

Faculdade de Engenharia da Universidade do Porto



Previsão de consumo a médio e longo prazo

André Marques Rodrigues

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

Orientador: Nuno Fidalgo
Coorientador: José Paulos

26 de julho de 2021

Resumo

O objetivo principal deste trabalho é desenvolver um modelo para previsão do consumo a médio e longo prazo. Pretende-se prever o diagrama de carga anual, com detalhe horário, de três países europeus: Portugal, Espanha e França. Assume-se que estes diagramas de carga podem ser decompostos em duas componentes distintas: distribuição e amplitude (consumo global anual). A amplitude é frequentemente estimada em termos de análise de tendências, que dependem sobretudo de fatores socioeconómicos. É bastante frequente ver projeções da evolução do consumo anual realizadas por entidades oficiais. Assim, o foco deste trabalho reside na componente de distribuição do consumo global anual por cada hora do ano de previsão.

Este trabalho começa com uma análise das variáveis disponíveis com o intuito de perceber o seu comportamento ao longo de vários períodos do ano, para possam ser escolhidas aquelas que melhor se relacionam com o consumo energético horário, de forma a integrarem um modelo de previsão a médio e longo prazo. São construídos diversos modelos com conjuntos de variáveis de input diferentes que são testados consoante o erro apresentado face aos consumos registados reais.

Finalmente, é incluída também a temperatura com o objetivo de entender a influência desta na previsão a médio e longo prazo, e através de testes com várias séries de temperaturas, definir margens de previsão para cada hora.

Página em branco

Abstract

The purpose of this dissertation is the elaboration of a medium and long term load forecasting model. For this, we analyze several variables and their behavior across different periods of time, with the intent of inserting them into the model.

Various models with different combinations of input variables are built along the paper and tested against the actual registered load of three countries: Portugal, Spain, and France.

After the best model is found, we include the temperature as a new variable, to assess and study its influence, and, through the use of series of temperatures, we can build a cluster of predictions for each hour of the year and, that way, give flexibility to our load forecasting model.

Página em branco

Índice

Resumo	iii
Abstract	v
Índice	vii
Lista de figuras	ix
Lista de tabelas	x
Capítulo 1	1
Introdução	1
Capítulo 2	3
Dados	3
Capítulo 3	5
Metodologia	5
Capítulo 4	14
Modelos e Resultados	14
Modelo 1	14
Modelo 2	16
Modelo 3	18
Modelo 4	19
Inclusão da Temperatura	20
Margens de incerteza da previsão	21
Capítulo 5	26
Conclusão	26

Referências.....28

Lista de figuras

Figura 1 - Consumo médio horário por ano em Portugal (em MW).....	7
Figura 2 - Consumo médio horário por ano em Espanha (em MW)	7
Figura 3 - Consumo médio horário por ano em França (em MW).....	7
Figura 4 - Consumo médio horário por mês em Portugal (em MW)	8
Figura 5 - Consumo médio horário por mês em Espanha (em MW)	9
Figura 6 - Consumo médio horário por mês em França (em MW)	9
Figura 7 - Consumo médio horário em Portugal (em MW)	10
Figura 8 - Consumo médio horário em Espanha (em MW)	11
Figura 9 - Consumo médio horário em França (em MW)	11
Figura 10 - Consumo médio horário por dia da semana em Portugal (em MW)	12
Figura 11 - Consumo médio horário por dia da semana em Espanha (em MW)	12
Figura 12 - Consumo médio horário por dia da semana em França (em MW)	13
Figura 13 - Exemplo de previsões (7 dias) para Portugal associadas às séries de temperatura (em MW)	22
Figura 14 - Exemplo de previsões (7 dias) para Espanha associadas às séries de temperatura (em MW)	22
Figura 15 - Exemplo de previsões (7 dias) para França associadas às séries de temperatura (em MW)	23
Figura 16 - Previsões máximas e mínimas associadas às séries de temperatura para Portugal (em MW)	24
Figura 17 - Previsões máximas e mínimas associadas às séries de temperatura para Espanha (em MW)	24
Figura 18 - Previsões máximas e mínimas associadas às séries de temperatura para França (em MW)	25

Lista de tabelas

Tabela 1 - Consumo médio horário por ano e por país (em MW).....	6
Tabela 2 - Consumo médio horário por mês e por país (em MW)	8
Tabela 3 - Consumo médio horário por país (em MW)	10
Tabela 4 - Consumo médio horário por dia da semana e por país (em MW).....	12
Tabela 5 - Resultados do modelo 1 (incluindo 2020) referente a Portugal	15
Tabela 6 - Resultados do modelo 1 referente a Portugal	15
Tabela 7 - Resultados do modelo 1 referente a Espanha	16
Tabela 8 - Resultados do modelo 1 referente a França	16
Tabela 9 - Resultados do modelo 2 referente a Portugal	17
Tabela 10 - Resultados do modelo 2 referente a Espanha	17
Tabela 11 - Resultados do modelo 2 referente a França.....	17
Tabela 12 - Resultados do modelo 3 referente a Portugal.....	18
Tabela 13 - Resultados do modelo 3 referente a Espanha	18
Tabela 14 - Resultados do modelo 3 referente a França.....	18
Tabela 15 - Resultados do modelo 4 referente a Portugal.....	19
Tabela 16 - Resultados do modelo 4 referente a Espanha	19
Tabela 17 - Resultados do modelo 4 referente a França.....	20
Tabela 18 - Resultados do modelo 4 após inclusão da temperatura referente a Portugal	20
Tabela 19 - Resultados do modelo 4 após inclusão da temperatura referente a Espanha	21
Tabela 20 - Resultados do modelo 4 após inclusão da temperatura referente a França	21

Capítulo 1

Introdução

A previsão do consumo ajuda as empresas responsáveis na operação e gestão da rede elétrica [1]. Os estudos de previsão dividem-se em três categorias consoante o seu horizonte de previsão. Mais concretamente, a previsão a curto prazo permite estimar o consumo desde a próxima hora até uma semana e, assim, decidir que unidades de geração despachar de forma a satisfazer o consumo antevisto. Já a médio prazo, o horizonte corresponde a um período que pode variar entre uma semana e um ano, e o objetivo passa por prever o consumo de modo a satisfazer as épocas de verão ou inverno e agendar períodos de manutenção. Finalmente, o horizonte da previsão a longo prazo vai desde um ano até aos 20 anos e permite perceber onde é necessário reforçar o sistema elétrico de energia de modo que possa responder às previsões futuras [2].

Com este trabalho pretende-se prever o diagrama de carga anual, com detalhe horário, a médio e longo prazo, de três países europeus: Portugal, Espanha e França. Estes diagramas de carga podem ser decompostos em duas componentes distintas: distribuição e amplitude. Em termos de distribuição, o consumo varia consoante os ciclos diários e sazonais [3] e outros fatores como feriados e pontes. No que toca à amplitude, o consumo de energia depende da situação socioeconómica do país e respetiva evolução [4]. Assim, este trabalho pretende focar mais atenção no estudo da distribuição deste consumo energético ao longo do ano, ou seja, o intuito principal deste trabalho é obter um processo fiável que determine a variação do consumo horário em diferentes países e em diferentes períodos do ano.

Concretamente, os objetivos desta dissertação passam, primeiro, pela criação de modelos de previsão de consumo horário a médio e longo prazo para cada um dos países

através de redes neurais e, depois, pelo estudo da influência da temperatura nos valores de consumo, através de modelos com séries de temperatura.

Sendo que a maioria dos trabalhos de previsão de longo prazo não separa a parte do nível de consumo e a sua distribuição ao longo do ano, uma das principais contribuições deste trabalho passa precisamente por esta distinção e no respetivo foco na distribuição. Além disso, é realizado um processamento de dados prévio que é inovador neste trabalho e que permitiu melhorar o desempenho.

