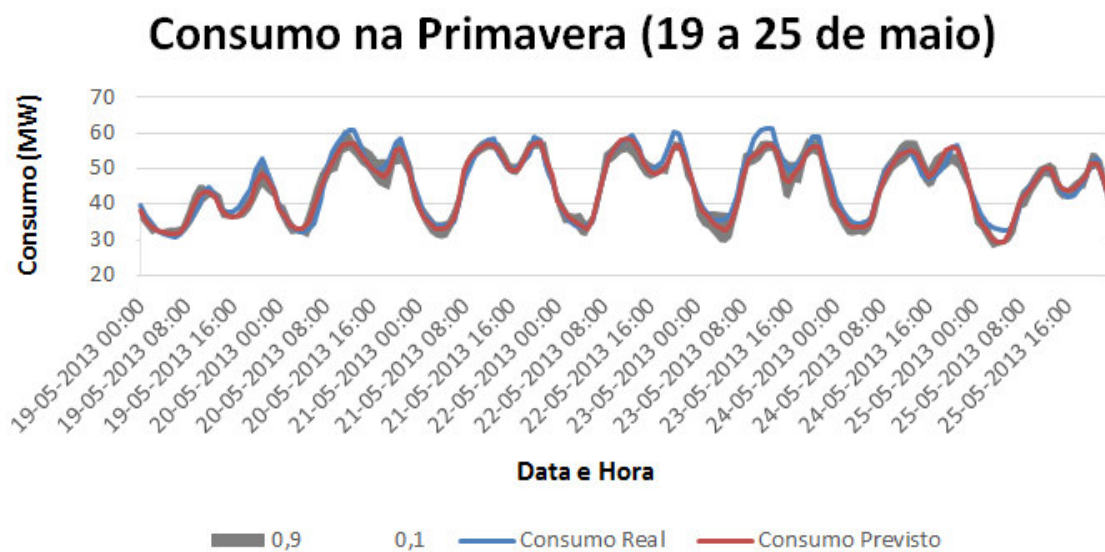


Figura B.18: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Hrakleio2.

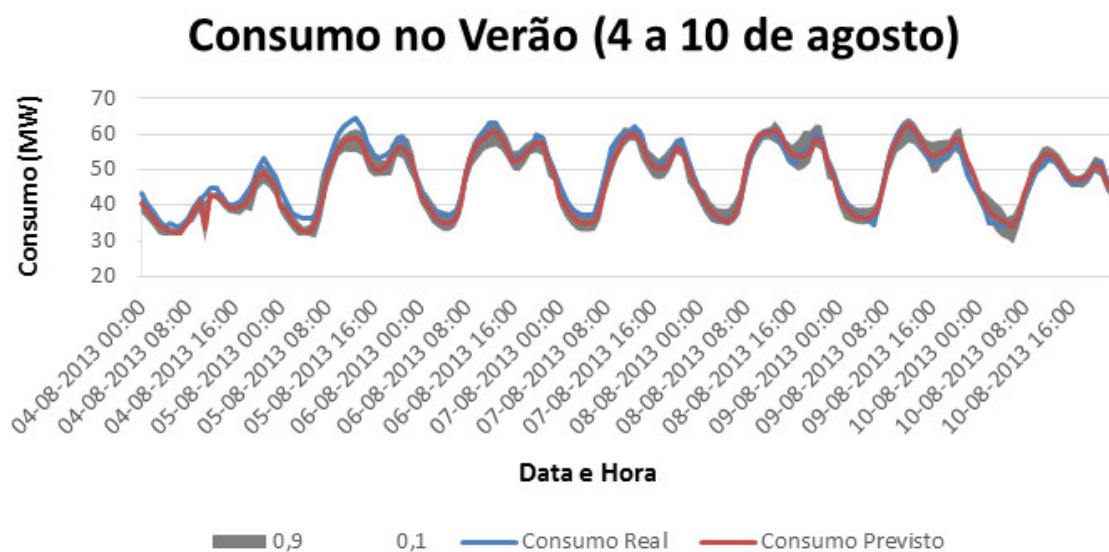


Figura B.19: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Hrakleio2.

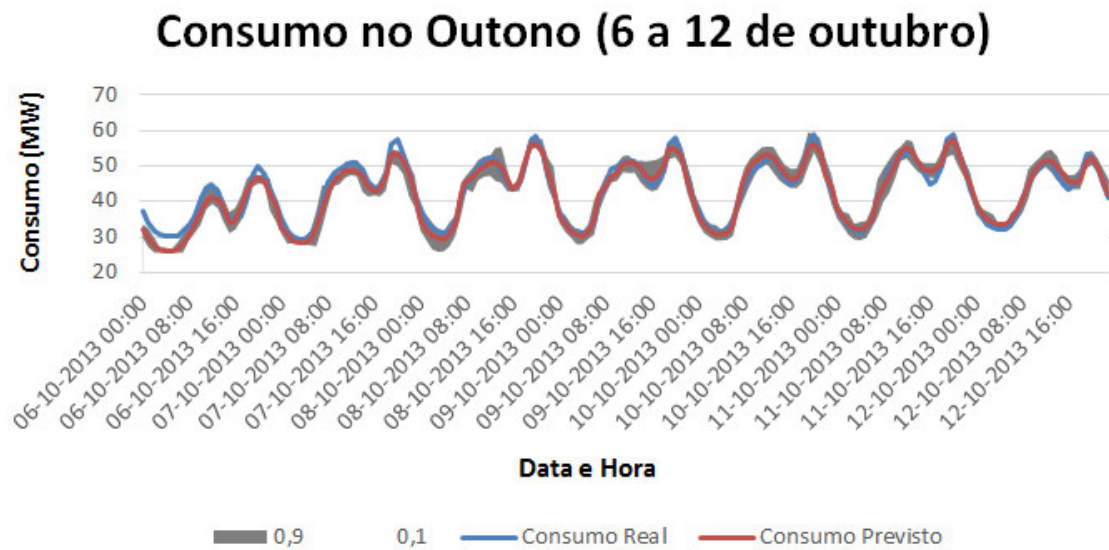


Figura B.20: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Hrakleio2.

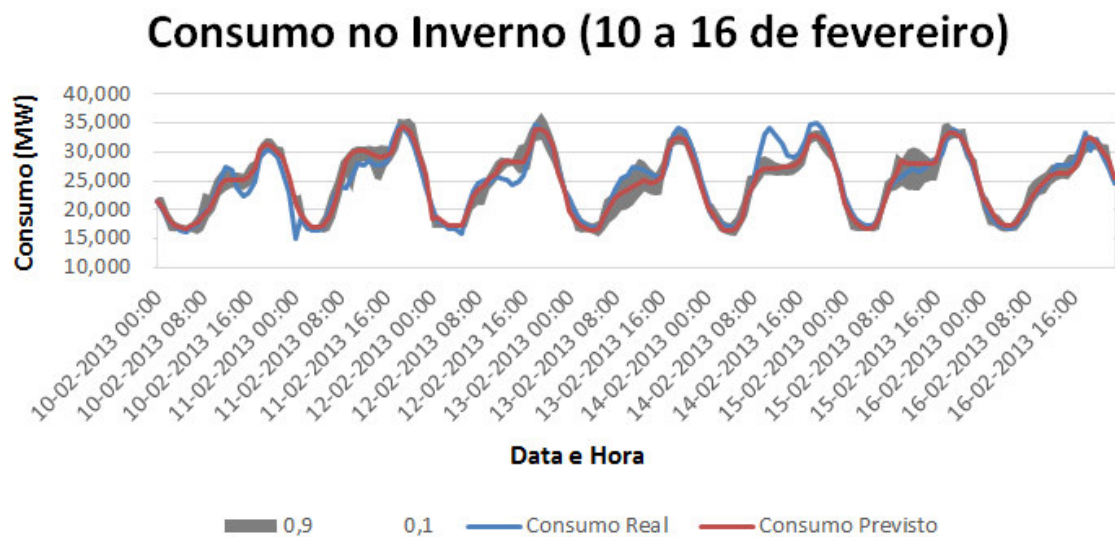


Figura B.21: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Hrakleio3.

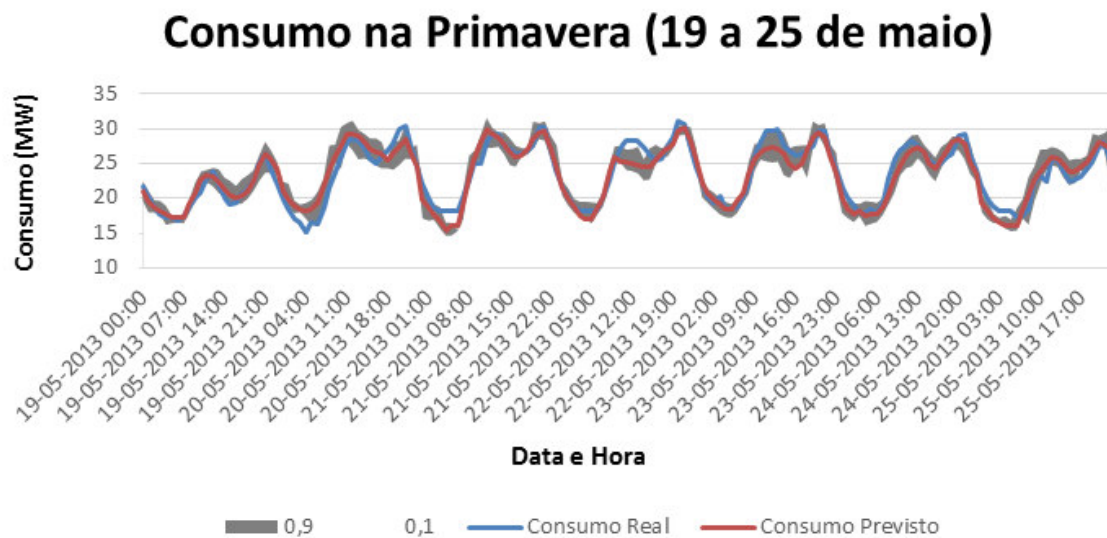


Figura B.22: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Hrakleio3.

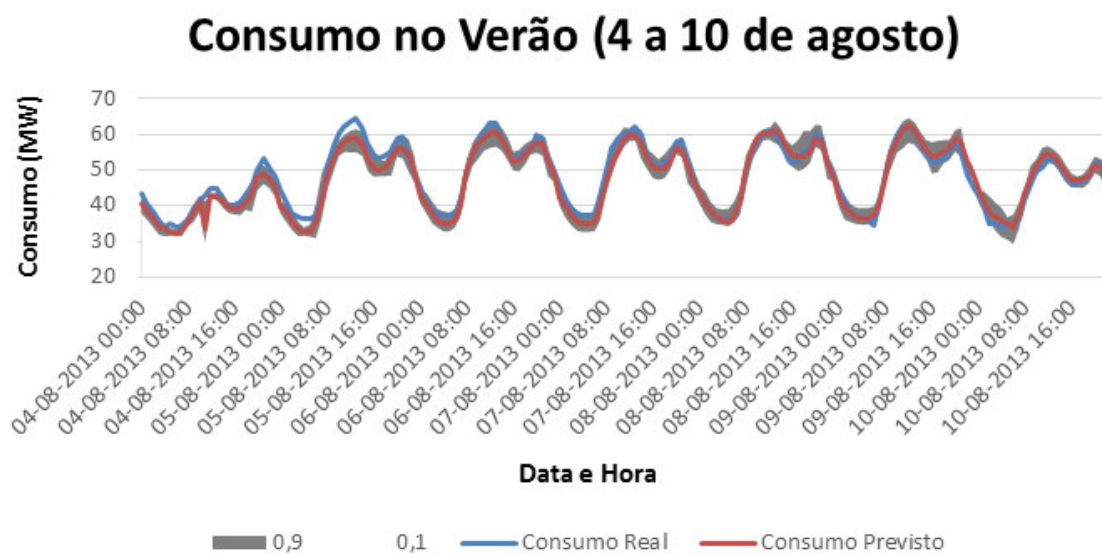


Figura B.23: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Hrakleio3.

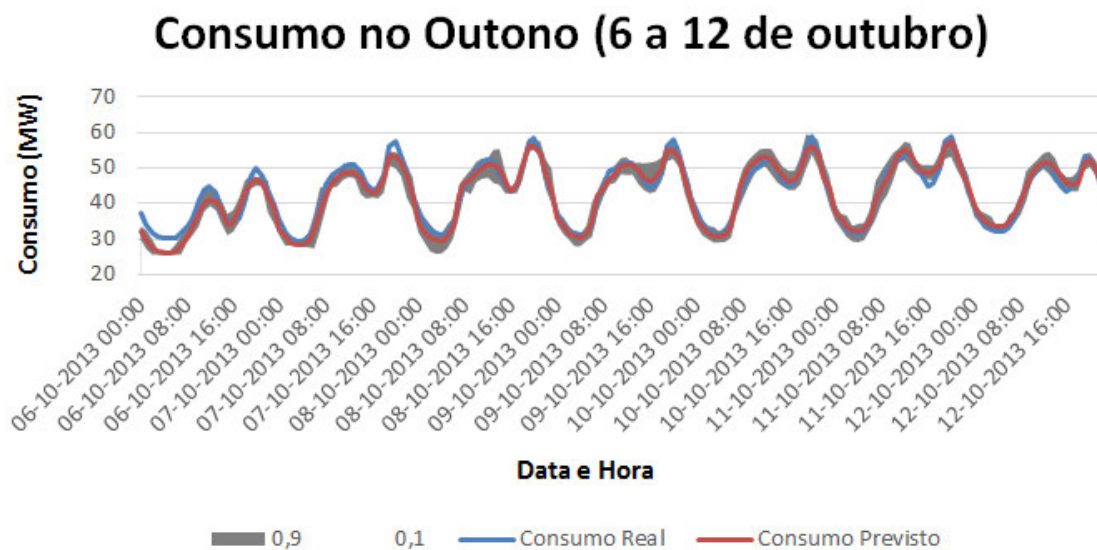


Figura B.24: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Hrakleio3.

