

Faculdade de Engenharia da Universidade do Porto



**Previsão probabilística de consumos agregados para a
optimização de estratégias de procura no mercado
MIBEL**

Marta Filipa Ferreira Simões

Dissertação realizada no âmbito do Mestrado Integrado em Engenharia Eletrotécnica e
de Computadores - Major Energia

Orientador: Professor Doutor Cláudio Monteiro

25 de junho de 2018

Resumo

O Mercado Ibérico de Energia Elétrica (MIBEL) caracteriza-se por ser um mercado concorrencial, no qual é necessário antever as necessidades da rede e apresentar propostas de compra e venda de energia. Tudo isto com o propósito de garantir a fiabilidade da rede elétrica, mas também, para que os desvios resultantes das diferenças entre as propostas de compra de energia e o consumo real sejam o mais baixos possível. Quanto mais elevados forem estes desvios maior será o prejuízo para os comercializadores e maiores serão os transtornos causados para toda a rede de consumidores uma vez que poderá haver consequência a nível de preços de eletricidade.

Neste âmbito, o tema desta dissertação surge com a finalidade de desenvolver um método de previsão de consumos de energia elétrica para o dia seguinte, baseado nos valores históricos de consumos anteriores, através do estudo de padrões de consumo, para que, assim, os comercializadores possam realizar propostas mais assertivas aquando da compra de energia para o dia seguinte e diminuam assim os custos associados aos desvios causados pelas grandes diferenças apresentadas entre energia consumida e energia comprada.

O uso da previsão probabilística permite obter uma incerteza associada à previsão. A informação sobre incerteza é a base para processos de decisão baseados em análise de risco.

O trabalho realizado consiste em modelos de previsão probabilísticos baseados em modelos de decomposição, decompondo as séries temporais históricas em: 1) componente de tendência; 2) componente sazonal de padrão semanal ao longo do ano; 3) componente de padrão diário ao longo da semana.

São utilizadas duas abordagens diferentes para a modelização da componente de tendência, aplicada sobre as pontas e vazios semanais. A modelização probabilística é feita sobre o padrão diário, utilizando modelização com base na distribuição beta.

O modelo foi testado para algumas semanas de 2018, apresentando-se os quantis de previsão e a avaliação de erros da previsão pontual e probabilística.

Palavras chave: Previsão de consumos de eletricidade, Incerteza da Previsão, Previsão Probabilística.

Abstract

The Iberian Electricity Market (MIBEL) is characterized for being a competitive market, in which it is necessary to foresee the needs of the network and to present proposals for the purchase and sale of energy. All this with the purpose of guaranteeing the reliability of the electricity network, but also, so that the deviations resulting from the differences between the proposals of purchase of the energy and the real consumption are as low as possible. The higher the deviations, bigger is the damage to the

energy distributors and bigger it is the disruption caused to the entire network of consumers, once it can effect the level of prices of electricity.

In this context, the theme of this dissertation comes with the purpose of developing a method of forecasting electricity consumption for the following day, based on the historical values of previous consumptions, through the study of patterns of consumption, so that distributors can make more assertive proposals when buying energy for the next day and reduce the costs associated with deviations caused by the large differences between consumed energy and purchased energy.

The use of the probabilistic forecast allows an uncertainty associated with the forecast. Information about uncertainty is the basis for decision-making based on risk analysis.

The work consists in probabilistic prediction models based on decomposition models, decomposing the historical time series into: 1) trend component; 2) seasonal component with a weekly pattern throughout the year; 3) Daily pattern component throughout the week.

Two different approaches are used for the modeling of the trend component, applied over tips and weekly voids. The probabilistic modeling is done over the daily pattern, using modeling based on the beta distribution.

The model was tested for a few weeks in 2018, presenting the forecast and the error evaluation of the point and probabilistic forecast.

Key words: Forecast of electrical consumption, Forecast uncertainty, Probabilistic forecast.

Agradecimentos

O caminho até aqui foi muito longo e cheio de altos e baixos. O fato de chegar até aqui devo a todos os que nunca deixaram que desistisse e que nunca desistiram de mim principalmente a minha mãe.

À minha querida avó Ermelinda agradeço de coração todo o esforço que sempre fez para me ajudar a finalizar esta etapa tão importante da minha vida, sem ela não teria sido possível.

No entanto, se não fosse a grande amiga que esta aventura me deu, a Ana Barros, possivelmente não teria chegado até aqui. Foi amiga, companheira, o meu braço direito ao longo destes anos, e a pessoa que mais birras e choros aturou não deixando que eu virasse costas a este sonho.

Agradeço também aos grandes AMIGOS que a vida me foi dando pelo apoio incondicional e força que sempre me deram.

Ao professor Cláudio Monteiro devo esta dissertação, foi sem dúvida a pessoa que mais contribuiu para que chegasse ao fim, e quero agradecer todo o tempo perdido comigo durante estes meses, sem ele não teria sido possível. Obrigado de coração.

Por fim, quero agradecer à pessoa que “chegou por último” à minha vida, mas que foi sem dúvida quem mais me “chateou” para que eu me dedicasse e que lutasse por tudo o que sempre quis, o Miguel, obrigado pelas horas a aturar-me, a ajudar e pelo bom amigo em que te tornaste.

Este fim foi doloroso, mas consegui e em grande parte graças a todos vocês.

Um GRANDE OBRIGADO!

Marta Simões

Índice

| | |
|---|-----------|
| Capítulo 1 | 1 |
| Introdução | 1 |
| 1.1 Contextualização e Motivação..... | 1 |
| 1.2 Objetivos..... | 2 |
| 1.3 Estrutura da dissertação | 2 |
| Capítulo 2 | 3 |
| Estado da Arte..... | 3 |
| 2.1 Fatores que influenciam os padrões de consumo..... | 3 |
| 2.2 Horizonte Temporal..... | 4 |
| 2.3 Previsão Pontual | 4 |
| 2.3.1 Modelos Probabilísticos..... | 5 |
| 2.3.2 Modelos Multiagente..... | 6 |
| 2.3.3 Inteligência Artificial..... | 6 |
| 2.3.3.1 Redes Neurais Artificiais..... | 6 |
| 2.3.3.2 Support Vector Machine..... | 7 |
| 2.4 Previsão Probabilística..... | 7 |
| 2.4.1 Intervalos Estatísticos | 8 |
| 2.4.2 Regressão Linear de Quantis | 9 |
| 2.4.3 Quantile Regression Forests | 10 |
| 2.4.4 Modelos Aditivos de Regressão de Quantis por <i>Boosting</i> | 12 |
| 2.4.5 Redes Neurais para Regressão de Quantis..... | 12 |
| 2.4.6 Regressão de Quantis no Reproducing Kernel Hilbert Space | 13 |
| 2.4.7 KDE (Kernel Density Estimation)..... | 14 |
| 2.4.7.1 Função densidade de probabilidade da distribuição beta..... | 15 |
| 2.4.7.2 Funcionamento do KDE | 16 |
| 2.5 Medidas de Avaliação..... | 16 |
| 2.5.1 Coeficiente de determinação e Coeficiente de correlação de Pearson..... | 17 |
| 2.5.2 Índices de Erro..... | 17 |
| 2.5.3 Reliability e Sharpness | 18 |
| Capítulo 3 | 23 |
| Análise das Séries de Consumo | 23 |
| 3.1 Descrição dos Dados utilizados | 23 |
| 3.2 Tendência de Crescimento..... | 24 |
| 3.3 Padrão Sazonal | 26 |
| 3.4 Padrão Semanal e Diário | 29 |
| 3.4.1 Normalização dos dados | 30 |
| 3.4.2 Resíduo Semanal | 32 |
| 3.4.3 Padrão Diário..... | 33 |
| 3.4.4 Caracterização probabilística do padrão diário..... | 36 |
| 3.4.4.1 Cálculo da Variância..... | 37 |

| | |
|--|-----------|
| 3.4.4.2 Cálculo dos parâmetros α e β | 38 |
| 3.4.5 Distribuição Beta vs Distribuição Normal..... | 39 |
| Capítulo 4..... | 43 |
| Previsão Probabilística baseada em Modelos de Decomposição | 43 |
| 4.1 Modelo - Variante I..... | 44 |
| 4.1.1 Resultados obtidos | 45 |
| 4.2 Modelo - Variante II | 46 |
| 4.2.1 Resultados Obtidos | 48 |
| 4.2.2 Teste do Modelo | 49 |
| Capítulo 5..... | 53 |
| Conclusões e Estudos Futuros | 53 |
| 5.1 Cumprimento dos Objetivos | 53 |
| 5.2 Trabalhos Futuros | 54 |
| Referências..... | 55 |

Lista de figuras

| | |
|---|----|
| Figura 1 - Exemplo ilustrativo da diferença entre intervalo de confiança (área cinza claro) e intervalo de previsão (área cinza escuro) [25]..... | 9 |
| Figura 2 - Função densidade de probabilidade da distribuição beta para valores diferentes de alfa e beta. [33] | 16 |
| Figura 3 - Histogramas representativos dos diferentes cenários de <i>reliability</i> | 19 |
| Figura 4 - Etapas do processo de análise de dados. | 24 |
| Figura 5 - Consumos registados e respetivas retas de tendência de vazio e ponta, do período em estudo..... | 24 |
| Figura 6 – Gráfico dos consumos de julho a dezembro de 2015. | 25 |
| Figura 7 – Gráfico dos consumos de 2016..... | 25 |
| Figura 8 – Gráfico dos consumos de 2017..... | 25 |
| Figura 9 – Gráfico de consumos com realce nas semanas onde existe alteração significativa de consumos – agosto (a vermelho) e dezembro (a amarelo)..... | 26 |
| Figura 10 – Zoom do gráfico de consumos de 2016 para o período entre 25 de julho a 23 de setembro com realce nas semanas de agosto em que ocorre a diminuição de consumos... | 27 |
| Figura 11 - Zoom do gráfico de consumos de 2015 para o período entre 25 de julho a 23 de setembro com realce nas semanas de agosto em que ocorre a diminuição de consumos... | 27 |
| Figura 12 - Zoom do gráfico de consumos de 2017 para o período entre 25 de julho a 23 de setembro com realce nas semanas de agosto em que ocorre a diminuição de consumos... | 27 |
| Figura 13 - Zoom do gráfico de consumos de 2016 e 2017 para o período entre 15 de dezembro de 2016 a 10 de janeiro de 2017 com realce da semana em que ocorre a diminuição de consumos. | 28 |
| Figura 14 - Zoom do gráfico de consumos de 2015 e 2016 para o período entre 15 de dezembro de 2015 a 10 de janeiro de 2016 com realce da semana em que ocorre a diminuição de consumos. | 28 |
| Figura 15 - Zoom do gráfico de consumos de 2017 e 2018 para o período entre 15 de dezembro de 2017 a 10 de janeiro de 2018 com realce da semana em que ocorre a diminuição de consumos. | 28 |
| Figura 16 – Gráfico de consumos da semana de 2 a 8 de outubro de 2016. | 29 |

| | |
|--|----|
| Figura 17 - Gráfico de consumos da semana de 23 a 29 de abril de 2017. | 29 |
| Figura 18 - Curvas de tendência de pontas e vazio dos consumos utilizados para previsão. | 31 |
| Figura 19 – Gráfico dos resíduos de vazio e de ponta para cada semana do ano. | 32 |
| Figura 20 – Padrão diário dos consumos em estudo. | 33 |
| Figura 21 – Padrão Diário de dezembro de 2015. | 34 |
| Figura 22 - Padrão Diário de dezembro de 2016. | 34 |
| Figura 23 - Padrão Diário de dezembro de 2017. | 34 |
| Figura 24 - Padrão Diário de agosto de 2015. | 35 |
| Figura 25 - Padrão Diário de agosto de 2016. | 35 |
| Figura 26 - Padrão Diário de agosto de 2017. | 35 |
| Figura 27 - Gráfico da variância dos consumos ao longo de cada hora, para cada dia da semana. | 37 |
| Figura 28 – Gráfico representativos dos valores de α calculados para todas as horas de cada dia da semana. | 38 |
| Figura 29 - Gráfico representativos dos valores de α calculados para todas as horas de cada dia da semana. | 38 |
| Figura 30 - Gráfico comparativo do comportamento das duas distribuições à hora 6 de segunda-feira. | 39 |
| Figura 31 - Gráfico comparativo do comportamento das duas distribuições à hora 6 de domingo. | 39 |
| Figura 32 - Gráfico comparativo do comportamento das duas distribuições à hora 12 de quinta-feira. | 40 |
| Figura 33 - Gráfico comparativo do comportamento das duas distribuições à hora 12 de sexta-feira. | 40 |
| Figura 34 - Gráfico comparativo do comportamento das duas distribuições à hora 20 de terça-feira. | 40 |
| Figura 35 - Gráfico comparativo do comportamento das duas distribuições à hora 20 de sábado. | 40 |
| Figura 36 - Fluxograma explicativo do modelo aplicado para previsão de consumos de eletricidade. | 44 |
| Figura 37 – Gráfico comparativo dos valores dos consumos reais e os valores da previsão do Q50. | 45 |
| Figura 38 – Gráfico da Previsão de diferentes Quantis vs Consumo real para a semana de 18 a 24 de fevereiro de 2018. | 46 |

| | |
|--|----|
| Figura 39 – Gráfico de <i>reliability</i> para avaliação do sucesso da previsão. | 46 |
| Figura 40 - Fluxograma explicativo da variante II do modelo aplicado para previsão de consumos de eletricidade..... | 47 |
| Figura 41 - Gráfico comparativo dos valores dos consumos reais e os valores da nova previsão do Q50..... | 48 |
| Figura 42 - Gráfico da Previsão de diferentes Quantis da Variante II vs Consumo real para a semana de 18 a 24 de fevereiro de 2018..... | 49 |
| Figura 43 - Gráfico de <i>reliability</i> para avaliação do sucesso da previsão..... | 49 |
| Figura 44 - Gráficos dos quantis de previsão em teste para uma semana..... | 50 |
| Figura 45 – Gráfico dos consumos previstos para testar o modelo implementado..... | 50 |
| Figura 46 – Gráficos dos consumos de 1 de março a 30 de junho de 2017..... | 50 |

Lista de tabelas

| | |
|---|----|
| Tabela 1 - Associação entre dia da semana e número auxiliar..... | 30 |
| Tabela 2 - Feriados identificados no período em estudo..... | 30 |

Capítulo 1

Introdução

A presente dissertação foi realizada no âmbito da obtenção do Mestrado Integrado em Engenharia Eletrotécnica e de Computadores – Major Energia da Faculdade de Engenharia da Universidade do Porto.

O estudo realizado consiste na previsão probabilística de consumos agregados para a otimização de estratégias de procura no mercado MIBEL.

A Península Ibérica possui o seu próprio mercado de energia elétrica tal como acontece noutros agregados de países, o Mercado Ibérico da Energia Elétrica – MIBEL.

Este mercado é resultado da cooperação dos Governos de Portugal e Espanha a fim promover a integração dos sistemas elétricos dos dois países e poder obter benefícios para os consumidores, garantindo a todos os interessados igualdade, transparência e objetividade nas condições de acesso.

1.1 Contextualização e Motivação

Os mercados de eletricidade baseados nas curvas da procura e oferta, caracterizam-se sobretudo pela grande volatilidade dos preços da mesma, sendo possível, de uma hora para a outra, verificarem-se variações na ordem das dezenas de euros. Posto isto, o risco de mercado é bastante elevado, logo, torna-se essencial para os agentes de mercado a realização de previsões de consumos que possibilitem a avaliação do risco das suas propostas de compra de energia.

A previsão de consumos de energia elétrica consiste em prever, com a melhor exatidão possível, os valores de consumo de energia para diferentes horizontes temporais. É um processo complexo visto que os consumos podem ser influenciados por diferentes variáveis: dia da semana, do mês, do ano, temperatura, etc., o que faz com que esteja associada uma incerteza a cada previsão efetuada.

Visto que as incertezas associadas a este processo são grandes, é necessário efetuar previsões a curto prazo com a maior precisão possível, para otimizar as ofertas de quantidade de compra em mercado e minimizar os desvios e as correspondentes penalizações.

A previsão probabilística de consumos é essencial para o processo de decisão de risco na compra e venda de eletricidade efetuada com um dia de antecedência.

Deve também ter-se em atenção o risco associado a estas previsões, visto que, se os desvios

entre os valores contratados e os valores reais de consumos forem muito elevados, a influência no mercado será significativa. No caso de o consumo real ultrapassar a quantidade de energia adquirida, será necessário recorrer ao mercado para cobrir a energia em falta, o que levará a sobrecustos desnecessários

É então visível a importância de se realizar uma avaliação cuidada de todas as variáveis que devem integrar a previsão para aproveitar a maior quantidade possível de informação útil quanto aos padrões de consumo. As previsões com maior precisão levarão à obtenção de um lucro mais elevado, pois evitam que hajam entradas em serviço desnecessárias das centrais controláveis e é reduzida a necessidade de recorrer aos mercados para compensações inesperadas. Torna-se vantajoso para os agentes de mercado conhecerem uma estimativa da incerteza associada às previsões, sabendo o grau de dispersão dos valores previstos e as suas probabilidades. Sendo assim, tem-se vindo a revelar que é cada vez mais importante efetuar previsões probabilísticas, que serão alvo de estudo nesta dissertação.

1.2 Objetivos

O objetivo principal desta dissertação passa pelo desenvolvimento e construção de um modelo de previsão probabilística de consumos de energia elétrica para cada uma das 24 horas do dia seguinte.

Para a realização destas previsões, é necessário efetuar uma análise prévia dos dados de modo a entender os padrões de consumo e eliminar todos os resíduos que possam ser prejudiciais e, calcular as variáveis fulcrais para implementar de uma distribuição probabilística que levará à previsão dos consumos.

Na construção do modelo serão considerados essencialmente dados históricos de consumos e é expectável que se cumpram os seguintes objetivos:

- Tratamento adequado de dados para que não existam variáveis prejudiciais à previsão.
- Otimização dos modelos utilizados para realizar a previsão, encontrando o melhor para efetuar a previsão em questão.
- Avaliação do desempenho do melhor modelo de previsão.

1.3 Estrutura da dissertação

A presente dissertação encontra-se dividida em 5 capítulos distintos.

O capítulo 1 é um capítulo introdutivo, onde é feita a contextualização, é descrita a motivação que justifica o presente documento e os objetivos espetáveis.

No capítulo 2 é apresentado o estado da arte, onde são apresentadas abordagens já realizadas no contexto do tema em análise.

No capítulo 3 é realizada a apresentação dos dados que serão utilizados e análise detalhada dos mesmos.

No capítulo 4 é apresentado método utilizado para previsão no contexto dos dados e resultados obtidos.

No capítulo 5, são apresentadas as devidas conclusões e feitas as considerações necessárias.

Capítulo 2

Estado da Arte

A previsão de consumos é um processo crucial no planeamento e funcionamento dos sistemas elétricos, é por isso, que diversas entidades têm investido na investigação e desenvolvimento de modelos precisos e eficazes. Isto é resultado da existência de diferentes metodologias de previsão que divergem em horizonte temporal, erro de previsão e complexidade computacional.

Posto isto, neste capítulo irão ser abordados e analisados, fatores teóricos que devem ser considerados quando são realizadas previsões pontuais de consumo de eletricidade (*Point Forecast*) e probabilísticas (*Probabilistic Forecast*).

Será feita uma abordagem inicial aos fatores que influenciam os padrões de consumo e à classificação das previsões de consumo. Posteriormente, serão apresentados modelos que podem ser utilizados nos dois tipos de previsão.

2.1 Fatores que influenciam os padrões de consumo

As redes de energia elétrica são constituídas por um conjunto de vários pontos de consumo, o que dá origem a diferentes padrões de carga devido às circunstâncias de cada um desses pontos. Torna-se, então, importante realizar um estudo dos fatores que influenciam os padrões de consumo, por forma a serem apresentadas soluções que minimizem a diferença em o consumo real e o consumo previsto.

Existe uma vasta lista de fatores que podem influenciar os padrões de consumo [1], tais como: meteorológicos, sociais, classe de consumidores, etc.

Os fatores podem ser divididos em dois grupos principais [2]: eventos especiais – greves, datas festivas, eventos desportivos; e eventos diários – condições meteorológicas, classe de consumidores.

Os eventos especiais são caracterizados por, geralmente, o momento em que se inicia e a sua duração ser conhecido (períodos festivos como Natal, Páscoa, Carnaval), greves, eleições, eventos desportivos.

Os eventos diários, constituem as características da conduta humana, condições meteorológicas, sendo que estas últimas são de grande importância, visto que devemos ter em

conta temperatura real e aparente e humidade [3][4]. Os comportamentos humanos devem ter-se em grande atenção, pois a necessidade de conforto térmico leva à utilização de equipamentos reguladores de temperatura que irão influenciar os padrões de consumo.

Existem ainda outros fatores que não devem ser desprezados, como o tipo de consumidor (doméstico, comercial, industrial) que levam a diferentes regulações de consumo.

2.2 Horizonte Temporal

Quando existe necessidade de realizar previsões uma das primeiras coisas a fazer após definir a variável a prever é definir o horizonte temporal. Este é definido como sendo o período temporal entre o momento em que a previsão é efetuada e o instante para o qual se quer prever uma determinada variável.

Normalmente, tendo em conta o horizonte temporal, é possível falar de três tipos de previsão: longo, médio e curto prazo.

O que diferencia as previsões destes três tipos de horizontes temporais são, fundamentalmente, o tipo de técnicas e variáveis de influência que são utilizadas, as aplicações e avaliações de desempenho.

Segundo [5][6], é possível definir estes três tipos de previsão da seguinte forma:

Previsões a longo prazo: utilizadas para horizontes temporais que vão desde meses até mesmo a anos, cuja previsão depende essencialmente de previsões meteorológicas. São utilizadas para planeamento e investimentos. Os resultados apresentam-se através de distribuição de probabilidade e recorre-se a métodos de uso final (*end-use*) e abordagens ecométricas;

Previsões a médio prazo: para intervalos que podem ir de alguns dias até alguns meses. Utilizam-se para planeamento financeiro, manutenção de equipamentos, cálculo de balanços, na realocação de recursos e realização de contratos bilaterais. Os métodos são semelhantes aos das previsões a longo prazo;

Previsões a curto prazo: são as mais utilizadas. Estabelecidas para previsões que vão desde uma hora até no máximo uma semana e são de grande importância nas operações diárias dos mercados, para as propostas de compra e venda e em funções de operação de sistemas energéticos. Utilizam-se para prevenir sobrecargas e estimar fluxos de carga, que levam à melhoria da fiabilidade da rede e à redução de ocorrências de falhas nos equipamentos da rede.