Este documento está organizado por capítulos. O primeiro é a introdução, que faz o enquadramento do trabalho e apresenta os respetivos objetivos. O segundo capítulo menciona a proveniência dos dados utilizados ao longo do trabalho e a sua organização. O terceiro capítulo descreve a metodologia, de modo a aprofundar os processos e conceitos envolvidos no trabalho. O quarto capítulo apresenta os modelos e os resultados obtidos, procedendo à sua análise e comparação. Finalmente, o quinto capítulo é dedicado à apresentação das principais conclusões.

Capítulo 2

Dados

Os dados que vão ser utilizados e analisados ao longo deste trabalho foram obtidos de uma base de dados do INESC TEC [5] e estão divididos em dois tipos de documentos que correspondem, respetivamente, à informação energética e à meteorológica dos três países já mencionados: Portugal, Espanha e França, para os anos entre 2015 e 2020. Adicionalmente, foram também proporcionados dados relativos às temperaturas registadas em cada hora desde 2010 até 2020, em Portugal, Espanha e França. Especificamente, o primeiro documento sintetiza a informação energética de cada país através das seguintes variáveis para cada data e hora:

- consumo (MW);
- produção fotovoltaica (MW);
- eólica (MW);
- nuclear (MW);
- hidroelétrica (reserva, fio de água e bombeada) (MW);
- gás natural (MW);
- carvão (MW);
- biomassa (MW);
- preço SPOT (€/MWh).

No outro documento, estão organizados por data e hora os dados referentes à irradiância solar, temperatura e velocidade do vento.

Antes de ser iniciado o treino de cada rede neuronal, é necessário dividir as amostras em três conjuntos: treino, validação e teste. Como se pretende treinar esta rede com perspetiva de previsão futura, o conjunto de treino e validação foi reservado para as amostras dos anos de 2015 a 2018 e o conjunto de teste para as amostras de 2019 e 2020.

Isto significa que a rede neuronal será exposta ao conjunto das amostras entre 2015 e 2018 com o intuito de treinar e validar os seus processos internos e, de seguida, será testada pelas amostras de 2019 e 2020, comparando os resultados das previsões da rede com os valores reais obtidos nestes dois anos.

Capítulo 3

Metodologia

Nesta secção, irá ser realizada a análise dos dados disponibilizados, de forma a identificar que relações existem entre as variáveis e, por sua vez, efetuar a seleção das variáveis conforme a sua influência na saída. De seguida, constroem-se várias hipóteses para modelos de previsão consoante os resultados da análise anterior, referindo as combinações de entradas e arquiteturas utilizadas em cada rede neuronal.

Um dos principais desafios na previsão, além do método, passa pela escolha do grupo de variáveis que melhor se adequa e relaciona com o elemento a prever. Enquanto em modelos a curto prazo, deve ser trabalhada a correlação desse elemento com um grupo de variáveis de cariz recente, em modelos a longo prazo, é necessário utilizar variáveis menos voláteis no tempo e que transmitam a informação do comportamento a longo prazo. A previsão de médio a longo prazo refere-se a horizontes tipicamente entre 1 e 10 anos no futuro. Para obter previsões plausíveis, é necessário estudar o impacto que determinadas variáveis têm nos valores do consumo e concluir se terão a mesma importância no futuro e, assim, obter um conjunto de variáveis fiáveis para este trabalho. Neste aspeto, as variáveis de calendário são candidatas naturais, pois representam as relações históricas ao longo de cada ano, mês, hora, etc. Para além de representarem os efeitos sazonais (diários, semanais, etc.), são grandezas que se encontram antecipadamente definidas, ao contrário, por exemplo, da temperatura que pode afetar o consumo [6], mas é impossível prever detalhadamente a temperatura a longo prazo.

A primeira abordagem à previsão do consumo foi feita através do treino e aplicação de redes neuronais. Dado que as redes neuronais são uma técnica bastante divulgada [7], não se inclui nesta dissertação a sua descrição.

Em resumo, a abordagem adotada inclui os seguintes passos fundamentais:

1. Análise dos dados disponibilizados.
2. Seleção de variáveis.
3. Realização de testes de previsão e seleção dos melhores modelos.
4. Inclusão da temperatura e testes de previsão com diferentes séries de temperatura.

Dado que as fases 2 a 4 são bastante interativas, optou-se por incluir a sua descrição no capítulo de Resultados.

Análise dos dados

Foram facultados três conjuntos de dados referentes à produção e consumo de energia, entre os anos de 2015 e 2020, de três países distintos: Portugal, Espanha e França.

Sendo o principal objetivo a previsão a médio e longo prazo, optou-se por focar inicialmente na análise das variáveis de calendário. Para isso, foram criadas amostras que agrupem as variáveis de calendário com o consumo horário de cada país, ou seja, para cada país, cada amostra apresenta as seguintes variáveis: ano, mês, dia, dia da semana, hora e o consumo correspondente. Além disso foi também acoplada uma variável binária externa que representa se o dia em questão é considerado feriado oficial.

Previamente ao treino da rede neuronal, foi também feito um estudo ao conjunto de dados proporcionado, para ser possível perceber como varia o consumo com as variáveis de calendário e, assim, ter uma perspetiva mais consolidada em relação ao impacto destas variáveis. Para isso foram construídos gráficos que comparem os consumos médios registados para cada variável de calendário.

Começou-se pelo estudo do consumo médio horário de cada um dos três países para os anos de 2015 a 2020:

Tabela 1 - Consumo médio horário por ano e por país (em MW)

Ano	Portugal	Espanha	França
2015	5589	28185	54089
2016	5609	28508	55272
2017	5660	28795	54664
2018	5816	29066	53057
2019	5738	28550	53231
2020	5650	26931	55037

Para ilustrar a variação deste valor ao longo dos 6 anos, foram construídos os seguintes gráficos de barras:

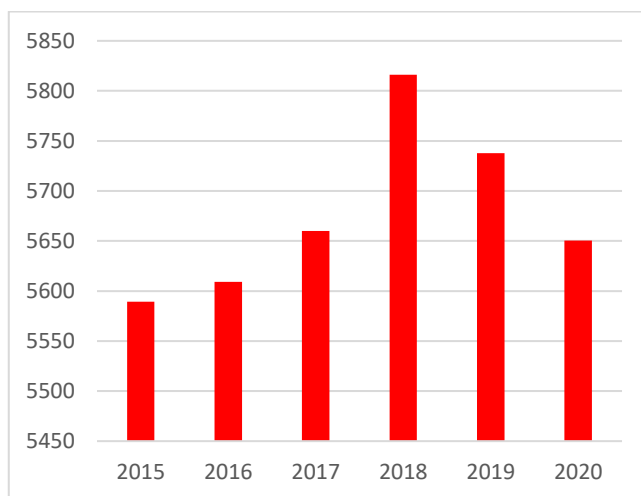


Figura 1 - Consumo médio horário por ano em Portugal (em MW)

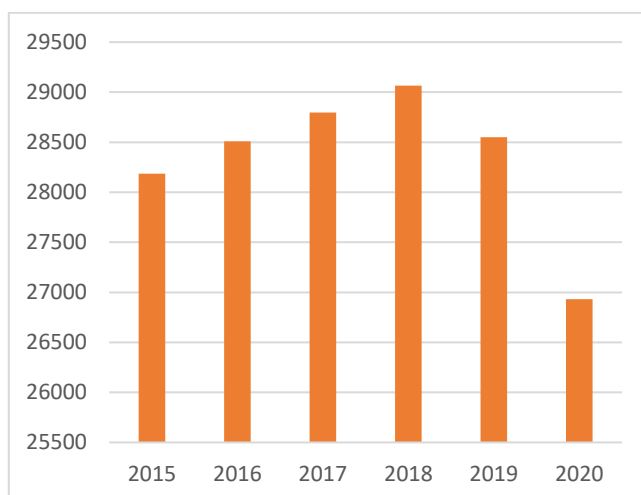


Figura 2 - Consumo médio horário por ano em Espanha (em MW)