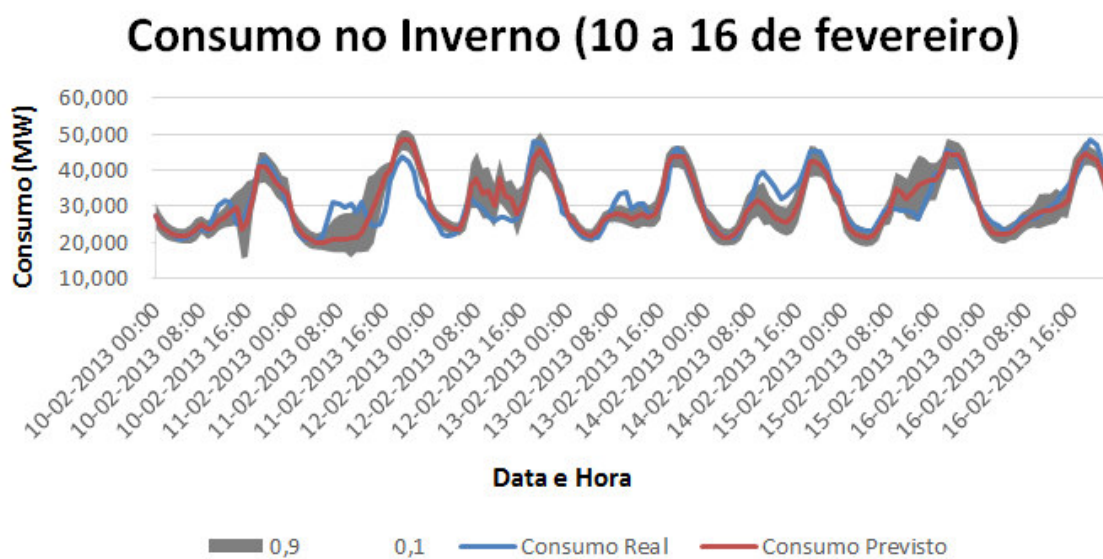


Figura B.25: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Rethymno.

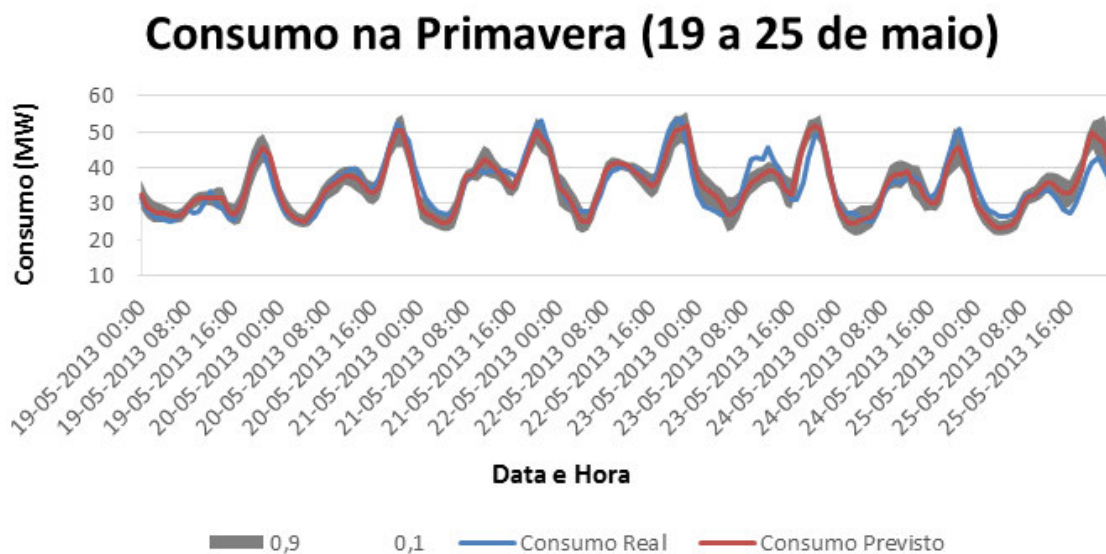


Figura B.26: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Rethymno.

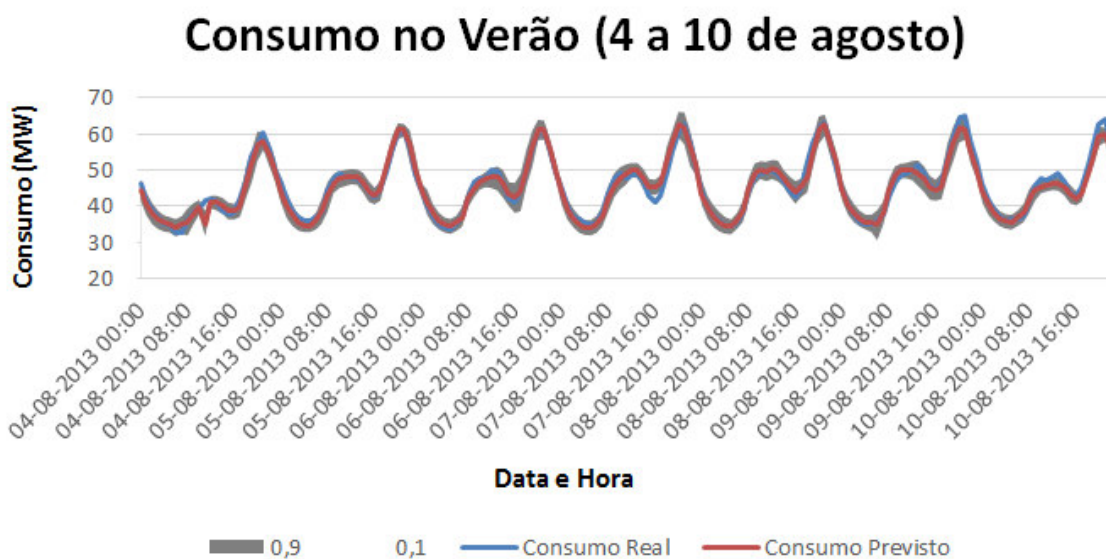


Figura B.27: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Rethymno.

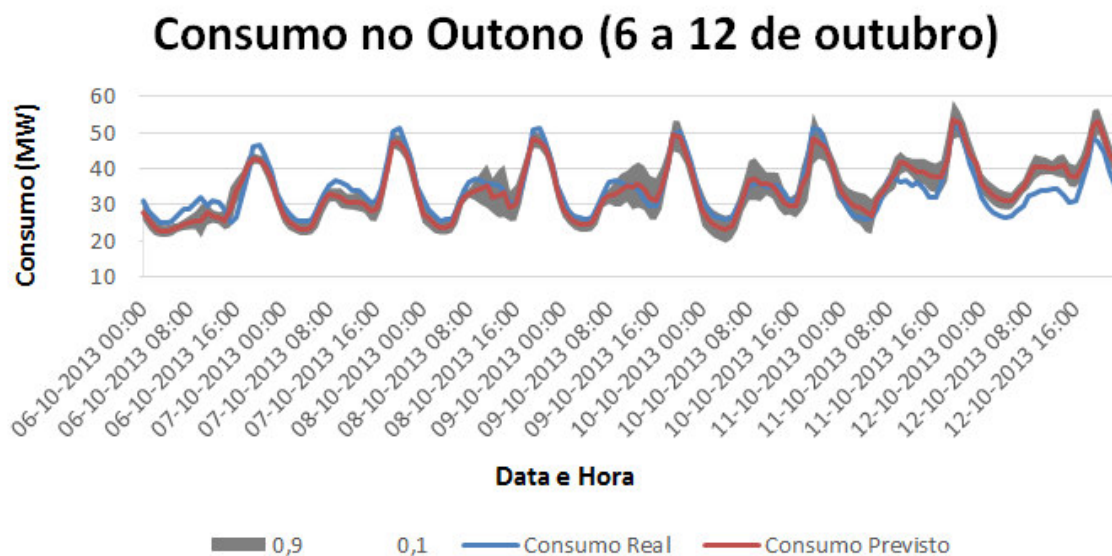


Figura B.28: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Rethymno.

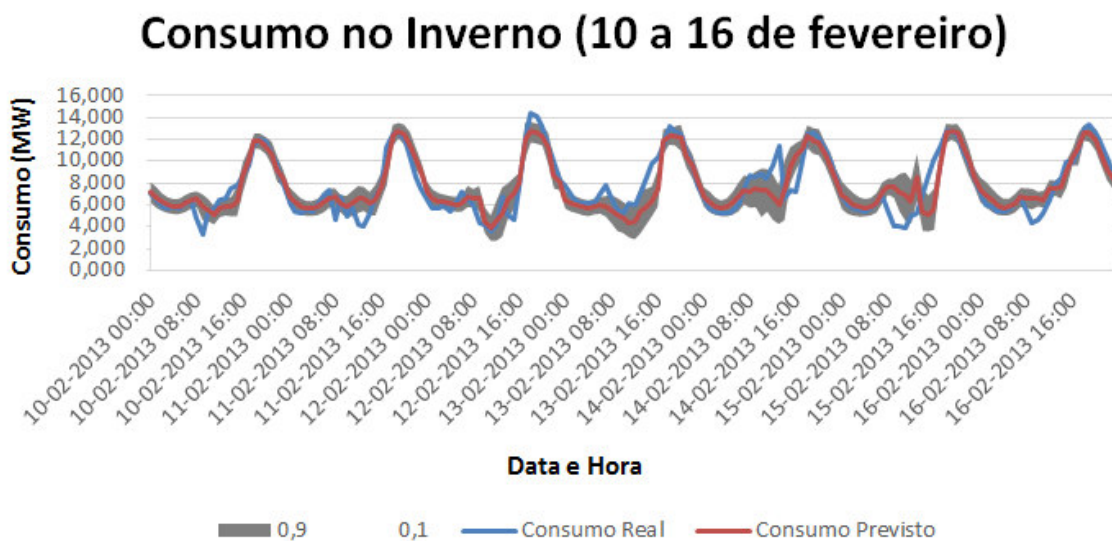


Figura B.29: Relação entre o Consumo Previsto e o Consumo Real, para o período de inverno, para a subestação Sitia.

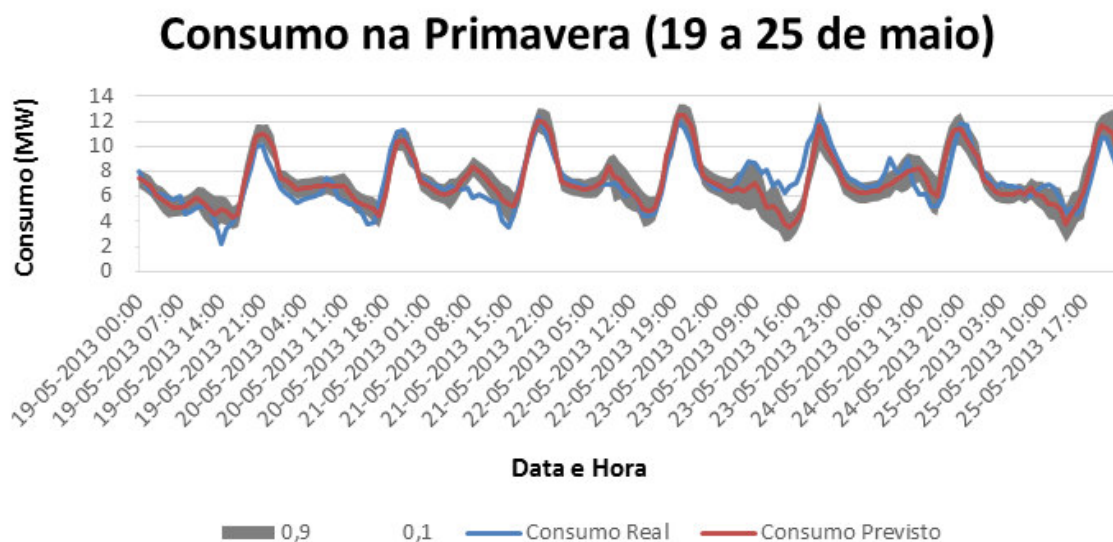


Figura B.30: Relação entre o Consumo Previsto e o Consumo Real, para o período de primavera, para a subestação Sitia.

