

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA



## **Participação da geração renovável no mercado de reservas de um sistema eléctrico**

João Pedro Passagem dos Santos

**Mestrado em Engenharia da Energia e Ambiente**

Dissertação orientada por:  
Professora Doutora Ana Estanqueiro  
Doutor Hugo Algarvio

2023



# Resumo

O resumo deve conter a informação relevante do trabalho. Uma frase de enquadramento (importância da dissertação), uma frase referente ao objetivo e método e uma frase final de conclusões.

**Palavras chave:** até 6 palavras diferentes do título e que ajudem a procurar assuntos relacionados (o equivalente aos # das redes sociais)

# Abstract

Abstract should have the relevant information of the developed work. One sentence with the framework of the research. One sentence with the goal and method. One sentence with the main conclusions.

**Keywords:** up to six different from the title and helpful to find related subjects (like # in social media)

# Agradecimentos

Opcional, embora no caso de dissertações que decorram no âmbito de projetos financiados por exemplo pela FCT ou programas Europeus devem ser mencionados aqui a referência e nome do projeto, e mais alguma informação de acordo com as regras de publicitação do projeto em questão.

Nome do Autor

# Nomenclatura

Lista de siglas, acrónimos, abreviaturas e simbologia apresentadas por ordem alfabética.

## Abreviaturas

(A/F)	Relação mássica ar/combustível
pme	Pressão média efectiva
vol	Volume

## Siglas e acrónimos

ODS	Objetivos de Desenvolvimento Sustentável
vRES	variable Renewable Energy Systems
MIBEL	Mercado Ibérico de Eletricidade
APA	Agência Portuguesa do Ambiente
EDP	Eletricidade De Portugal
RNN	Recurrent neural network
CNN	Convolutional neural network
LSTM	Long short-term memory
AR	dec
MA	dec
ARMA	dec
ARIMA	dec
SARIMA	dec

## Simbologia

A	Área
$\eta$	Eficiência
p	Pressão
T	Temperatura

# Índice

<b>Resumo</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>Agradecimentos</b>	<b>iii</b>
<b>Nomenclatura</b>	<b>iv</b>
<b>List of Figures</b>	<b>vii</b>
<b>List of Tables</b>	<b>viii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Enquadramento . . . . .	1
1.2 Objetivos e perguntas de investigação . . . . .	2
1.3 Organização do documento . . . . .	2
<b>2 Revisão bibliográfica</b>	<b>3</b>
<b>3 Contextos</b>	<b>4</b>
3.1 Mercados de Sistema . . . . .	4
3.2 MIBEL . . . . .	4
<b>4 Dados</b>	<b>5</b>
4.1 Dados Utilizados . . . . .	5
4.1.1 Aquisição dos Dados . . . . .	5
4.2 Estudo dos dados . . . . .	5
4.2.1 Correlações . . . . .	7
4.2.1.1 Correlações entre atributos . . . . .	7
4.2.1.2 Correlações Temporais . . . . .	9
4.2.2 Agrupamento . . . . .	10
4.3 Tratamento dos dados . . . . .	12
4.4 Considerações adicionais . . . . .	14
<b>5 Architecturas de Modelos</b>	<b>15</b>
5.1 Blocos . . . . .	15
5.1.1 Bloco Dense . . . . .	15
5.1.2 Blcoo CNN . . . . .	16
5.1.3 Blcoo LSTM . . . . .	16

5.2	Arquiteturas . . . . .	16
5.2.1	Vanilla . . . . .	16
5.2.2	Stacked . . . . .	16
5.2.3	MultiHead . . . . .	16
5.2.4	MultiTail . . . . .	16
5.2.5	UNET . . . . .	17
5.3	Considerações adicionais . . . . .	17
<b>6</b>	<b>Métodos</b>	<b>18</b>
6.1	Benchmark . . . . .	18
6.2	Modelos estatísticos . . . . .	19
6.2.1	MA . . . . .	19
6.2.2	ARMA . . . . .	20
6.2.3	ARIMA . . . . .	21
6.2.4	SARIMA . . . . .	21
6.3	Forecat . . . . .	22
6.3.1	Construtor de modelos . . . . .	22
6.3.2	subsubsection Gerador de dados (depth 2) . . . . .	22
6.4	Treino e Resultados . . . . .	22
6.4.1	Arquiteturas e numeros de epocas . . . . .	22
6.4.2	Funções de Perda (Loss) . . . . .	23
6.4.3	Hiperparametrização . . . . .	23
6.4.3.1	Activação . . . . .	23
6.4.3.2	Optimizadores . . . . .	23
6.4.4	Janelas Temporais . . . . .	23
6.4.5	Classificação . . . . .	24
6.4.6	Pesos . . . . .	24
6.4.6.1	Modelos lineares . . . . .	24
6.4.6.2	Modelos Lineares e de Classificação . . . . .	24
6.5	Considerações adicionais . . . . .	24
<b>7</b>	<b>Resultados e discussão</b>	<b>26</b>
7.1	Métricas . . . . .	26
7.2	Experiências . . . . .	26
7.2.1	Arquiteturas e numeros de epocas . . . . .	27
7.2.1.1	Arquiteturas . . . . .	27
7.2.1.2	Epocas . . . . .	27
7.2.2	Funções de Perda (Loss) . . . . .	27
7.2.3	Funções de Perda (Loss) . . . . .	27
<b>8</b>	<b>Conclusões e sugestões futuras</b>	<b>28</b>
<b>9</b>	<b>Referências</b>	<b>29</b>
<b>A</b>	<b>Anexos</b>	<b>31</b>



# List of Figures

<b>Figura 1.1</b> Deve aparecer por baixo da figura. Se a figura for feita pelo aluno não necessita de referência. Caso seja retirada de uma fonte bibliográfica deve ser pedida autorização para a sua reprodução/cópia. Caso seja retirada de um website <i>freevector</i> deve ser creditada de acordo com as regras estipuladas no mesmo. Por exemplo pode encontrar todos os ODS com reprodução livre em <a href="https://www.un.org/sustainabledevelopment/news/communications-material/">https://www.un.org/sustainabledevelopment/news/communications-material/</a> . . .	1
<b>Figura 1.2</b> Exemplo de duas imagens numa figura. . . . .	2
<b>Figura 4.1</b> Serie Temporal dos dados alvo . . . . .	6
<b>Figura 4.2</b> Janelas Temporais dos dados alvo . . . . .	6
<b>Figura 4.3</b> Frequência dos dados alvos . . . . .	7
<b>Figura 4.4</b> Correlação entre atributos . . . . .	8
<b>Figura 4.5</b> Valores de correlação entre atributos . . . . .	9
<b>Figura 4.6</b> Serie Temporal dos dados alvo . . . . .	9
<b>Figura 4.7</b> Serie Temporal dos dados alvo . . . . .	10
<b>Figura 4.8</b> Serie Temporal dos dados alvo . . . . .	11
<b>Figura 4.9</b> Histograma das classes . . . . .	12
<b>Figura 4.10</b> Outliers . . . . .	13
<b>Figura 4.11</b> Dados em falta . . . . .	14
<b>Figura 5.1</b> Bloco Densse . . . . .	15
<b>Figura 6.1</b> Serie Temporal do benchmark . . . . .	18
<b>Figura 6.2</b> Serie Temporal do benchmark 2021 . . . . .	19
<b>Figura 6.3</b> Previsões 2021 com modelo MA . . . . .	20
<b>Figura 6.4</b> Previsões 2021 com modelo ARMA . . . . .	20
<b>Figura 6.5</b> Previsões 2021 com modelo ARIMA . . . . .	21
<b>Figura 6.6</b> Previsões 2021 com modelo SARIMA . . . . .	22
<b>Figura 7.1</b> Exemplo de como considerar um gráfico . . . . .	27
<b>Figura A.1</b> Exemplo de como considerar um gráfico nos anexos. . . . .	31