Estes três horizontes temporais podem ser utilizados em previsões pontuais ou previsões probabilísticas.

2.3 Previsão Pontual

Ao longo das últimas décadas têm sido desenvolvidos diferentes modelos de previsão pontual, entre os quais estão os modelos Probabilísticos, Multiagente e de Inteligência Artificial, e ainda têm sido cada vez mais utilizados os modelos híbridos, que resultam da junção de modelos diferentes.

A previsão pontual é um tipo de previsão determinista em que o resultado da previsão é um valor de previsão único estimado em cada instante.

2.3.1 Modelos Probabilísticos

Os modelos de previsão probabilísticos, ao contrário da previsão pontual, apresentam como resultado um a representação de múltiplas possibilidades de consumos, na forma de intervalo, quantis ou densidade de probabilidade.

Estes modelos permitem a construção de previsões a partir de uma combinação de valores históricos de consumo com valores atuais ou até valores anteriores de outras variáveis, como dados meteorológicos. As técnicas mais utilizadas são as séries temporais e modelos GARCH.

Este tipo de modelos é muito atrativo visto ser possível ligar interpretações físicas aos seus componentes. As interações entre as variáveis explanatórias são visíveis, ao contrário do que acontece nos modelos de inteligência computacional, o que permite aos operadores de sistema perceberem os seus comportamentos o que as torna ideias para definir que modelos se vão aplicar quanto às variáveis utilizadas.

As séries temporais têm em conta a natureza aleatória dos dados e a sua correlação com o tempo. Os modelos mais utilizados são os *Autoregressive* e *Autoregressive Moving Average*.

A ARMA é uma técnica suportada por séries temporais padrão, que tem em conta a origem aleatória dos dados e a relação temporal do fenómeno em estudo. Os consumos são então previstos através da sua relação linear com variáveis autorregressivas, componente AR (*autoregressive*), e de um termo de ruído, componente média móvel (*moving average*)[7].

Esta técnica impõe que a série temporal seja estacionária, logo é necessário que seja verificada a condição de homoscedasticidade, ou seja, que a média seja um valor nulo e a variância e a covariância sejam valores constantes [8], se esta imposição não se verificar, é necessário proceder à transformação das séries não estacionárias em séries estacionárias. Foi devido a esta exigência que *Box and Jenkins*, através de diferenciação, criaram em 1976 um modelo para superar esta barreira: o *Autoregressive Integrated Moving Average* (ARIMA) ou modelo *Box-Jenkins*. Deve ter-se em conta que se a diferenciação for realizada com *lags* de magnitude superior a 1 em vez de ARIMA o modelo passará a ser SARIMA (*Seasonal Autoregressive Integrated Moving Average*) [7].

Crespo et al. [9] utilizaram várias variantes de modelos AR e ARMA para previsões a curto prazo no mercado alemão EEX, concluindo que a modelação individual de cada hora do dia, bem como o envolvimento de processos probabilísticos simples levavam a uma melhoria dos resultados. Weron e Misiorek [10] realizaram um estudo semelhante, mas para o mercado da Califórnia concluindo que modelos AR com *lags* de 24, 48 e 168 horas, nos quais cada hora é modelada individualmente, conduzem a melhores resultados do que um único modelo geral (SARIMA).

Atualmente, os modelos AR e ARIMA têm sido utilizados em modelos híbridos combinados com outras técnicas mais complexas. Foi proposto por Shafie-khah et al. [11], um método em que primitivamente é utilizada uma transformação *wavelet* para fazer com que a série fique mais estável, de seguida um modelo ARIMA para conceber previsões e por fim uma Rede neuronal com funções de base radial para corrigir a estimação do erro pela previsão *wavelet*- ARIMA.

2.3.2 Modelos Multiagente

Os modelos multiagente apresentam uma maleabilidade muito importante para a análise do comportamento do mercado elétrico. Contudo, focam-se maioritariamente numa análise qualitativa e não quantitativa, pelo que permitem compreender o que influencia a variação dos consumos ou o que determina a estratégia dos diferentes agentes do mercado, porém quando se pretendem resultados quantitativos precisos, como no contexto da previsão de consumos, revelam fraquezas. Têm sido pouco utilizados e normalmente como componentes de modelos híbridos. Uma aplicação de modelos multiagente pode ser consultada em [12].

2.3.3 Inteligência Artificial

Os modelos de inteligência artificial destacam-se pela agilidade e capacidade de lidar com a complexidade e não-linearidade dos dados. Os resultados nas previsões a curto prazo têm-se verificado muito bons, destacando-se especialmente a utilização de redes neuronais artificiais.

Estes modelos são muito flexíveis e são capazes de lidar com problemas complexos e não lineares. Os modelos principais utilizados nesta classe são as redes neuronais, os SVC (*support vector machine*) e *data-mining*.

2.3.3.1 Redes Neuronais Artificiais

As redes neuronais artificiais são uma estrutura composta por unidades computacionais extremamente conectadas (neurónios) com capacidade de aprendizagem. Este processo visa simular o comportamento do cérebro humano, já que o seu conhecimento é adquirido através de interações com ambiente (aprendizagem) e é armazenado nas conexões entre neurónios.

Os neurónios estão organizados num grupo muito pequeno de camadas (ligadas entre si) que se localizam entre os inputs e outputs da rede [13]. Às ligações entre neurónios está associado um fator multiplicativo (peso), sendo que durante o processo de aprendizagem o peso das conexões é adaptado até se atingir um determinado objetivo ou estado de conhecimento da rede.

Sendo assim, [14] a operação das Redes Neuronais Artificiais é iniciada com um processo de treino, onde são fornecidos à rede inputs (variáveis de influência) e outputs (variáveis a prever, que correspondem aos respetivos inputs inseridos), iniciando-se assim o processo de aprendizagem onde a rede vai determinar quais são as melhores relações entre os inputs e outputs, por forma a que o erro médio quadrático entre os outputs produzidos pela rede e os outputs reais sejam mínimos.

Numa fase seguinte, inicia-se a fase de teste de RNA, onde são aplicados os pesos (encontrados na fase de treino) a um conjunto de inputs (com características iguais às dos utilizados na fase de treino) não fornecidos anteriormente à rede, para que possam ser calculados um conjunto de outputs, obtendo-se assim os valores da variável que se pretende prever. Posteriormente realiza-se a avaliação de desempenho da rede.

Existem diferentes tipos de redes neuronais [13] que são classificadas consoante a sua arquitetura e algoritmo de aprendizagem, a primeira refere-se à forma como se organizam as ligações entre neurónios e a segunda, diz respeito à forma como a rede neuronal ajusta os pesos

das suas conexões. Existem dois grupos de RNA: as redes neuronais *feedforward* e as redes neuronais *feedback* (recorrentes). O primeiro grupo pode dividir-se em três sub-grupos: redes de uma só camada, redes multicamada e redes com função de base radial.

2.3.3.2 Support Vector Machine

Durante a aplicação de uma rede neuronal, é necessário proceder-se à otimização de alguns parâmetros da rede (número de camadas escondidas, número de neurónios nas camadas escondidas, etc.). Contudo, não há nenhum algoritmo que permita otimizar estes parâmetros de forma automática, sendo necessário muita pesquisa sobre o tema e adotar um método tentativa-erro para se conseguir chegar à otimização ideal, o que se pode tornar bastante demoroso.

[15] O *Support Vector Machine* (SVM) é considerado uma evolução das redes neuronais por ser um instrumento, que permite tornar este processo muito mais rápido e menos dependente da experiência do utilizador. Este método, a contrário das redes neuronais, faz um mapeamento não linear dos dados (inputs) para um espaço dimensional superior (para tornar o sistema linearmente separável) através de uma função de *kernel*, ou seja, ordena os dados em diferentes grupos e, depois aplica funções lineares simples para criar limites de decisão lineares no novo espaço.

O SVM possui então algumas vantagens, visto ter sido projetado para otimizar os seus parâmetros em função dos inputs utilizados na fase de treino e, porque também é realizada uma otimização quadrática que tem apenas uma solução, não apresentando várias soluções possíveis, logo a qualidade dos resultados pode ser considerada superior relativamente às redes neuronais.

As técnicas de *data-mining*, consistem em processos de análise dos inputs, com o objetivo de organizá-los em categorias (*clustering*) mediante a identificação de padrões. Este método ainda não foi muito analisado e estudado, contudo tem vindo a ganhar popularidade [16].

2.4 Previsão Probabilística

A literatura sobre previsão probabilística de consumos elétricos é muito reduzida quando comparada com previsões probabilísticas em geral [17] ou por exemplo comparando com previsão probabilística de energia eólica [18][19]. No entanto, a previsão probabilística de consumos deve ser considerada tão importante quanto a de energia eólica principalmente no setor público. Segundo o exemplo dado em [1], um utilizador médio dos EUA com um consumo anual de 1GW-10GW, num típico dia de folga, o erro de previsão de carga é de cerca de 3%, enquanto que num dia normal o erro de previsão de energia eólica é de cerca de 15%. Se a penetração do vento for cerca de 20%, então, em média, o erro absoluto das previsões de consumos é semelhante ao da previsão de energia eólica. Como foi discutido por Hong, T. [20], um erro de previsão de consumo de 1% em termos de percentagem de erro absoluto (MAPE) pode traduzir-se em prejuízos de centenas de milhares de dólares por GW para uma empresa.

Contudo, já há alguns anos que a previsão de consumos tem vindo a ser estudada. Em 1987, foi publicado um pequeno tutorial por Gross, G. e Galiana, F. D. [21] sobre previsões de consumos a curto prazo onde fazia referência a aplicações, fatores que influenciam as previsões, técnicas, entre outros.

Mais recentemente, Hong, T. [22] publicou um artigo onde apresenta a evolução da previsão

probabilística na indústria de energia, onde é dada grande importância à previsão de consumos, enfatizando a importância da realização de testes rigorosos para que as previsões sejam o mais precisas possível para evitar prejuízos agravados às empresas.

A previsão probabilística consiste na determinação da previsão de uma variável aleatória indicando a incerteza que lhe está associada. Pode ser expressa através de medidas probabilísticas como a função densidade de probabilidade, função distribuição acumulada, conjunto de quantis ou momentos de distribuição. Existem duas possíveis abordagens para estas técnicas: paramétrica e a não paramétrica. Atualmente, existem diferentes métodos não paramétricos de regressão por quantis dos quais se destacam a Regressão Linear de Quantis, Modelos Aditivos de Regressão de Quantis por *Boosting*, *Quantile Regression Forests*, Redes Neurais e Regressão no RKHS.

A regressão linear de quantis é o método clássico de regressão por quantis e prevê uma relação linear entre as variáveis explanatórias e a variável a prever. Porém, em aplicações reais são raros os processos precisamente lineares, pelo que normalmente é considerada uma extensão da regressão linear a modelos mais flexíveis, os modelos aditivos de regressão de quantis. Estes combinam a regressão de quantis com uma estrutura aditiva que permite modelar de uma forma mais flexível o efeito das variáveis explanatórias na variável a prever.

Outra opção para a estimação dos modelos aditivos de regressão por quantis é a utilização de algoritmos de *boosting*, particularmente o *component-wise gradient boosting*, que se destaca por efetuar uma seleção automática de variáveis no processo de construção dos modelos.

O início das restantes técnicas vem de algoritmos de inteligência artificial e são capazes de lidar com problemas de elevada dimensão e modelar diferentes tipos de relações. A *quantile regression forests* é caracterizada por estimar diretamente a função distribuição acumulada pelo que não se verifica o problema da ocorrência de sobreposição de quantis. As redes neurais para regressão por quantis emergem da substituição da função custo do algoritmo tradicional por uma aproximação da função custo utilizada na regressão linear de quantis. Por último a regressão por quantis no RKHS tem a sua origem nos SVM. Nesta técnica é assumido que a função de regressão pertence a um RKHS, sendo penalizada por uma norma definida no mesmo de modo a prevenir a ocorrência de *overfitting*.

2.4.1 Intervalos Estatísticos

Um intervalo de confiança provê uma medida de confiança para um determinado estimador de um parâmetro da distribuição. Assume-se que o parâmetro é desconhecido e o intervalo de confiança é construído a partir de uma amostra. Assim os intervalos de confiança apenas avaliam o erro intrínseco ao processo de determinar um parâmetro populacional a partir de uma amostra, não fornecendo indicação sobre a distribuição de valores individuais [23][24].

Para estimar a incerteza de $E(Y|X = x)$ devemos construir um intervalo de confiança, todavia esse intervalo irá desprezar a incerteza quando se assume \hat{y}_{t+k} como estimativa de $Y_{t+k}|(X = x)$. Para isso é então necessário considerar intervalos de previsão, que também consideram a variabilidade da distribuição condicional $Y|(X = x)$.

O intervalo de previsão $\hat{I}_{t+k|t}^{(\alpha)}$ estimado no instante t para o horizonte $t + k$, é o intervalo que se espera que contenha y_{t+k} com uma probabilidade $(1 - \alpha)$, designada por cobertura nominal,

$$P\left(y_{t+k} \in \hat{I}_{t+k|t}^{(\alpha)}\right) = P\left(y_{t+k} \in \left[\hat{L}_{t+k|t}^{(\alpha)}, \hat{U}_{t+k|t}^{(\alpha)}\right]\right) = 1 - \alpha \quad (2.1)$$

em que $\hat{L}_{t+k|t}^{(\alpha)}$ e $\hat{U}_{t+k|t}^{(\alpha)}$ são respetivamente o extremo inferior e superior do intervalo de previsão.

Normalmente, os intervalos de previsão são intervalos de previsão centrais, isto é, a probabilidade de obter um valor acima ou abaixo dos limites do intervalo é igual. Posto isto, os extremos do intervalo corresponderão aos quantis com proporções $\left(\frac{\alpha}{2}\right)$ e $\left(1 - \frac{\alpha}{2}\right)$ da função de distribuição estimada, estando centrados relativamente à mediana da distribuição estimada.

$$\hat{L}_{t+k|t}^{(\alpha)} = \hat{Q}_{t+k|t}^{(\frac{\alpha}{2})}, \quad P\left(y_{y+k} < \hat{L}_{t+k|t}^{(\alpha)}\right) = \frac{\alpha}{2}$$

$$\hat{U}_{t+k|t}^{(\alpha)} = \hat{Q}_{t+k|t}^{(1-\frac{\alpha}{2})}, \quad P\left(y_{y+k} < \hat{U}_{t+k|t}^{(\alpha)}\right) = 1 - \frac{\alpha}{2}$$

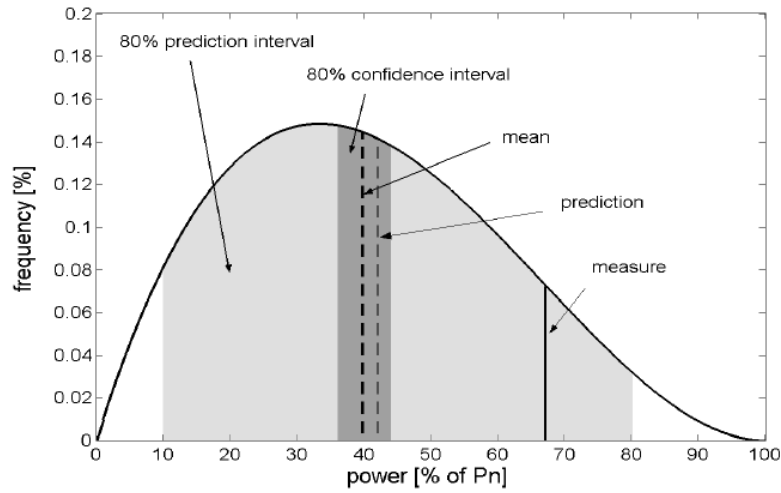


Figura 1 - Exemplo ilustrativo da diferença entre intervalo de confiança (área cinza claro) e intervalo de previsão (área cinza escuro) [25].

2.4.2 Regressão Linear de Quantis

Introduzida por Basset e Koenker [26], a regressão linear de quantis é uma técnica estatística para estimação de quantis de uma distribuição condicional.

Para uma variável aleatória Y , e um vetor x de variáveis explanatórias, a função distribuição condicional $F_y(y | X = x)$ é dada por:

$$F(y|X = x) = P(Y \leq y | X = x)$$

Logo, o quantil condicional de ordem τ , com $0 < \tau < 1$, pode ser definido como,

$$Q(\tau, x) = \inf\{y: F(y|X = x) \geq \tau\}$$

$Q(\tau, x)$ pode ser expresso como uma combinação linear das variáveis explanatórias com coeficientes desconhecidos, sendo o quantil condicional de ordem τ modelado por:

$$Q(\tau, x) = \beta_0(\tau) + \beta_1(\tau)x_1 + \dots + \beta_p(\tau)x_p \quad (2.2)$$

x e p são as variáveis explanatórias e $\beta_i(\tau)$ coeficientes desconhecidos dependentes de τ , e a estimar a partir das observações $(y_i, x_{i,1}, \dots, x_{i,p})$, $i=1, \dots, N$.

Utilizando a função consumo ρ_τ definida por:

$$\rho_\tau(\varepsilon) = \begin{cases} \tau\varepsilon & \text{se } \varepsilon \geq 0 \\ (\tau - 1)\varepsilon & \text{se } \varepsilon < 0 \end{cases}$$

e considerando $\beta(\tau)$ o vetor que contém os coeficientes, obtém-se que o melhor estimador de $\beta(\tau)$ é a solução do problema de otimização:

$$\hat{\beta}(\tau) = \arg \min_{\beta} \sum_{i=1}^N \rho_\tau[y_i - (\beta_0(\tau) + \beta_1(\tau)x_{i,1} + \dots + \beta_p(\tau)x_{i,p})] \quad (2.3)$$

Dado que em aplicações reais os processos estritamente lineares são poucos é vantajoso permitir alguma flexibilidade ao modelo, para tal uma das soluções pode ser considerar $Q(\tau, x)$ como um modelo aditivo. Este tipo de modelo pode ser expresso através de:

$$y = \alpha + f_1(x_1) + f_2(x_2) + \dots + f_p(x_p) + \varepsilon \quad (2.4)$$

em que α é uma constante e $f_j(x_j)$ funções que podem ser aproximadas por uma combinação linear de funções base da variável explanatória correspondente, ou seja,

$$f_j(x_j) = \sum_{k=1}^{n_j} b_{jk}(x_j)\theta_{jk} \quad (2.5)$$

onde $b_j(x_j)$ são funções base e θ_j coeficientes desconhecidos.

Comparando com (2.2), $Q(\tau, x)$ pode ser modelo por:

$$Q(\tau, x) = \alpha(\tau) + \sum_{j=1}^p f_j(x_j; \tau) = \alpha(\tau) + \sum_{j=1}^p \sum_{k=1}^{n_k} b_{jk}(x_j)\theta_{jk}(\tau) \quad (2.6)$$

com funções $f_j(\cdot)$, $j=1, \dots, p$ sujeitas à restrição $f_j(0) = 0$.

2.4.3 Quantile Regression Forests

O método proposto por Meinshausen [27], Quantile Regression Forests, consiste na regressão por quantis baseado na técnica random forests.

Random forests são um conjunto de árvores que crescem com N observações independentes $(y_i, x_{i,1}, \dots, x_{i,p})$, $i = 1, \dots, N$.

Para cada árvore considera-se apenas uma parte dos dados do conjunto de treino, e um subconjunto aleatório das variáveis explanatórias para realizar as divisões nos nós.

Sendo θ o vetor de parâmetros aleatórios que determina quais as variáveis consideradas para divisão em cada um dos nós, $T(\theta)$ corresponde à árvore e $B \subseteq \mathbb{R}^p$, onde p é o número de variáveis explanatórias, o espaço dimensional que contém x . Cada folha $l = 1, \dots, L$ da árvore acomoda um subespaço retangular de B , designado por R_l . Para cada $X \in B$, existe uma e uma só

folha tal que $x \in R_l$. Para cada $T(\theta)$ designa-se essa folha por $l(x, \theta)$.

A previsão de uma única árvore $T(\theta)$ para uma nova observação é obtida a partir da média dos valores observados na folha $l(x, \theta)$. Seja $\omega_i(x, \theta)$, $i = 1, \dots, N$ um vetor de pesos definido por uma constante positiva se x_i pertence à folha $l(x, \theta)$ e zero caso contrário. A soma dos pesos será igual a um e assim resulta:

$$\omega_i(x, \theta) = \frac{1_{\{x_i \in R_{l(x, \theta)}\}}}{\#\{j: X_j \in R_{l(x, \theta)}\}} \quad (2.7)$$

A previsão para uma única árvore, dado $X = x$ será a média ponderada das observações originais, y_i , $i = 1, \dots, N$,

$$\hat{\mu}(x) = \sum_{i=1}^N \omega_i(x, \theta) y_i \quad (2.8)$$

No *random forests* a média condicional $E(Y|X = x)$ é estimada pela média das previsões de k árvores, construídas a partir de vetores i.i.d θ_t , $t = 1, \dots, k$. Se $\omega_1(x)$ for a média dos $\omega_i(\theta)$ das k árvores,

$$\omega_1(x) = k^{-1} \sum_{i=1}^N \omega_i(x, \theta_t) \quad (2.9)$$

A previsão da *random forests* será então,

$$\hat{\mu}(x) = \sum_{i=1}^N \omega_i(x) y_i \quad (2.10)$$

A distribuição condicional de Y , dado $X = x$ é dada por:

$$F(y|X = x) = P(Y \leq y | X = x) = E(1_{\{Y \leq y\}} | X = x) \quad (2.11)$$

Realizando uma analogia com a expressão da média condicional, é possível definir uma previsão para $E(1_{\{Y \leq y\}} | X = x)$ como uma média ponderada das observações $1_{\{Y \leq y\}}$, usando os pesos como em (2.9),

$$\hat{F}(y|X = x) = \sum_{i=1}^N \omega_i(x) 1_{\{Y \leq y\}} \quad (2.12)$$

O algoritmo proposto para estimar $\hat{F}(y|X = x)$ é então:

- i. Construir k árvores $T(\theta_t)$, $t = 1, \dots, k$ como no algoritmo das *random forests*. No entanto, para cada folha deve tomar-se nota de todas as observações e não apenas da média.
- ii. Para um certo $X = x$, propaga-lo através de cada uma das k árvores. Determinar $\omega_i(x, \theta_t)$ da observação $i \in \{1, \dots, N\}$ para cada árvore como definido em (2.7). Calcular $\omega_i(x)$ para toda a observação $i \in \{1, \dots, N\}$ como uma média sobre $\omega_i(x, \theta_t)$, $t = 1, \dots, k$ como em 2.9.
- iii. Calcular o estimador para a função distribuição como definido em ii).