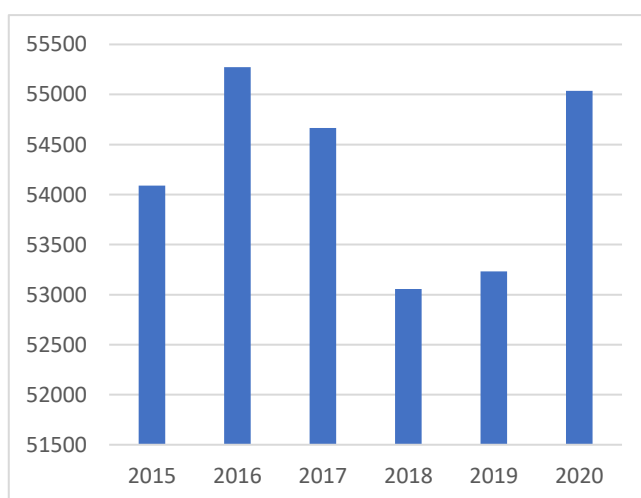


Figura 3 - Consumo médio horário por ano em França (em MW)

Consegue-se observar um crescimento gradual contínuo do consumo, tanto em Portugal como em Espanha, até ao ano de 2018, sendo que, de seguida, nota-se uma descida nos anos de 2019 e 2020. Já o consumo em França apresenta um comportamento mais errático, já que este aumenta nos dois primeiros anos, diminui em 2017 e 2018 e, finalmente, cresce novamente nos dois últimos anos do conjunto de dados.

O estudo seguinte refere-se ao consumo médio diário por mês:

Tabela 2 - Consumo médio horário por mês e por país (em MW)

Mês	Portugal	Espanha	França
1	6317	30665	69204
2	6128	30162	66931
3	5791	28454	60539
4	5328	26232	49365
5	5251	26171	45325
6	5479	28539	45307
7	5707	30387	45888
8	5392	28905	42584
9	5576	28212	45475
10	5434	27014	49586
11	5740	28417	58798
12	5924	28680	63906

Da mesma forma, elaborou-se um gráfico para cada país para facilitar a compreensão da influência que a época do ano tem no consumo de energia a nível nacional:

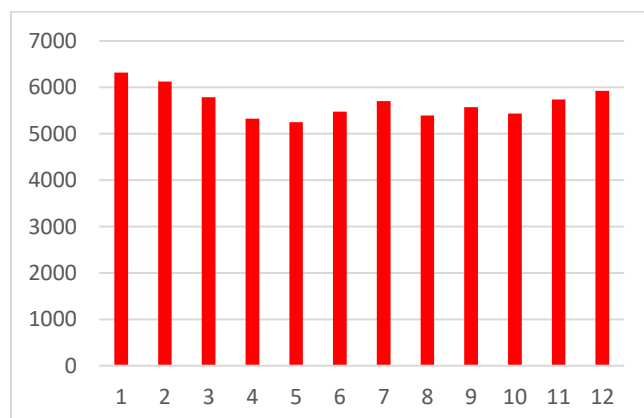


Figura 4 - Consumo médio horário por mês em Portugal (em MW)

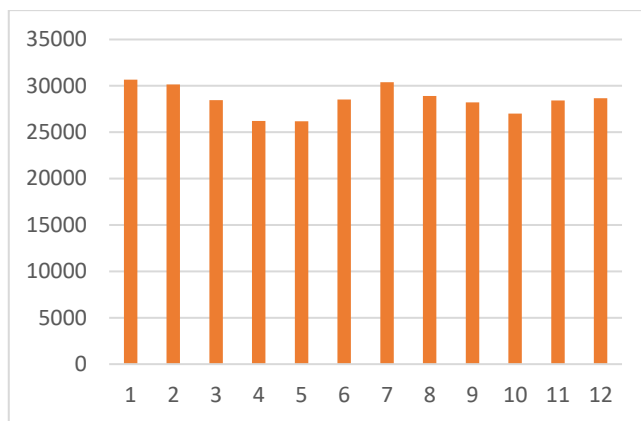


Figura 5 - Consumo médio horário por mês em Espanha (em MW)

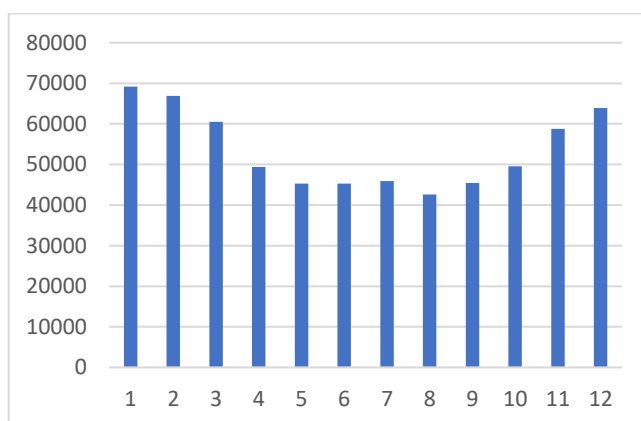


Figura 6 - Consumo médio horário por mês em França (em MW)

Neste caso, os três gráficos apresentam comportamentos mais aproximados, sendo possível identificar que o consumo médio alcança o seu valor máximo nos meses de Inverno, nomeadamente janeiro, fevereiro e dezembro, como seria de esperar, e também se nota a ponta de verão em julho, particularmente em Portugal e Espanha.

Avançando para a análise da distribuição do consumo médio por hora por país:

Tabela 3 - Consumo médio horário por país (em MW)

Hora	Portugal	Espanha	França
0	4940	23759	49318
1	4690	22959	47372
2	4543	22589	45759
3	4475	22634	46021
4	4473	23584	48384
5	4536	25473	52110
6	4830	27750	55449
7	5350	29575	57432
8	5830	30868	58416
9	6113	31616	58882
10	6276	31959	59481
11	6315	32019	58968
12	6242	31629	57555
13	6222	30785	55783
14	6213	30200	54215
15	6149	29913	53624
16	6066	29816	55134
17	6122	30427	57595
18	6417	31446	57552
19	6624	32110	55397
20	6550	31064	54711
21	6219	28949	55409
22	5793	26928	53919
23	5316	25151	51125

Para visualizar a variação do consumo médio a cada hora, foram construídos os três gráficos de barras a comparar a situação nos três países:

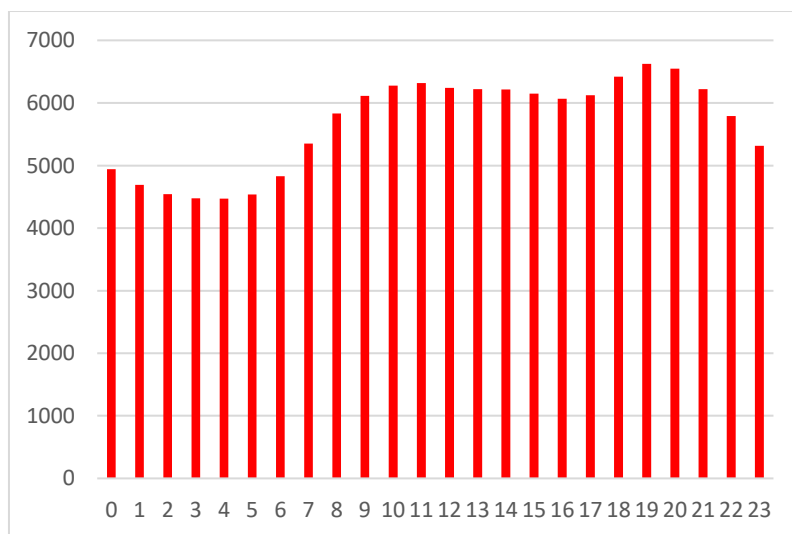


Figura 7 - Consumo médio horário em Portugal (em MW)