# List of Tables

<b>Tabela A.1</b> Isto é um exemplo de uma tabela. Se fôr igual(copiada) a outro autor deve ser pedido autorização para reproduzir. . . . .	31
--	----



# Introdução

## 1.1 Enquadramento

Esta dissertação encontra-se no âmbito do projecto [TradeRES](#), o qual estuda um sistema de mercado eléctricos que consiga dar resposta às necessidades da sociedade num sistema quase todo renovável. Tendo as características para se integrar nos [ODS](#) (ver Figura 1.1).

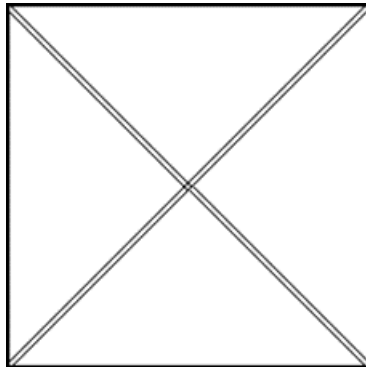
O estudo da acessibilidade das energias renováveis ao mercado vigente integra-se nos ODS nº7, “Energia Renováveis e Acessíveis”, indo directamente de encontro a um dos pontos deste objectivo: 7.2.1 “Peso das energias renováveis no consumo total final de energia”. Por meio deste objectivo, a participação das renováveis no mercado faz também cumprir, embora indirectamente, o objectivo nº8 “Trabalho Digno e Crescimento Económico”, através do ponto 8.4, onde é dada primazia à eficiência dos recursos globais no consumo e na produção. Indirectamente, pois, ao haver um melhor uso das renováveis, o uso de energias não limpas vai diminuir, melhorando a gestão de recursos, e baixando o consumo de recursos naturais não renováveis.

Por último podemos incluir o objectivo nº13, “Acção Climática”, no qual, referimos de novo a diminuição de consumo de recursos finitos, mas mais importante, a melhor gestão de recursos renováveis. Promovendo o planeamento e estratégias de combate a emissões de gases de efeito estufa.

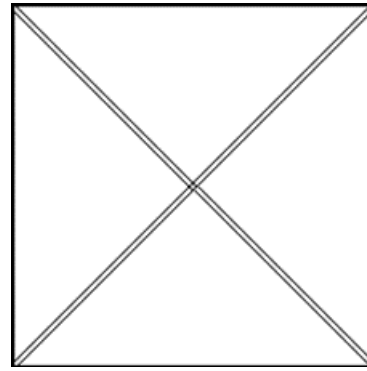


Figure 1.1: Deve aparecer por baixo da figura. Se a figura for feita pelo aluno não necessita de referência. Caso seja retirada de uma fonte bibliográfica deve ser pedida autorização para a sua reprodução/cópia. Caso seja retirada de um website *freevector* deve ser creditada de acordo com as regras estipuladas no mesmo. Por exemplo pode encontrar todos os ODS com reprodução livre em <https://www.un.org/sustainabledevelopment/news/communications-material/>

## 1. INTRODUÇÃO



(a) Explicar o que é Fig A



(b) Explicar o que é Fig B

Figure 1.2: Exemplo de duas imagens numa figura.

### 1.2 Objetivos e perguntas de investigação

Foram aprovadas a nível europeu (2020), medidas de alteração aos serviços de sistema, que serão seguidas pelos Estados-Membros. Nesta dissertação far-se-á a aplicação dessas medidas, identificando as melhorias face ao desenho actual e, avaliando se as novas medidas serão suficientes para assegurar a operação de um sistema eléctrico 100% renovável, eventualmente identificando acções adicionais que garantam a robustez e segurança do sistema eléctrico sem recurso a combustíveis fósseis.

- a) É positivo para as vRES participar no mercado de reserva?
- b) Como configurar essa participação para otimizar o lucro do ponto de vista das vRES?
- c) Essa participação é positiva para o sistema eléctrico num todo?

### 1.3 Organização do documento

Explicar a lógica da organização em termos do conteúdo de cada secção. Por exemplo, no capítulo 2 mostram-se os estudos referentes a. . . . . Na secção 1.2 mostra-se. . . . . Na secção 1.1 apresentam-se. . . . . Finalmente, no capítulo 8 são respondidas as perguntas de investigação, descrevem-se as limitações do estudo e recomendam-se. . . . . para estudos futuros.

# Revisão bibliográfica

A revisão bibliográfica deve recorrer a normas, livros, artigos científicos e, obrigatoriamente, deve indicar a bibliografia consultada de forma correta com recorrência a um reference manager. A maneira mais usual é adotar o sistema numerado por ordem de aparecimento do texto.

Isto é um exemplo de uma equação:

$$a = 1 \tag{2.1}$$

A tabela deve ter uma legenda por cima da mesma, tal como o exemplo em baixo. Se os valores da tabela não são calculados pelo autor e referem-se a valores de outros autores tem de constar as respetivas referências aos seus trabalhos.

# Contextos

## 3.1 Mercados de Sistema

[1] [2] [3] Foram aprovadas a nível europeu (2020), medidas de alteração aos serviços de sistema, que serão seguidas pelos Estados-Membros. Nesta dissertação far-se-á a aplicação dessas medidas, identificando as melhorias face ao desenho actual e, avaliando se as novas medidas serão suficientes para assegurar a operação de um sistema eléctrico 100% renovável, eventualmente identificando acções adicionais que garantam a robustez e segurança do sistema eléctrico sem recurso a combustíveis fósseis.

## 3.2 MIBEL

[4] [5] [6] [7] Foram aprovadas a nível europeu (2020), medidas de alteração aos serviços de sistema, que serão seguidas pelos Estados-Membros. Nesta dissertação far-se-á a aplicação dessas medidas, identificando as melhorias face ao desenho actual e, avaliando se as novas medidas serão suficientes para assegurar a operação de um sistema eléctrico 100% renovável, eventualmente identificando acções adicionais que garantam a robustez e segurança do sistema eléctrico sem recurso a combustíveis fósseis.