2.4.4 Modelos Aditivos de Regressão de Quantis por *Boosting*

Os modelos aditivos de regressão de quantis podem ser estimados utilizando algoritmos de *boosting*, dos quais se destaca o *componente-wise gradient boosting*. Esta técnica de aprendizagem automática para otimização de previsões ou estimação de modelos estatísticos utiliza o método do gradiente descendente. Uma das características mais importantes deste método é o fato de durante o processo de construção dos modelos se proceder a uma seleção de variáveis sem depender de técnicas heurísticas.

Este método tem como objetivo a estimação da função de previsão ótima, f^* , definida por:

$$f^* = \operatorname{argmin}_f E_{Y,X}[\rho(y, f(x))] \quad (2.13)$$

Onde a função custo ρ é diferenciável relativamente a f . Para regressão de quantis a função custo é a definida em (2.3). Apesar de não ser diferenciável em zero, na prática, uma vez que a variável resposta é contínua, podemos ignorar esse fato definindo,

$$-\frac{\partial \rho_t(\varepsilon)}{\partial f} = \begin{cases} \tau & \varepsilon \geq 0 \\ \tau - 1 & \varepsilon < 0 \end{cases}$$

Visto que, na prática o valor esperado definido em (2.13) é desconhecido, os algoritmos de *boosting* minimizam o erro quadrático médio observado, $\mathcal{R} = \sum_{i=1}^N \rho(y_i f(x_i))$.

O algoritmo utilizado para minimizar \mathcal{R} é [28]:

- i. Inicializar a função $\hat{f}^{[0]}$ com valores offset. De notar que $\hat{f}^{[0]}$ é um vetor de comprimento N . No seguimento do método, $\hat{f}^{[m]}$, refere-se ao vetor estimado na iteração m .
- ii. Distinguir um conjunto de *base-learners*. *Base-learners* são estimadores simples de regressão com um conjunto fixo de variáveis de entrada e uma resposta univariada. O conjunto de variáveis de entrada pode variar nos diferentes *base-learners* e normalmente são subconjunto, de pequena dimensão, das variáveis explanatórias. Seja P o número de *base-learners* e defina-se $m = 0$.
- iii. Incrementar m por 1, sendo m o número de iterações.
- iv. Calcular o gradiente negativo $-\frac{\partial \rho}{\partial f}$ da função custo e avaliá-lo em $\hat{f}^{[m-1]}(x_i)$, $i = 1, \dots, N$, o que produz o vetor gradiente negativo,

2.4.5 Redes Neurais para Regressão de Quantis

A utilização de redes neuronais para a regressão de quantis foi introduzida por Taylor [29]. Consiste na adoção da função ρ_τ , definida em (2.3), como um caso especial da função custo do algoritmo clássico de redes neuronais. Contudo, o processo de otimização das redes neuronais hipoteticamente requer uma função diferenciável em todos os pontos, o que não se verifica para a função ρ_τ que não é diferenciável na origem, sendo que não é, ainda, claro se poderão existir problemas de convergência.

Cannon [30] propõe a substituição da função ρ_τ por uma diferenciável em todos os pontos, utilizando para tal uma aproximação sugerida inicialmente por Chen [31].

Dada a função Huber, para um dado *threshold* δ , definida por,

$$h(\varepsilon) = \begin{cases} \frac{\varepsilon^2}{2\delta} & \text{se } 0 \leq |\varepsilon| \leq \delta \\ |\varepsilon| - \frac{\delta}{2} & \text{se } |\varepsilon| > \delta \end{cases} \quad (2.14)$$

Então a função ρ_τ pode ser aproximada por,

$$\rho_\tau^* = \begin{cases} \tau h(\varepsilon) & \text{se } \varepsilon \geq 0 \\ (\tau - 1)h(\varepsilon) & \text{se } \varepsilon < 0 \end{cases} \quad (2.15)$$

2.4.6 Regressão de Quantis no Reproducing Kernel Hilbert Space

O algoritmo da regressão de quantis no RKHS tem como objetivo encontrar a função f que minimiza o risco esperado

$$E_{Y,X}[\rho_\tau(y - f(x))]$$

em que $\rho_\tau(\cdot)$ é a função custo definida em (2.2).

Visto que na prática este valor esperado é desconhecido, para solucionar este problema, pode minimizar-se o funcional regularizado,

$$\frac{1}{n} \sum_{i=1}^n \rho_\tau(y_i - f(x_i)) + \frac{\lambda}{2} \|g\|_{\mathcal{H}}^2 \quad (2.16)$$

onde $f = g + b$ com $b \in \mathbb{R}$ e $g \in \mathcal{H}$, \mathcal{H} é um RKHS gerado por um kernel $k(x, x') = \langle \phi(x), \phi(x') \rangle$, $\|\cdot\|_{\mathcal{H}}$ uma norma no RKHS e λ um parâmetro de regularização.

Para que possa ser realizada uma implementação computacional mais eficiente é necessário formular o problema de otimização de forma dual. Usando a conexão entre RKHS e o espaço de características (feature space), podemos definir $f(x) = \langle \phi(x), w \rangle + b$ obtend o seguinte problema de otimização,

$$\min_{w,b,\varepsilon_i^*} C \sum_{i=1}^m \tau \varepsilon_i + (1 - \tau) \varepsilon_i^* + \frac{1}{2} \|w\|^2, \text{ em que } C = \frac{1}{\lambda m}$$

$$\text{Com as restrições: } \begin{cases} \varepsilon_i, \varepsilon_i^* \geq 0 \\ y_i - \langle \phi(x_i), w \rangle - b \leq \varepsilon_i \\ \langle \phi(x_i), w \rangle + b - y_i \leq \varepsilon_i^* \end{cases}$$

A forma dual deste problema pode ser obtida diretamente usando multiplicadores de Lagrange α_i , obtendo-se:

$$\min_{\alpha} \frac{1}{2} \alpha^T K \alpha - \alpha^T y \quad (2.17)$$

$$\text{Com as restrições: } \begin{cases} \mathbf{1}^T \alpha = 0 \\ C(\tau - 1) \leq \alpha_i \leq C\tau \text{ para todo } 1 \leq i \leq m \end{cases}$$

A função f é obtida através da expansão da família do *kernel*, $w = \sum_i \alpha_i \phi(x_i)$ ou equivalente $f(x) = \sum_i \alpha_i k(x_i, x) + b$.

Podem ser consultados mais pormenores sobre este método em [32].

2.4.7 KDE (Kernel Density Estimation)

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

A estimação é caracterizada pelo *kernel* [29] que determina a forma da função de peso (*weighting function*), ou seja, o peso a ser conferido à vizinhança dos dados, bem como a largura de banda, que representa a distância para a qual é efetuada a procura entre dados vizinhos na função de peso. A esta largura está relacionada com a quantidade de suavização (*smoothing*) que será necessário fazer.

Então, em primeiro lugar, é necessário verificar alguns pontos: assumir uma fdp (função normal ou beta); estimar parâmetros (esta estimação é realizada com base no tipo de *kernel* e na largura de banda, por forma a otimizar as propriedades de $\hat{f}(x)$); efetuar um teste de precisão, visto que se trata de uma estimação, é necessário o teste precisão do método para que, caso a largura de banda não tenha sido a mais adequada, o processo seja reinicializado pela escolha de uma nova distribuição.

Nos problemas de previsão, a estimação da função de densidade condicional futura tem um papel relevante visto ser ela que descreve a relação existente entre as variáveis explanatórias e as variáveis alvo.

Pode ver-se esta estimativa como uma generalização da regressão devido ao fato de a estimativa de densidade condicional ter como objetivo a obtenção de uma função densidade probabilidade conjunta $f_{y|x}(y|x)$ (2.18), enquanto que a regressão tem como objetivo a estimação média condicional $E(y|x)$ [30].

$$f_{y|x}(y|x) = \frac{f_{y,x}(y|x)}{f_x(x)} \quad (2.18)$$

$f_x(x)$, função densidade marginal.

A fdp que se obtém nestes problemas, pode ser utilizada para representar a incerteza.

Para os problemas de previsão apresentados em (2.19) a estimação da fdp condicional futura é realizada para uma certa variável aleatória, num horizonte temporal $t+k$, tendo em conta um conjunto de aprendizagem com N amostras compostas por pares de (X_n, P_N) resumindo toda a informação histórica disponível até ao instante t . Cada par constitui um conjunto de variáveis explanatórias X_n e no valor medido correspondente da variável a prever. É também assumido que para o instante t , é conhecido um conjunto de variáveis explanatórias $X_{t+k/t}$ para o horizonte temporal onde se quer prever.

$$\hat{f}_p(p_{t+k}|x_{t+k|t}) = \frac{f_{p,x}(p_{t+k}|x_{t+k|t})}{f_x(x_{t+k|t})} \quad (2.19)$$

p_{t+k} , é a produção prevista para o horizonte temporal $t+k$.

Visto que as densidades conjuntas e marginais não são conhecidas, as expressões vão ser estimadas através do procedimento do KDE, para uma variável, usando as amostras históricas disponíveis, através de:

$$\hat{f}_h(x) = \frac{1}{N \times h} \sum_{i=1}^N K \frac{x-X_i}{h} \quad (2.20)$$

Com,

- h : largura de banda.
- N : número de amostras.
- K : função de *kernel*.
- X_i : amostras num dado instante i .

Relativamente à função de *kernel*, existem vários tipos [31] e, podemos seleccionar qualquer um, uma vez que o impacto na qualidade da previsão é mínimo independentemente de qual se escolha quando comparado com a largura de banda escolhida. Contudo, o *kernel* mais utilizado é o gaussiano:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \quad (2.21)$$

2.4.7.1 Função densidade de probabilidade da distribuição beta

Após ser aplicado o KDE, o resultado que se obtém será aplicado a uma distribuição beta (2.22), pois trata-se de uma boa aproximação para a modelização de variáveis, com um limite máximo (A) e um limite mínimo (B).

$$f(x, \alpha, \beta) = \left(\frac{x^{\alpha-1} \times (1-x)^{\beta-1}}{B[\alpha, \beta]} \right) \quad (2.22)$$

Com,

- x : definido no intervalo $[0,1]$, que garante o domínio entre $[0, 1]$.
- B : função beta.
- α : coeficiente de forma da função beta, maior que 0.
- β : coeficiente de forma da função beta, maior que 0.

Os coeficientes de forma, α e β definem a forma da distribuição beta e podem ser determinados através das expressões apresentadas a seguir.

$$\alpha = \mu \left(\frac{\mu(1-\mu)}{\nu} - 1 \right) \quad (2.23)$$

$$\beta = (1 - \mu) \left(\frac{\mu(1-\mu)}{\nu} - 1 \right) \quad (2.24)$$

Onde, μ e ν são os valores da média e variância originais, respetivamente.

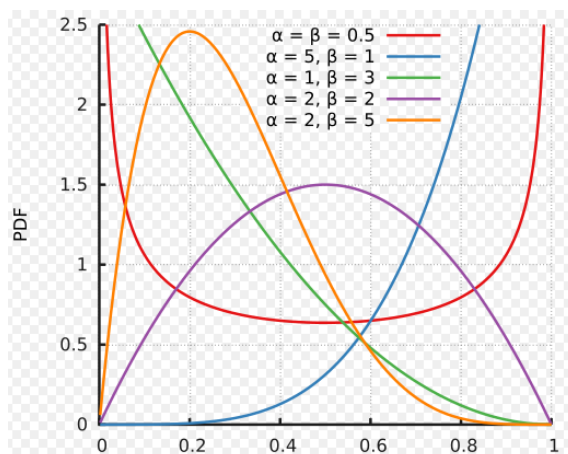


Figura 2 - Função densidade de probabilidade da distribuição beta para valores diferentes de alfa e beta. [33]

2.4.7.2 Funcionamento do KDE

O KDE faz uma seleção dos pontos próximos e cria uma gaussiana artificial, concedendo mais peso aos vizinhos mais próximos. O intervalo das variáveis não pode ser muito pequeno, pois isso pode originar um sobre ajuste do modelo aos dados de treino (*overfitting*) existindo então a necessidade de se determinar a largura ótima para cada variável. Contudo, 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 então executado um processo iterativo por forma a diminuir o erro do modelo e da incerteza, por alteração dos parâmetros das diferentes variáveis.

Após execução do programa obtém-se os valores, a seguir apresentados, que serão usados para estudos da precisão do modelo e da incerteza:

C – Consumo previsto.

α e β – fatores 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.

2.5 Medidas de Avaliação

No que diz respeito à avaliação de desempenho de previsões probabilísticas, pode falar-se em avaliação de duas propriedades: *reliability* e *sharpness*.

Para que os modelos sejam eficazes e precisos, deve efetuar-se uma escolha correta das variáveis a serem utilizadas, pelo que é necessário também, numa fase inicial, perceber se existe correlação entre os dados.

2.5.1 Coeficiente de determinação e Coeficiente de correlação de *Pearson*

O coeficiente de determinação, R^2 , mede a porcentagem da variância de uma das variáveis que pode ser explicada pela variância da outra. Pode variar entre 0 e 1. O valor 0 corresponde à inexistência de relação e 1 corresponde à existência e relação linear ótima entre as variáveis [34]. Este coeficiente pode ser calculado através de:

$$R^2 = 1 - \frac{\frac{1}{m} \sum_{t=1}^m (P_t - \hat{P}_t)^2}{\sum_{t=1}^m (P_t - P)^2} \quad (2.25)$$

Onde:

$$P = \frac{1}{m} \sum_{t=1}^m P_t \quad (2.26)$$

Esta variância também pode ser calculada pelo quadrado do coeficiente de correlação de *Pearson* (R) que mostra a influência que o valor de uma variável aleatória tem sobre os seus vizinhos, o que a torna uma ferramenta muito importante na determinação de padrões. Este coeficiente varia entre -1 e 1, sendo que:

$R = 1$ – mostra que existe uma relação perfeita positiva entre duas variáveis, ou seja, que são diretamente proporcionais;

$R = 0$ – mostra que não existe uma relação linear entre as variáveis, embora possam ter outro tipo de relação;

$R = -1$ – mostra a presença de uma relação perfeita negativa entre as duas variáveis, ou seja, as variáveis são inversamente proporcionais.

O coeficiente de correlação é calculado através de:

$$R(k) = \frac{E[(Y_t - \mu)(Y_{t+k} - \mu)]}{\sigma^2} \quad (2.27)$$

Em que,

- Y_t , variável aleatória, discreta e estacionária;
- k , distância com que se pretende medir a variável Y (se $k=0$, o coeficiente de correlação é 1, o que significa que a variável está correlacionada com ela própria);
- μ , valor da média.
- E , valor esperado.
- σ^2 , variância ou quadrado do desvio-padrão.

2.5.2 Índices de Erro

Para avaliar a performance de um modelo em termos de previsões pontuais as medidas mais utilizadas são o *Mean Absolute Error* (MAE) e o *Root Mean Square Error* (RMSE), [34] definidas

por:

$$MAE = \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t| \quad e \quad RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2}$$

em que y_t e \hat{y}_t são respetivamente o valor observado e previsto para o instante t .

Em [35] é possível encontrar descrito detalhadamente um procedimento para avaliação de previsões probabilísticas. O conjunto de teste é formado por uma série de quantis estimados, para uma ou várias coberturas nominais.

É possível utilizar também outras medidas para a avaliação da performance, como o *Mean Absolute Percentage Error* (MAPE) e o *Mean Square Error* (MSE):

$$MAPE = \frac{\frac{1}{n} \sum_{t=1}^n |C_t - \hat{C}_t|}{\frac{1}{n} \sum_{t=1}^n |C_t|} \quad (2.28)$$

$$MSE = \frac{1}{n} \sum_{t=1}^n (\hat{C}_t - C_t)^2 \quad (2.29)$$

Em que:

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

Do conjunto de medidas possíveis para avaliação, são de destacar a calibração, a *sharpness* e o *continuous ranked probability score* (CRPS) que serão descritos nos parágrafos seguintes.

2.5.3 Reliability e Sharpness

A *reliability* (ou *calibration*) refere-se à consistência estatística entre distribuições de previsão e observações. Divulga a capacidade da previsão probabilística de corresponder eficazmente a uma avaliação a longo prazo, ou seja, se se prevê que a uma determinada hora a probabilidade do consumo de eletricidade atingir um determinado valor é 95%, é espetável que a longo prazo, para casos idênticos, 95% das vezes o consumo atinja esse valor. [19][36] É possível quantificar esta propriedade através do uso de histogramas de ranking como proposto por Talagrande et al. [37], que são utilizados para avaliar previsões de conjuntos discretos.

Em previsões contínuas pode utilizar-se a *probability integral transform* (PIT), que é traduzida por:

$$p_t = F_t(x_t)$$

Onde F_t representa a distribuição da previsão probabilística e x_t é um número aleatório da função G_t , que representa a distribuição real da probabilidade e é um conceito hipotético. Se a previsão for ideal (numa previsão perfeita $F_t = G_t$) e F_t for contínuo, p_t terá uma distribuição uniforme, tornando-se assim possível avaliar a previsão. Contudo, segundo [8] e [37] este método de avaliação tem algumas falhas, pois embora uma distribuição calibrada demonstre uma PIT

uniforme, uma PIT uniforme não implica necessariamente uma distribuição calibrada.

Quando se realiza a calibração de uma previsão existem quatro cenários possíveis:

Overdispersed Forecast – Neste cenário o histograma terá a forma de uma lomba. Existe uma dispersão elevada na Função Densidade de Probabilidade, o que incita uma elevada frequência de casos reais previstos na zona central da distribuição. Para resolver esta questão, uma possível solução passa por aumentar os valores de α e β da distribuição (equações 2.23 e 2.24, respetivamente).

Underdispersed Forecast – Neste caso o histograma terá a forma de U. A Função Densidade de Probabilidade apresenta uma baixa dispersão, que provoca uma frequência elevada de casos reais previstos nos extremos da distribuição e uma baixa frequência na zona central. Aqui a solução passaria por diminuir os valores de α e β , ao contrário do caso anterior.

Calibrated Forecast – A Função Densidade Probabilidade apresenta uma distribuição uniforme, previsão ideal, e acompanha a frequência de casos reais. Sendo assim, não é necessário alterar os valores de α e β .

Points Out of Distribution – Estas situações só ocorrem caso a distribuição seja limitada e, numa situação destas a distribuição é incapaz de caracterizar parte dos valores reais.

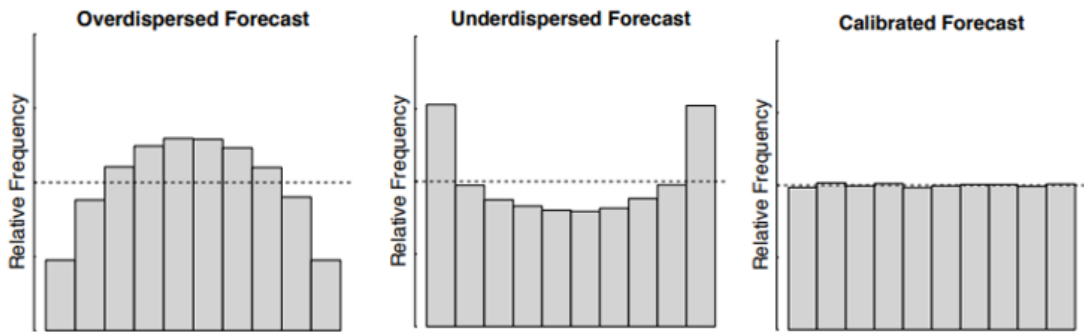


Figura 3 - Histogramas representativos dos diferentes cenários de *reliability*.