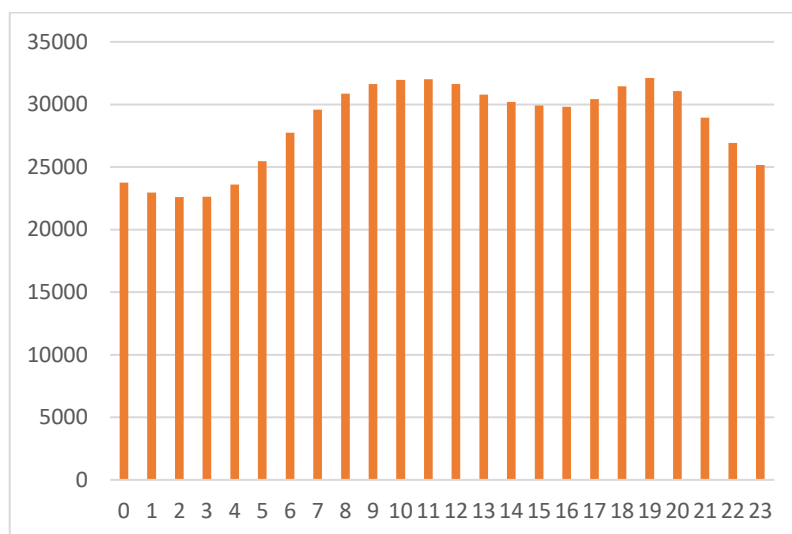


Figura 8 - Consumo médio horário em Espanha (em MW)

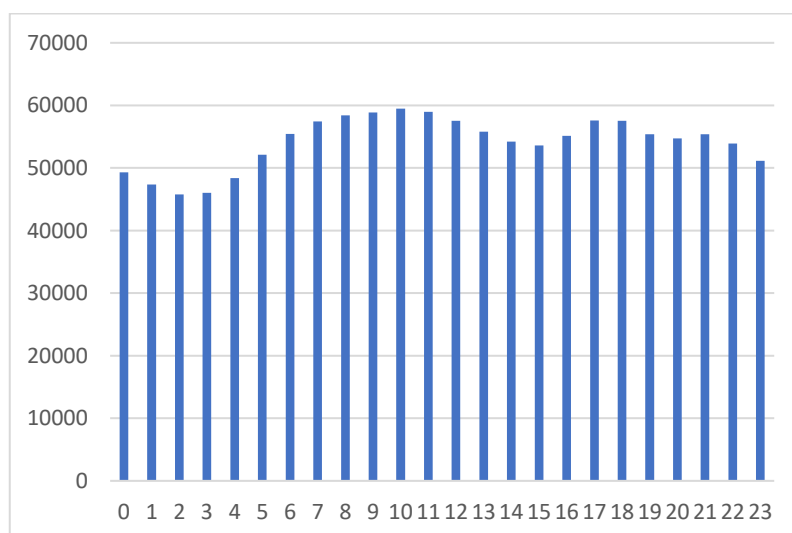


Figura 9 - Consumo médio horário em França (em MW)

As três figuras exibem uma distribuição horária do consumo muito semelhante, o que mostra que esta é uma variável de alto nível de confiança no que toca à previsão a longo prazo.

Finalmente, realizou-se o mesmo estudo para o dia da semana, ou seja, calculou-se o consumo médio horário de cada dia da semana para cada país:

Tabela 4 - Consumo médio horário por dia da semana e por país (em MW)

Dia da Semana		Portugal	Espanha	França
Domingo	(1)	4916	24696	48239
Segunda	(2)	5798	29122	55104
Terça	(3)	5937	29824	56548
Quarta	(4)	5978	29886	56563
Quinta	(5)	5963	29745	56458
Sexta	(6)	5901	29387	55613
Sábado	(7)	5249	26510	50533

A seguir são traduzidos estes resultados em gráficos de barras:

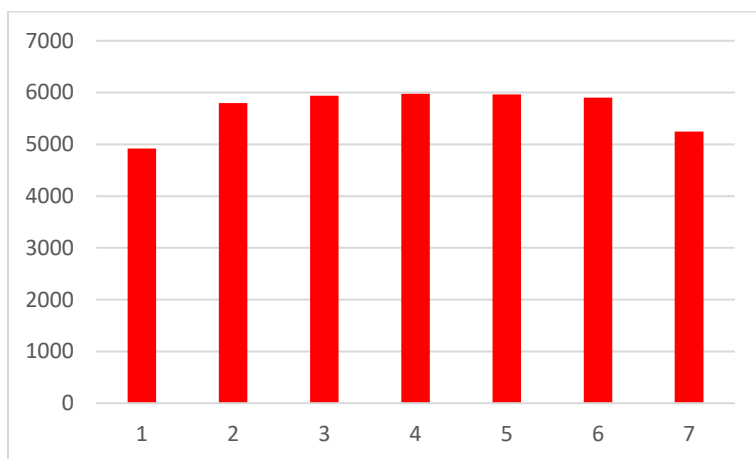


Figura 10 - Consumo médio horário por dia da semana em Portugal (em MW)

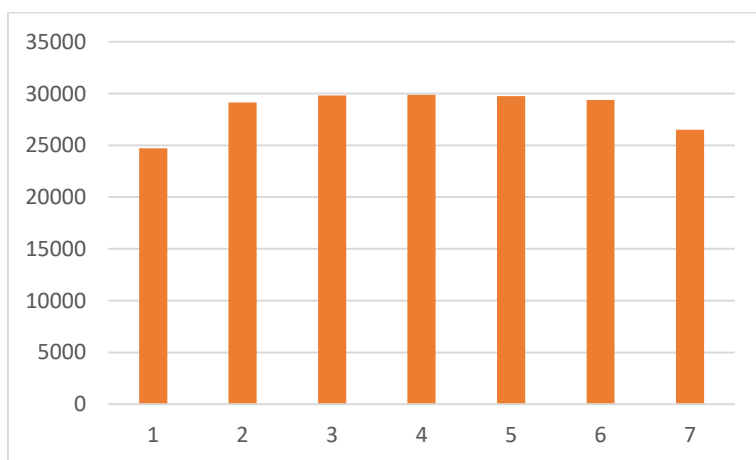


Figura 11 - Consumo médio horário por dia da semana em Espanha (em MW)

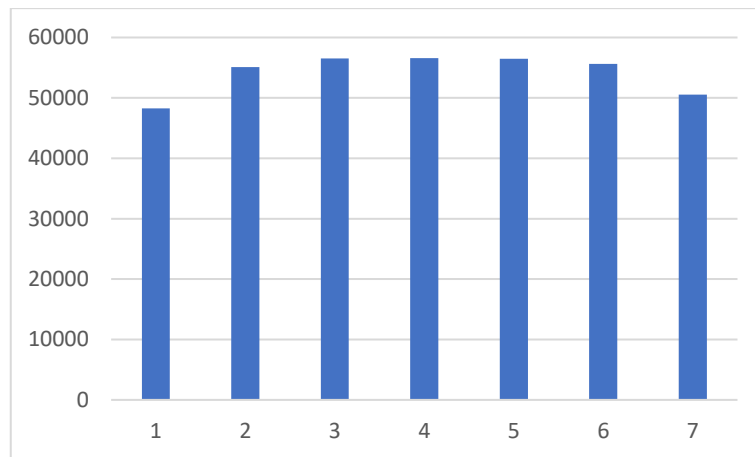


Figura 12 - Consumo médio horário por dia da semana em França (em MW)

Novamente, a distribuição do consumo apresenta as mesmas características nos três países, realçando o consumo médio nos dias de fim de semana (no gráfico representados por 1 e 7), em que, logicamente, o valor decresce. Assim, esta é também uma variável segura a utilizar na previsão a longo a prazo.

Capítulo 4

Modelos e Resultados

Através da interface de criação de redes neuronais do software *Matlab*, foram criadas várias redes neuronais, sempre com um rácio de validação de 25% para o conjunto de dados correspondente aos anos entre 2015 e 2018, sendo necessário definir o número de neurónios escondidos para cada rede.

Para cada modelo, foram construídas e testadas várias redes, de raiz, com um número diferente de neurónios escondidos, sendo guardadas, nas tabelas a seguir, as quatro redes com os melhores resultados de teste.