# Dados

## 4.1 Dados Utilizados

Os dados em estudo são do mercado energético espanhol, retirados do site da [ESIOS](#).

indicators	names	units
632	SecondaryReserveAllocationAUpward	MW
633	SecondaryReserveAllocationADownward	MW
680	UpwardUsedSecondaryReserveEnergy	MWh
681	DownwardUsedSecondaryReserveEnergy	MWh
1777	WindD+1DailyForecast	MWh
1779	PhotovoltaicD+1DailyForecast	MWh
1775	DemandD+1DailyForecast	MWh
10258	TotalBaseDailyOperatingSchedulePBFGeneration	MWh
14	BaseDailyOperatingSchedulePBFSolarPV	MWh
10073	BaseDailyOperatingSchedulePBFWind	MWh
10186	BaseDailyOperatingShedulePBFTotalBalanceInterconnections	MWh

### 4.1.1 Aquisição dos Dados

No âmbito da automatização destes dados foi modificado o repositório [ESIOS](#) para ser usado como uma biblioteca de python, aberta, em pypi. Sendo uma ferramenta mais facilmente acessível para a extrair dados do mercado espanhol, [pyesios](#):

## 4.2 Estudo dos dados

Os dados que propunho a prever são: "UpwardUsedSecondaryReserveEnergy",  
"DownwardUsedSecondaryReserveEnergy"



#### 4. DADOS

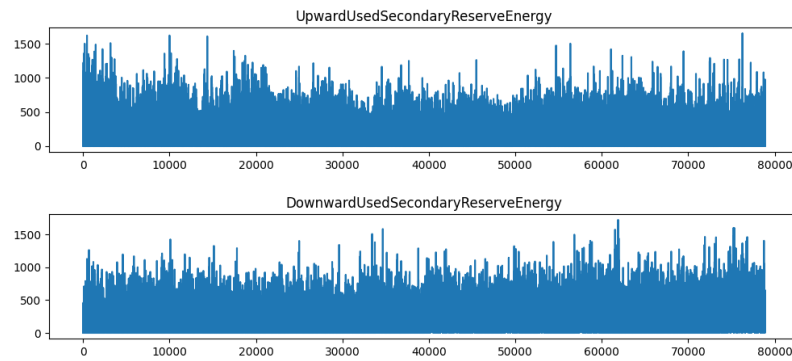


Figure 4.1: Serie Temporal dos dados alvo

Para termos uma melhor percepção dos mesmos segue algumas janelas temporais mais pequenas.

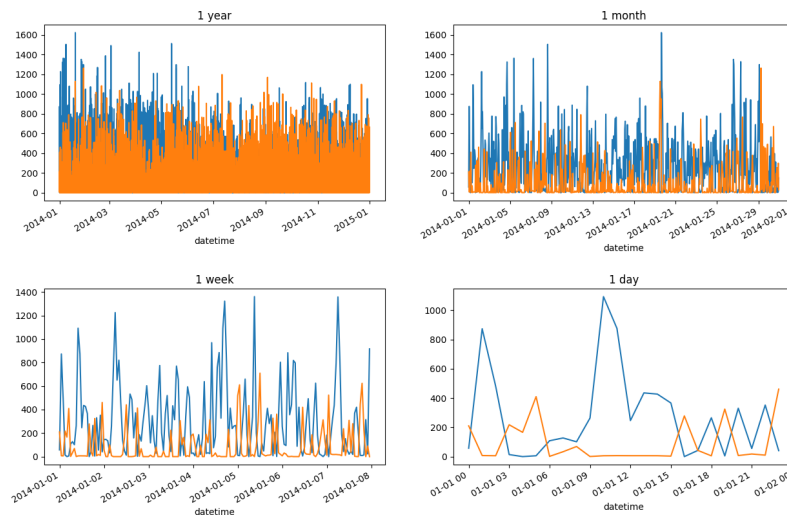


Figure 4.2: Janelas Temporais dos dados alvo

Estas mostram claramente que ambos os atributos mantêm um comportamento tanto discreto, como linear, isto é, que ou existe algum valor, ou é zero, e se existe valor este tem comportamento linear.

A distribuição destes dados é claramente exponencial. O que é importante para a escolha de alguns parametros no modelação

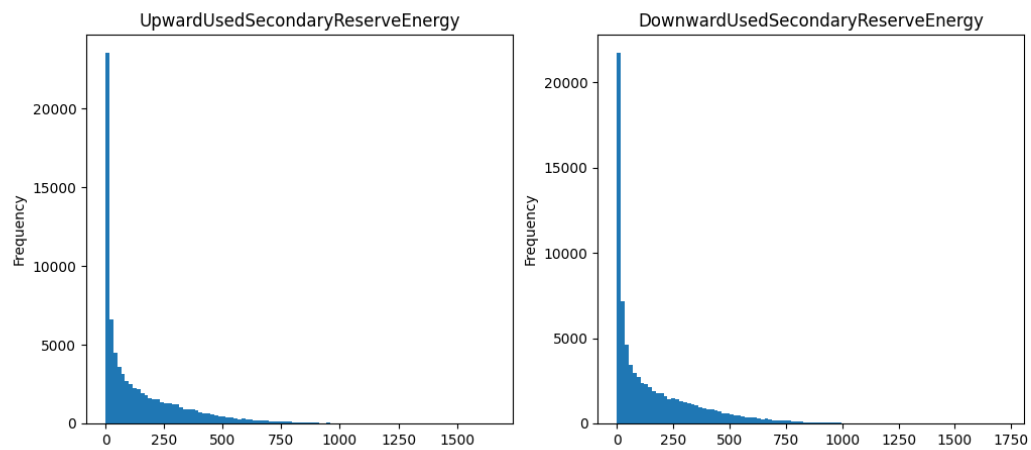


Figure 4.3: Frequência dos dados alvos

## 4.2.1 Correlações

### 4.2.1.1 Correlações entre atributos

Os modelos vão depender bastante de correlação entre variáveis. Nesta secção queremos tentar identificar se há visíveis relações entre as variáveis, e se há relações temporais visíveis nas colunas alvo.