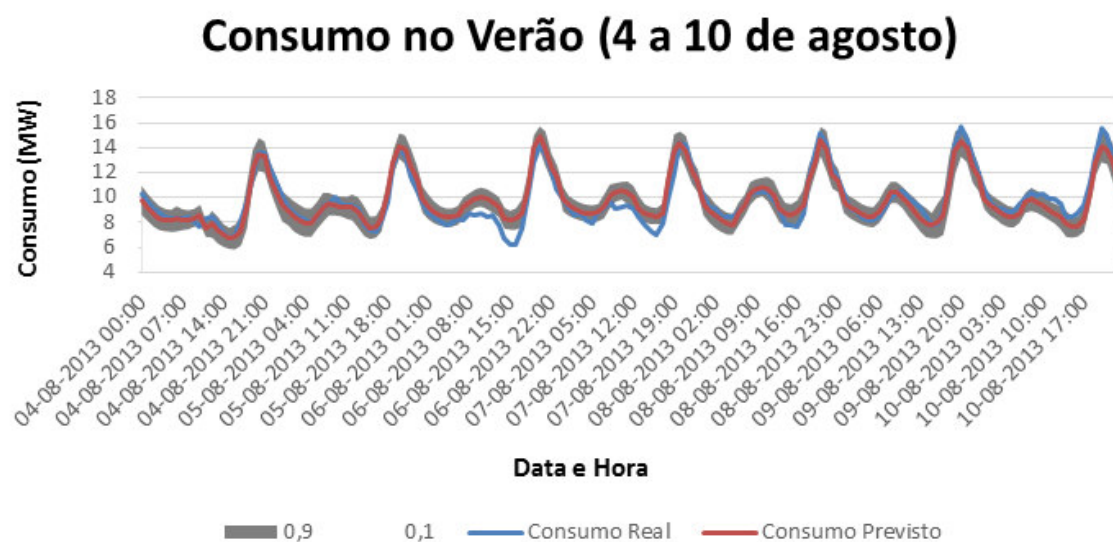


Figura B.31: Relação entre o Consumo Previsto e o Consumo Real, para o período de verão, para a subestação Sitia.

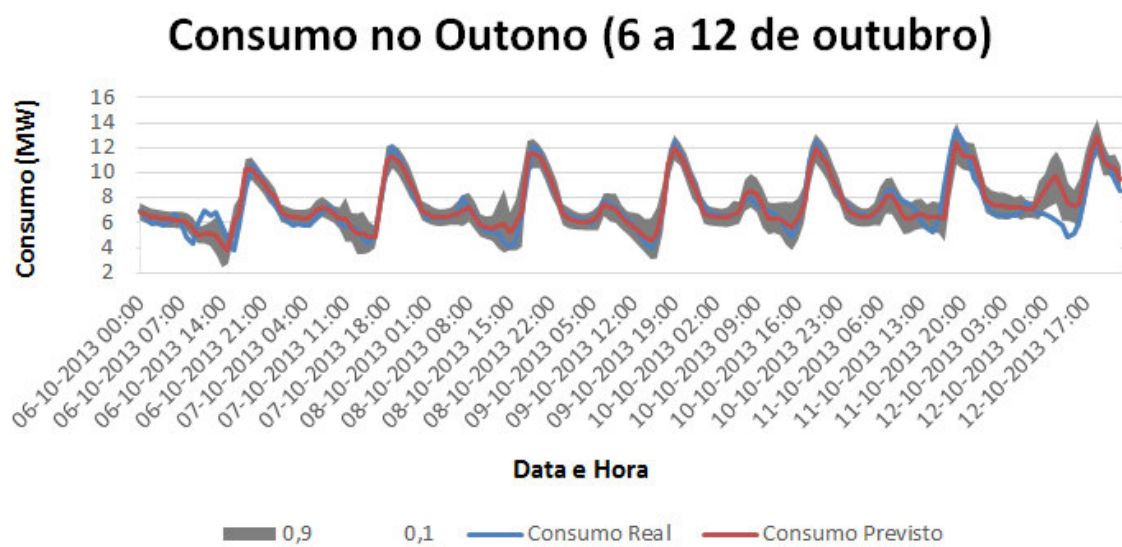


Figura B.32: Relação entre o Consumo Previsto e o Consumo Real, para o período de outono, para a subestação Sitia.

Anexo C

Previsão com interdependência geográfica

C.1 Correlação

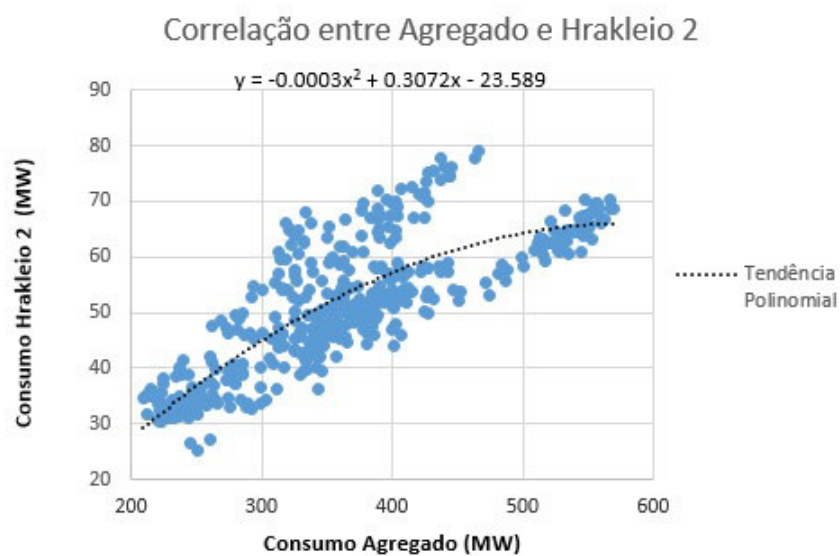


Figura C.1: Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Hrakleio2, para o conjunto 1 (dias úteis).

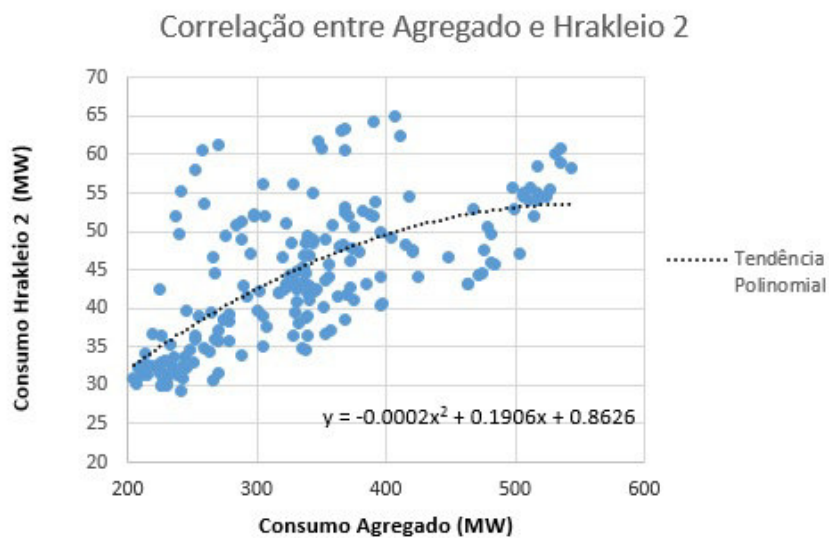


Figura C.2: Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Hrakleio2, para o conjunto 2 (restantes dias).

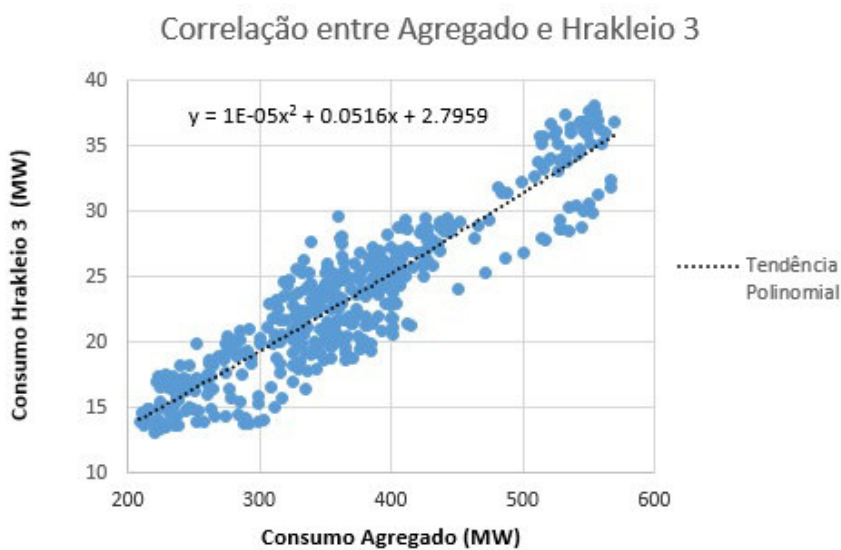


Figura C.3: Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Hrakleio3, para o conjunto 1 (dias úteis).

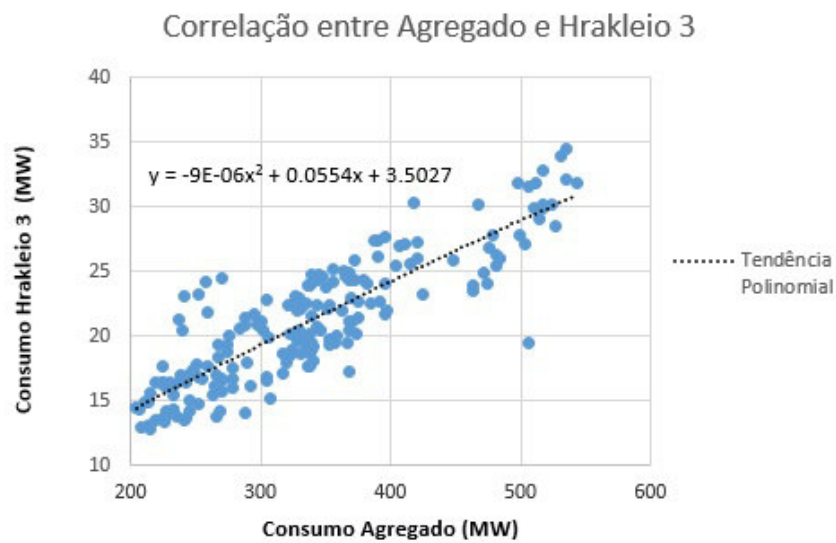


Figura C.4: Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Hrakleio3, para o conjunto 2 (restantes dias).

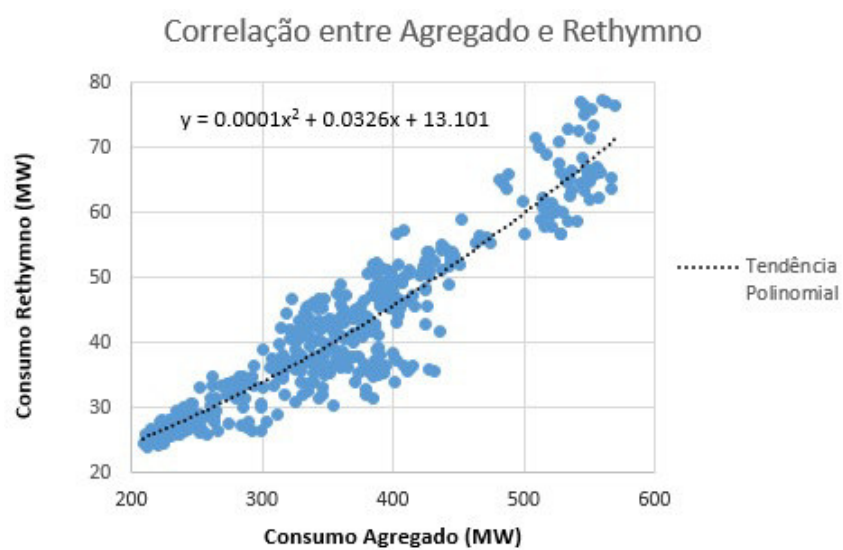


Figura C.5: Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Rethymno, para o conjunto 1 (dias úteis).

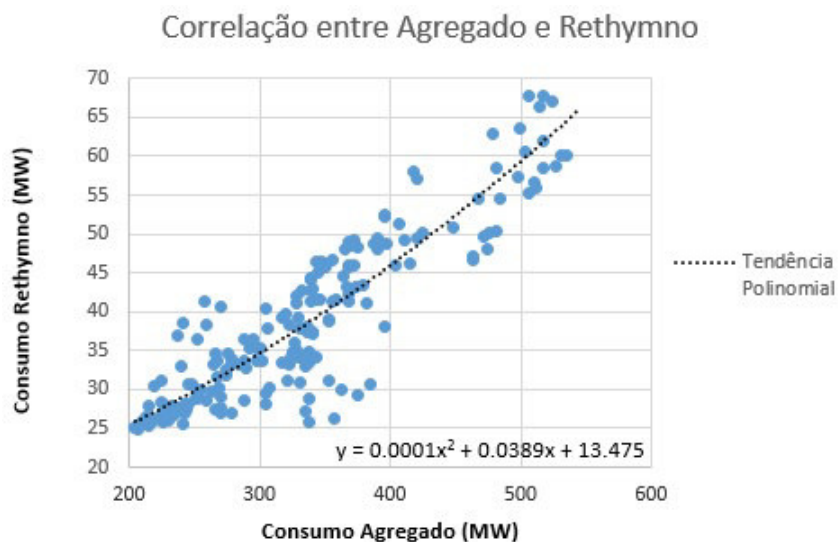


Figura C.6: Gráfico de Dispersão que relaciona o Consumo Agregado com o Consumo da subestação Rethymno, para o conjunto 2 (restantes dias).