Os resultados serão sumarizados em tabelas nas quais serão exibidos três diferentes tipos de medidas que demonstram a performance da rede respetiva, além do número de neurónios escondidos definido e o número de iterações alcançado no processo. O primeiro é o coeficiente R, também conhecido por coeficiente de correlação de *Pearson*, que avalia a correlação linear entre dois conjuntos de dados, neste caso, os valores de previsão a que a rede chegou e os valores reais de consumo. O segundo é o RMSE, *Root Mean Square Error*, ou seja, a raiz do quadrado do erro médio de previsão, e o terceiro é o MAPE, *Mean Absolute Percentage Error*, que indica o erro médio absoluto em percentagem. Serão usados indicadores para o conjunto de treino e teste separadamente, para distinguir o desempenho nessas duas fases. Nestas tabelas, cada coluna corresponde a uma rede diferente, ou seja, uma inicialização e treino diferentes.

Modelo 1

Para a primeira tentativa de previsão, foram escolhidas as variáveis de calendário regulares, ou seja, os valores diretos do ano, mês, dia, dia da semana, hora e um variável

binária que indica se o dia corresponde a um feriado oficial, ou seja, 1 se for um feriado e 0 caso não seja.

Inicialmente este processo foi feito com o grupo de variáveis de teste a incluir os dados dos anos de 2019 e 2020, e o resultado para Portugal foi o seguinte:

Tabela 5 - Resultados do modelo 1 (incluindo 2020) referente a Portugal

N ESCONDIDOS	30	13	11	14
ITERAÇÕES	269	105	63	136
R	0.940	0.931	0.915	0.914
RMSE TREINO	288.84	313.53	349.61	359.92
RMSE TESTE	457.74	484.85	538.80	523.31
MAPE TREINO	3.74%	4.21%	4.69%	4.81%
MAPE TESTE	6.02%	6.24%	7.05%	6.96%

Observou-se que, nos dados de 2020, estava em falta informação relativa aos meses entre junho e outubro, o que impede de ser feita a previsão para esse período, e pode levar ao aumento dos erros. Para evitar isto, foi repetido este processo apenas com os dados de 2019, no que toca às amostras de teste.

O primeiro modelo de previsão apresentou os seguintes resultados:

Tabela 6 - Resultados do modelo 1 referente a Portugal

N ESCONDIDOS	15	16	18	20
ITERAÇÕES	201	213	75	72
R	0.938	0.931	0.941	0.943
RMSE TREINO	325.39	347.98	317.82	300.39
RMSE TESTE	416.58	404.83	405.51	434.45
MAPE TREINO	4.34%	4.64%	4.21%	3.91%
MAPE TESTE	5.44%	5.18%	5.20%	5.46%

Concluiu-se, de imediato, que os resultados são melhores. Esta situação pode ser explicada, também, com o facto de 2020 ter sido um ano de pandemia, resultando numa situação atípica do ponto de vista energético. Logo, foi decidido utilizar apenas os dados de 2019 como conjunto de teste, nos restantes modelos.

Tabela 7 - Resultados do modelo 1 referente a Espanha

N ESCONDIDOS	12	13	14	15
ITERAÇÕES	305	236	86	27
R	0.926	0.928	0.922	0.922
RMSE TREINO	1674.90	1673.95	1739.63	1757.61
RMSE TESTE	1895.76	1771.69	1884.06	1842.72
MAPE TREINO	4.52%	4.48%	4.69%	4.69%
MAPE TESTE	5.01%	4.71%	5.04%	4.96%

Tabela 8 - Resultados do modelo 1 referente a França

N ESCONDIDOS	14	12	10	12
ITERAÇÕES	453	298	128	174
R	0.934	0.923	0.936	0.931
RMSE TREINO	3971.15	4402.84	4073.33	4243.11
RMSE TESTE	5107.84	4873.09	4298.72	4431.25
MAPE TREINO	5.56%	6.29%	5.74%	5.92%
MAPE TESTE	7.10%	6.90%	6.06%	6.29%

Os resultados aparentam ser positivos. O coeficiente R mostra valores elevados em todos os países, principalmente em Portugal, já o MAPE varia algo de país para país, registrando-se, como se vê pela tabela 7, o valor mais alto em França.

Modelo 2

Neste modelo o único input alterado foi a variável do ano que foi substituído por uma variável que representa o consumo horário médio registado em cada ano, denominada *PYear*. Depois foi eliminada a variável Dia, visto esta ter pouca influência no consumo registado. Portanto, as variáveis de entrada são: *PYear*, mês, dia da semana, hora e feriado.

Tabela 9 - Resultados do modelo 2 referente a Portugal

N ESCONDIDOS	15	14	12	14
ITERAÇÕES	40	38	163	47
R	0.936	0.934	0.934	0.929
RMSE TREINO	337.43	346.41	343.71	357.94
RMSE TESTE	387.57	394.88	378.89	381.16
MAPE TREINO	4.45%	4.60%	4.68%	4.83%
MAPE TESTE	5.00%	5.11%	4.99%	4.91%

Tabela 10 - Resultados do modelo 2 referente a Espanha

N ESCONDIDOS	15	14	14	16
ITERAÇÕES	128	54	71	565
R	0.938	0.924	0.927	0.933
RMSE TREINO	1566.78	1736.15	1702.59	1610.75
RMSE TESTE	1642.50	1741.98	1712.92	1713.04
MAPE TREINO	4.14%	4.72%	4.57%	4.27%
MAPE TESTE	4.38%	4.70%	4.72%	4.58%

Tabela 11 - Resultados do modelo 2 referente a França

N ESCONDIDOS	10	11	10	11
ITERAÇÕES	710	126	296	185
R	0.923	0.926	0.922	0.929
RMSE TREINO	4531.34	4380.18	4543.13	4257.35
RMSE TESTE	4391.01	4586.94	4473.81	4607.17
MAPE TREINO	6.28%	6.23%	6.34%	5.99%
MAPE TESTE	6.20%	6.54%	6.48%	6.65%

Depois de analisar as três tabelas anteriores, verifica-se que as alterações introduzidas neste modelo 2 permitiram alcançar melhores resultados, principalmente em Portugal e Espanha. Já o caso francês apresenta resultados similares, podendo ser observado facilmente, através da análise ao coeficiente R, que a rede neuronal não conseguiu encontrar uma relação tão forte entre as novas variáveis e o consumo registado.

Modelo 3

Seguindo a lógica do modelo anterior, são alteradas outras três variáveis de calendário para as suas respetivas versões em termos de consumo médio horário. Especificamente, as variáveis do mês, dia da semana e hora são transformadas nas variáveis de entrada *PMonth*, *PWeekday* e *PHour*.

Tabela 12 - Resultados do modelo 3 referente a Portugal

N ESCONDIDOS	10	15	16	12
ITERAÇÕES	21	17	14	28
R	0.931	0.938	0.930	0.932
RMSE TREINO	350.87	345.31	354.03	345.53
RMSE TESTE	391.59	390.19	395.26	396.04
MAPE TREINO	4.64%	4.56%	4.69%	4.64%
MAPE TESTE	5.15%	5.07%	5.14%	5.18%

Tabela 13 - Resultados do modelo 3 referente a Espanha

N ESCONDIDOS	10	11	12	11
ITERAÇÕES	24	82	55	28
R	0.922	0.926	0.927	0.924
RMSE TREINO	1750.80	1686.77	1691.83	1726.35
RMSE TESTE	1758.29	1805.63	1753.68	1765.84
MAPE TREINO	4.73%	4.53%	4.54%	4.69%
MAPE TESTE	4.78%	4.98%	4.74%	4.81%

Tabela 14 - Resultados do modelo 3 referente a França

N ESCONDIDOS	5	6	7	8
ITERAÇÕES	347	96	436	69
R	0.925	0.925	0.926	0.926
RMSE TREINO	4496.33	4467.89	4427.30	4425.16
RMSE TESTE	4370.35	4403.29	4397.16	4495.44
MAPE TREINO	6.29%	6.22%	6.16%	6.15%
MAPE TESTE	6.27%	6.32%	6.30%	6.45%

Este modelo apresenta melhorias ligeiras em França. Neste modelo, a transformação destas variáveis nunca iria trazer os mesmos benefícios da transformação realizada no modelo 2, pois o ajuste ao consumo horário esperado em cada ano já foi feito, logo o

aprofundamento deste processo para o mês, semana e hora não poderia resultar em retornos superiores.