#### 4. DADOS

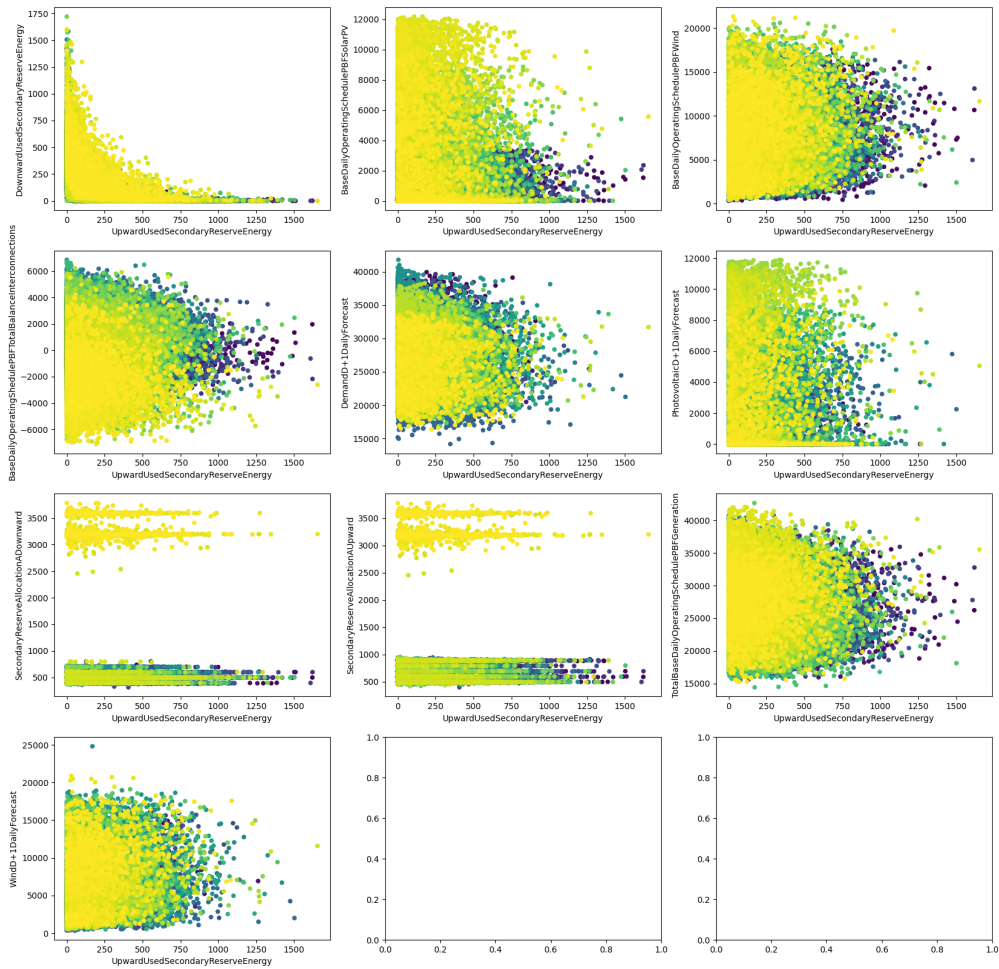


Figure 4.4: Correlação entre atributos

As correlações entre variáveis parecem muito escassas o que apresenta já que a previsão deste dados usando estas variáveis vai ser um problema difícil. Por norma é feito uma seleção de atributos baseado nestas correlações, eliminando assim os atributos que ajudam menos, ou até prejudicam os modelos. Segue os valores de correlação onde podemos ver numericamente que existe muito pouca correlação entre os atributos.