A calibração deve ser calculada para cada um dos quantis estimados e mede o quanto estes diferem dos quantis nominais a que se referem. Para estimar a calibração de um quantil é necessário inserir uma variável indicadora, $\xi_{t,k}^{(\tau)}$. Dado um quantil estimado no instante t para $t+k$, $\hat{Q}_{t+k|t}^{(\tau)}$, e o consumo observado, y_{t+k} , então,

$$\xi_{t,k}^{(\tau)} = 1_{t+k < \hat{Q}_{t+k|t}^{(\tau)}} = \begin{cases} 1, & \text{se } y_{t+k} < \hat{Q}_{t+k|t}^{(\tau)} \\ 0, & \text{caso contrário} \end{cases} \quad (2.30)$$

Um estimador para a cobertura atual é definido por:

$$\hat{a}_k^\tau = E[\xi_{t,k}^\tau] = \frac{1}{N} \sum_{t=1}^N \xi_{t,k}^\tau \quad (2.31)$$

Pelo que o desvio relativamente a uma calibração perfeita, r_k^τ é dado por,

$$r_k^\tau = \tau_k - \hat{a}_k^\tau. \quad (2.32)$$

A *sharpness* apresenta a capacidade do modelo para prever acontecimentos de uma forma explícita, em vez de os prever com incertezas demasiado elevadas, ou seja, previsões que têm uma distribuição uniforme e que têm pouca utilidade para quem tomará decisões baseadas nas mesmas. Ao contrário da *reliability* esta característica é avaliada independentemente da comparação dos valores obtidos na previsão com os valores reais, mas sim pela observação das concentrações das distribuições probabilísticas da previsão – quanto menos dispersa for a distribuição mais precisa é a previsão. [19] [36]

Em [37] é proposto que seja realizada a avaliação da *sharpness* através de gráficos em forma de diagramas de caixas, que serão denominados por diagramas de *sharpness*. Por outro lado, em [38] foi proposta a avaliação do *sharpness* através de um indicador que representa a média das distâncias entre quantis da FDP, (2.33).

$$\bar{\delta}_k^\beta = \frac{1}{N} \sum_{t=1}^N \delta_{t,k}^{(\beta)} = \frac{1}{N} \sum_{t=1}^N \left(\hat{Q}_{t+k|t}^{(1-\frac{\beta}{2})} - \hat{Q}_{t+k|t}^{(\frac{\beta}{2})} \right). \quad (2.33)$$

Com,

$$\delta_{t,k}^{(\beta)} = \hat{Q}_{t+k|t}^{(1-\frac{\beta}{2})} - \hat{Q}_{t+k|t}^{(\frac{\beta}{2})}$$

Para se obter uma previsão como boa performance, é apresentado em [37] como objetivo a maximização da *sharpness* da distribuição da previsão, condicionada pela *reliability*, ou seja, diminuir ao máximo as incertezas da previsão sem colocar em risco a fiabilidade da mesma. Sendo assim, a fiabilidade da previsão probabilística não pode ser avaliada apenas por apenas uma destas propriedades, mas sim pelas duas em conjunto. Neste contexto, é possível inserir as *skill score* que consistem num único critério de avaliação que contem informação relativa às propriedades *reliability* e *sharpness*. Estas são definidas para que a previsão com melhor qualidade tenha a melhor pontuação e têm como principal vantagem a compactação de informação, contudo, a contribuição de cada uma das suas propriedades é impercetível. [19] [38]

Existem diversas medidas de pontuação para representação de incertezas, mas, as *logarithmic score* são bastante sensíveis a eventos extremos (picos de valores), no entanto, a *continuous ranked probability score* (CRPS – 2.34) é mais robusta a lidar com casos destes. [19] Desta, surgem ainda as *average continuous ranked probability score* (\overline{CRPS} – 2.35) e a *continuous ranked probability skill score* (CRPSS – 2.36). [39]

$$CRPS = \int_{-\infty}^{+\infty} (\hat{F}(x) - F_0(x))^2 dx \quad (2.34)$$

Onde, $F_0(x)$ representa a função Heaviside que pode ser expressa por:

$$F_0(x) = H(x - x_0) = \begin{cases} 0, & \text{se } x - x_0 < 0 \\ 1, & \text{se } x - x_0 \geq 0 \end{cases}$$

O \overline{CRPS} será então:

$$\overline{CRPS} = \frac{1}{N} \sum_{i=1}^N CRPS_i \quad (2.35)$$

E o CRPSS é definido por:

$$CRPSS = 1 - \frac{\overline{CRPS}_{model}}{\overline{CRPS}_{ref}} \quad (2.36)$$

Onde \overline{CRPS}_{model} é o valor médio do CRPS do modelo a ser avaliado e o \overline{CRPS}_{ref} é o valor médio do CRPS de um modelo de referência escolhido.

Para terminar, a *pinball loss function* e os *interval score* são medidas utilizadas para avaliar as previsões feitas através de quanties e intervalos, respetivamente.

Capítulo 3

Análise das Séries de Consumo

Neste capítulo é feita uma análise das séries temporais de consumo agregado de uma empresa comercializadora. Nesta análise serão identificados os comportamentos de tendência de longo prazo, os comportamentos sazonais, os padrões anuais semanais e diários. Esta análise é essencial para identificar e modelar as componentes da série temporal, componentes estas que serão modeladas nos modelos de previsão desenvolvidos no Capítulo 4.

3.1 Descrição dos Dados utilizados

Os dados utilizados como base de realização desta dissertação foram históricos de consumos relativos a cada uma das 24 horas do período entre 1 de julho de 2015 até 28 de fevereiro de 2018 fornecidos por uma comercializadora de energia elétrica que iniciou a sua atividade no início de 2015. Os dados fornecidos começaram na data de fundação, contudo, devido aos baixos valores de consumos que se verificavam à data, achou-se que para obtenção de melhores resultados seria melhor excluí-los.

Este capítulo foca-se numa análise de decomposição de informação dado que os dados de consumos apresentam padrões anuais, semanais e diários. Posto isto, foi fulcral analisar estes padrões e as tendências de mínimos e máximos de modo a que fosse possível eliminar os efeitos residuais que podem prejudicar a previsão e, manter apenas a informação que interessa para o estudo em questão.

Previamente será realizada uma análise de tendência de crescimento de vazios e pontas de modo a perceber qual o comportamento do padrão em estudo e posteriormente será feita a análise do padrão semanal e diário para obter uma normalização dos dados e podermos decompor a informação em variáveis necessárias para a previsão.

Foram seguidas as etapas apresentadas na **Figura 4**.



Figura 4 - Etapas do processo de análise de dados.

3.2 Tendência de Crescimento

A Figura 5 mostra o comportamento dos consumos registados durante o período em estudo. É possível observar que no início da série o nível de consumos era ainda muito baixo, devido a, como já referido anteriormente, se tratar do primeiro ano em que a empresa entrou no mercado e o número de clientes ser ainda baixo. Dada esta situação padrão geral de consumos apresenta uma tendência crescente, começando a estabilizar no final de 2017.

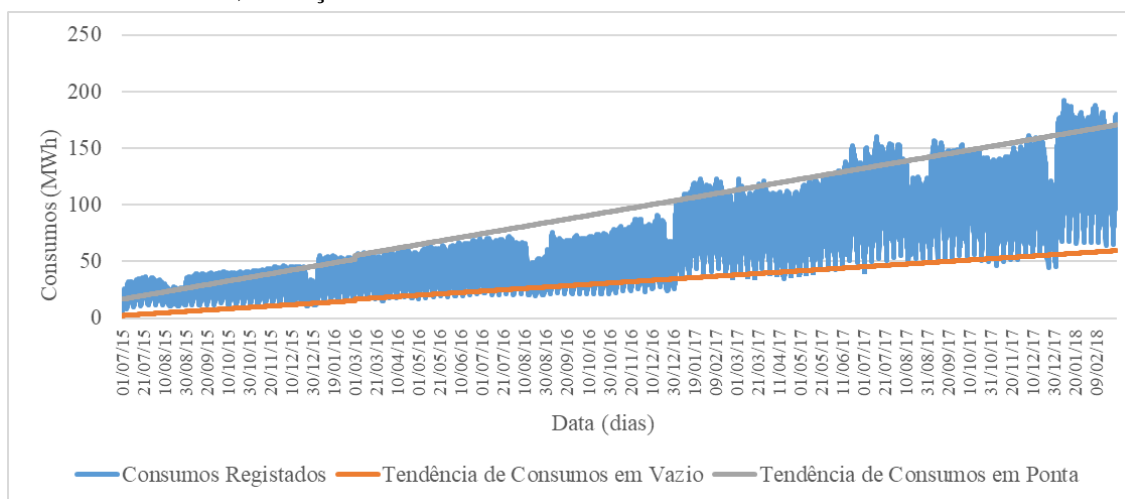


Figura 5 - Consumos registados e respetivas retas de tendência de vazio e ponta, do período em estudo.

Através da análise do gráfico da figura anterior, é possível verificar que realmente existe uma tendência de crescimento tanto de pontas como de vazios, com zonas em que a tendência anda acima ou abaixo dos valores reais. Estes desvios devem-se ao comportamento incerto dos consumos em algumas épocas do ano.

Foi necessário analisar os dados com pormenor pois, como podemos reparar existem áreas no gráfico com comportamentos diferentes (descidas e subidas de consumos acentuadas), e são estes comportamentos que irão fazer com que a tendência se altere.

As figuras seguintes, mostram os gráficos dos consumos de cada ano em separado para que se possa ter uma noção mais detalhada das alturas em que ocorrem estas variações acentuadas.

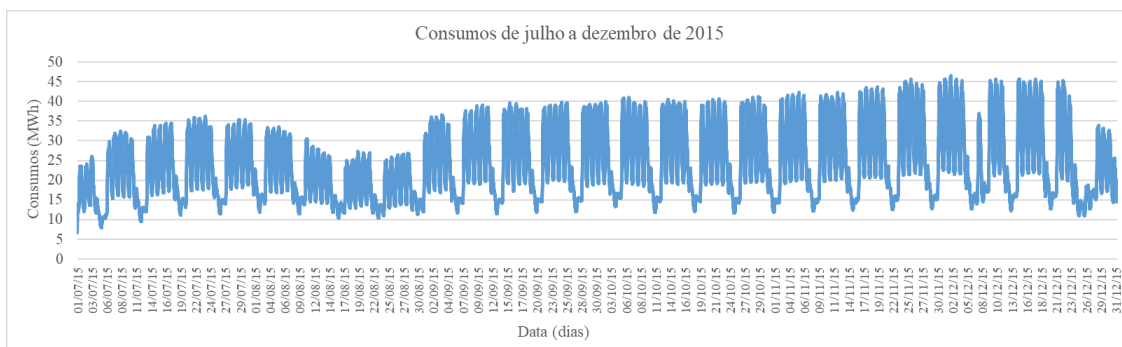


Figura 6 – Gráfico dos consumos de julho a dezembro de 2015.

Dado que em 2015 só existem dados desde 1 de julho é normal que nos anos seguintes a forma do gráfico seja ligeiramente diferente daí, ser necessário visualizar atentamente o eixo da data para percebermos em que altura do ano ocorrem as alterações significativas que se repetem de ano para ano. Podem facilmente identifica-se duas épocas mais críticas – agosto e dezembro. Quanto à análise de tendência, é importante observar os valores de consumos que tal como se vai notar vão sofrendo um aumento gradual.

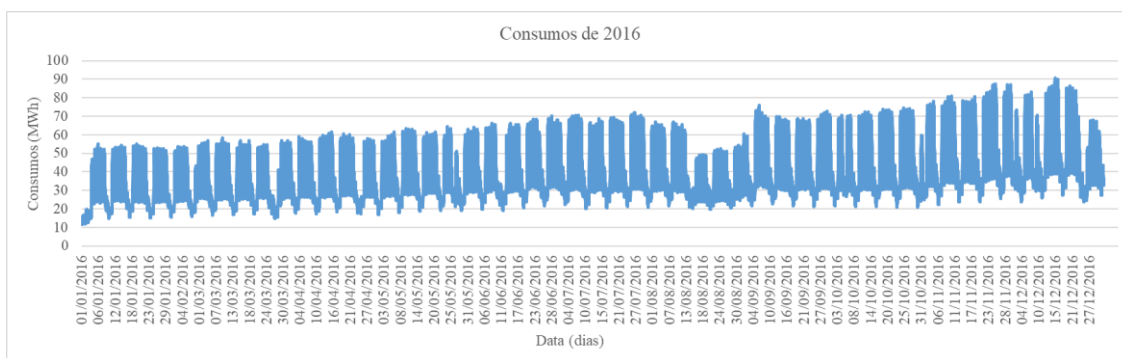


Figura 7 – Gráfico dos consumos de 2016.

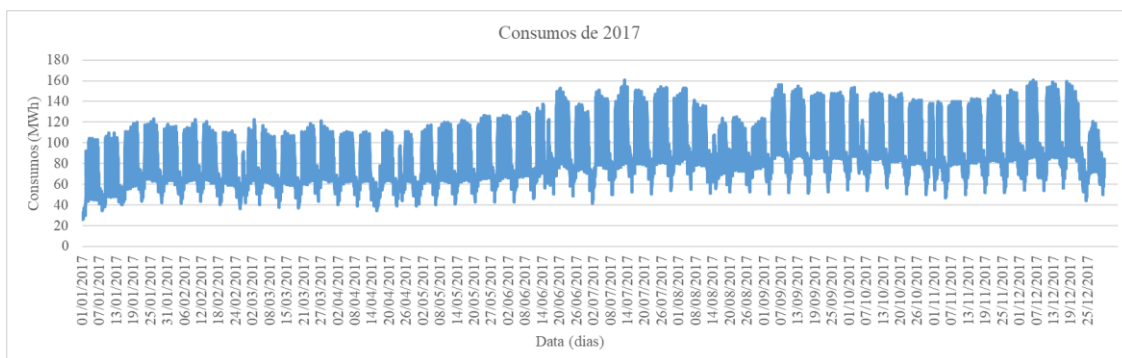


Figura 8 – Gráfico dos consumos de 2017.

Após a visualização dos gráficos referentes aos consumos de 2016 e 2017, **Figura 7** e **Figura 8**, respetivamente, é de notar que realmente existe um padrão anual de consumos com subidas e descidas dos mesmos que se verifica sempre por volta das mesmas datas – agosto e dezembro (comportamento que será analisado posteriormente). Relativamente ao aumento de consumos verificado no decorrer do tempo, pode concluir-se que o valor dos consumos vai aumentando quase para o dobro de ano para ano. Em 2015 registou-se um máximo que ronda os 45MWh perto do final do ano; em 2016 o máximo ocorre também no final do ano e alcança valores de cerca de

90MWh e em 2017 registou-se um máximo perto dos 160MWh.

Estes aumentos devem-se, sobretudo, à tendência de aumento do número de clientes que, com o avançar do tempo, deverá começar a estabilizar.

Aproveitando ainda alguma informação dada pelos gráficos anteriores, uma das diferenças que podem identificar-se facilmente, é o aumento dos mínimos, ou seja, os valores de consumo em vazio (mínimos) quando comparados em 2016 e 2017 subiu consideravelmente durante a semana, aumentando a diferença entre destes valores relativamente aos fins-de-semana.

Relativamente a 2018, os dados são apenas de janeiro e fevereiro, o que não permite que se possa realizar uma comparação equilibrada. Contudo, os valores de consumos já começam a apresentar uma forma mais estável.

Esta análise é importante para que se possa entender bem o comportamento dos dados e analisar todas as alterações que possam ser críticas e interferir com o sucesso da previsão. Para obter melhores resultados será necessário excluir estas tendências e por isso é realizada, no ponto seguinte, a análise do padrão sazonal para começar a identificar pontos comuns ao longo do ano e perceber a qual a melhor abordagem de tratamento de dados.

3.3 Padrão Sazonal

O padrão sazonal é um padrão de variação ao longo do ano com discretização semanal, ou seja, é uma análise anual com foco nas semanas em que o comportamento dos consumos difere muito do restante ano, tentando perceber o porquê de isso acontecer.

Através do gráfico da **Figura 5** é facilmente perceptível que ao longo do ano existe um padrão que se repete de ano para ano. Este padrão forma-se devido às diferenças de carga que existem de semana para semana ao longo do ano, que fazem com que as demandas sazonais se vão alterando. Quer isto dizer que, por exemplo, que tanto no Inverno como no Verão a tendência do aumento de uso de equipamentos de regulação de temperatura é maior, quer seja para aquecer quer para refrescar, o que faz com que haja variação de consumos.

Como já foi referido, existem duas variações consideráveis de consumos durante o ano que se repetem anualmente. Analisando agora os gráficos das figuras seguintes relativos aos vários anos, verifica-se que em algumas semanas de agosto e dezembro há uma diminuição significativa de consumos (áreas assinaladas no gráfico) e que em janeiro de cada ano há um aumento de consumos também elevado.

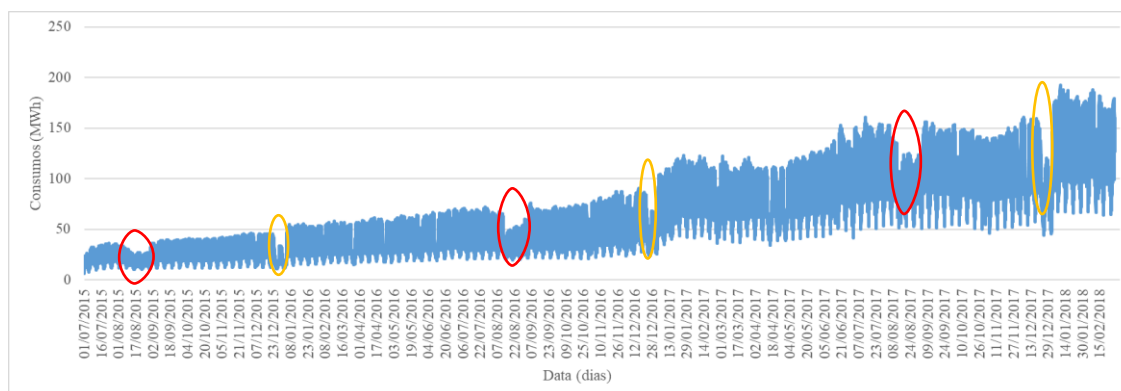


Figura 9 – Gráfico de consumos com realce nas semanas onde existe alteração significativa de consumos – agosto (a vermelho) e dezembro (a amarelo).

Através do gráfico anterior comprova-se que a situação anteriormente descrita se repete realmente todos os anos. A diminuição de consumos verificada durante as semanas observadas ocorre devido a épocas de férias de verão e natal onde os consumidores normalmente saem mais da sua habitação permanente, logo é normal que ocorram nestas datas variações do padrão mensal que se repete durante o resto do ano.

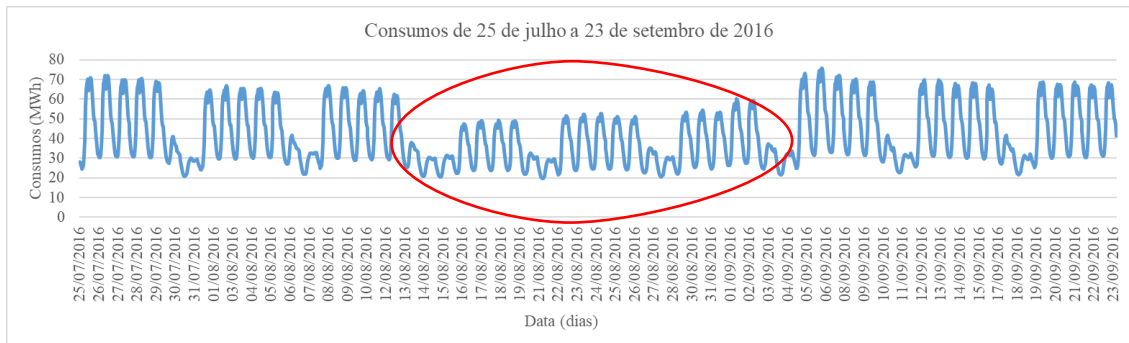


Figura 10 – Zoom do gráfico de consumos de 2016 para o período entre 25 de julho a 23 de setembro com realce nas semanas de agosto em que ocorre a diminuição de consumos.

Agora com mais pormenor consegue perceber-se através da **Figura 10** que a maior quebra de consumos ocorre entre 13 de agosto e 4 de setembro, altura do ano que coincide com as férias de grande parte dos consumidores.

Nas figuras que se seguem pode verificar-se que isto acontece realmente para a mesma altura todos os anos podendo apenas variar um dia ou outro.

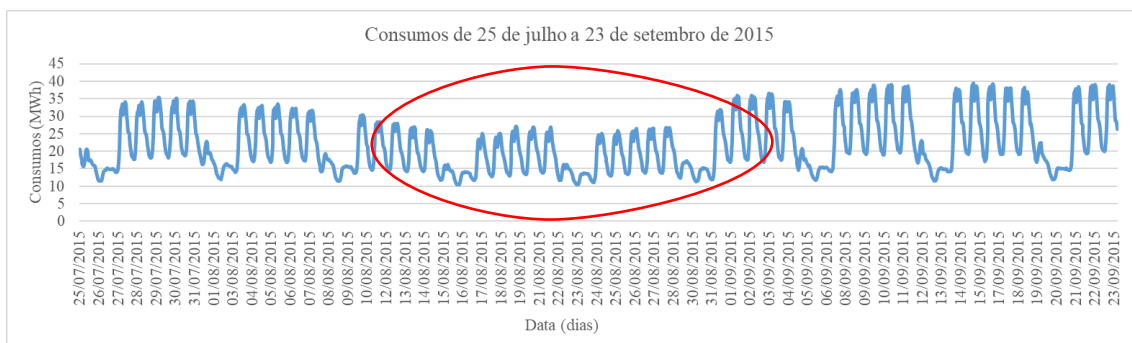


Figura 11 - Zoom do gráfico de consumos de 2015 para o período entre 25 de julho a 23 de setembro com realce nas semanas de agosto em que ocorre a diminuição de consumos.

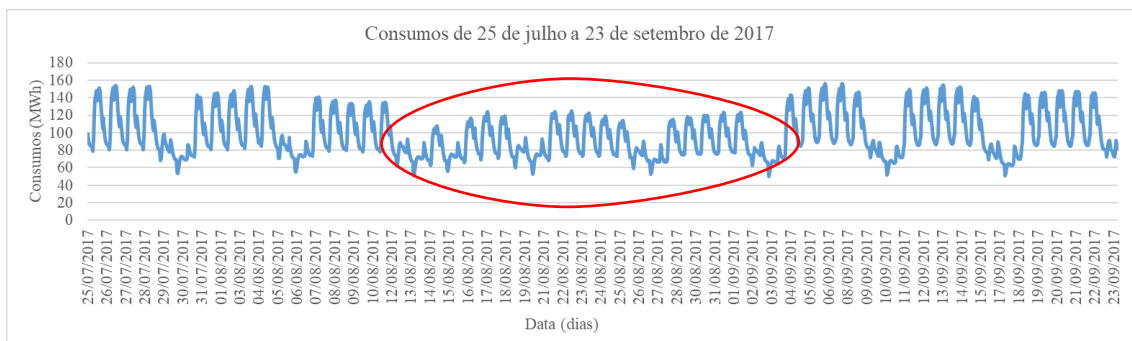


Figura 12 - Zoom do gráfico de consumos de 2017 para o período entre 25 de julho a 23 de setembro com realce nas semanas de agosto em que ocorre a diminuição de consumos.

No gráfico da **Figura 13**, serão analisadas as restantes zonas de padrões anómalos que

ocorrem em dezembro. Através da análise do gráfico da **Figura 9**, consegue perceber-se que a quebra de consumos ocorre junto ao final do mês de dezembro. Posto isto, realizou-se um zoom das últimas semanas de dezembro e primeiras de janeiro para ser possível uma melhor visualização do que acontece.