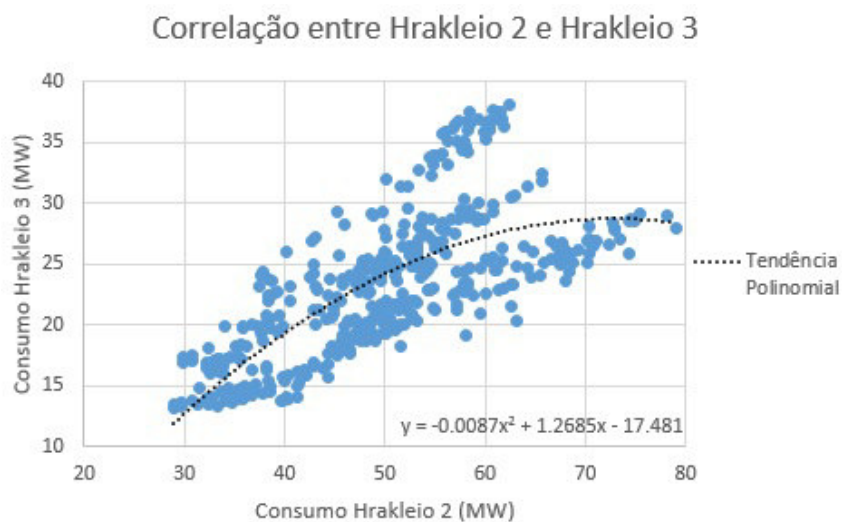


Figura C.7: Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio2 com o Consumo da subestação Hrakleio3, para o conjunto 1 (dias úteis).

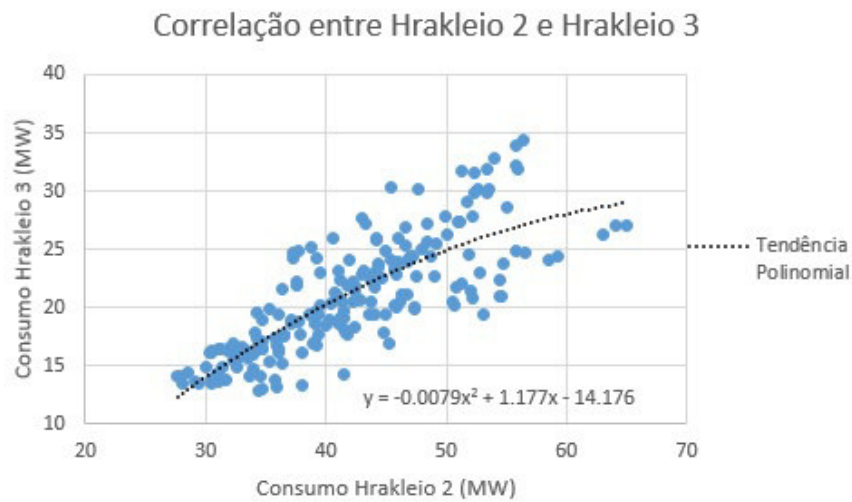


Figura C.8: Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio2 com o Consumo da subestação Hrakleio3, para o conjunto 2 (restantes dias).

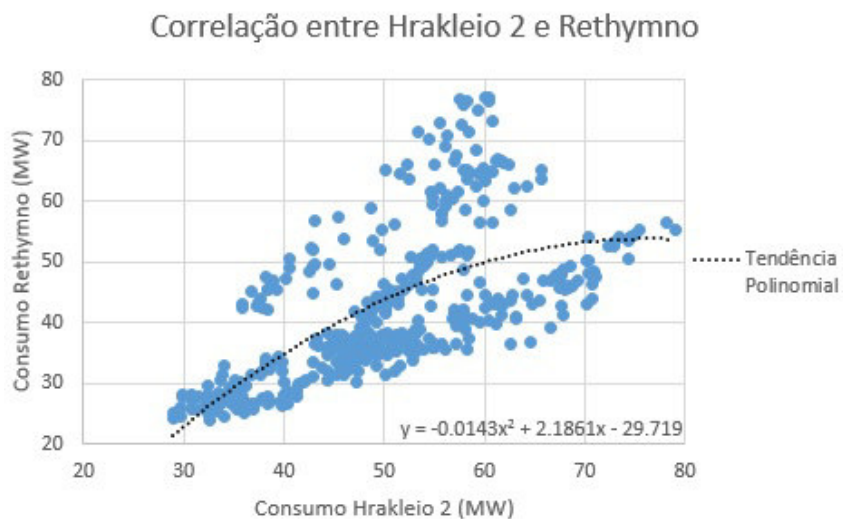


Figura C.9: Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio2 com o Consumo da subestação Rethymno, para o conjunto 1 (dias úteis).

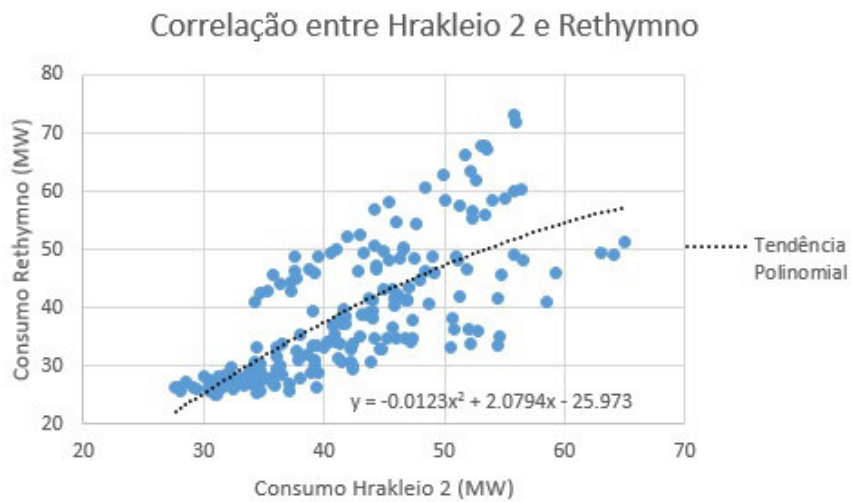


Figura C.10: Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio2 com o Consumo da subestação Rethymno, para o conjunto 2 (restantes dias).

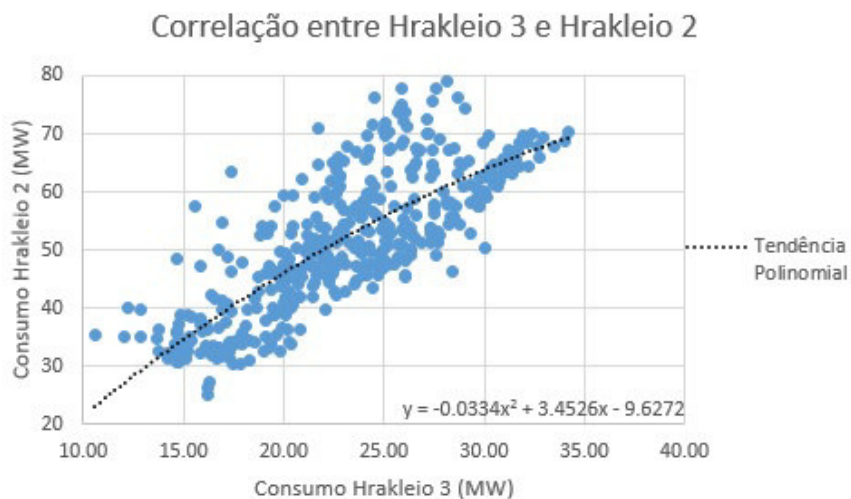


Figura C.11: Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio3 com o Consumo da subestação Hrakleio2, para o conjunto 1 (dias úteis).

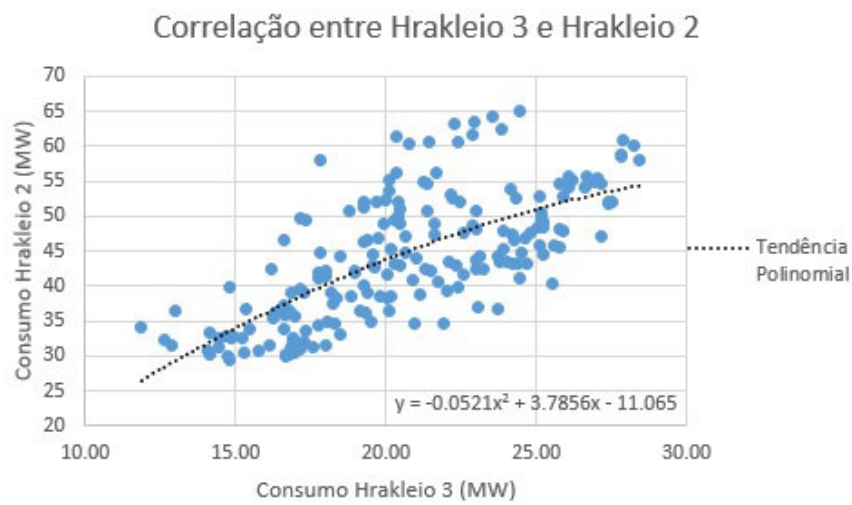


Figura C.12: Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio3 com o Consumo da subestação Hrakleio2, para o conjunto 2 (restantes dias).

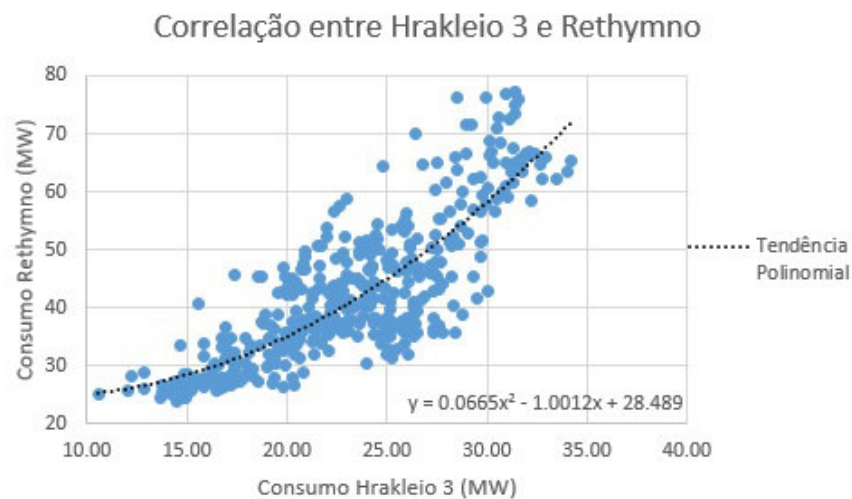


Figura C.13: Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio3 com o Consumo da subestação Rethymno, para o conjunto 1 (dias úteis).

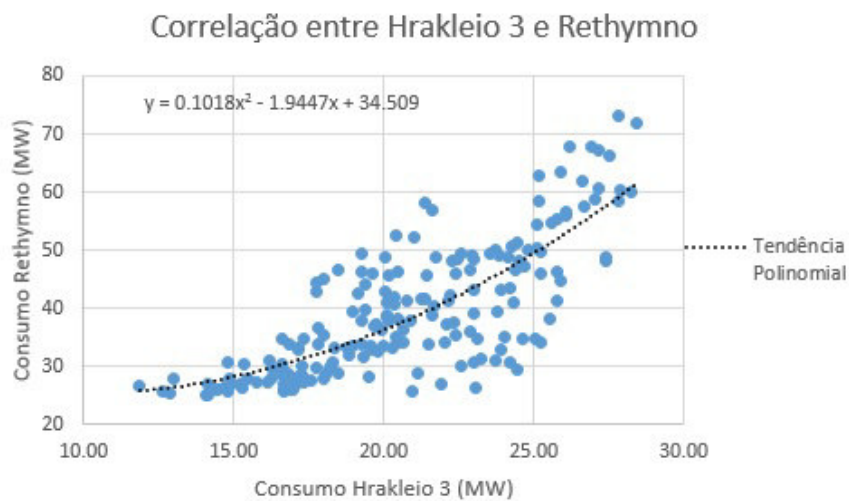


Figura C.14: Gráfico de Dispersão que relaciona o Consumo da subestação Hrakleio3 com o Consumo da subestação Rethymno, para o conjunto 2 (restantes dias).

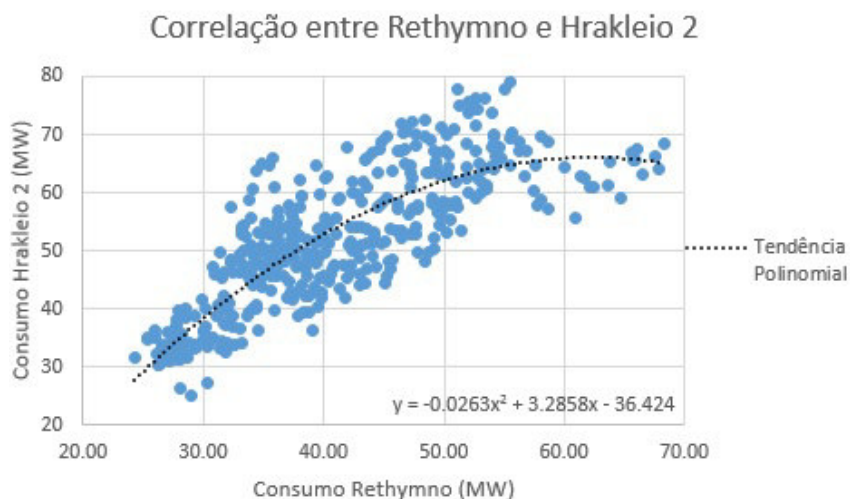


Figura C.15: Gráfico de Dispersão que relaciona o Consumo da subestação Rethymno com o Consumo da subestação Hrakleio2, para o conjunto 1 (dias úteis).