Modelo 4

Depois de analisar os resultados dos modelos anteriores, foi feita uma revisão dos erros absolutos em cada hora e observou-se que existe uma tendência, em todos os modelos e todos os países, de subestimar o valor do consumo em determinados dias particulares do ano. Para mitigar este comportamento, foi criada uma nova variável denominada *Specialday*, atribuída ao dia 24 de dezembro e aos dias entre 26 e 31 de dezembro. Além disso, foi decidido substituir a variável *PHour* pela variável designada *PWeekhour*. Esta variável permite obter, para cada hora, o seu consumo médio horário em cada semana do ano e, assim, continuar a cumprir a função da variável anterior, resultando numa rede mais eficaz e menos redundante. Assim, para o modelo 4, as variáveis de entrada são: *PYear*, *PMonth*, *PWeekday*, *PWeekhour*, *feriado* e *Specialday*.

Tabela 15 - Resultados do modelo 4 referente a Portugal

N ESCONDIDOS	6	7	6	8
ITERAÇÕES	44	41	75	127
R	0.955	0.956	0.953	0.955
RMSE TREINO	288.88	284.60	293.12	282.02
RMSE TESTE	307.64	313.83	330.53	333.11
MAPE TREINO	3.84%	3.76%	3.87%	3.74%
MAPE TESTE	4.07%	4.10%	4.37%	4.47%

Tabela 16 - Resultados do modelo 4 referente a Espanha

N ESCONDIDOS	8	6	7	9
ITERAÇÕES	53	61	240	21
R	0.951	0.950	0.950	0.947
RMSE TREINO	1380.83	1386.47	1392.30	1439.93
RMSE TESTE	1522.17	1519.84	1485.26	1520.56
MAPE TREINO	3.63%	3.68%	3.68%	3.82%
MAPE TESTE	4.10%	4.14%	4.01%	4.11%

Tabela 17 - Resultados do modelo 4 referente a França

N ESCONDIDOS	5	6	5	6
ITERAÇÕES	50	27	595	59
R	0.939	0.943	0.943	0.945
RMSE TREINO	4017.09	3858.11	3863.29	3772.00
RMSE TESTE	4115.34	4102.56	4098.54	4085.83
MAPE TREINO	5.44%	5.21%	5.23%	5.11%
MAPE TESTE	5.80%	5.94%	5.82%	5.77%

A adição da variável *Specialday* para este modelo teve apenas o intuito de diminuir os erros à volta do dia 25 de dezembro.

Depois de analisar os erros em cada hora, é visível uma melhoria dos resultados do dia 24 de dezembro, contudo o mesmo não se regista para os restantes dias.

Em relação aos indicadores gerais da performance da rede neuronal, o modelo 4 apresenta os melhores resultados até agora registados. Esta performance deve-se, em grande parte, à nova variável *PWeekhour* que sintetiza muito bem a informação do consumo horário ao longo do ano.

Inclusão da Temperatura

Neste capítulo pretende-se elaborar uma nova rede que tenha por base o melhor modelo obtido anteriormente, o modelo 4, e, que de seguida seja adicionada uma variável de input que represente a temperatura registada a cada hora, entre os anos de 2015 e 2018, para o processo de treino e validação. Os resultados das tabelas que se seguem tiveram como conjunto de teste os dados referentes ao ano de 2019.

Tabela 18 - Resultados do modelo 4 após inclusão da temperatura referente a Portugal

N ESCONDIDOS	8
ITERAÇÕES	298
R	0.962
RMSE TREINO	260.98
RMSE TESTE	298.61
MAPE TREINO	3.40%
MAPE TESTE	3.80%

Tabela 19 - Resultados do modelo 4 após inclusão da temperatura referente a Espanha

N ESCONDIDOS	8
ITERAÇÕES	198
R	0.955
RMSE TREINO	1324.80
RMSE TESTE	1398.32
MAPE TREINO	3.38%
MAPE TESTE	3.71%

Tabela 20 - Resultados do modelo 4 após inclusão da temperatura referente a França

N ESCONDIDOS	8
ITERAÇÕES	448
R	0.966
RMSE TREINO	2988.38
RMSE TESTE	3143.44
MAPE TREINO	4.09%
MAPE TESTE	4.44%

A inclusão da temperatura permitiu reduzir o erro nos três países. Isto é expectável pois as variações de temperatura, e em particular os extremos, têm influência na quantidade de tempo que a população passa nas suas casas e no consumo de energia.

Margens de incerteza da previsão

De seguida, com a rede já treinada, o processo de teste passa por associar ao ano de previsão (2019) as diferentes séries de temperatura (2010 a 2019), de modo a obter uma série anual de previsões para cada série de temperatura considerada.

Com os conjuntos de previsões obtidos, é possível retirar vários valores de previsão para cada hora e, desta maneira, obter uma margem de resultados.

Para facilitar a leitura e análise dos gráficos a seguir, foi escolhida, como exemplo, a semana entre 2 e 9 de setembro de 2019 para ilustrar e comparar cada previsão resultante de cada série de temperatura, em cada país.

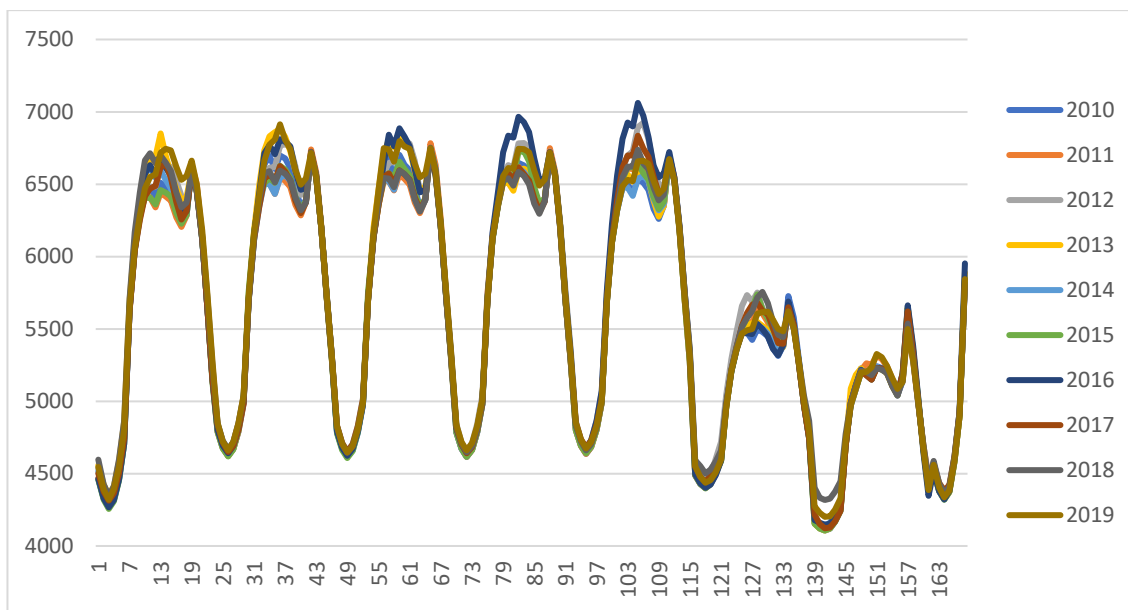


Figura 13 - Exemplo de previsões (7 dias) para Portugal associadas às séries de temperatura (em MW)