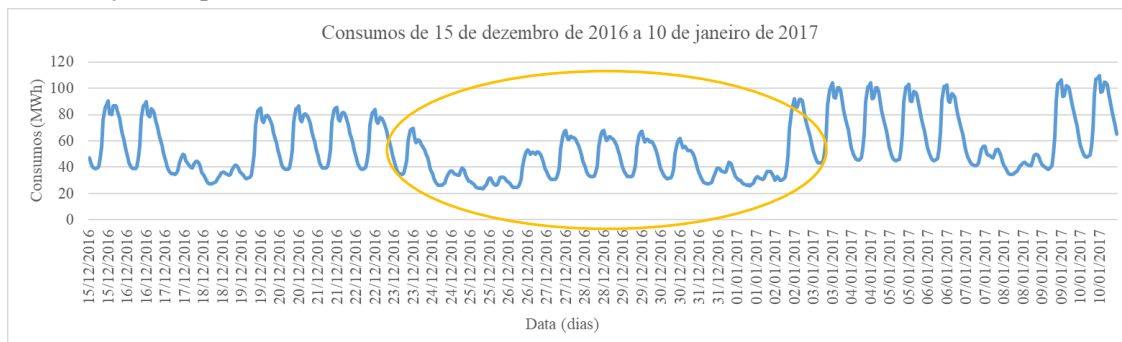


Figura 13 - Zoom do gráfico de consumos de 2016 e 2017 para o período entre 15 de dezembro de 2016 a 10 de janeiro de 2017 com realce da semana em que ocorre a diminuição de consumos.

Mais uma vez, observando com mais pormenor pode perceber-se que a diminuição de consumos se dá na semana entre o natal e a passagem de ano, época onde novamente os consumidores aproveitam as férias para se deslocarem para outro local que não a sua habitação. E, novamente, este registo repete-se anualmente como se verifica a seguir.

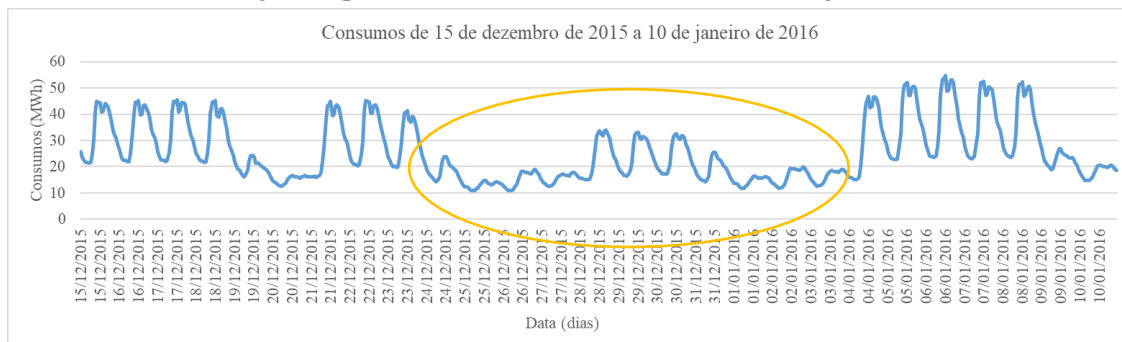


Figura 14 - Zoom do gráfico de consumos de 2015 e 2016 para o período entre 15 de dezembro de 2015 a 10 de janeiro de 2016 com realce da semana em que ocorre a diminuição de consumos.

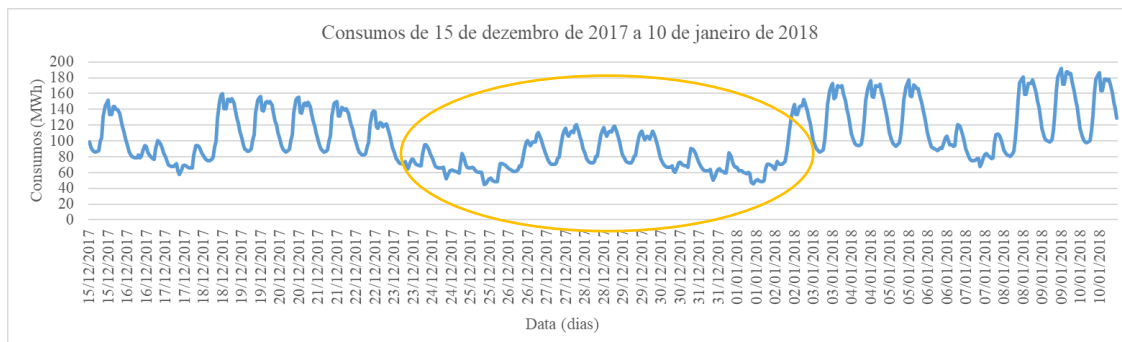


Figura 15 - Zoom do gráfico de consumos de 2017 e 2018 para o período entre 15 de dezembro de 2017 a 10 de janeiro de 2018 com realce da semana em que ocorre a diminuição de consumos.

Através da análise do comportamento do padrão anual foi possível perceber quais são as alturas do ano que requerem um ajuste do padrão.

Em seguida, será apresentada uma análise mais detalhada através do padrão semanal e diário que será a base dos modelos de previsão probabilística.

3.4 Padrão Semanal e Diário

A parte final desta análise é mais específica pois, é aqui que se começam a normalizar os dados que posteriormente serão a base da previsão.

Um dos principais fatores que explicam a variação dos padrões de consumo é o dia da semana, uma vez que caso se trate de um dia útil o consumo será maior do que ao fim de semana ou num feriado, uma vez que todos os setores estarão a contribuir em pleno para o consumo de energia elétrica e os consumidores estando a trabalhar têm muito mais atividade.

Posto isto, a distinção dos diferentes dias da semana é uma variável com elevada importância para efeitos de diferenciação, o que trará melhorias substanciais para os resultados das previsões. Através da análise dos diferentes padrões de consumo, conclui-se que o comportamento dos consumos nos feriados é diferente, a título de exemplo apresentam-se os dois gráficos seguintes, que representam a semana de 2 a 8 de outubro de 2016, e a semana de 23 a 29 de abril de 2017. Nestas duas semanas existem feriados, 5 de outubro e 25 de abril. O que se pretende analisar é o comportamento dos consumos nestes dias e identificar com que dia semana são mais parecidos.

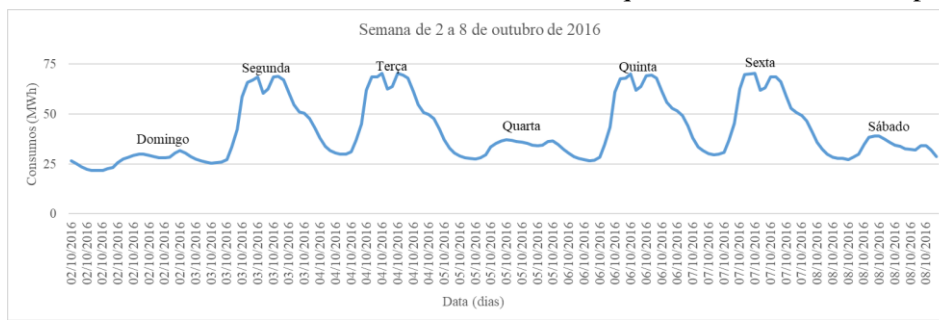


Figura 16 – Gráfico de consumos da semana de 2 a 8 de outubro de 2016.

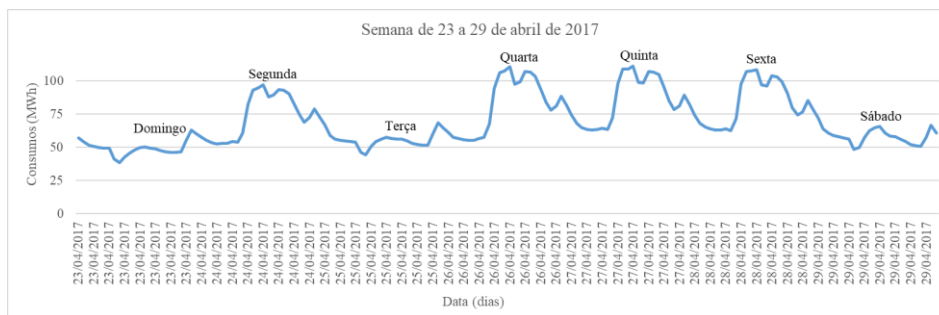


Figura 17 - Gráfico de consumos da semana de 23 a 29 de abril de 2017.

Ao analisar o comportamento dos dois feriados chegou-se à conclusão de que os consumos seriam mais parecidos com o sábado, portanto, atribuiu-se aos feriados o número que seria utilizado para identificar os sábados.

Com o auxílio do Microsoft Excel, foi possível associar a data ao dia da semana correspondente, atribuindo um número a cada dia para auxiliar nos cálculos (Tabela 1).

Tabela 1 - Associação entre dia da semana e número auxiliar.

| Dia da Semana | Número Atribuído |
|----------------------|-------------------------|
| Domingo | 1 |
| Segunda-feira | 2 |
| Terça-feira | 3 |
| Quarta-feira | 4 |
| Quinta-feira | 5 |
| Sexta-feira | 6 |
| Sábado | 7 |
| Feriado | 7 |

Ao longo do período de estudo foram selecionados todos os feriados que ocorreram, e encontram-se identificados na **Tabela 2** todos os que ocorrem em dias da semana, para depois serem identificados com o número correspondente ao sábado.

Tabela 2 - Feriados identificados no período em estudo.

| 2015 | 2016 | 2017 | 2018 |
|------------|------------|------------|------------|
| 05/10/2015 | 01/01/2016 | 14/04/2017 | 01/01/2018 |
| 01/12/2015 | 25/03/2016 | 25/04/2017 | |
| 08/12/2015 | 25/04/2016 | 01/05/2017 | |
| 25/12/2015 | 26/05/2016 | 15/06/2017 | |
| | 10/06/2016 | 15/08/2017 | |
| | 15/08/2016 | 05/10/2017 | |
| | 05/10/2016 | 01/11/2017 | |
| | 01/11/2016 | 01/12/2017 | |
| | 01/12/2016 | 08/12/2017 | |
| | 08/12/2016 | 25/12/2017 | |

3.4.1 Normalização dos dados

Tal como já foi referido, os dados de consumos disponíveis são referentes a diferentes anos civis, existindo uma tendência de crescimento no nível de carga, tanto nos valores de ponta como nos valores de vazio. Existe também a possibilidade de alteração no padrão semanal e diário, que serão analisados mais adiante.

Para proceder à decomposição dos fatores, separando a componente de tendência das componentes de padrões semanais e diários, foi necessário proceder à sua normalização por forma a conseguir obter padrões de semelhança entre os dados e para facilitar a determinação da previsão.

Para uma melhor análise de padrões e tendências de consumos, foi essencial numa primeira fase calcular consumos em cheias, pontas e vazios, através das expressões apresentadas a seguir, para se obter depois o valor de consumo normalização que será utilizado no cálculo dos padrões. Estes valores foram calculados utilizando um método de média móvel, ou seja, para calcular o

valor da hora h do dia D (D_h) recorre-se aos consumos desde dos três dias anteriores ($D-3$) até aos três dias posteriores ($D+3$).

$$\text{Cheias: } C(D_h) = \frac{D+3}{D-3} \bar{L} \quad (3.1)$$

$$\text{Pontas: } P(D_h) = \frac{D+3}{D-3} \max(L) \quad (3.2)$$

$$\text{Vazio: } V(D_h) = \frac{D+3}{D-3} \min(L) \quad (3.3)$$

Através dos valores obtidos foi possível traçar as curvas de tendência de pontas e vazio (**Figura 18**). Como já foi referido, comprovamos que os dados apresentam uma tendência crescente devido ao número de clientes que foi aumentando.

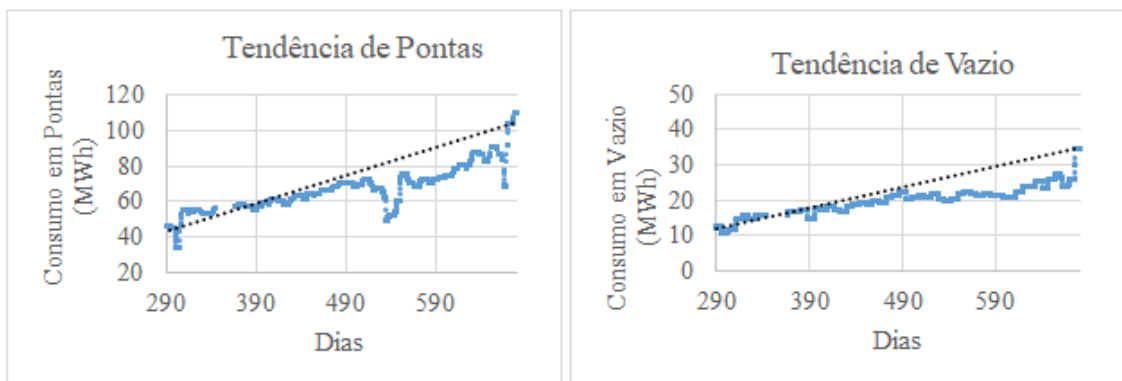


Figura 18 - Curvas de tendência de pontas e vazio dos consumos utilizados para previsão.

Destas curvas foi possível obter as respetivas equações da reta de tendência (equações 3.1 e 3.2) que serão úteis mais à frente.

$$PA(D) = 0,158 D - 2,5349 \quad (3.1)$$

$$VA(D) = 0,0588 D - 4,9106 \quad (3.2)$$

Onde,

- **PA** corresponde ao valor de Ponta Anual.
- **VA** corresponde ao valor de Vazio Anual.
- **D**, dia D , contado a partir da data de início da série em análise.

Após a obtenção destes valores, foi possível através da equação 3.3, calcular o valor normalizado do consumo para cada hora.

$$L_{norm} = \frac{L(D) - V(D)}{P(D) - V(D)} \quad (3.3)$$

Com,

- **L_{norm}** , é o valor do consumo normalizado.
- **L** , é o valor do consumo normal.
- **V** , é o valor de consumo em vazio calculado anteriormente.
- **P** , é o valor de consumo em ponta calculado anteriormente.

- D . é o dia para o qual se está a calcular o consumo normalizado.

Tendo os dados normalizados é possível obter os padrões semanal e diário que serão apresentados posteriormente.

3.4.2 Resíduo Semanal

O resíduo semanal é um valor que ajudará a limpar os valores dos vazios e de pontas.

Através das equações 3.1 e 3.2 calculou-se a ponta e o vazio anual para cada hora de modo a obter posteriormente o resíduo de ponta e de vazio (equações 3.4 e 3.5 respetivamente):

$$RP(D) = PA(D) - P(D) \quad (3.4)$$

$$RV(D) = VA(D) - V(D) \quad (3.5)$$

Obtidos estes valores foi possível calcular a média destes resíduos para cada semana do ano:

$$RP(S) = \frac{\sum RP(D_S)}{N} \quad (3.6)$$

$$RV(S) = \frac{\sum RV(D_S)}{N} \quad (3.7)$$

Com,

- $RP(S)$, resíduo de ponta da semana S .
- $RV(S)$, resíduo de vazio da semana S .
- $RP(D_S)$, resíduos de ponta de todos os dias da semana S .
- $RV(D_S)$, resíduos de vazio de todos os dias da semana S .
- N , número de dias do ano para a semana S .

O cálculo efetuado, calcula a média de todos os resíduos de ponta e vazio obtidos para a semana S – **Figura 19**.

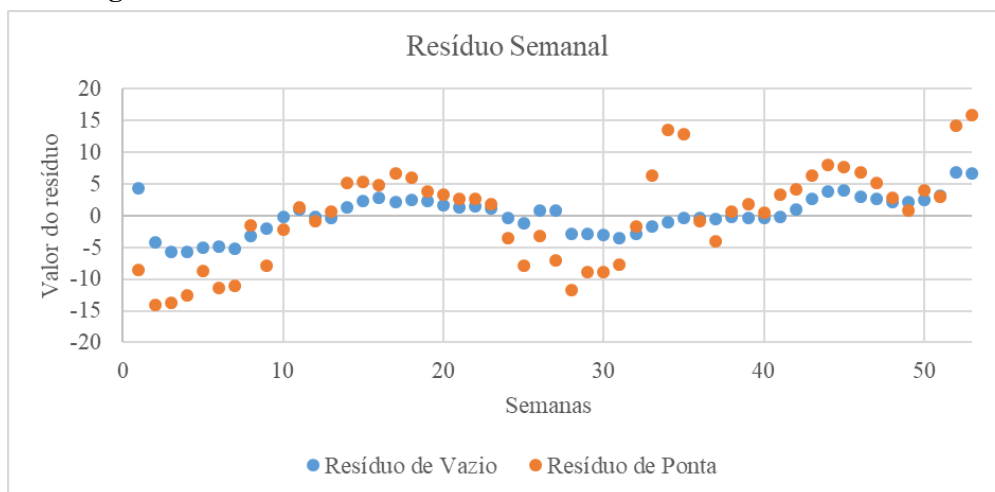


Figura 19 – Gráfico dos resíduos de vazio e de ponta para cada semana do ano.

Estes valores serão utilizados mais à frente nos cálculos auxiliares à previsão.

3.4.3 Padrão Diário

Este ponto da análise será o que mais contribuirá para a previsão visto que será a partir daqui que serão calculadas as variáveis que serão necessárias para aplicar nos modelos de previsão probabilística. É realizada uma caracterização probabilística do padrão diário para se perceber as incertezas existentes.

Para começar, foi necessário calcular o consumo médio diário e horário. Estes valores foram calculados já com base nos valores normalizados dos consumos aplicando a expressão:

$$\mu(D_h) = \frac{\sum L_{norm}(D_{Sh})}{N} \quad (3.8)$$

Onde,

- $\mu(D_h)$, é a média do valor do consumo normalizado do dia D à hora h .
- $L_{norm}(D_{Sh})$, é o valor do consumo normalizado, para o dia da semana D_s (1 a 7) à hora h (1 a 24).
- D_{Sh} corresponde ao dia da semana D_s (1 a 7) à hora h (1 a 24).
- N , número amostras do somatório (horas com igual dia da semana e hora).

Esta fórmula foi aplicada para calcular a média de valores para cada uma das 24 horas dos 7 dias da semana conseguindo-se assim obter uma representação gráfica do padrão diário dos dados em estudo apresentada na **Figura 20**.

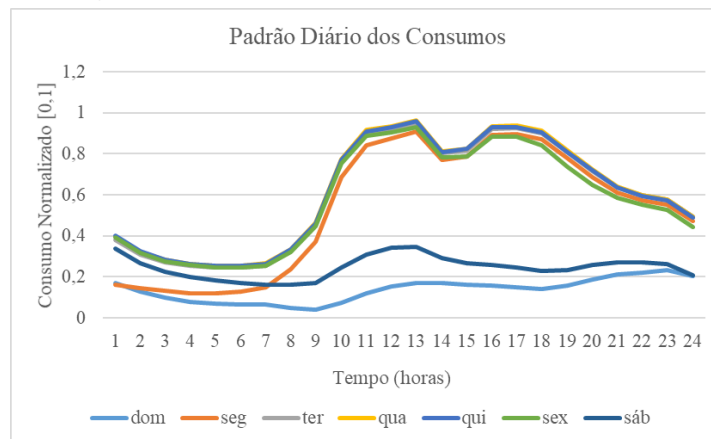


Figura 20 – Padrão diário dos consumos em estudo.

A Figura 19 vem comprovar o que já tinha sido referido anteriormente, que os valores de consumo aos fins-de-semana são bem mais baixos que durante a semana. Observa-se que nos dias de semana, de terça-feira a quinta-feira, os padrões de consumo são muito próximos. Na sexta-feira existe algum decaimento do consumo no final do dia e na segunda-feira existe uma transição de crescimento muito acentuada no início do dia. Para analisar as alterações de padrão ao longo dos anos, devido à alteração no agregado de consumidores, serão analisados de seguida os padrões diários para diferentes épocas do ano com distinção de cada ano, para que se possa perceber quais foram as alterações que o padrão foi sofrendo.

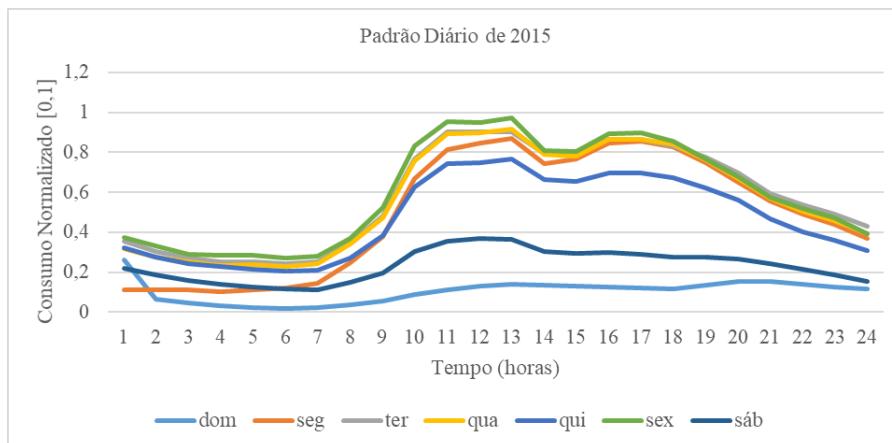


Figura 21 – Padrão Diário de dezembro de 2015.

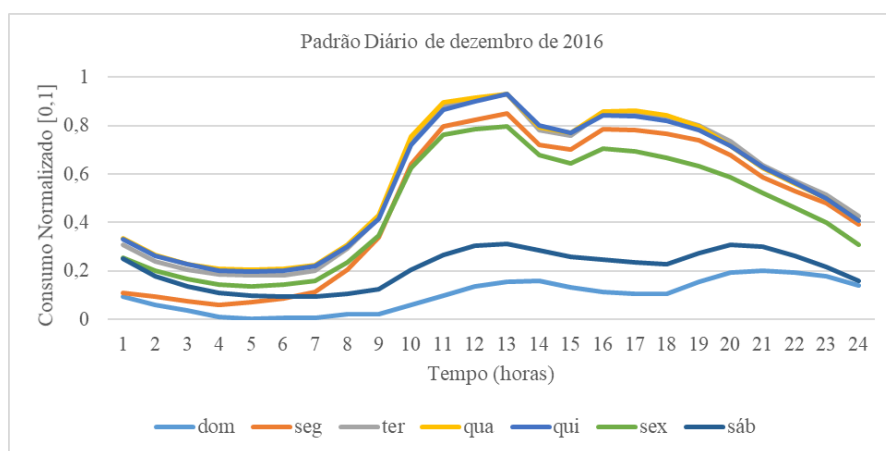


Figura 22 - Padrão Diário de dezembro de 2016.