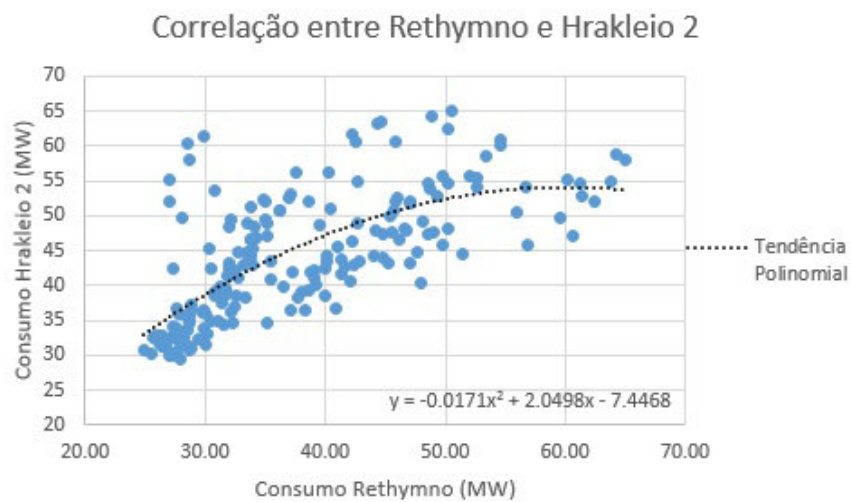


Figura C.16: Gráfico de Dispersão que relaciona o Consumo da subestação Rethymno com o Consumo da subestação Hrakleio2, para o conjunto 2 (restantes dias).

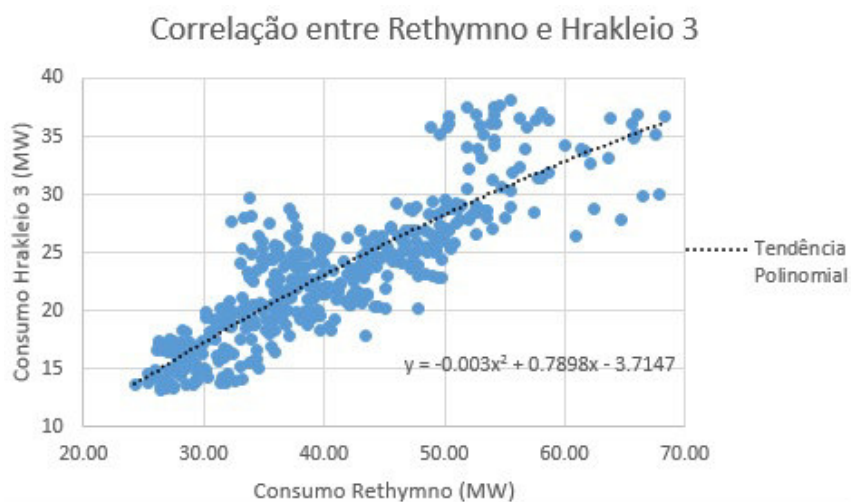


Figura C.17: Gráfico de Dispersão que relaciona o Consumo da subestação Rethymno com o Consumo da subestação Hrakleio3, para o conjunto 1 (dias úteis).

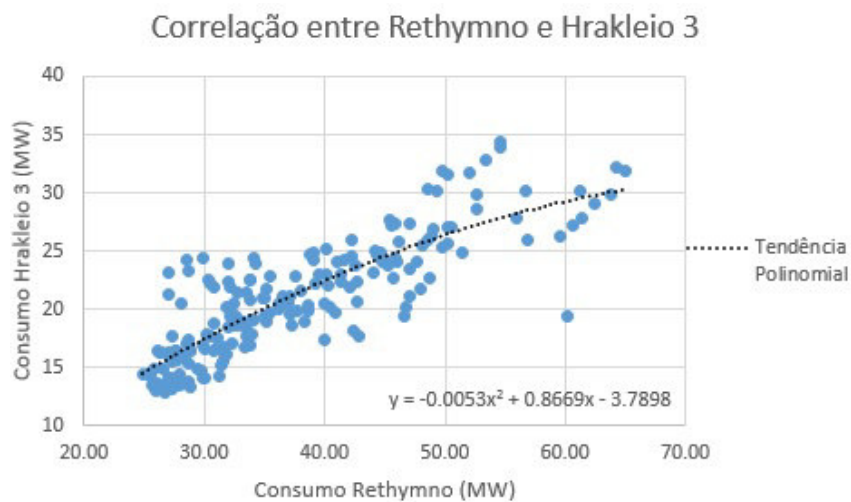


Figura C.18: Gráfico de Dispersão que relaciona o Consumo da subestação Rethymno com o Consumo da subestação Hrakleio3, para o conjunto 2 (restantes dias).

C.2 Previsão

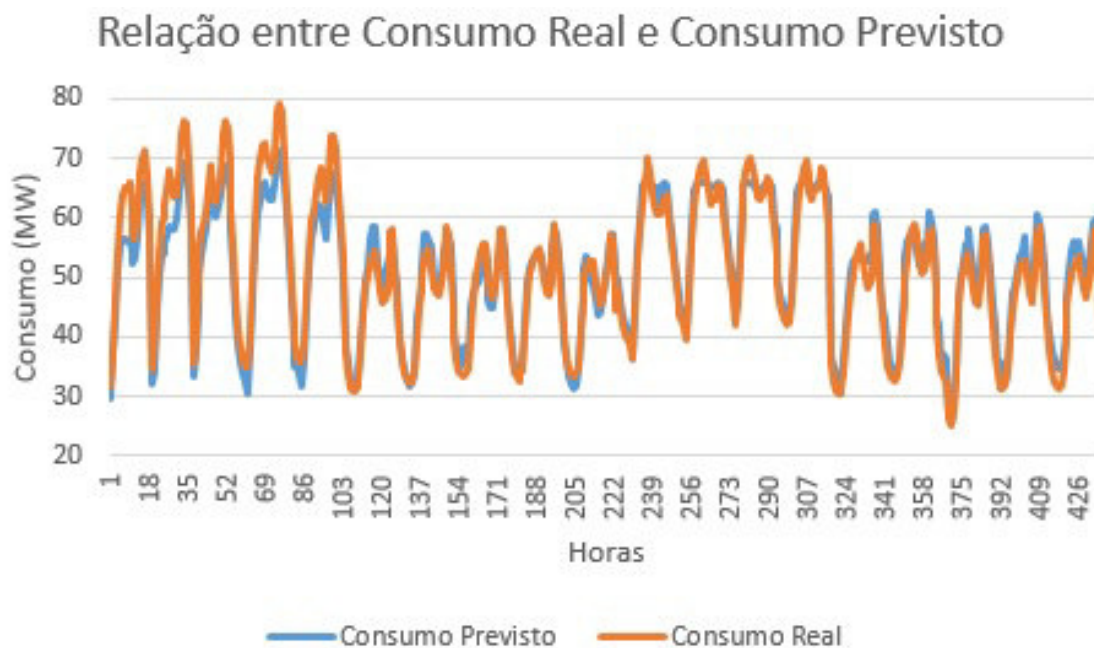


Figura C.19: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 1 (dias úteis), com base na correlação com o Agregado.

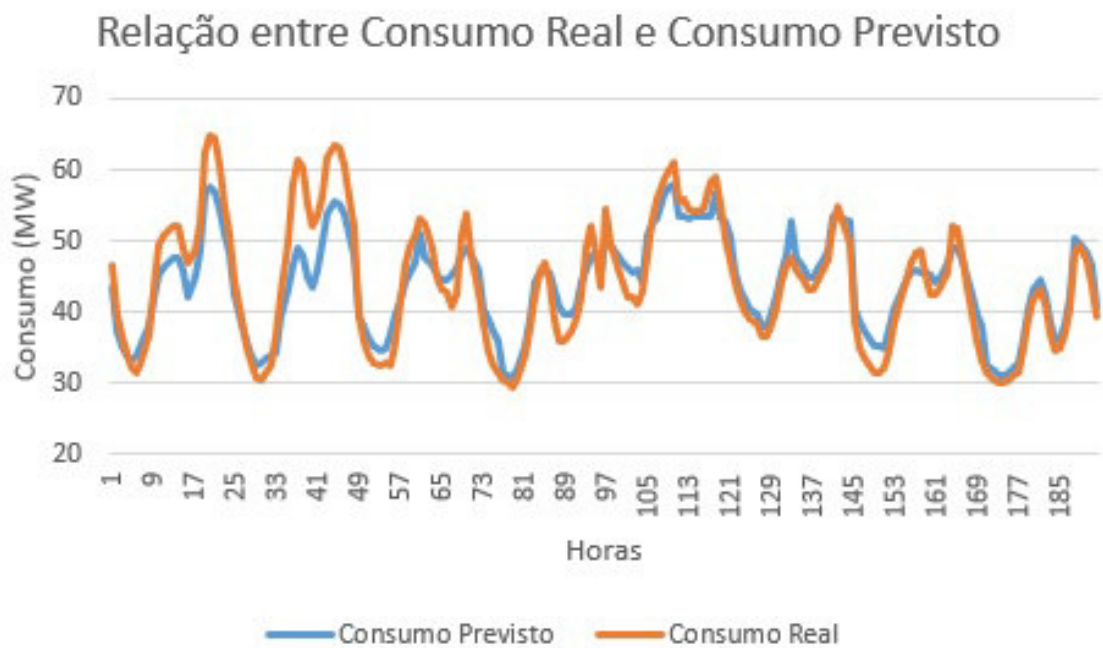


Figura C.20: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 2 (restantes dias), com base na correlação com o Agregado.

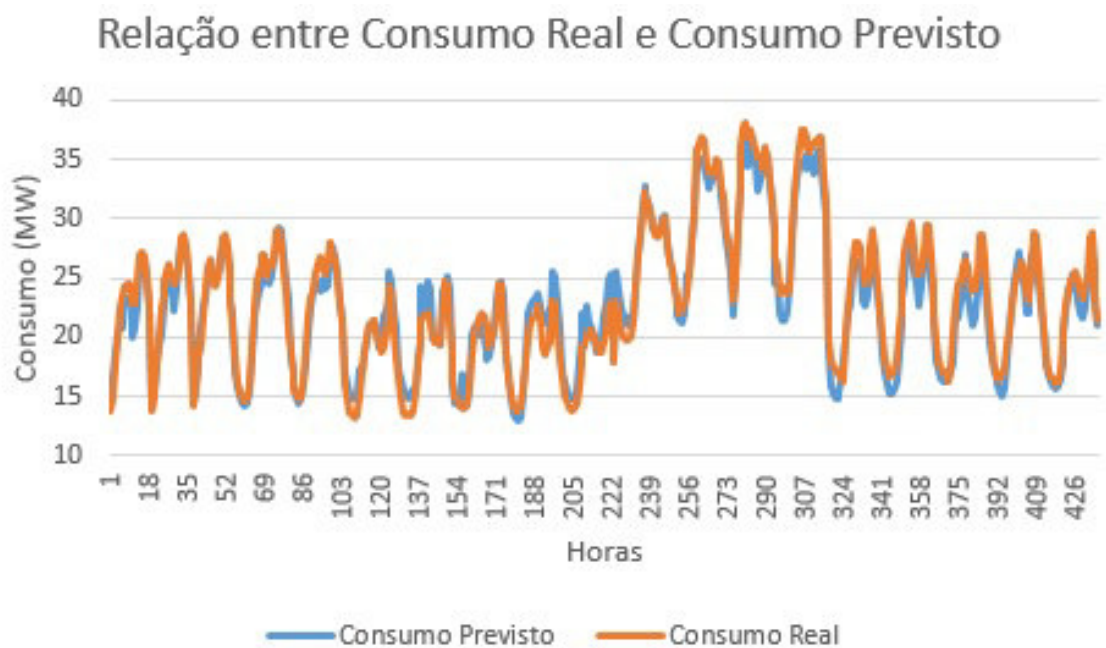


Figura C.21: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 1 (dias úteis), com base na correlação com o Agregado.

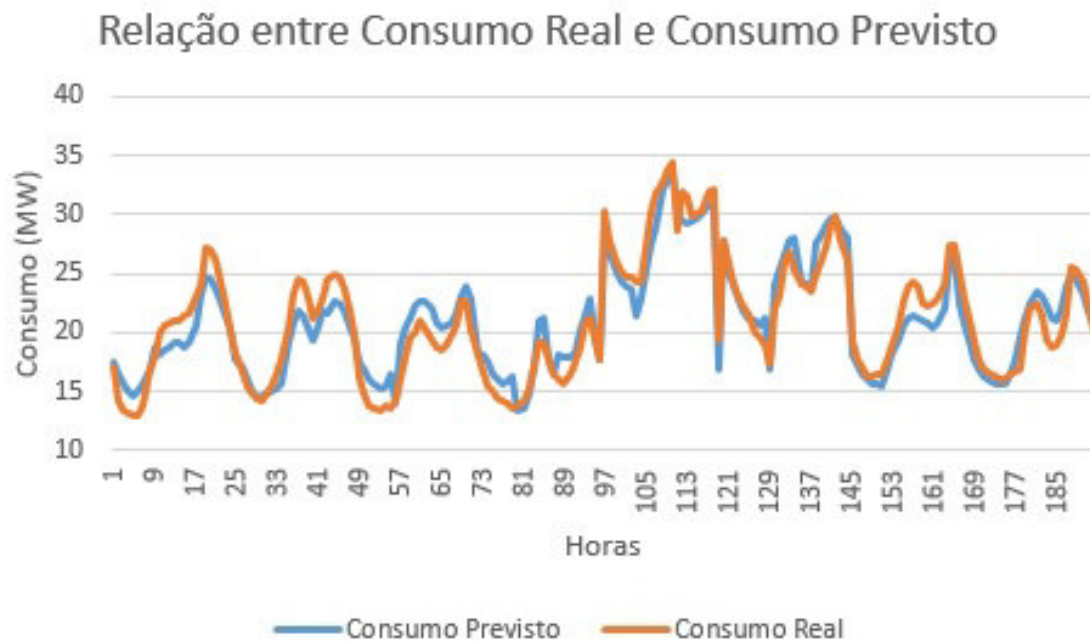


Figura C.22: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 2 (restantes dias), com base na correlação com o Agregado.