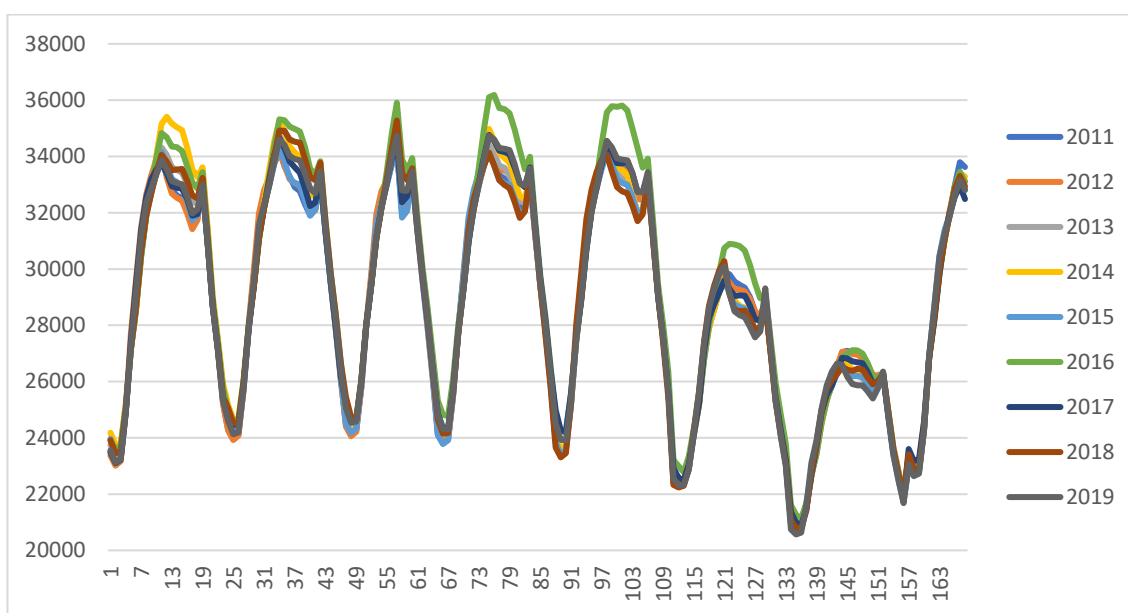


Figura 14 - Exemplo de previsões (7 dias) para Espanha associadas às séries de temperatura (em MW)

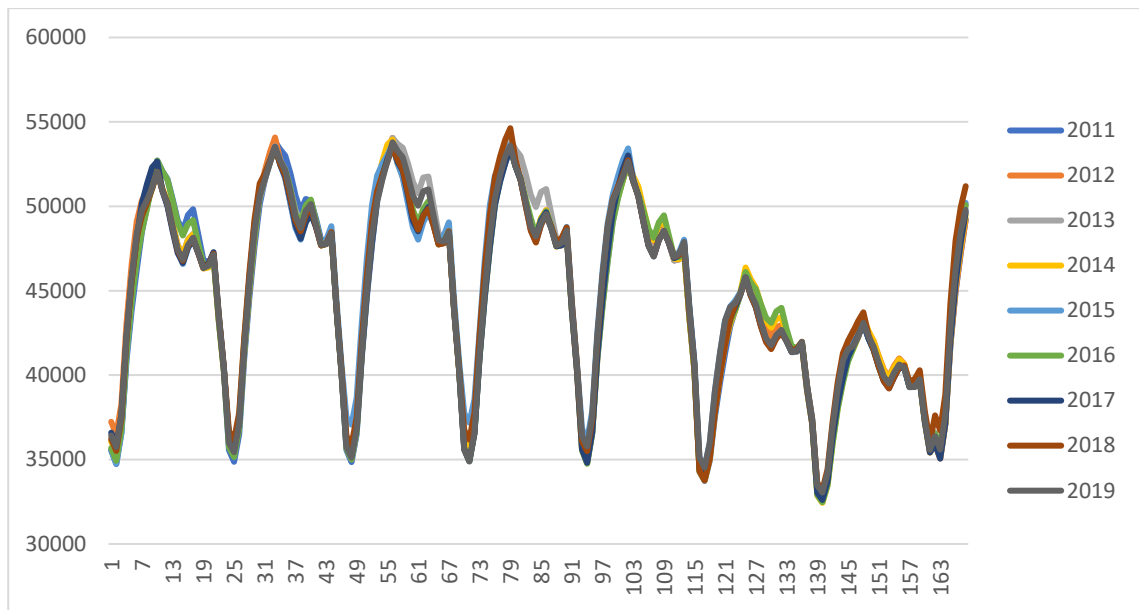


Figura 15 - Exemplo de previsões (7 dias) para França associadas às séries de temperatura (em MW)

É possível encontrar muitas semelhanças nas três figuras, contudo, para esta semana, as previsões para Portugal e Espanha são um pouco mais diferenciadas do que em França, que apresenta resultados mais consistentes. Isto pode dever-se ao facto de o consumo em França ter uma forte componente industrial [8], menos suscetível a variações de temperatura. Outra possível explicação será a eventual incapacidade de o modelo conseguir captar corretamente a influência da temperatura, dando-lhe pouco peso. Esta última hipótese não parece muito provável, dado que se verificou uma diminuição considerável do erro quando o modelo passou a incluir temperatura.

De seguida pretende-se destacar os limites máximos e mínimos obtidos com a introdução destas séries de temperatura, e comparar com o consumo real registado.

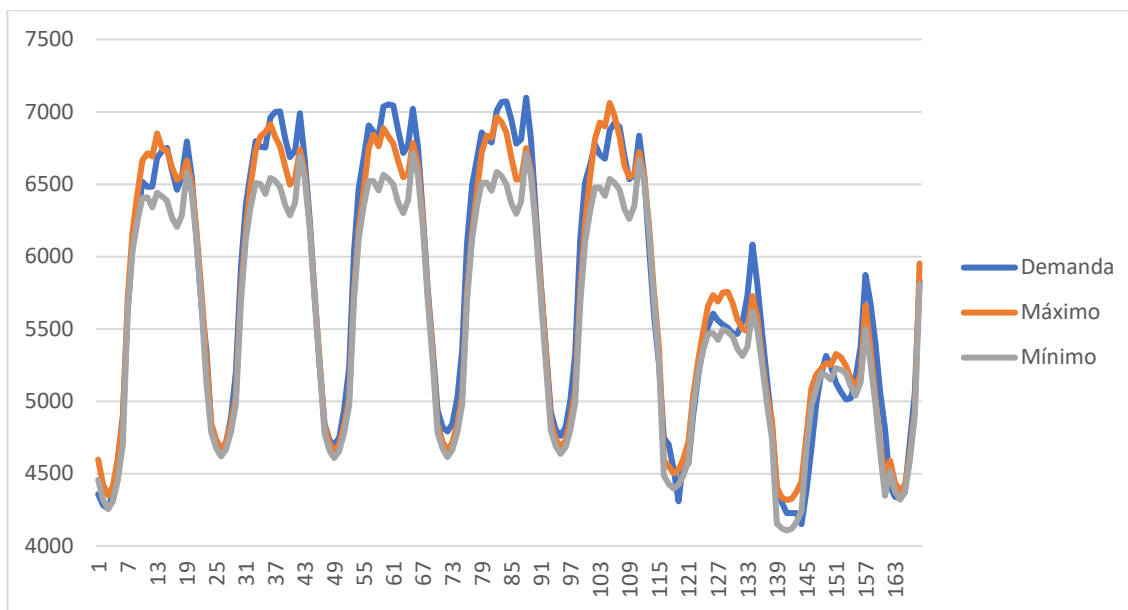


Figura 16 - Previsões máximas e mínimas associadas às séries de temperatura para Portugal (em MW)