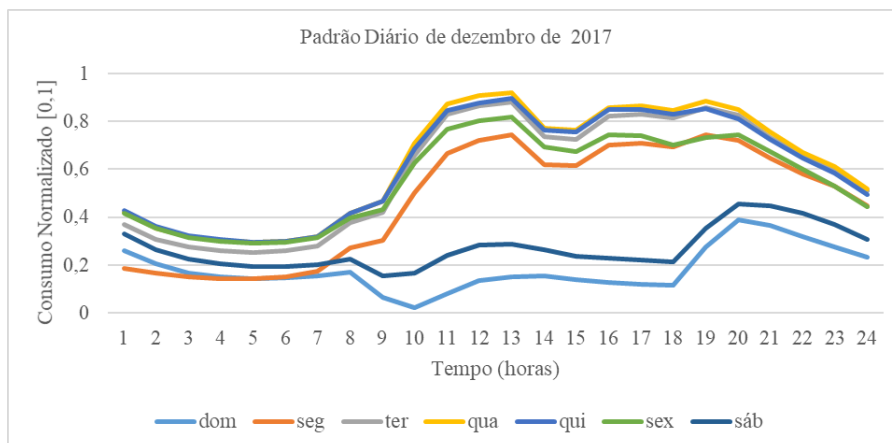


Figura 23 - Padrão Diário de dezembro de 2017.

Analisando os gráficos das figuras anteriores respeitantes aos padrões diários do mês de dezembro dos diferentes anos, as conclusões que se podem retirar é que as curvas do padrão diário sofrem algumas alterações, embora se possa assumir um padrão médio semelhante. De 2015 para 2016 é de notar a alteração de consumos que houve relativa à quinta e sexta-feira, visto que em 2015 o consumo de quinta era inferior ao de sexta o que no padrão diário dos restantes anos acontece de forma contrária, ou seja, é a sexta que apresenta valores mais baixos de consumos visto estar a aproximar-se o fim-de-semana. As curvas de sábado e domingo continuam a

apresentar consumos mais baixos ao longo da semana. Nos restantes dias pode ver-se que houve uma ligeira evolução, não muito significativa, de uns anos para os outros sendo que os dias em que a mudança foi maior, foram a segunda e a quinta em que as respetivas retas subiram um pouco mantendo-se mais alinhadas nos valores dos restantes dias.

Analisar-se-á a seguir a mesma situação para o mês de agosto também pelo mesmo motivo.

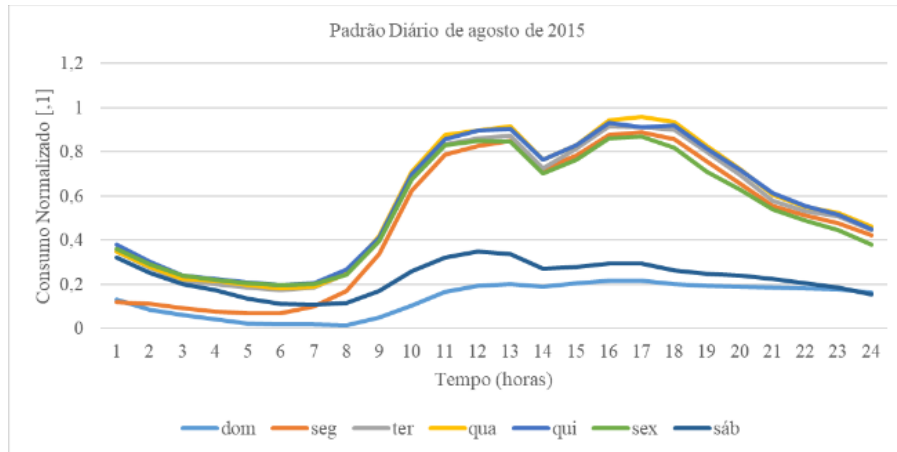


Figura 24 - Padrão Diário de agosto de 2015.

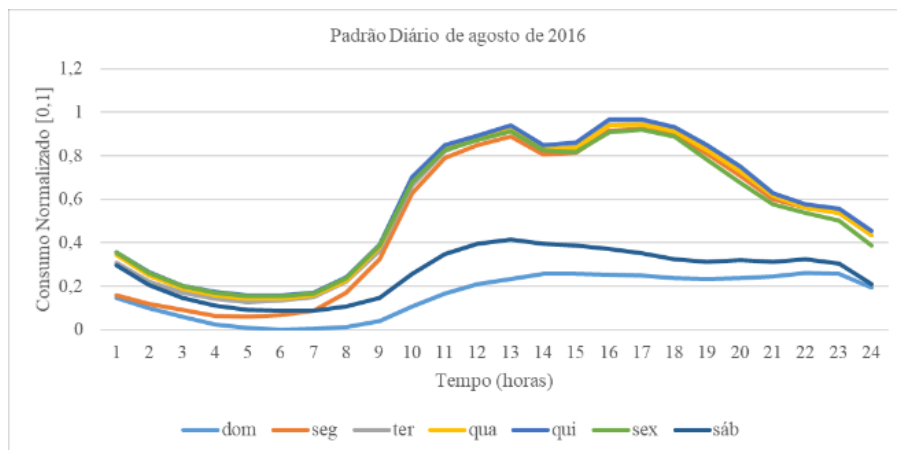


Figura 25 - Padrão Diário de agosto de 2016.

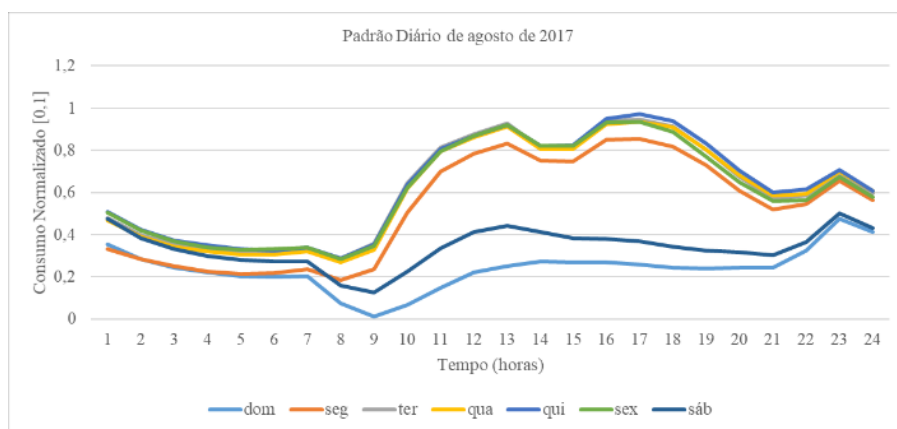


Figura 26 - Padrão Diário de agosto de 2017.

Mais uma vez estamos perante uma situação onde as alterações no padrão diário não são significativas. Apesar da situação que se verifica sempre no mês de agosto, período de baixa de consumos, podemos comprovar que esta situação não tem mais uma vez um impacto intenso nas curvas do padrão diário. Observam-se algumas alterações ligeiras de uns anos para os outros e é mais evidente a alteração do padrão no final do dia, devido a uma alteração para um agregado de consumos com mais consumidores domésticos. Observando a **Figura 26** relativa ao ano 2017 a única coisa que salta mais á vista é a curva de segunda-feira, em que existe um ligeiro decréscimo de consumos.

A etapa seguinte passou pela escolha da distribuição de probabilidade que se iria aplicar para efetuar a previsão. Depois de analisar diferentes opções, a função distribuição beta mostrou-se a mais adequada visto ser uma função com a capacidade de se adaptar a situações extremas devido à sua grande flexibilidade.

3.4.4 Caracterização probabilística do padrão diário

A função distribuição beta é uma função versátil com capacidade de modelizar assimetrias da distribuição e capaz de modelizar os limites físicos das variáveis estocásticas, mínimos e máximos, que é característico do padrão de consumo de eletricidade, limitado entre o limite mínimo, consumo de vazio, e limite máximo, consumo de ponta. É usada para modelação de proporção e percentagens dentro do intervalo $[0,1]$, através de dois parâmetros positivos, α e β , que aparecem com expoentes da variável aleatória x , definida em $[0,1]$, e controlam o formato da distribuição.

A função densidade de probabilidade (f.d.p) da distribuição beta, para $0 \leq x \leq 1$ e parâmetros α e $\beta > 0$ é definida por:

$$f(x; \alpha, \beta) = \text{constante} \times e^{\alpha-1}(x-1)^{\beta-1} \quad (3.9)$$

Onde,

$$\text{constante} = \frac{1}{B(\alpha, \beta)}$$

E B é a função beta, uma constante de normalização, que assegura que a probabilidade total integra a 1.

Para se poder aplicar esta função foi necessário o cálculo prévio de grandezas e variáveis integrantes da mesma.

Para obtenção dos parâmetros α e β , serão utilizados os valores calculados para o consumo médio normalizado para cada hora de cada dia da semana através da expressão já apresentada anteriormente (3.9) e valores de variância que fornecerão a componente de incerteza necessária para a previsão.

3.4.4.1 Cálculo da Variância

A variância foi calculada através da expressão seguinte:

$$v = E(x^2) - E(x)^2$$
$$v(D_h) = \left(\frac{\sum L_{norm}(D_{Sh})^2}{N_{Sh}} \right) - \left(\frac{\sum L_{norm}(D_{Sh})}{N_{Sh}} \right)^2 \quad (3.10)$$

Onde,

- $v(D_{Sh})$, valor da variância do dia da semana S.
- $E(x)$, valor esperado ou média μ .
- N_{Sh} , número amostras do somatório.

Na **Figura 27** apresentam-se os valores de variância $v(D_{Sh})$, traçados para cada hora de cada dia da semana.

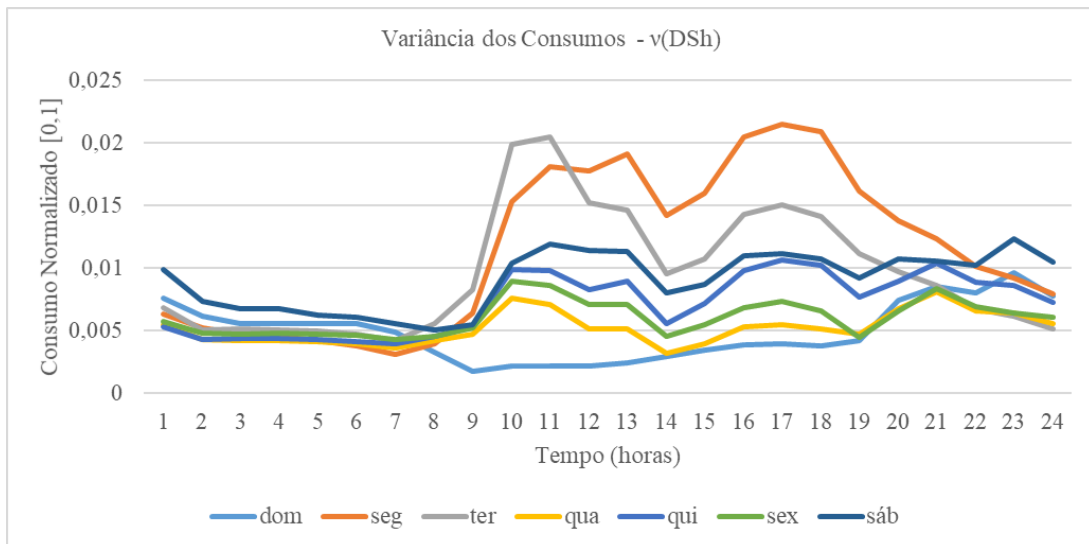


Figura 27 - Gráfico da variância dos consumos ao longo de cada hora, para cada dia da semana.

Este gráfico de variância fornece informação acerca da incerteza associada a cada hora de cada dia da semana. É possível observar quais são os dias mais estáveis e quais os que apresentam uma incerteza maior.

As horas que correspondem a períodos de vazio apresentam uma incerteza em geral baixa para todos os dias da semana, mas, a partir das 9h há um disparo na incerteza em quase todos os dias. Esta diferença nota-se no período em que há aumento da atividade dos consumidores que, como nem sempre é regular, pode levar a valores de incerteza mais elevados. Os dias que apresentam uma incerteza muito alta são a segunda e a terça-feira, quarta, quinta e sexta apresentam uma variância mais estável, não havendo grandes alterações ao longo do dia e ao domingo é quando a variância atinge os valores mais baixos.

Conclui-se que realmente é muito importante efetuar o estudo da incerteza associada aos consumos pois, como se pode verificar esta pode variar muito e levar a que se cometam erros aquando da realização de previsões de consumos de eletricidade.

3.4.4.2 Cálculo dos parâmetros α e β

Resta agora efetuar o cálculo dos valores de α e β para que todas as condições estejam reunidas, através das expressões:

$$\alpha(D_{Sh}) = \mu(D_{Sh}) \left(\frac{\mu(D_{Sh})(1-\mu(D_{Sh}))}{v(D_{Sh})} - 1 \right) \quad (3.11)$$

$$\beta(D_{Sh}) = (1 - \mu(D_{Sh})) \left(\frac{\mu(D_{Sh})(1-\mu(D_{Sh}))}{v(D_{Sh})} - 1 \right) \quad (3.12)$$

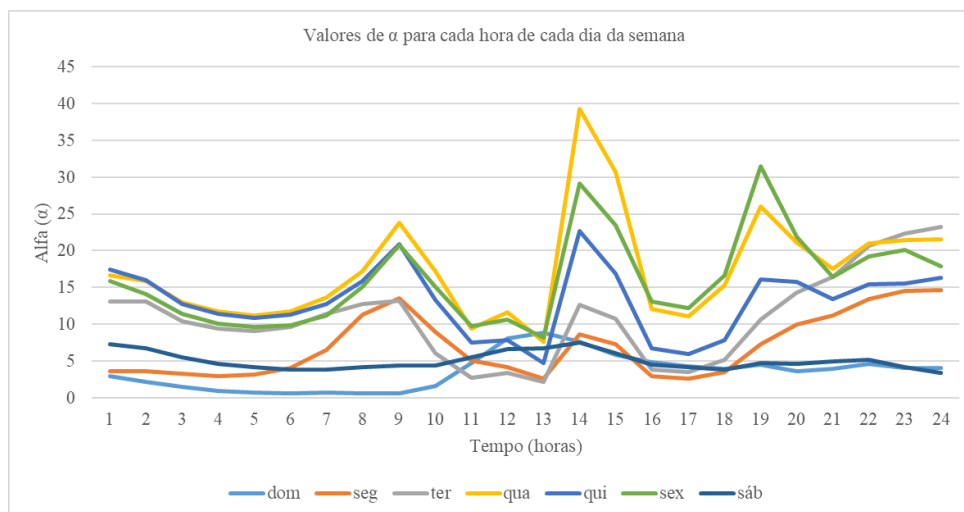


Figura 28 – Gráfico representativos dos valores de α calculados para todas as horas de cada dia da semana.

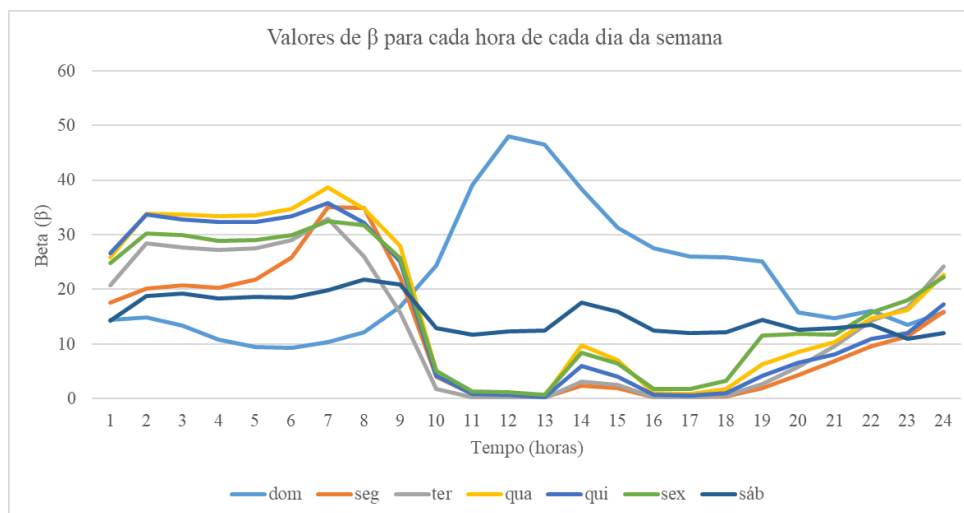


Figura 29 - Gráfico representativos dos valores de β calculados para todas as horas de cada dia da semana.

Na **Figura 28** e **Figura 29** estão representados os valores de α e β calculados para cada hora de cada dia da semana.

Valores de α e β mais elevados implicarão distribuições mais apertadas. A título de exemplo, perante estes gráficos pode verificar-se que nas quintas-feiras poderá ocorrer uma situação destas, visto que os valores dos dois parâmetros apresentam valores elevados em conjunto entre as 13h e as 16h.

Já para situações em que valores de α sejam superiores a β irá verificar-se uma assimetria nas

pontas e o caso contrário uma assimetria típica dos vazios.

3.4.5 Distribuição Beta vs Distribuição Normal

Aquando da escolha de qual seria a melhor modelização a utilizar para os modelos de previsão foram realizados testes com a distribuição normal para que se pudesse ter um termo de comparação para decidir qual a que se adequava melhor ao estudo em questão.

O fato de a escolha ter recaído sobre a distribuição beta deve-se às razões já descritas anteriormente, a capacidade de se moldar devido aos parâmetros α e β o que leva a uma maior precisão.

Foram aplicadas as duas distribuições e comparadas graficamente com a frequência real para que nos se pudesse analisar qual a acompanharia melhor.

Serão apresentados, a seguir, gráficos com testes realizados para diferentes dias da semana e várias horas, para se conseguir analisar estes comportamentos.

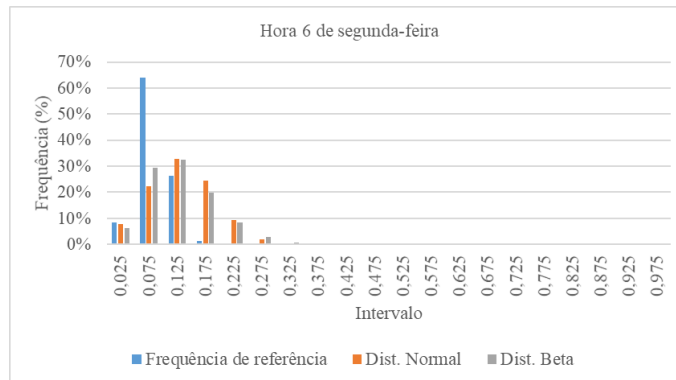


Figura 30 - Gráfico comparativo do comportamento das duas distribuições à hora 6 de segunda-feira.

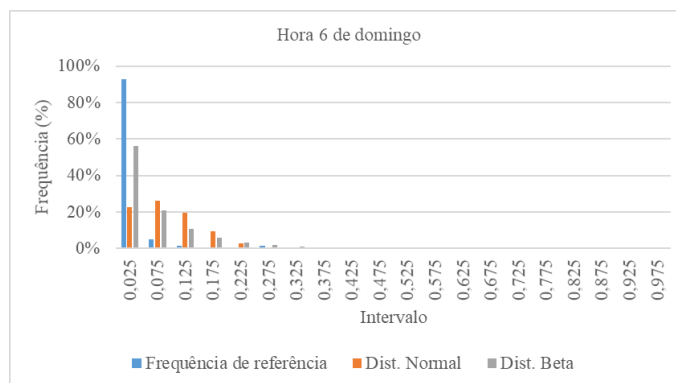


Figura 31 - Gráfico comparativo do comportamento das duas distribuições à hora 6 de domingo.

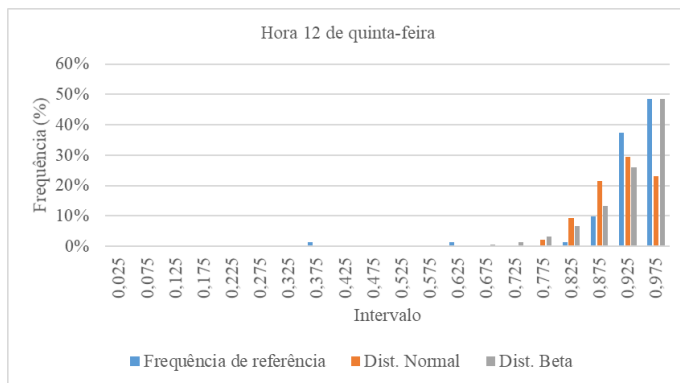


Figura 32 - Gráfico comparativo do comportamento das duas distribuições à hora 12 de quinta-feira.

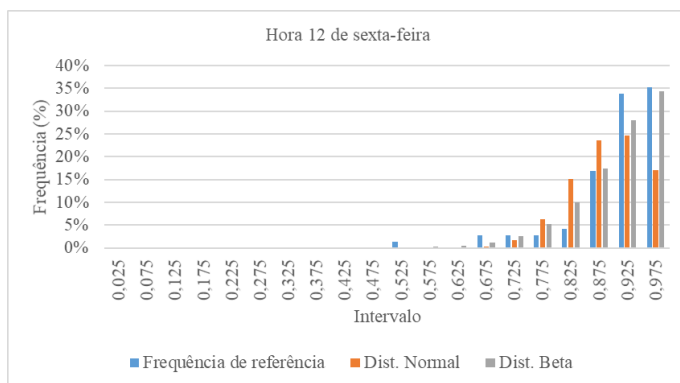


Figura 33 - Gráfico comparativo do comportamento das duas distribuições à hora 12 de sexta-feira.