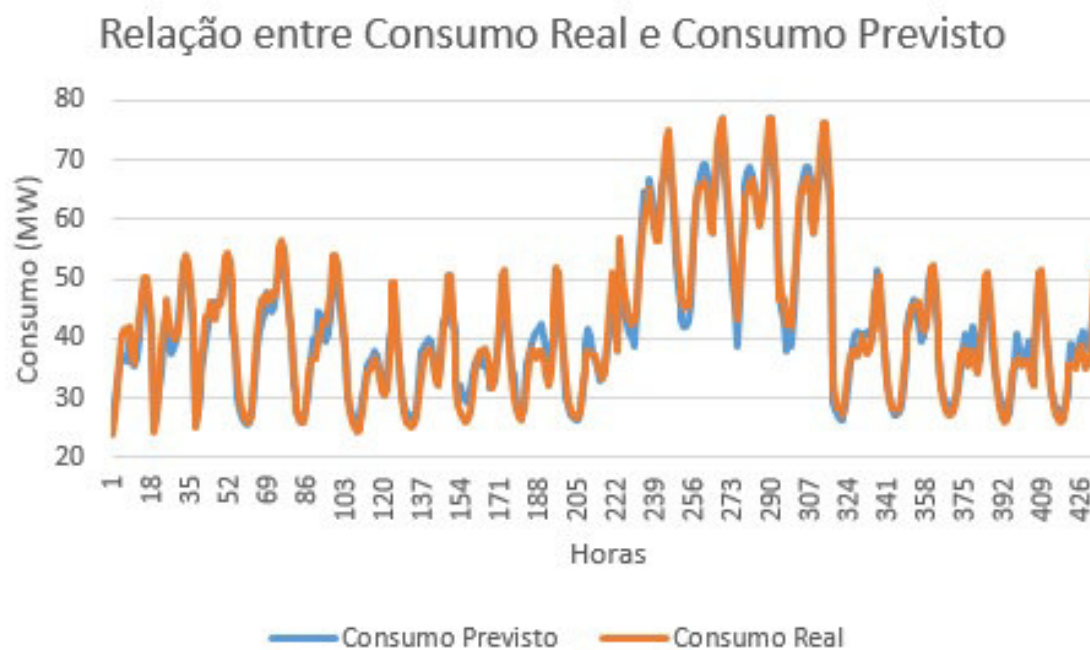


Figura C.23: Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 1 (dias úteis), com base na correlação com o Agregado.

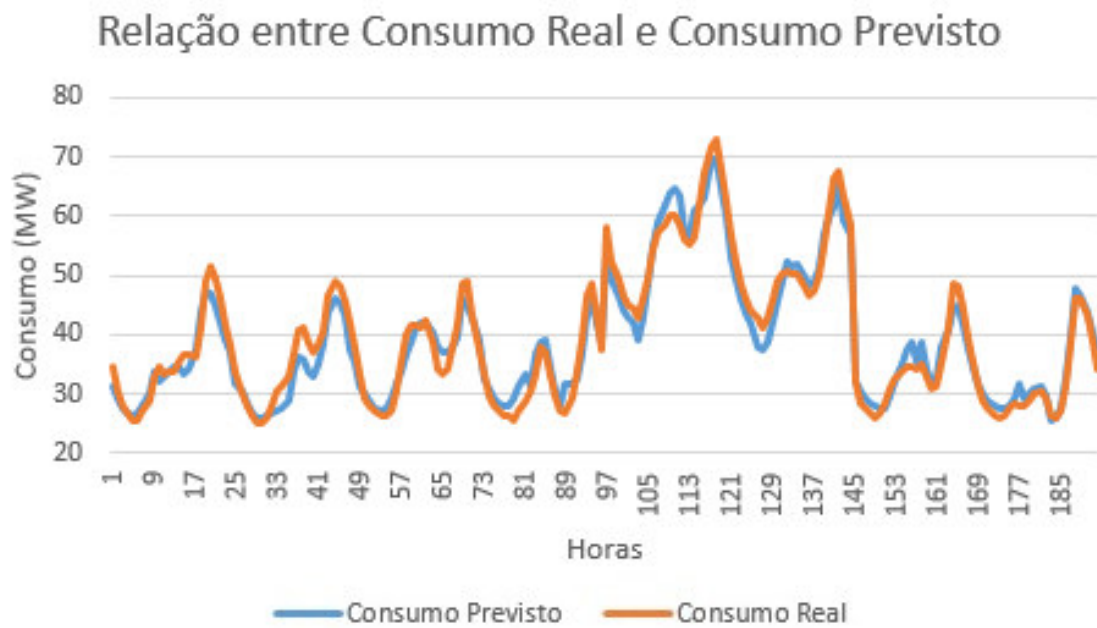


Figura C.24: Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 2 (restantes dias), com base na correlação com o Agregado.

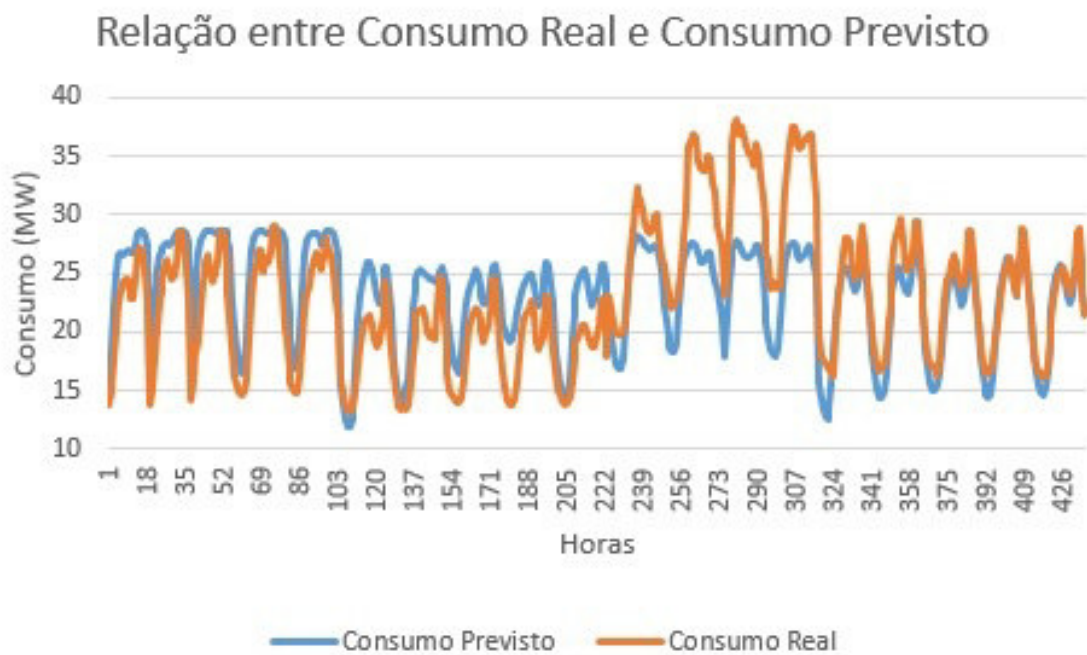


Figura C.25: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 1 (dias úteis), com base na correlação com a subestação Hrakleio2.

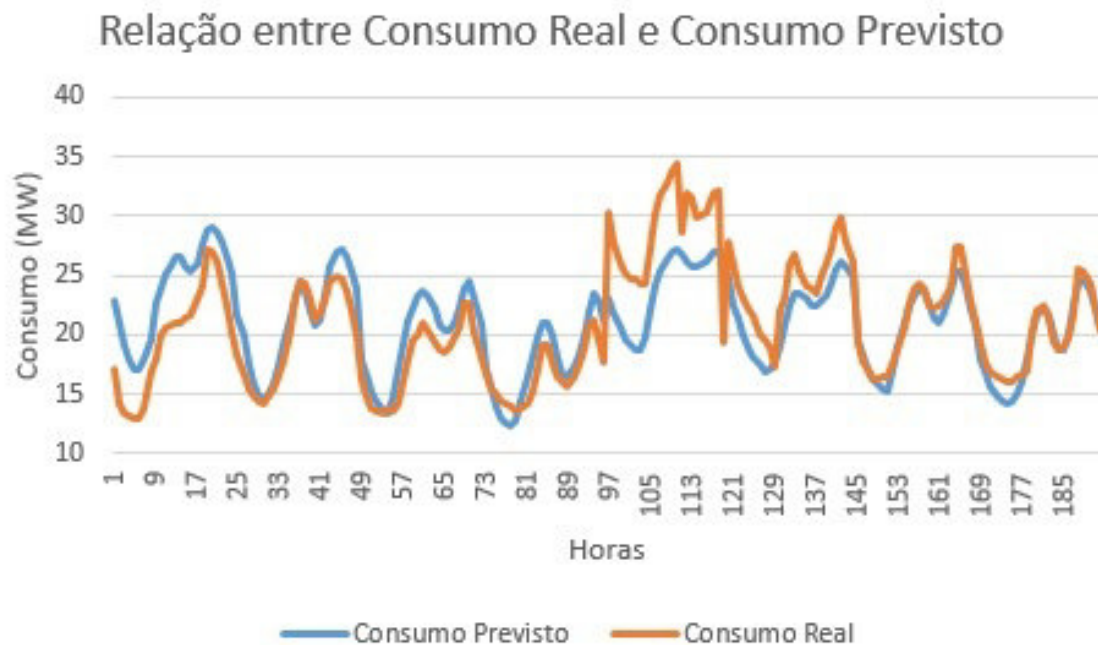


Figura C.26: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 2 (restantes dias), com base na correlação com a subestação Hrakleio2.

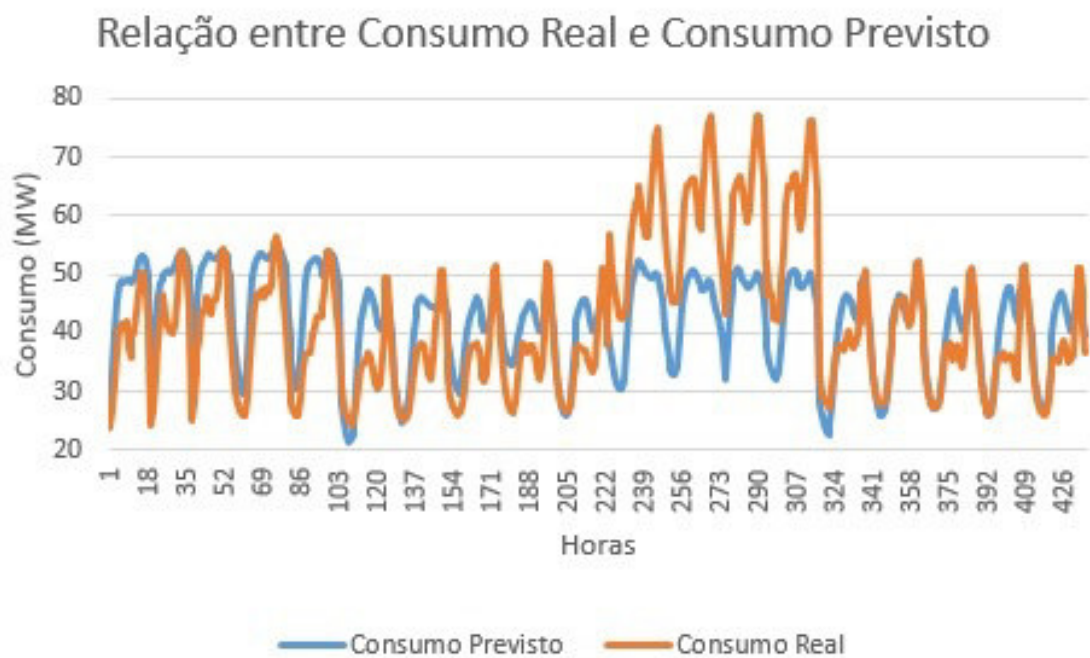


Figura C.27: Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 1 (dias úteis), com base na correlação com a subestação Hrakleio2.