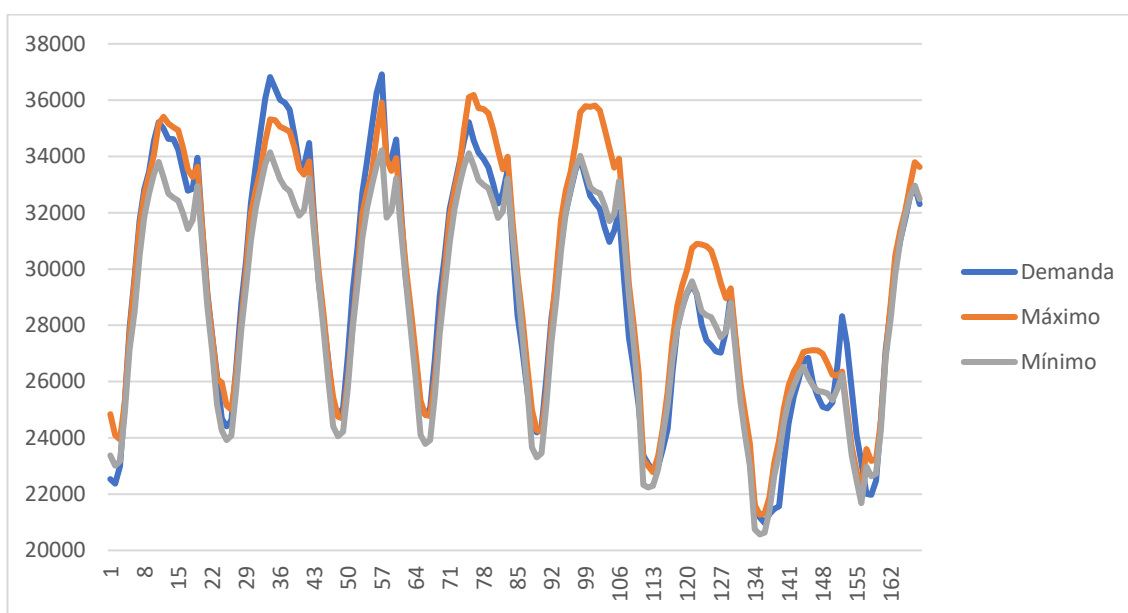


Figura 17 - Previsões máximas e mínimas associadas às séries de temperatura para Espanha (em MW)

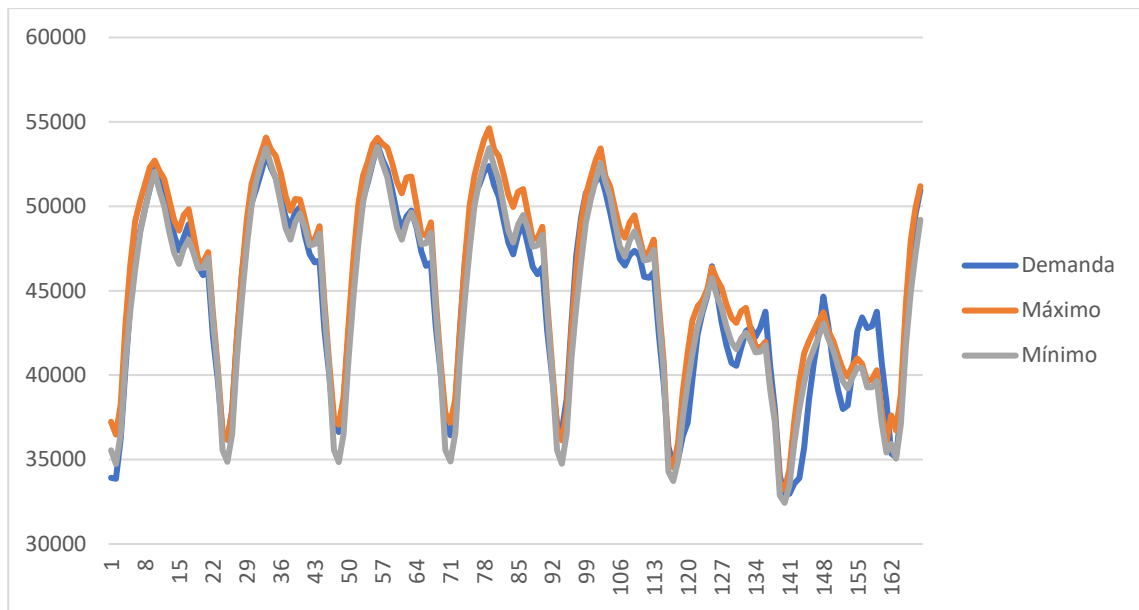


Figura 18 - Previsões máximas e mínimas associadas às séries de temperatura para França (em MW)

Nestas figuras, é possível visualizar o consumo real a azul, a previsão máxima a laranja e a previsão mínima a cinzento. Comparando com as figuras de 16 e 17, a figura 18 mostra mais claramente que, em França, a margem de diferença entre as previsões mínimas e máximas é menor que em Portugal e Espanha.

A introdução de séries de temperatura permite, como se vê nas figuras anteriores, obter um conjunto de previsões para cada hora ao invés de apenas uma, como tem sido o caso nos modelos anteriores.

Capítulo 5

Conclusão

O objetivo desta dissertação foi a construção de um modelo de previsão do consumo energético a médio e longo prazo. Para isso, foi necessário decidir as variáveis a integrar neste modelo. O estudo que foi feito às variáveis de entrada permitiu concluir que as variáveis de calendário (mês, dia da semana, hora, etc.) são as variáveis que melhor traduzem, de forma consistente, a informação necessária ao longo de vários anos, ou seja, este tipo de input é o mais indicado para modelos de previsão a médio e longo prazo, pois permite ao modelo encontrar relações entre o consumo energético e as diferentes épocas do ano (estações, férias, etc.) e as horas do dia.

De seguida, a criação dos modelos foi feita através de redes neuronais, sendo estas treinadas com os dados referentes aos anos entre 2015 e 2018, e posteriormente testadas com os dados de 2019. O ano de 2020 também fez parte do conjunto de teste, inicialmente, mas dada a quantidade de lacunas e como este se tratou de um ano de pandemia, o que implicava consumos diferentes dos esperados, foi retirado. Depois de vários modelos serem testados, com diferentes conjuntos de entrada, aquele que apresentou as melhorias mais acentuadas foi o modelo 4. Especificamente, este modelo teve como entradas as variáveis que representam o consumo médio horário em cada ano, mês e dia da semana, duas variáveis binárias que abrangem os feriados nacionais e dias especiais, e, finalmente, uma variável que representa, simultaneamente, o consumo médio horário em cada semana do ano e em cada hora do dia. Estas variáveis de entrada permitiram sintetizar a informação relativa à variação do consumo ao longo do ano e obter previsões fiáveis e resultados muito positivos.

A partir deste modelo, foi construído um novo modelo através da inclusão da temperatura. Além disso, foram também utilizadas séries de temperatura de 2010 a 2019 para obter um conjunto de previsões para cada hora. Isto permitiu não só reduzir o erro

de previsão associado ao modelo 4, já que a temperatura é um dado que se relaciona muito bem com os hábitos de consumo da população, como também adquirir uma margem de previsões, que dá uma flexibilidade adicional aos resultados deste trabalho.

Referências

- [1] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, H. Zareipour, “Energy Forecasting: A Review and Outlook”, 9 de outubro de 2020.
- [2] M. Hammad, B. Jereb, B. Rosi, D. Dragan, “Methods and Models for Electric Load Forecasting: A Comprehensive Review”, fevereiro de 2020.
- [3] P. Nystrup, H. Madsen, E. Blomgren, G. de Zotti, “Clustering commercial and industrial load patterns for long-term energy planning”, 20 de março de 2021.
- [4] K. Gaur, H. Rathour, P. Agarwal, K. Baba, S. Soonee, “Analysing the Electricity Demand Pattern”, dezembro de 2016.
- [5] Center for Power and Energy Systems of INESC TEC, "Sentinel", base de dados com informação de consumos, produções e dados meteorológicos, incluindo previsões. Consulta em janeiro de 2021.
- [6] X. Meng, S. Zhang, J. Wang, S. Wu, “Load forecasting method considering temperature effect for distribution network”, 2016.
- [7] H. Park, B. Lee, J. Son, H. Ahn, “A comparison of neural network-based methods for load forecasting with selected input candidates”, março de 2017.
- [8] Statista Research Department 2021, acessado a 20 de julho de 2021, www.statista.com/statistics/771171/final-electricity-consumption-in-france-sector/.