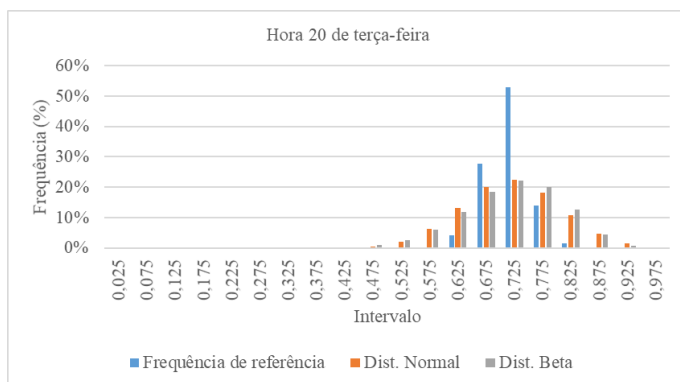


Figura 34 - Gráfico comparativo do comportamento das duas distribuições à hora 20 de terça-feira.

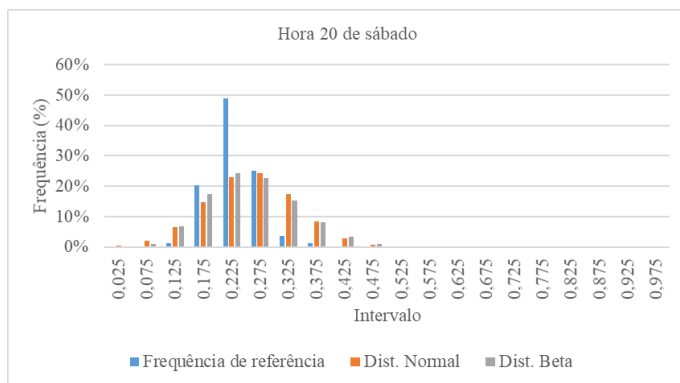


Figura 35 - Gráfico comparativo do comportamento das duas distribuições à hora 20 de sábado.

Nestes gráficos interessa observar qual a distribuição que melhor acompanha a “frequência de referência”, pois será a que se adequa mais para o estudo em questão. Tanto a distribuição normal como a beta estão próximas do que se pretende, havendo casos em que a normal se pode adequar melhor. Contudo nota-se que nos extremos de vazão e ponta, a distribuição beta é mais adequada que a distribuição normal. Assim, para este tipo de modelização de previsão adaptou-se a distribuição beta.

Capítulo 4

Previsão Probabilística baseada em Modelos de Decomposição

Neste capítulo será explicada a metodologia seguida para a realização da previsão probabilística de consumos a curto prazo.

Serão descritas duas variantes do mesmo modelo baseado em decomposição de padrões.

Numa primeira fase foi realizada uma primeira previsão com base integral na análise de dados realizada previamente. Contudo, no final os resultados obtidos não foram os que se desejavam e apresentavam um erro mais elevado, o que levou a que fossem realizadas algumas alterações para que a previsão fosse ajustada e para se obter um erro menor.

Ambos os modelos desenvolvidos, Variante I e Variante II, são baseados na decomposição numa componente de tendência, numa componente sazonal e numa componente de padrão horário ao longo da semana. A previsão da componente de tendência é aplicada aos valores de ponta e vazio semanais, com uma janela móvel semanal. A componente sazonal é aplicada numa base semanal, com um ajuste da ponta e vazio para cada semana do ano. A componente de padrão horário é aplicada a cada hora de cada dia da semana. No padrão horário a modelização é probabilística, proporcionando uma modelização dos parâmetros da distribuição beta em cada hora da semana.

O modelo da Variante I, consiste numa abordagem de modelo de decomposição aditiva em que a componente de tendência, do vazio e ponta semanal, é modelizada por uma tendência linear. A estas duas tendências lineares são adicionados ou subtraídos valores de ajuste sazonal, calculados com base no histórico. A componente probabilística de padrão horário semanal é integrada no modelo por ajuste de escala dos valores normalizados para a escala entre o mínimo (vazio) e máximo (ponta) semanal.

O modelo da Variante II, consiste numa abordagem de modelo de decomposição multiplicativa em que a componente de tendência, do vazio e ponta semanal, é modelizada como uma componente autoregressiva, que usa as pontas e vazios observados por média móvel na semana anterior. À tendência linear da ponta e vazio são multiplicados os valores de ajuste sazonal. Tal como no modelo da Variante I, a componente probabilística de padrão horário semanal é integrada no modelo por ajuste de escala dos valores normalizados para a escala entre o mínimo (vazio) e máximo (ponta) semanal.

4.1 Modelo - Variante I

O fluxograma da **Figura 36** serve de orientação para explicar a metodologia de previsão passo a passo.

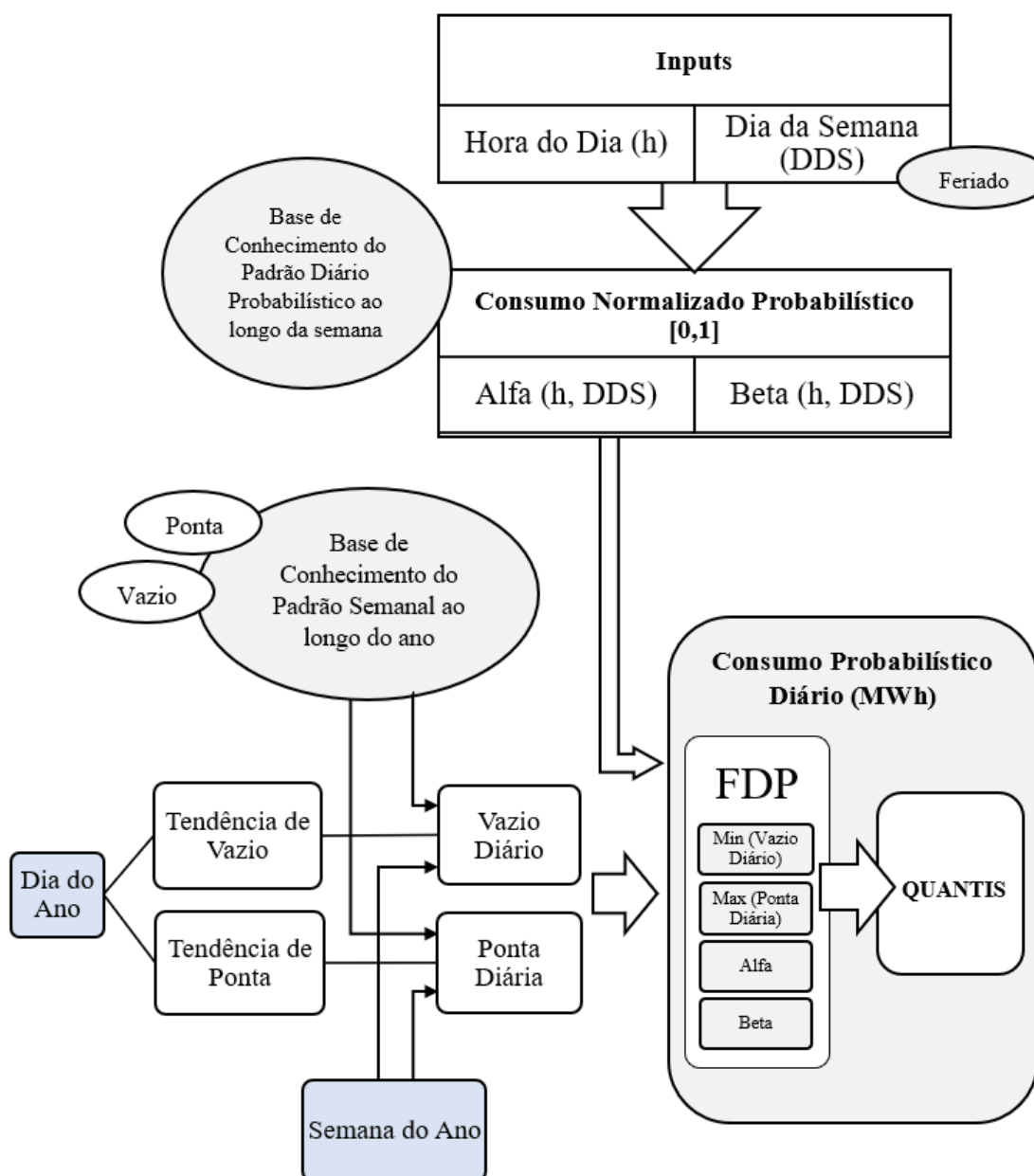


Figura 36 - Fluxograma explicativo do modelo aplicado para previsão de consumos de eletricidade.

1. **Inputs e Base de Conhecimento:** Na Variante I os Inputs do modelo foram: hora do dia, dia da semana, dia do ano e semana do ano. As bases de conhecimento: tabelas do padrão diário probabilístico ao longo da semana, tabelas do padrão sazonal em vazio e ponta e tabelas de valores de α e β para cada hora da semana.
2. **Cálculo da tendência do vazio e tendência da ponta:** Através das equações apresentadas na análise de dados foram calculados os valores das tendências do vazio (VA) e de ponta (PA) para cada D - dia do ano.

$$PA(D) = 0,158 D - 2,5349 \quad (4.1)$$

$$VA(D) = 0,0588 D - 4,9106 \quad (4.2)$$

3. **Modelo aditivo:** Aqui calcularam-se os valores de máximos (pontas) e mínimos (vazios) que resultam da soma do resíduo de ponta e de vazio (tabelados na base de conhecimento) às tendências calculadas no ponto 2:

$$\max(D) = PA(D) - RP(S) \quad (4.3)$$

$$\min(D) = VA(D) - RV(S) \quad (4.4)$$

4. **Integração do padrão horário semanal,** através da associação de cada hora da semana aos valores de α e β respetivos que estão na base de conhecimento.
5. **Cálculo dos quantis:** Por fim, procedeu-se ao cálculo dos quantis da previsão, onde

$$QX(u) = \inf\{x: F(x) \geq u\}$$

Em que, $QX(u)$ é o Quantil o consumo previsto (X) e (u) é a probabilidade associada ao quantil (0,1; 0,2; ...; 0,9). $F(x)$ é a função inversa da função acumulada Beta com parâmetros (min, max, α , β , x) – equação 4.5.

$$F(x) = \frac{1}{B(\alpha,\beta)} \int_{\min}^{\max} x^{\alpha-1}(1-x)^{\beta-1} dx \quad (4.5)$$

4.1.1 Resultados obtidos

Depois de aplicada a equação 4.5 a todos os dados obtiveram-se os valores de todos os quantis de consumos. No gráfico da **Figura 37** apresenta-se apenas como exemplo os valores do Q50 sobrepostos aos consumos reais.

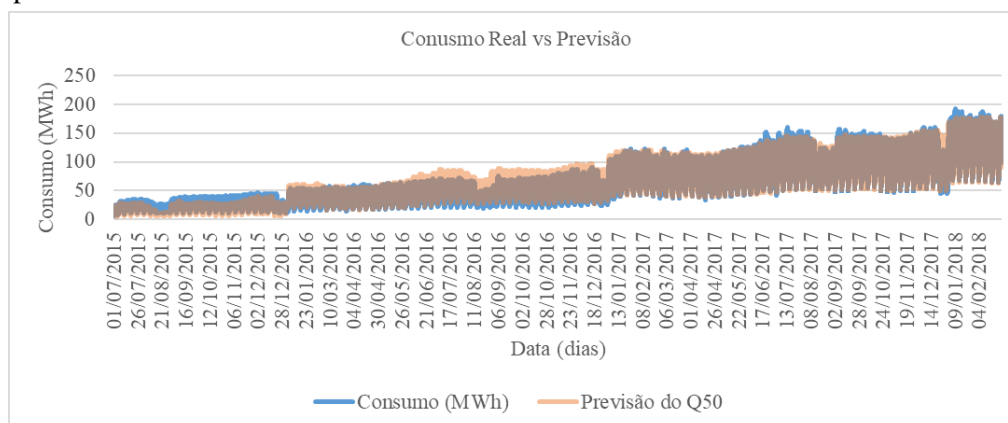


Figura 37 – Gráfico comparativo dos valores dos consumos reais e os valores da previsão do Q50.

Para uma melhor perceção dos resultados o gráfico da **Figura 38** é mais adequado porque mostra a previsão de vários quantis para uma semana, para assim se conseguir ter uma noção mais precisa do quanto se aproximam ou não dos consumos reais.

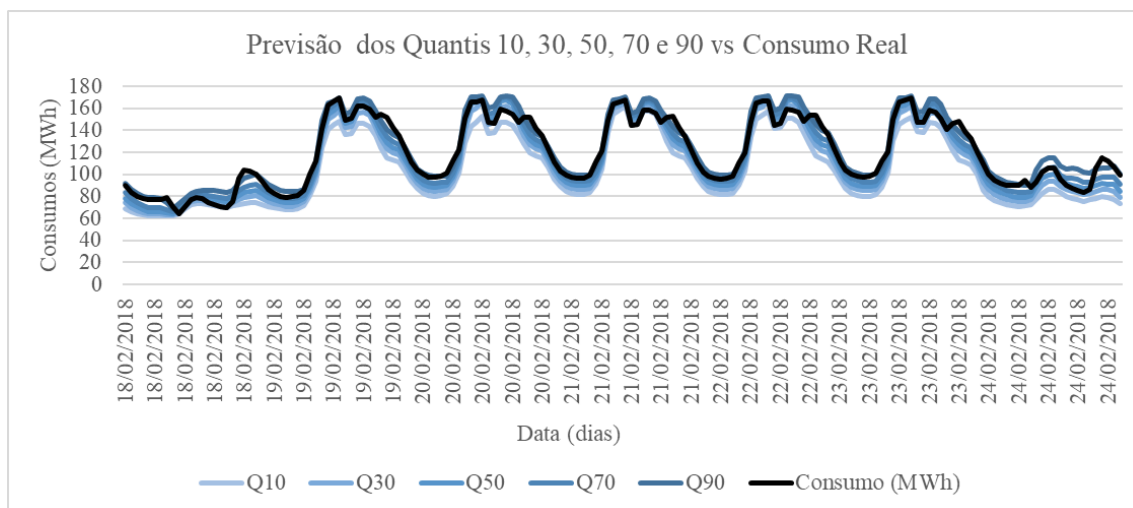


Figura 38 – Gráfico da Previsão de diferentes Quantis vs Consumo real para a semana de 18 a 24 de fevereiro de 2018.

Como se pode ver no gráfico, os valores dos quantis de previsão aproximam-se dos consumos reais, contudo, é visível que existe um erro de previsão que não pode ser descartado.

Para se conseguir ter uma noção mais exata de que o resultado não foi o que deveria apresentar-se a seguir o gráfico de *reliability* (**Figura 39**).

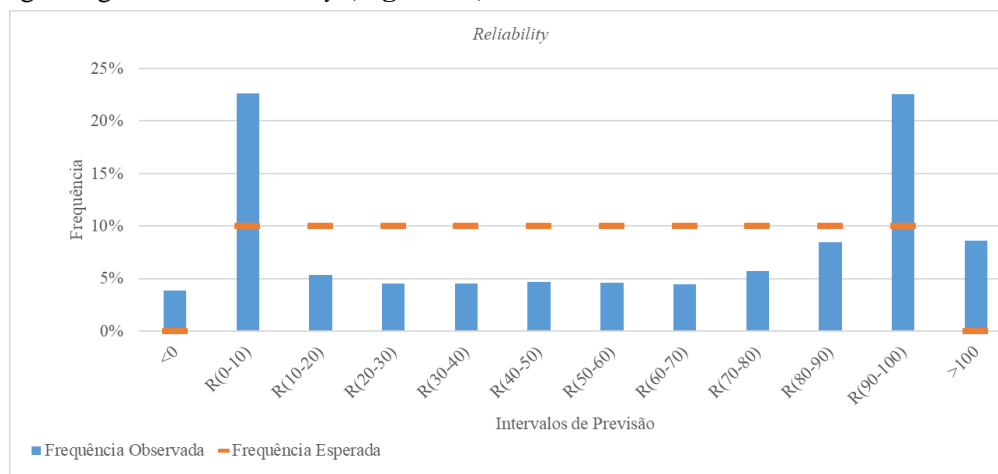


Figura 39 – Gráfico de *reliability* para avaliação do sucesso da previsão.

O gráfico apresenta uma situação de *Underdispersed Forecast* pois os valores da previsão encontram-se centrados nos extremos do gráfico e não respeitam os limites, ou seja, não vão de encontro com os valores dos consumos reais o que indica que a previsão falhou.

Foi também calculado o MAPE (equação 2.28) para posterior comparação com a Variante II, realizada para corrigir a previsão e diminuir o erro. O erro MAPE obtido foi de 17,2%.

4.2 Modelo - Variante II

Visto a Variante I ter falhado em alguns pontos, procedeu-se à alteração do modelo para que a previsão fosse mais precisa e para tentar diminuir o erro.

A nova variante está representada no fluxograma da **Figura 40**.

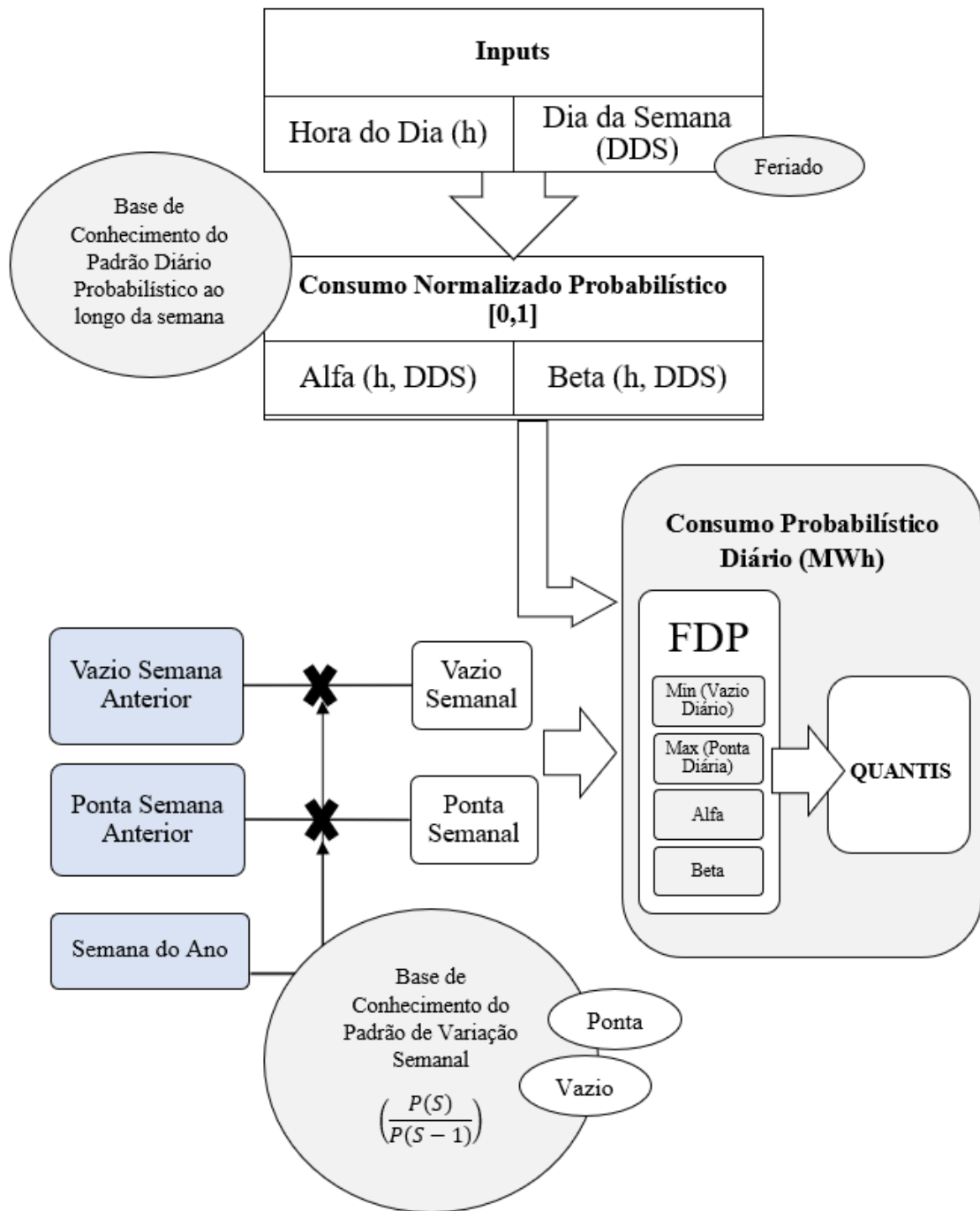


Figura 40 - Fluxograma explicativo da variante II do modelo aplicado para previsão de consumos de eletricidade.

- 1. Inputs e Base de Conhecimento:** Na Variante II os Inputs do modelo foram: hora do dia, dia da semana, consumos de vazio e ponta relativos à semana anterior à previsão e semana do ano. As bases de conhecimento: tabelas do padrão diário probabilístico ao longo da semana, tabelas do padrão de variação semanal em vazio e ponta e tabelas de valores de α e β para cada hora da semana.
- 2. Integração dos valores de ponta e vazio:** através da associação das horas a prever foram integrados os valores dos consumos de vazio e ponta relativos à semana anterior à previsão

3. **Modelo multiplicativo:** Aqui calcularam-se os valores de máximos (pontas) – equação 4.6 e mínimos (vazios) – equação 4.7, que resultam da multiplicação do fator de ponta (FP) e de vazio (FV) semanais (tabelados na base de conhecimento) pelos consumos integrados no ponto 2:

$$\max(D) = P(D) \times FP(S) \quad (4.8)$$

$$\min(D) = V(D) \times FV(S) \quad (4.9)$$

4. **Integração de variação semanal,** através da associação de cada hora da semana aos valores de α e β respetivos que estão na base de conhecimento.
5. **Cálculo dos quantis:** Por fim, procedeu-se ao cálculo dos quantis da previsão, onde

$$QX(u) = \inf\{x: F(x) \geq u\}$$

Em que, $QX(u)$ é o Quantil o consumo previsto (X) e (u) é a probabilidade associada ao quantil (0,1; 0,2; ...; 0,9). $F(x)$ é a função inversa da função acumulada Beta com parâmetros ($\min, \max, \alpha, \beta, x$) – equação 4.5.

$$F(x) = \frac{1}{B(\alpha, \beta)} \int_{\min}^{\max} x^{\alpha-1} (1-x)^{\beta-1} dx \quad (4.5)$$

4.2.1 Resultados Obtidos

Depois de aplicada a equação 4.5 a todos os dados obtiveram-se os novos valores de todos os quantis de consumos. No gráfico da **Figura 37** apresenta-se apenas como exemplo os valores do Q50 sobrepostos aos consumos reais.