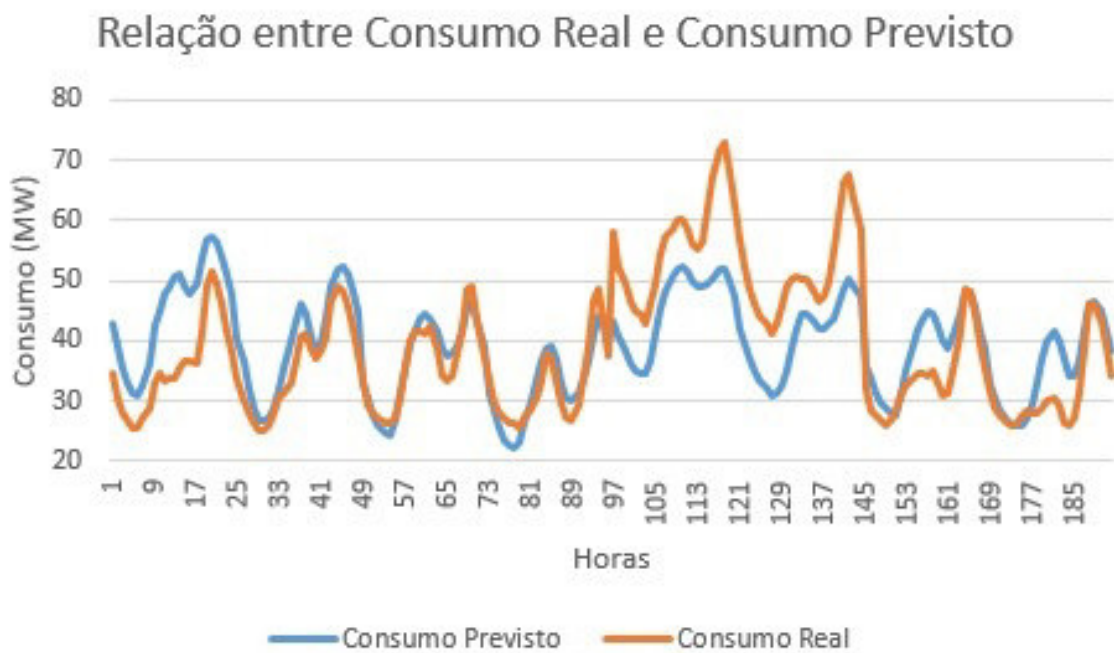


Figura C.28: Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 2 (restantes dias), com base na correlação com a subestação Hrakleio2.

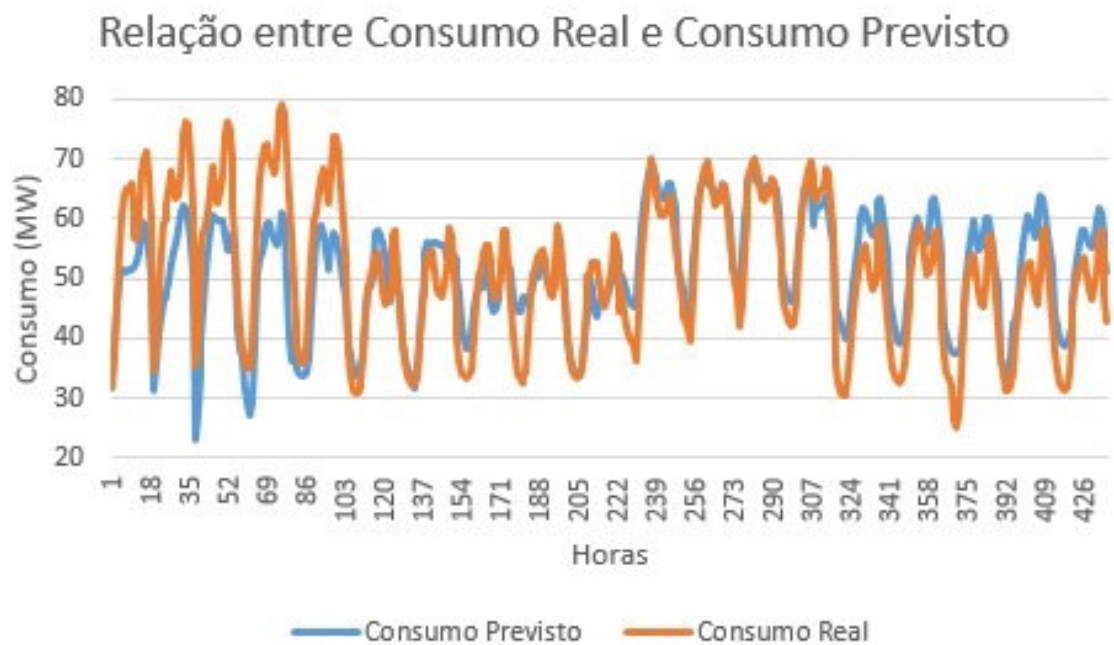


Figura C.29: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 1 (dias úteis), com base na correlação com a subestação Hrakleio3.

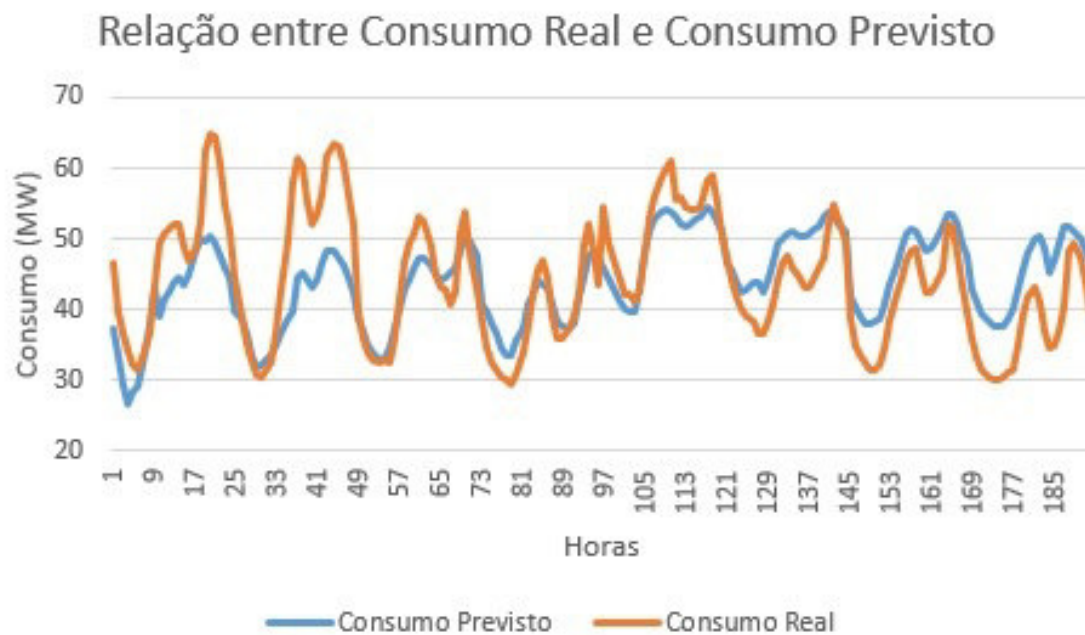


Figura C.30: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 2 (restantes dias), com base na correlação com a subestação Hrakleio3.

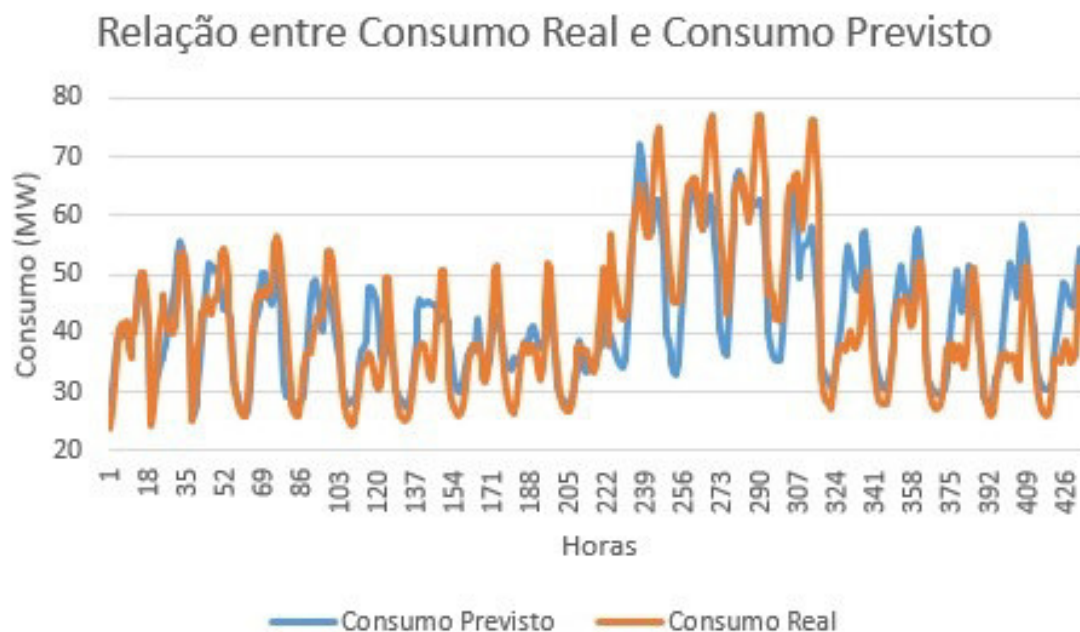


Figura C.31: Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 1 (dias úteis), com base na correlação com a subestação Hrakleio3.

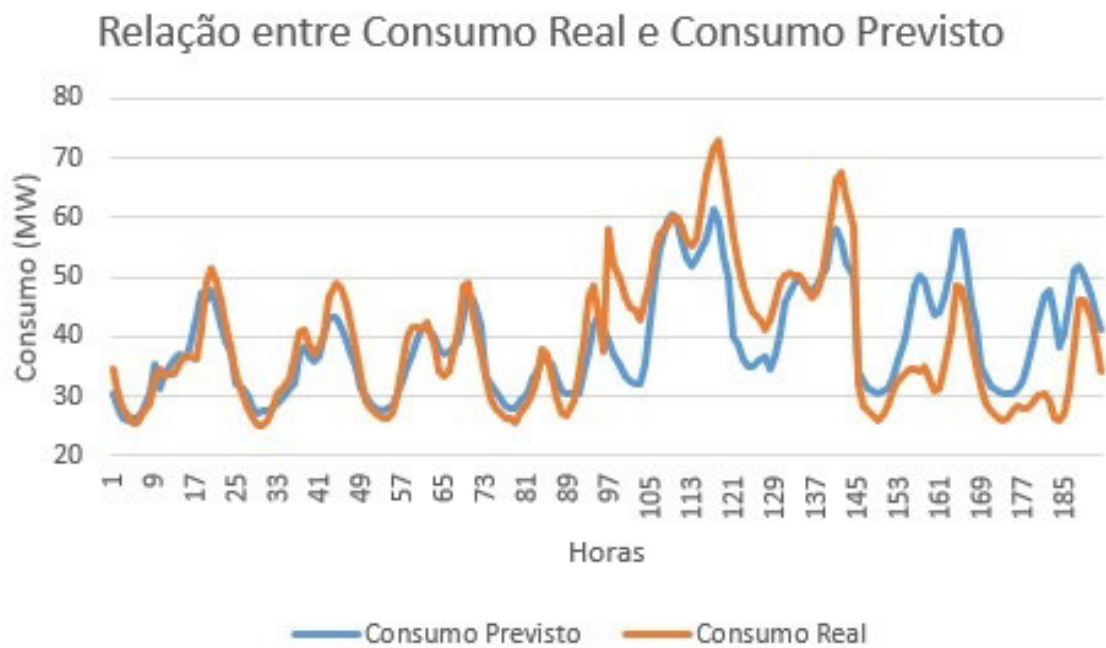


Figura C.32: Relação entre o Consumo Real e Consumo Previsto para a subestação Rethymno, no conjunto 2 (restantes dias), com base na correlação com a subestação Hrakleio3.

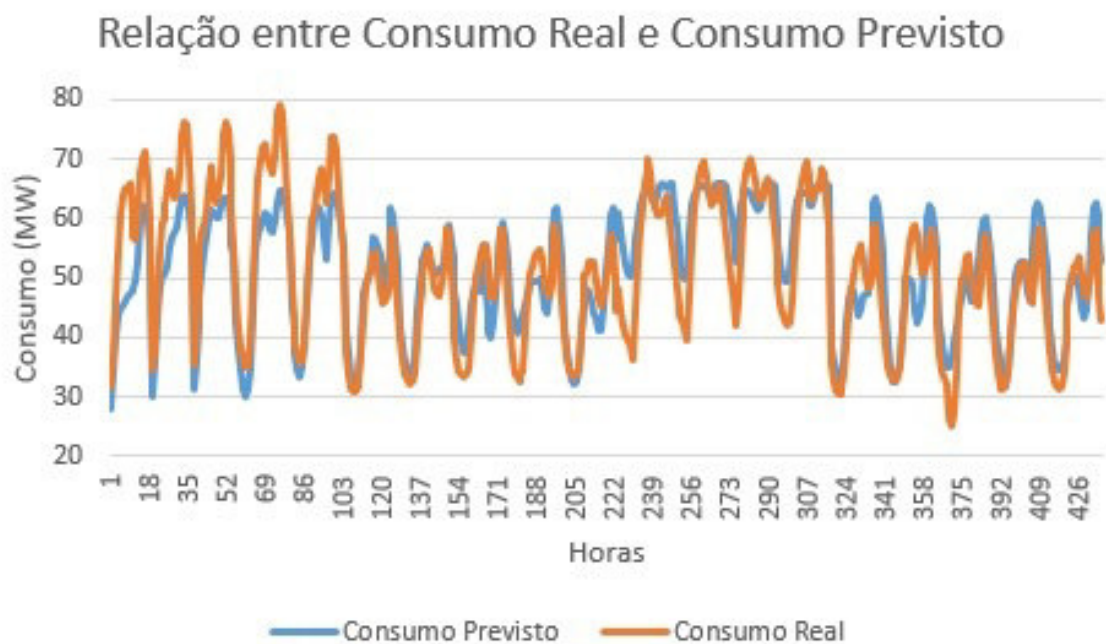


Figura C.33: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 1 (dias úteis), com base na correlação com a subestação Rethymno .