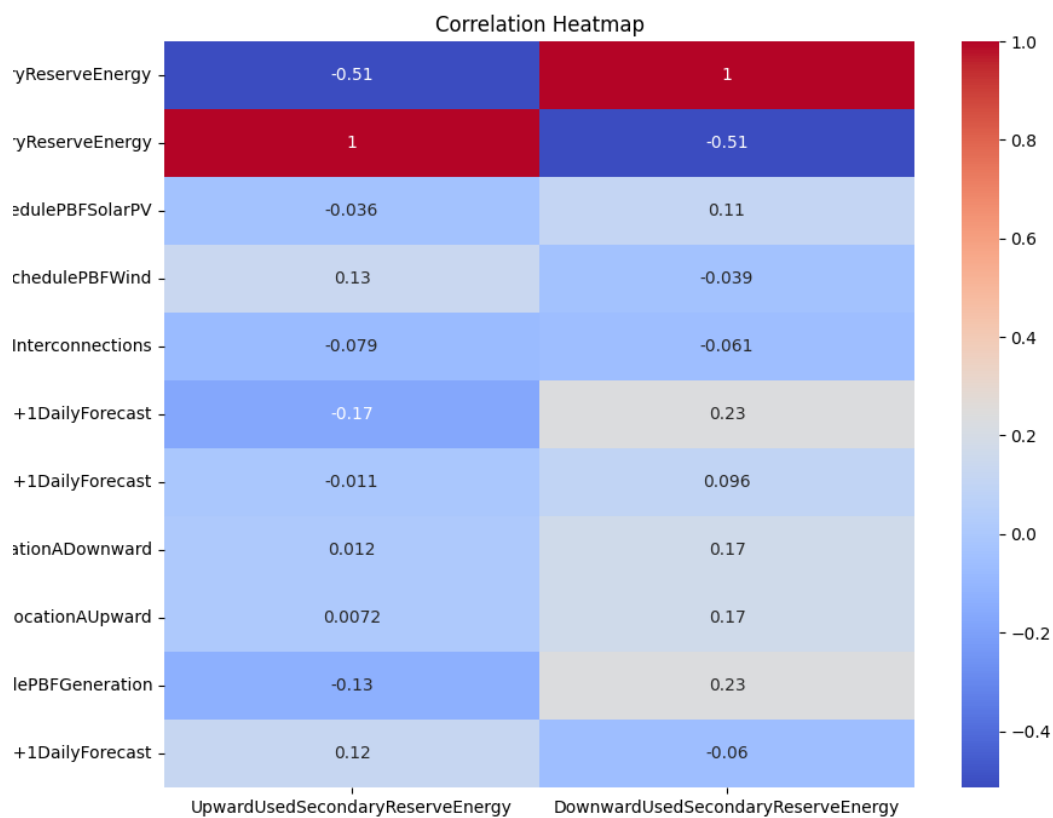


Figure 4.5: Valores de correlação entre atributos

4.2.1.2 Correlações Temporais

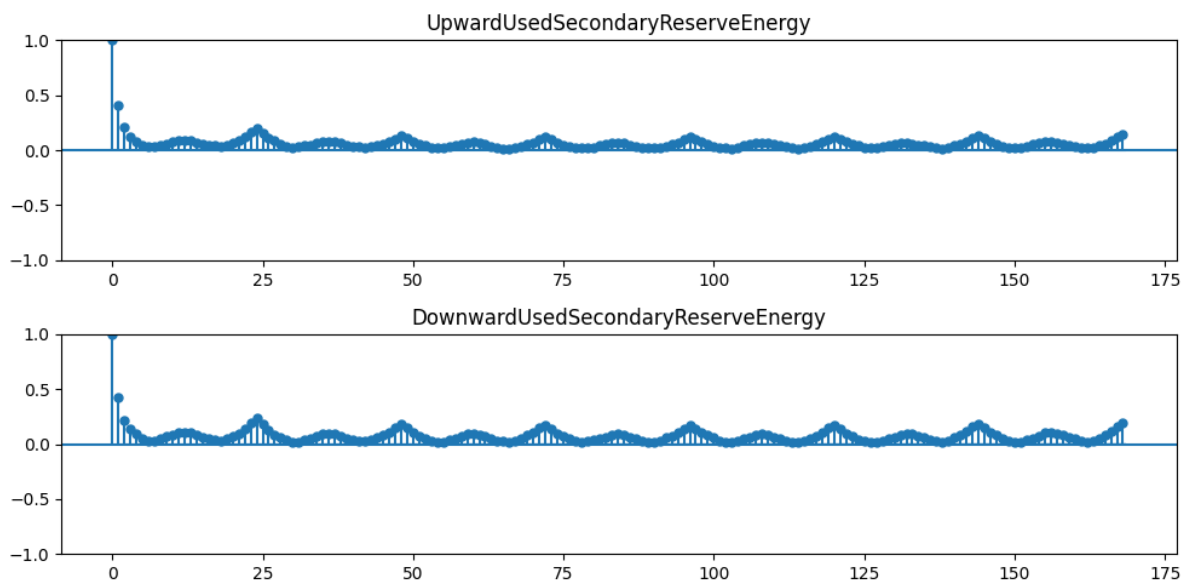


Figure 4.6: Serie Temporal dos dados alvo

#### 4. DADOS

A autocorrelação, em ambos os "targets", é mais forte nas 3 horas mais próximas, e nos pontos com diferença de 12 e 24 horas. É de notar que estes valores são baixos, prometendo já também uma baixa regressividade temporal. Outro ponto a denotar é que os objectos não têm um comportamento completamente linear, i.e., parece existir um comportamento discreto na questão ser alocado ou não esta reservas secundárias, e caso seja alocado, aí existir alguma linearidade. Logo qualquer tipo de modelação terá de resolver primeiramente este problema.

Estas relações mostram que em termos de atributos usados vai ser um desafio complicado para qualquer tipo de modelo.

No âmbito desta dissertação queremos verificar a qualidade das previsões usando estes mesmo atributos, logo, não será feita seleção dos mesmos.

A nível da relação temporal, a maior parte dos modelos que iremos testar aplica um janela na dimensão temporal, usando todos os valores nessa janela, e aplicando os pesos nessas distancias que mais se enquadram. Logo também não é relevante escolher apenas as distancias temporais com maior correlação, pois os modelos vão fazer essa pesagem.

#### 4.2.2 Agrupamento

Uma das possibilidades na modelção será a utilização de grupo de valores, classes, em conjunto com a linearidade. Devido ao comportamento não exclusivamente linear (tenho de procurar um nome para isto) é também estudado as possíveis agregações (clustering) em que podemos dividir os valores em classe. Tendo por base que uma das classes é o valor zero, devido ao comportamento não linear desta série, vamos apenas testas quantos, e quais, as melhores classes em que podemos dividir os dados. É realizado o teste de silhoeta (ref) e o teste do cotovelo (ref), que se baseiam nos resultados de silhoeta do modelo GMM (ref) e nos valores de inercia do modelo K-Means (ref).

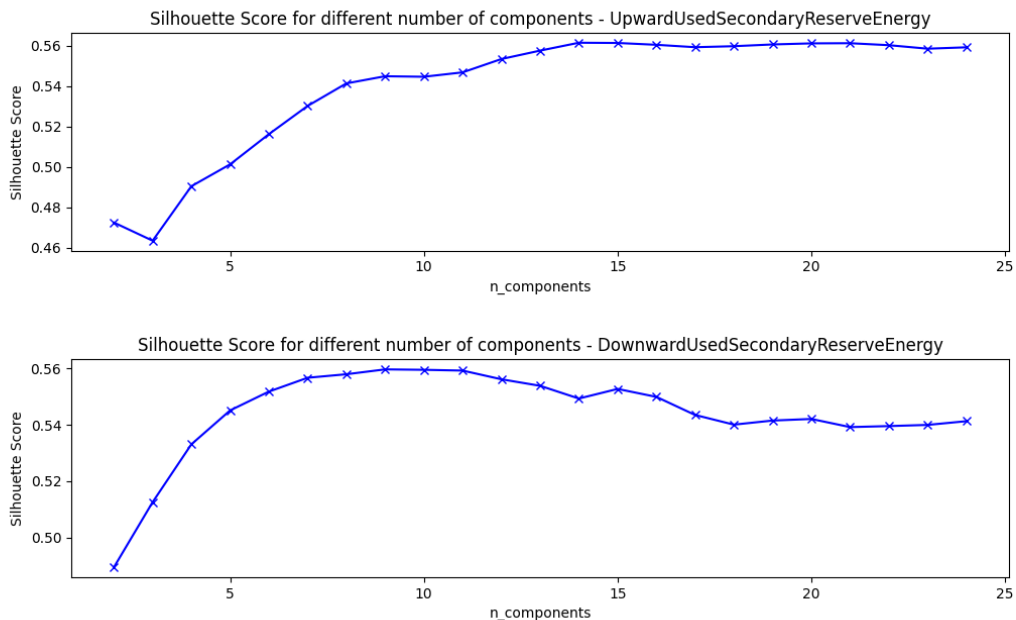


Figure 4.7: Serie Temporal dos dados alvo

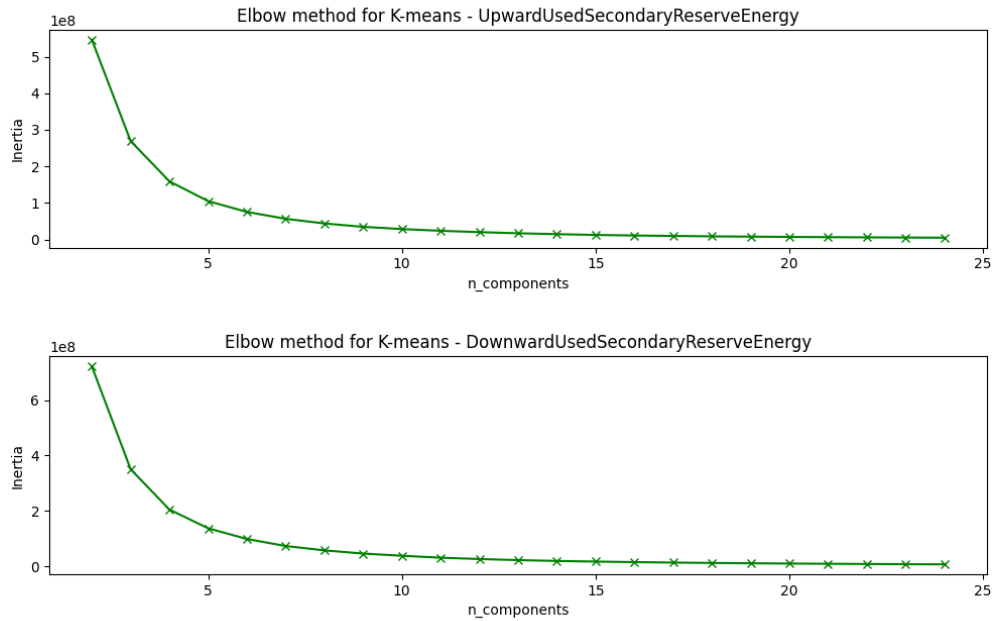


Figure 4.8: Serie Temporal dos dados alvo

Ambos os casos apontam um assintota na relação interna dos clusters, a partir de cerca de 5 clusters, sendo que o melhor valor dos verificados seria com 14 clusters em "UpwardUsedSecondaryReserveEnergy" e 9 em "DownwardUsedSecondaryReserveEnergy". Para a nossa questão, queremos algumas classes, mas quanto menos classes mais fácil será para os modelos correctamente identificar a que classe pertence. Logo para os valores apresentados, escolhemos 5 clusters, sendo que este já pode ser um numero elevado de classes, logo usemos 3 clusters se os modelos tiverem muuita dificuldade com 5. O método do cotovelo apresenta como numero ideal de clusters 4, mas sendo que K-Means é um metodo mais apropriado para distribuições normais, em que a distancia nos limites dos clusters não varia, deixamos apenas informativo.