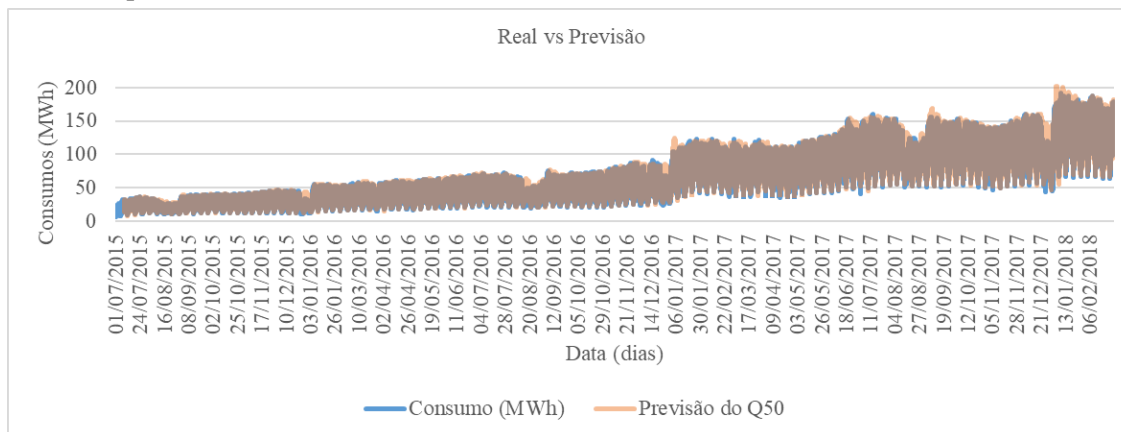


Figura 41 - Gráfico comparativo dos valores dos consumos reais e os valores da nova previsão do Q50.

Mais uma vez para uma melhor perceção dos resultados apresenta-se o gráfico com os quantis e o consumo real na **Figura 38** para uma melhor comparação.

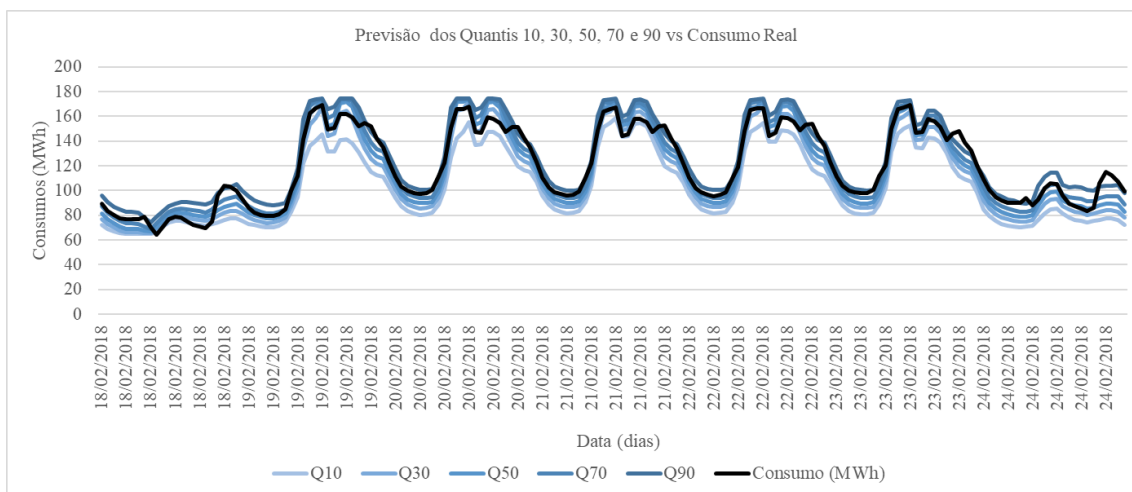


Figura 42 - Gráfico da Previsão de diferentes Quantis da Variante II vs Consumo real para a semana de 18 a 24 de fevereiro de 2018.

Neste gráfico é perceptível que os quantis da nova previsão acompanham muito melhor a curva de consumos reais, ou seja, as alterações realizadas na Variante II surtiram o efeito desejado e notável a melhoria nos valores da previsão, embora apresente ainda algum erro.

Através do gráfico de *reliability* apresentado na figura seguinte, pode concluir-se que desta vez o cenário é diferente pois a forma obtida que vai de encontro a uma previsão muito mais precisa.

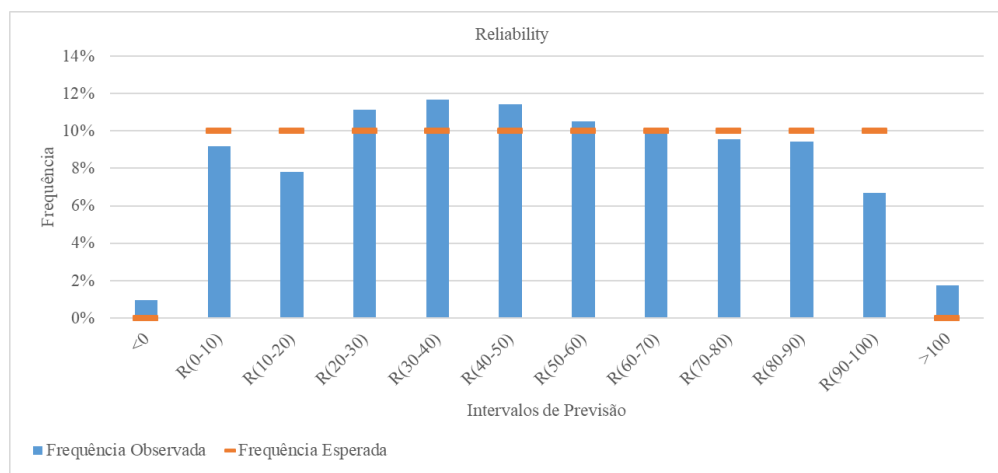


Figura 43 - Gráfico de *reliability* para avaliação do sucesso da previsão.

Agora, os valores obtidos são sem dúvida muito mais próximos dos valores que seriam esperados o que vem confirmar também o valor do erro MAPE obtido, 8,4%, cerca de metade do obtido na Variante I do mesmo modelo.

Posto isto, estão reunidas condições para efetuar um teste de previsão, prevendo valores de consumos para datas que vão além dos dados em estudo.

4.2.2 Teste do Modelo

Como forma de testar o modelo criado foi efetuada uma previsão de consumos de eletricidade

para o período de 1 de março a 30 de junho de 2018. Mostra-se de seguida o gráfico (**Figura 44**) dos quantis de previsão e, como forma de comparação mostra-se posteriormente os gráficos das **Figura 45** que mostra apenas os consumos previstos pelo quantil 50 para 3 meses e a **Figura 46** referente aos mesmos meses do ano anterior para se poder comparar o padrão de consumos.

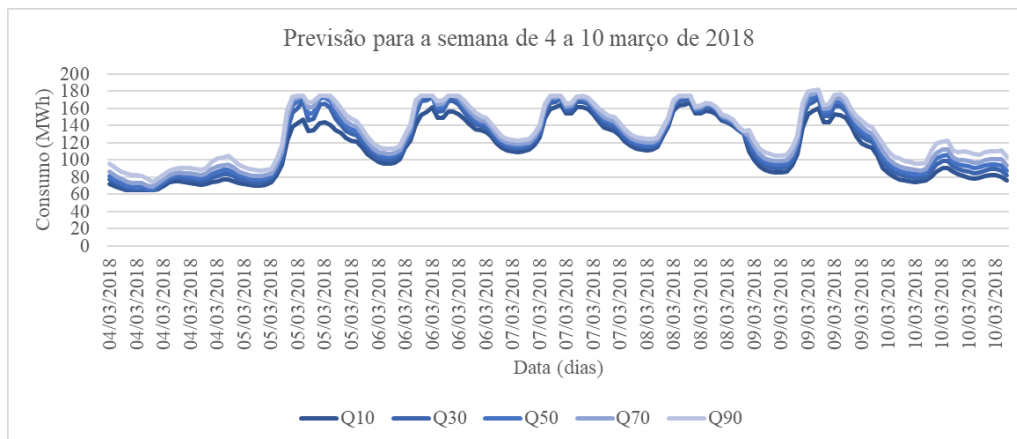


Figura 44 - Gráficos dos quantis de previsão em teste para uma semana

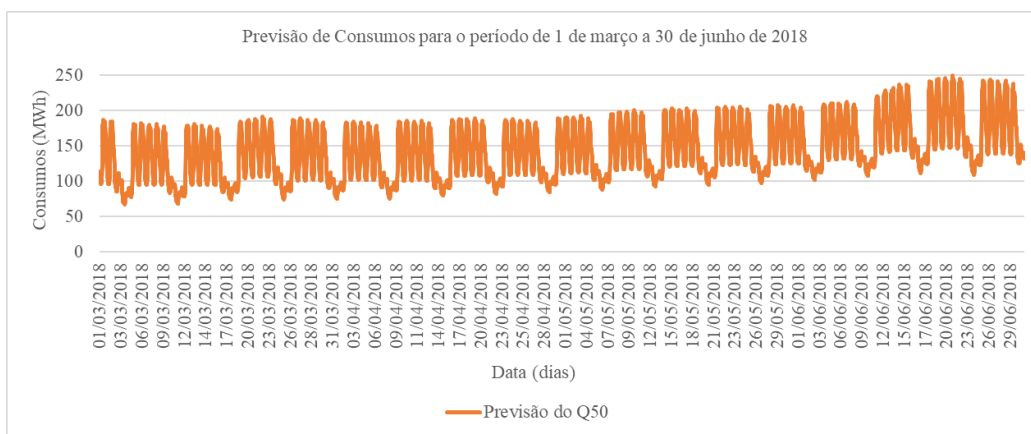


Figura 45 – Gráfico dos consumos previstos para testar o modelo implementado.

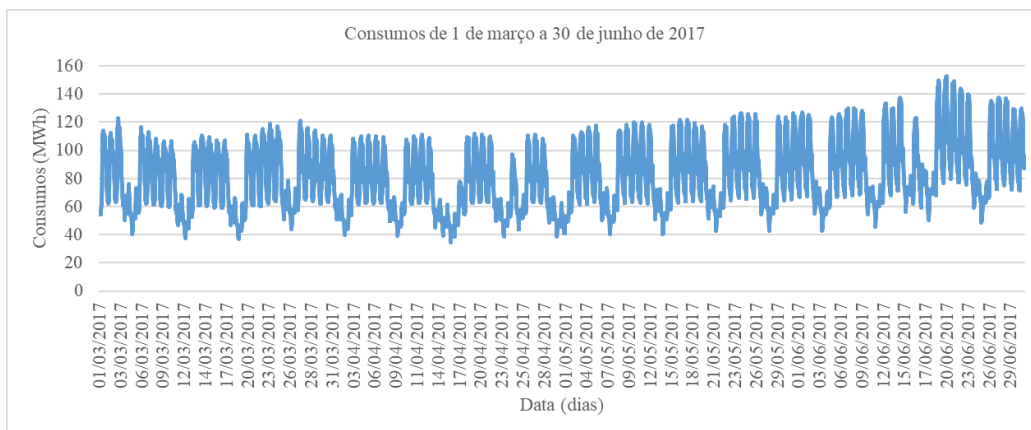


Figura 46 – Gráficos dos consumos de 1 de março a 30 de junho de 2017.

Embora as curvas observadas nas duas figuras não sejam perfeitamente idênticas, se analisarmos em termos de padrão, podemos constatar que não existe uma diferença muito grande entre elas podendo daqui concluir-se que a previsão não terá falhado muito. Existe apenas um

aumento dos consumos devido à componente de tendência existente.

Capítulo 5

Conclusões e Estudos Futuros

Neste capítulo, pretende-se analisar o cumprimento dos objetivos propostos e, com base no trabalho desenvolvido e no estudo geral realizado, são apresentadas algumas propostas de possíveis trabalhos futuros que permitam o desenvolvimento desta área de conhecimento.

5.1 Cumprimento dos Objetivos

Os objetivos desta dissertação passavam pelo tratamento dos dados de consumos elétricos que foram fornecidos por uma comercializadora e utilizá-los para a construção de um modelo de previsão probabilística baseado em decomposição de padrões.

Os dados em análise possuíam uma tendência crescente devido ao fato de a empresa comercializadora ser nova e o número de consumidores ir aumentando gradualmente. Posto isto, esta tendência teve de ser analisada pois podia ser uma componente crítica para a previsão.

Realizada a devida análise, ajustadas as tendências e feitos os devidos ajustes em todas as variáveis para que a previsão funcionasse, realizou-se a previsão dos quantis dos consumos elétricos baseada na distribuição beta. A escolha desta distribuição de probabilidade foi sem a mais acertada devido à sua capacidade de adaptação a qualquer situação.

Ao longo da construção do modelo surgiram duas variantes, a I e a II, pois achou-se que as pequenas falhas resultantes da Variante I podiam ser suavizadas obtendo-se assim valores mais precisos. A Variante II apresenta apenas algumas alterações relativamente à I de modo a melhorar a previsão efetuada, mas segue o mesmo modelo de decomposição de padrões.

Posto isto, conclui-se que o modelo da Variante I, poderá ser interessante para previsões de médio prazo (previsão para o ano seguinte), desde que a tendência anual se mantenha linear. No entanto para previsões de muito curto prazo, dia seguinte ou dias da semana seguinte, o modelo da Variante II resulta melhor uma vez que é muito mais adaptativo, apresenta melhor desempenho e não apresenta a necessidade de ajuste das tendências.

5.2 Trabalhos Futuros

Após a conclusão desta dissertação, existem trabalhos futuros e melhorias que se considera pertinente sugerir:

- Estudo e análise mais detalhada dos dias especiais: na presente dissertação os feriados foram tidos em consideração, contudo, no futuro os dias de acontecimentos importantes, tais como eventos desportivos ou eleições, deveriam ser considerados uma vez que causam um impacto significativo nas atividades dos consumidores.
- Seria também interessante a utilização de novas variáveis que trouxessem mais informação útil para a previsão, como a temperatura, humidade, etc.
- De modo mais generalizado, seria interessante aplicar outras técnicas de previsão referidas no capítulo do estado da arte para posteriormente comparar resultados obtidos com os dos modelos de decomposição.
- Por fim, a exploração de estratégias de oferta/procura em mercado com base em análise de risco que utilize previsão probabilística.

Referências

- [1] S. Khatoon and A. K. Singh, “Effects of various factors on electric load forecasting: An overview,” *Alexandria Eng. J.*, 2014.
- [2] K. R. Godden and S. J. L. J. Snyman, “ELECTRIC LOAD FORECASTING FOR HOLIDAY PERIODS by,” no. November, 1997.
- [3] C. L. Hor, S. J. Watson, and S. Majithia, “Analyzing the impact of weather variables on monthly electricity demand,” *IEEE Trans. Power Syst.*, 2005.
- [4] S. Ružić, A. Vučković, and N. Nikolić, “Weather Sensitive Method for Short Term Load Forecasting in Electric Power Utility of Serbia,” *IEEE Trans. Power Syst.*, 2003.
- [5] T. Hong and S. Fan, “Probabilistic Electric Load Forecasting: A Tutorial Review,” *Int. J. Forecast.*, vol. 32, pp. 1–32, 2014.
- [6] J. Á. González Ordiano, S. Waczowicz, V. Hagenmeyer, and R. Mikut, “Energy forecasting tools and services,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, p. e1235, 2017.
- [7] A. Misiorek, S. Trueck, and R. Weron, “Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models,” *Stud. Nonlinear Dyn. Econom.*, vol. 10, no. 3, 2006.
- [8] K. Stenshorne, “A Framework for Constructing and Evaluating Probabilistic Forecasts of Electricity Prices - A Case Study of the {N}ord {P}ool Market,” no. June, 2011.
- [9] J. Crespo Cuaresma, J. Hlouskova, S. Kossmeier, and M. Obersteiner, “Forecasting electricity spot-prices using linear univariate time-series models,” *Appl. Energy*, vol. 77, no. 1, pp. 87–106, 2004.
- [10] R. Weron and A. Misiorek, “Forecasting Spot Electricity Prices with Time Series Models,” *Int. Conf. "The Eur. Electr. Mark. EEM-05"*, p. 8, 2005.
- [11] M. Shafie-Khah, M. P. Moghaddam, and M. K. Sheikh-El-Eslami, “Price forecasting of day-ahead electricity markets using a hybrid forecast method,” *Energy Convers. Manag.*, vol. 52, no. 5, pp. 2165–2169, 2011.
- [12] K. C. Chatzidimitriou, A. C. Chrysopoulos, A. L. Symeonidis, and P. A. Mitkas, “Enhancing agent intelligence through evolving reservoir networks for predictions in Power Stock Markets,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2012, vol. 7103 LNAI, pp. 228–247.
- [13] R. Weron, “Electricity price forecasting: A review of the state-of-the-art with a look into the future,” *Int. J. Forecast.*, vol. 30, no. 4, pp. 1030–1081, 2014.
- [14] J. P. S. Catalão, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira, “Short-term electricity prices forecasting in a competitive market: A neural network approach,” *Electr. Power Syst. Res.*, vol. 77, no. 10, pp. 1297–1304, 2007.
- [15] D. Sansom, T. Downs, and T. Saha, “Evaluation of support vector machine based forecasting tool in electricity price forecasting for Australian national electricity market participants,” *J. Electr. ...*, vol. 22, no. 3, pp. 227–233, 2003.
- [16] J. H. Zhao, Z. Y. Dong, Z. Xu, and K. P. Wong, “A statistical approach for interval forecasting of the electricity price,” *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 267–

- 276, 2008.
- [17] T. Gneiting and M. Katzfuss, “Probabilistic Forecasting,” *Annu. Rev. Stat. Its Appl.*, 2014.
 - [18] P. Pinson, “Wind Energy: Forecasting Challenges for Its Operational Management,” *Stat. Sci.*, vol. 28, no. 4, pp. 564–585, 2013.
 - [19] Y. Zhang, J. Wang, and X. Wang, “Review on probabilistic forecasting of wind power generation,” *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.
 - [20] T. Hong, “Crystal Ball Lessons in Predictive Analytics,” *Energybiz*, vol. 12, no. 2, pp. 35–37, 2015.
 - [21] G. Gross and F. D. Galiana, “SHORT-TERM LOAD FORECASTING.,” *Proc. IEEE*, vol. 75, no. 12, pp. 1558–1573, 1987.
 - [22] T. Hong, “Energy Forecasting: Past, Present, and Future,” *Foresight Int. J. Appl. Forecast.*, no. 32, pp. 43–48, 2014.
 - [23] R. Weron and A. Misiorek, “Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models,” *Int. J. Forecast.*, vol. 24, no. 4, pp. 744–763, 2008.
 - [24] K. Maciejowska, J. Nowotarski, and R. Weron, “Probabilistic forecasting of electricity spot prices using Factor Quantile Regression Averaging,” *Int. J. Forecast.*, vol. 32, no. 3, pp. 957–965, 2016.
 - [25] P. Pinson, “Estimation of the uncertainty in wind power forecasting,” *Cent. Énergétique Procédés–Ecole des Mines*, p. 266, 2006.
 - [26] G. W. Bassett and R. W. Koenker, “Strong consistency of regression quantiles and related empirical processes,” *Econom. Theory*, vol. 2, no. 2, pp. 191–201, 1986.
 - [27] N. Meinshausen, “Quantile Regression Forests,” *J. Mach. Learn. Res.*, vol. 7, pp. 983–999, 2006.
 - [28] B. Hofner, A. Mayr, N. Robinzonov, and M. Schmid, “Model-based boosting in R: A hands-on tutorial using the R package mboost,” *Comput. Stat.*, vol. 29, no. 1–2, pp. 3–35, 2014.
 - [29] J. W. Taylor, “A Quantile Regression Neural Network Approach to Estimating the Conditional Density of Multiperiod Returns,” *J. Forecast.*, vol. 19, no. 0, pp. 299–311, 2000.
 - [30] A. J. Cannon, “Quantile regression neural networks: Implementation in R and application to precipitation downscaling,” *Comput. Geosci.*, vol. 37, no. 9, pp. 1277–1284, 2011.
 - [31] C. Chen, “A finite smoothing algorithm for quantile regression,” *J. Comput. Graph. Stat.*, vol. 16, no. 1, pp. 136–164, 2007.
 - [32] I. Takeuchi, Q. V. Le, T. Sears, and A. J. Smola, “Nonparametric Quantile Regression,” *J. Mach. Learn. Res.*, vol. 7, pp. 1001–1032, 2005.
 - [33] “No Title.” [Online]. Available: https://en.wikipedia.org/wiki/Beta_distribution. [Accessed: 25-Jun-2018].
 - [34] H. T. C. Pedro and C. F. M. Coimbra, “Assessment of forecasting techniques for solar power production with no exogenous inputs,” *Sol. Energy*, vol. 86, no. 7, pp. 2017–2028, 2012.
 - [35] P. Pinson, H. A. Nielsen, J. K. Møller, H. Madsen, and G. N. Kariniotakis, “Non-parametric probabilistic forecasts of wind power: Required properties and evaluation,” *Wind Energy*, vol. 10, no. 6, pp. 497–516, 2007.
 - [36] J. Juban, L. Fugon, and G. Kariniotakis, “Uncertainty estimation of wind power forecasts: Comparison of Probabilistic Modelling Approaches,” in *Proc. 2008 European Wind Energy Conference and Exhibition*, 2008, p. 10.
 - [37] T. Gneiting, F. Balabdaoui, and A. E. Raftery, “Probabilistic forecasts, calibration and sharpness,” *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 2007.
 - [38] R. J. Bessa, V. Miranda, A. Botterud, J. Wang, and E. M. Constantinescu, “Time Adaptive Conditional Kernel Density Estimation for Wind Power Forecasting,” *IEEE Transactions on Sustainable Energy*, vol. 3, no. 4, pp. 660–669, 2012.
 - [39] T. Jónsson, P. Pinson, H. Madsen, and H. A. Nielsen, “Predictive densities for day-ahead

electricity prices using time-adaptive quantile regression,” *Energies*, vol. 7, no. 9, pp. 5523–5547, 2014.