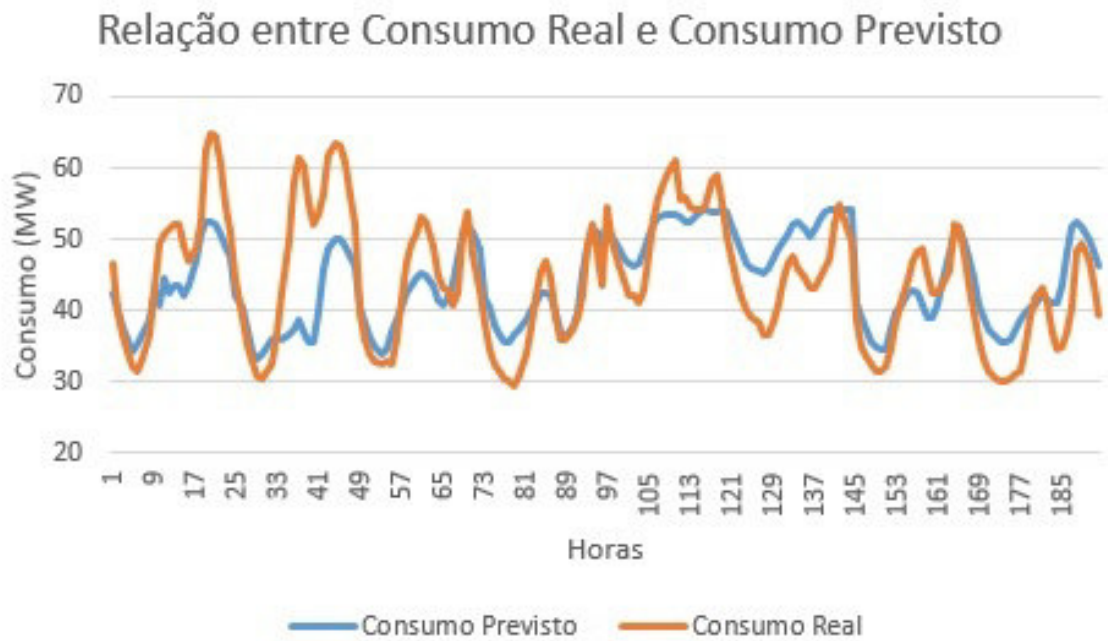


Figura C.34: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio2, no conjunto 2 (restantes dias), com base na correlção com a subestação Rethymno .

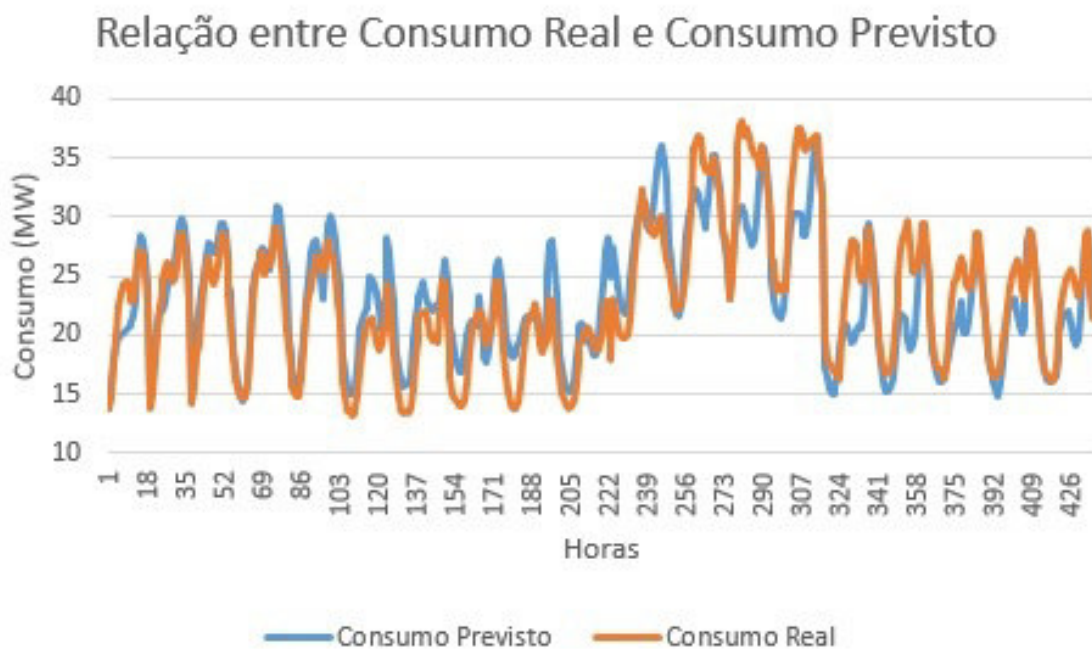


Figura C.35: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 1 (dias úteis), com base na correlção com a subestação Rethymno .

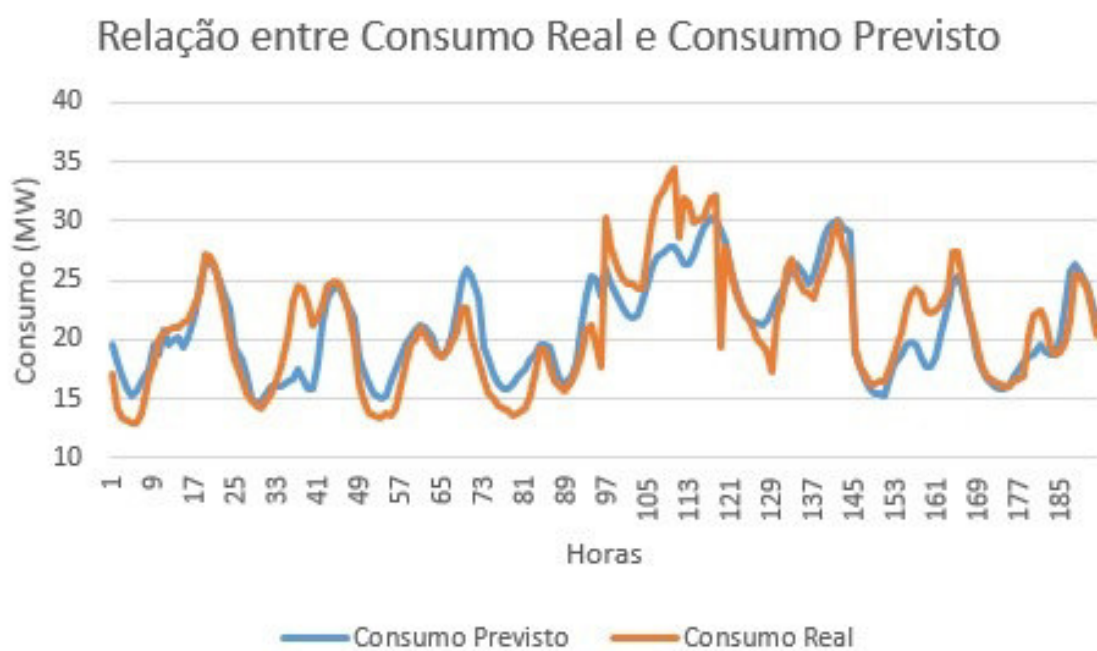


Figura C.36: Relação entre o Consumo Real e Consumo Previsto para a subestação Hrakleio3, no conjunto 2 (restantes dias), com base na correlção com a subestação Rethymno .

Referências

- [1] SiNGULAR. Smart and sustainable insular electricity grids under large-scale renewable integration. URL: <http://www.singular-fp7.eu/home/>.
- [2] Kieran Richard Godden. Electric load forecasting for holiday periods. master's thesis, faculty of science at the rand afrikaans university. 1997.
- [3] Antti Laukkanen. The use of special day information in a demand forecasting model for nordic power market. 2004.
- [4] Roberto Napoli G. Chicco e Federico Piglione. Load pattern clustering for short-term load forecasting of anomalous days. in power tech proceedings, iee porto, volume 2. 2001.
- [5] M. Sforza M. Caciotta R. Lamedica, A. Prudenzi e V.O. Cencelli. A neural network based technique for short-term forecasting of anomalous load periods. power systems, iee transactions on. 1996.
- [6] S.J. Watson Ching-Lai Hor e S. Majithia. Analyzing the impact of weather variables on monthly electricity demand. power systems, iee transactions on. 2005.
- [7] A. Vuckovic S. Ruzic e N. Nikolic. Weather sensitive method for short term load forecasting in electric power utility of serbia. power systems, iee transactions on. 2003.
- [8] G. Gross e F.D. Galiana. Short-term load forecasting. proceedings of the iee. 1987.
- [9] C.E. Pedreira H.S. Hippert e R.C. Souza. Neural networks for short-term load forecasting a review and evaluation. power systems, iee transactions on. 2001.
- [10] F. F. Wu J. H. Chow e J. A. Momoh. Applied mathematics for restructured electric power systems: Optimization, control, and computational intelligence. 2003.
- [11] S. Khatoon A. K. Singh e D. K. Chaturvedi. Load forecasting techniques and methodologies: A review, 2nd international conference on power, control and embedded systems. 2012.
- [12] H. K. Alfares e M. Nazeeruddin. Electric load forecasting: Literature survey and classification of methods, international journal of systems science. 2002.
- [13] C. Canizares H. Chen e A. Singh. Ann-based short-term load forecasting in electricity markets, iee conference publications. 2001.
- [14] S. Rahman. Formulation and analysis of a rule-based short-term load forecasting algorithm, proceedings of the iee. 1990.
- [15] C. Mustafa R. F. Engle e J. Rice. Modelling peak electricity demand, journal of forecasting. 1992.

- [16] R. Shoults M. Manry C. Kwan F. Lewis K. Liu, S. Subbarayan e J. Naccarino. Comparison of very shortterm load forecasting techniques, power systems. 1996.
- [17] J. Garcia Moreno G. Juberias, R. Yunta e C. Mendivil. A new arima model for hourly load forecasting, transmission and distribution conference. 1999.
- [18] J. Y. Fan e J. D. McDonald. A real-time implementation of short-term load forecasting or distribution power systems, iee transactions on power systems, vol. 9. 1994.
- [19] M. Aggoune M. Damborg, M. El-Sharkawi e R. I. Marks. Potential of artificial neural networks in power system operation, circuits and systems. 1990.
- [20] J.N. Fidalgo e J.A.P. Lopes. Load forecasting performance enhancement when facing anomalous events. power systems, iee transactions on. 2005.
- [21] J. Wang J. Valenzuela J. Sumaili R. Bessa H. Keko A. Botterud, Z. Zhou e V. Miranda. Unit commitment and operating reserves with probabilistic wind power forecasts. Paper accepted for presentation at the 2011 IEEE Trondheim PowerTech.
- [22] Walter Zucchini. Applied smoothing techniques, part 1: Kernel density estimation. temple university, philadelphia, pa. 2003.
- [23] António dos Santos Pires. Improving electrical production forecasting using gpu. master's thesis, faculdade de engenharia da universidade do porto. 2013.
- [24] C. Soares J. Catalão G. Osorio J. Rojas R. Nebot Medina G. Sannino A. Carillo C. González García e J. Contreras Sanz B. Santos, C. Monteiro. Report on the mathematical formulation of the different forecasting models and data analysis. deliverable d2.1. 2013.
- [25] SAS Institute Inc. Base sas ®9.3 procedures guide: Statistical procedures. 2011.
- [26] F. Y. Shih G. Fu e H. Wang. A kernel-based parametric method for conditional density estimation, pattern recognition, vol. 44. February 2011.
- [27] L. Fugon J. Juban e G. Kariniotakis. Probabilistic short-term wind power forecasting based on kernel density estimators. European Wind Energy Conference, May 2007.
- [28] A. Botterud R. J. Bessa, V. Miranda e J. Wang. Time adaptive conditional kernel density estimation for wind power forecasting,. IEEE Transactions on Sustainable Energy, October 2012.
- [29] Wikipedia. Beta distribution. URL: http://en.wikipedia.org/wiki/Beta_distribution.
- [30] L. Alfredo FernandezJimenez e A. Muñoz Jimenez Cláudio Monteiro, Ignacio J. Ramirez-Rosado. Short - term forecasting models for photovoltaic plants: Analytical vs soft computing techniques. energies. maio 2013.
- [31] L. Alfredo Fernandez-Jimenez e A. Muñoz-Jimenez Cláudio Monteiro, Ignacio J. Ramirez-Rosado. Short - term forecasting models for photovoltaic plants: Analytical vs softcomputing techniques. Maio 2013.
- [32] Hugo T.C. Pedro e Carlos F.M. Coimbra. Assessment of forecasting techniques for solar power production with no exogenous inputs.

- [33] ISLEPACT. Island sustainable energy action plan, island of crete. novembro 2011.
- [34] timeanddate.com. Time changes in year 2012 for greece - crete. URL: <http://www.timeanddate.com/time/change/greece/iraklion?year=2012>.
- [35] timeanddate.com. Time changes in year 2013 for greece - crete. URL: <http://www.timeanddate.com/time/change/greece/iraklion?year=2013>.
- [36] timeanddate.com. Time changes in year 2014 for greece - crete. URL: <http://www.timeanddate.com/holidays/greece/2012>.
- [37] timeanddate.com. Holidays and observances in greece 2012. URL: <http://www.timeanddate.com/holidays/greece/2013>.
- [38] timeanddate.com. Holidays and observances in greece 2013. URL: <http://www.timeanddate.com/holidays/greece/2014>.
- [39] timeanddate.com. Holidays and observances in greece 2014. URL: <http://www.timeanddate.com/time/change/greece/iraklion?year=2014>.
- [40] L.-M. Liu J. M. Liu, R. Chen e J. L. Harris. Asemi-parametric time series approach in modeling hourly electricity loads, journal of forecasting, vol. 25, no. 8, p. 537–559. December 2006.