O histograma das divisões pode ser visto em baixo. O valor é a sua própria classe para além das apresentadas.

## 4. DADOS

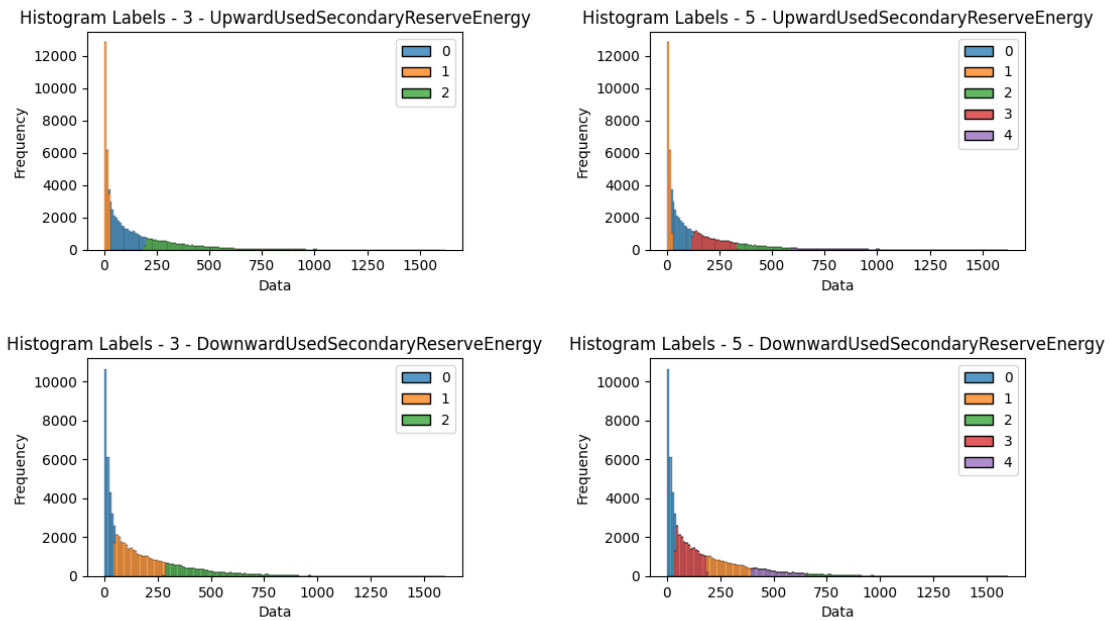


Figure 4.9: Histograma das classes

Os limites retirados são:

Atributo	Nº Clusters	Classe 1	Classe 2	Classe 3	Classe 4	Classe 5
UpwardUsedSecondaryReserveEnergy	3	26.2	"26.3-194.9"	195.0		
UpwardUsedSecondaryReserveEnergy	5	20.2	"20.3-119.2"	"119.3-331.9"	"332.0-592.6"	592.7
DownwardUsedSecondaryReserveEnergy	3	47.0	"47.1-281.9"	282.0		
DownwardUsedSecondaryReserveEnergy	5	38.9	"39.0-186.9"	"187.0-393.1"	"393.2-658.0"	658.2

## 4.3 Tratamento dos dados

**Normalização** A normalização foi deixada por ser aprendida nos modelos, sendo que todos têm como segunda camada, uma de normalização.

### Limpeza

Podemos ver pelos graficos seguintes que a existem alguns outliers, sendo estes definidos como 3 desvios padrão de distância à média. Estes graficos mostram também que existe uma variação do que são os valores normais de cada atributo a nível temporal. Logo um método de limpeza não se poderia basear apenas numa definição geral de outliers, mas teria de ser feito em janelas temporais. Pelo mesmo argumento e visto que os outliers fazem parte do que queremos também descobrir, não é aplicada nenhum método de remoção dos mesmo, sendo os dados passados a cru para os modelos.

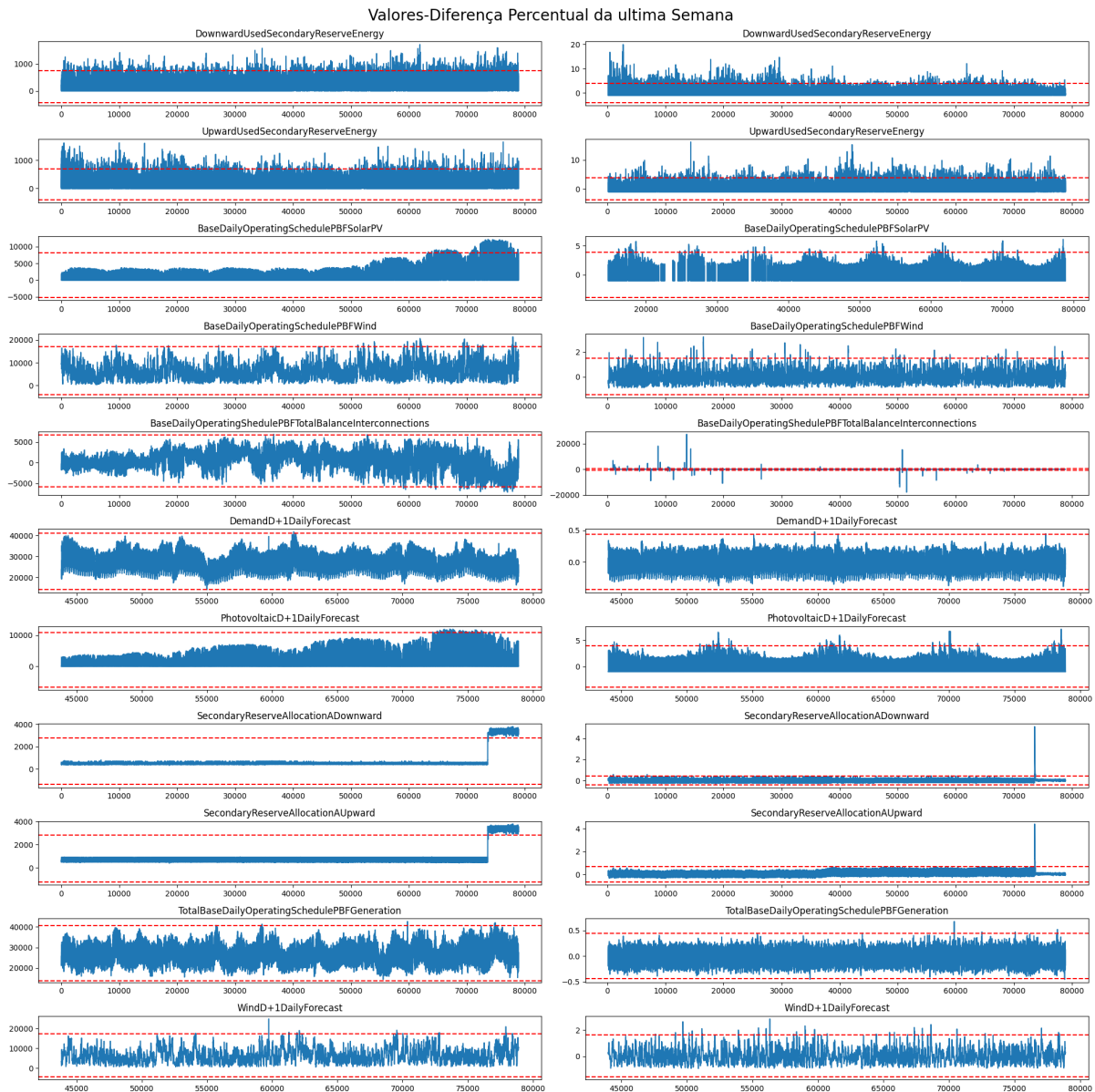


Figure 4.10: Outliers

Outra análise desta variação dos atributos a nível temporal leva-nos a que qualquer divisão dos dados para treino e teste deva levar as variações em consideração. Isto sendo que o treino deve ter representatividade de todas, ou maior parte, das condições diferentes. Visto estarmos a usar dados de 2014 a 2022, argumento até que podemos nem fazer divisão no treino, e usar como teste e validação dados de 2023.

Estudemos também o caso de dados em falta. Alguns destes atributos têm certas entradass vazias, e como podemos ver alguns não têm alguns anos inteiros. Como queremos usar o maximo de dados possiveis iremos usar tecnicas de imputing nesses dados. Podemos ver que temos dados em falta de varios anos, em três atributos, e um tem algumas horas esporadicas em falta nos primeiros anos.



## 4. DADOS

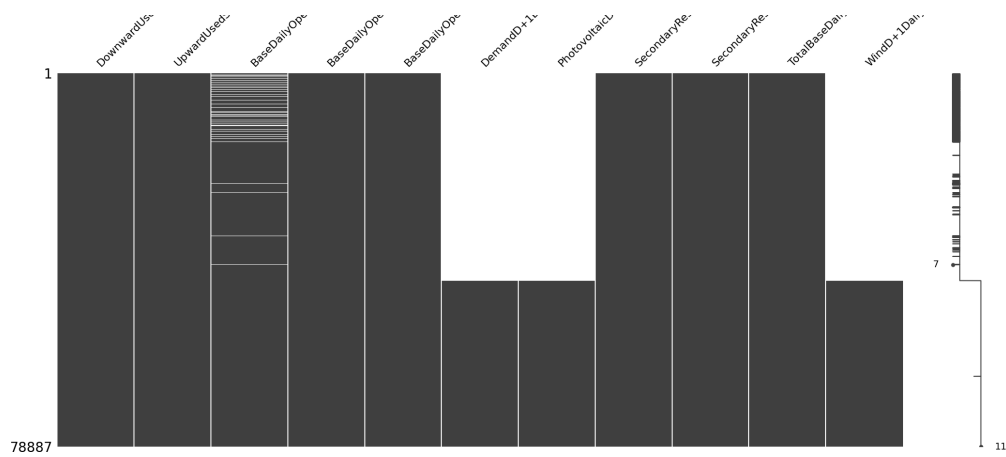


Figure 4.11: Dados em falta

Vamos aplicar o método experimental [IterativeImputer](#) da biblioteca de python [sklearn](#). Este método é baseado nos trabalhos de [8] e de [9]

Por último foi adicionado aos dados mais atributos, sendo eles todos de cariz temporal. É adicionado atributos correspondentes à hora, ao dia do ano, ao dia da semana, ao dia do mês, mês, ano. TODO

### 4.4 Considerações adicionais

Talvez aqui uma secção para finalizar e mostrar algumas coisas

# Arquitecturas de Modelos

Grande parte da literatura sobre previsões em modelos de aprendizagem apresenta as mesmas arquiteturas, sendo que são depois aprimoradas consoante os dados e o problema.

Apresento aqui as arquiteturas mais usadas em previsões, como também algumas usadas noutros ramos tentando prever a compatibilidade neste problema. As arquiteturas irão seguir um esquema lógico comum, um bloco de camadas de entrada, um bloco principal e um por fim um bloco interpretativo. As dimensionalidades destas camadas é o que irá formar as diferentes arquiteturas em estudo.

## 5.1 Blocos

Todas as arquiteturas em análise irão ter por base um bloco de camadas neuronais. A formação dessas arquiteturas passa pelas diferentes maneiras que se pode utilizar o bloco principal. Repetições em série ou em paralelo são um exemplo.

### 5.1.1 Bloco Dense

O bloco dense sendo ele o mais simples é formado por duas camadas Dense, em que a primeira apresenta um número maior de filtros que a segunda. Estas camadas não são mais do que uma criação de filtros aleatórios combinando as entradas, para criar todos os filtros de saída. São a base das camadas interpretativas. A acumulação em série (stacked) de camadas de dense está ligada a melhorias nas capacidades predictivas dos modelos [10].

Exemplo ilustrativo do nosso bloco básico onde entrariam 16 filtros na primeira camada e para finalizar o bloco com 2 filtros

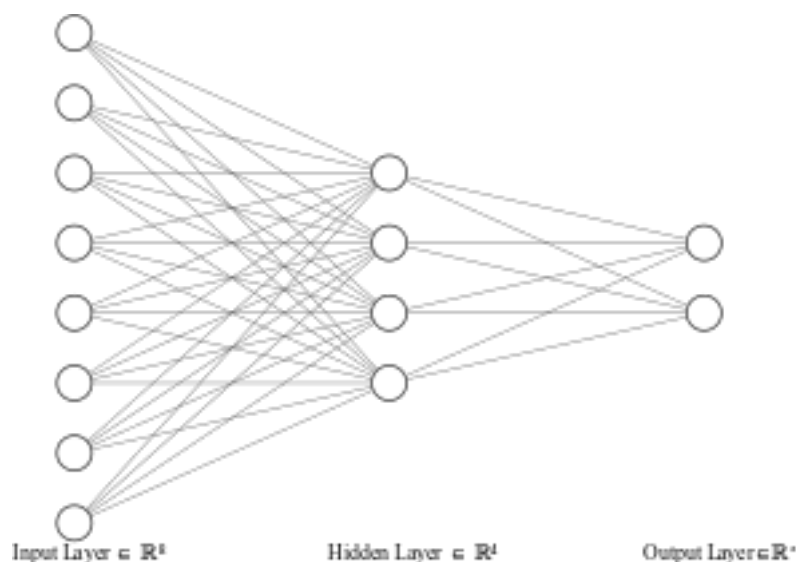


Figure 5.1: Bloco Dense

### 5.1.2 Bloco CNN

Bloco de CNN é aqui definido como uma convolução na dimensão temporal seguido de camadas para combater o overfitting, MaxPooling e Dropout. Normalmente usada em processamentos de imagens, o uso de convoluções temporais é também por si mesmo uma ideia forte.

explicar om que e CNN imagem

Usada também as ideias de attention, residual e o que eu chamei broad

### 5.1.3 Bloco LSTM

O uso de LSTM para previsões é uma área comum, mas aqui é seguido através das ideias partilhadas em [11], e reforçado pelo uso em previsões energéticas demonstrados em [12] O bloco LSTM é a aplicação das RNN, aqui sendo apenas definido como uma camada de LSTM. Estes blocos mantêm dentro de si ligações a diferentes camadas temporais, e cada filtro criado, mantêm uma "memória" dos filtros passados. Bastante utilizado em modelação de linguagem.

imagem

## 5.2 Arquiteturas

### 5.2.1 Vanilla

O termo "Vanilla" aqui é aplicado para arquiteturas que apenas usam um bloco de cada, um de entrada, um principal, e um interpretativo. Como exemplo a arquitetura de "VanillaCNN"

imagem da mesma

### 5.2.2 Stacked

Stacked refere-se a "amontoado" onde se utiliza o bloco principal várias vezes em série. E apenas um bloco de entrada e um interpretativo. Como exemplo a arquitetura de "StackedCNN"

imagem da mesma

### 5.2.3 MultiHead

Multihead é o termo para quando os blocos de entrada e principais são repetidos paralelamente, um caminho para cada atributo, ou uma outra paralelização à escolha. Sendo depois concatenadas essas camadas e passadas juntas para a camada interpretativa. Aqui foi usado sempre a paralelização por atributos, e ao invés de fazer Multihead no sentido de múltiplas entradas, para simplicidade de programação, foi feita uma paralelização interna no modelo, após a camada de entrada, onde a mesma é repetida para cada atributo. Foi testado a diferença, e para os dados usados não havia diferenças de qualidade, mas sim em tempo de treino, logo a mais rápida foi a escolhida. Como exemplo a arquitetura de "MultiheadCNN"

imagem da mesma

### 5.2.4 MultiTail

Esta arquitetura tem o mesmo conceito que a anterior a nível de paralelização, mas neste caso esta é feita apenas na camada interpretativa. Sendo que o resultado do bloco principal é repetido para criar a

paralelização. Neste caso foi paralelizado com o numero de tempos a prever, 24 horas, 24 objectos de saída destas modelos. A grande diferença desta arquitectura para a "Vanilla" que preve 24 horas, é que aqui cada hora tem o seu proprio valor de função de perda, logo o modelo como que está a treinar 24 modelos diferentes, e no caso "Vanilla" a função de perda é única e é a media do erro das horas todas. Como exemplo a arquitetura de "MultiTailCNN"

imagem da mesma

### 5.2.5 UNET

Normalmente usando em modelção de imagens, a arquitectura UNET passa por criar uma rede de expansão dos filtros, usando convoluções, e de seguida uma rede de contracção dos mesmo, até aos tamanhos pretendidos. O bloco principal contextualmente o mesmo que o CNN. Nas suas ligações UNET junta informação de filtros passados (não de nivel temporal mas de rede neuronal) para realçar informação já trabalhada, e assim identificar padrões de vários contextos diferentes. É habitual tambem adicionar aos blocos principais portões de atenção, portões residuais. Estas duas tecnicas são tambem estudadas aqui. É chamada assim pois é uma rede (NET) que forma um U na sua expansão e contracção.

Como exemplo a arquitetura de "UNET"

imagem da mesma

## 5.3 Considerações adicionais

Aqui e dizer que os modelos utilizados para teste sao as combinacoes deste blocos nestas aquiteturas. Imagens de layers criadas com dense <http://alexlenail.me/NN-SVG/index.html>

# Métodos

Neste capítulo percorremos as experiências realizadas. Estas foram feitas através do uso dos programas criados para o efeito, disponíveis no repositório GitHub do [projecto](#).

## 6.1 Benchmark

Como modelo benchmark iremos usar a alocação feita. Pois são estes valores que procuramos melhorar no caso prático.

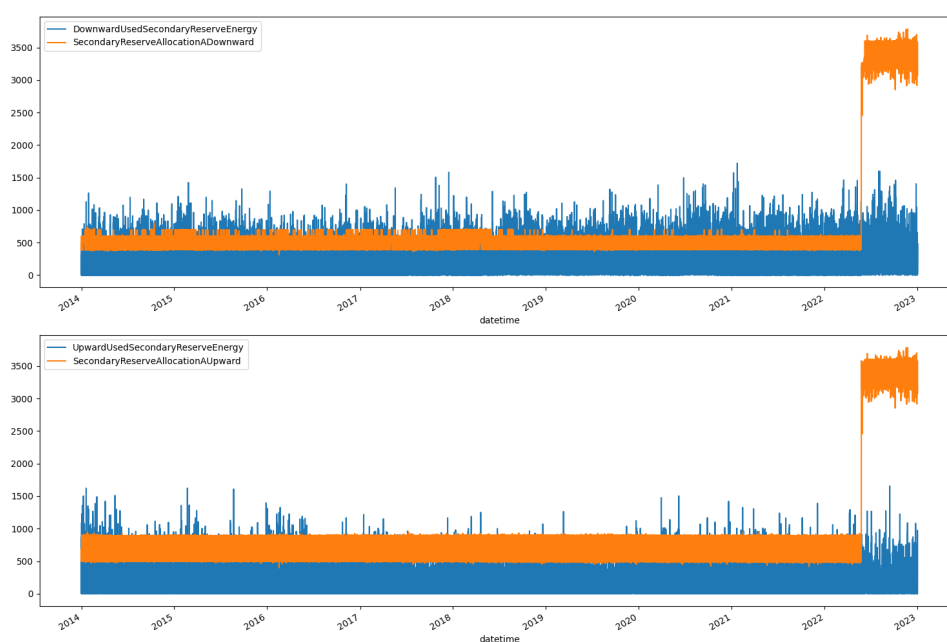


Figure 6.1: Serie Temporal do benchmark

rmse	abs erro	erro comp	r2 score	mape score	alloc missing	alloc surplus
963.1909591866699	53254091.2	False	-26.147637046658637	90.44161200506936	430902.00000000006	52823189.2

Para validação dos mesmos, vamos usar o ano 2021, devido aquele salto nos valores de alocação em 2022.

Para esses temos os seguintes dados:

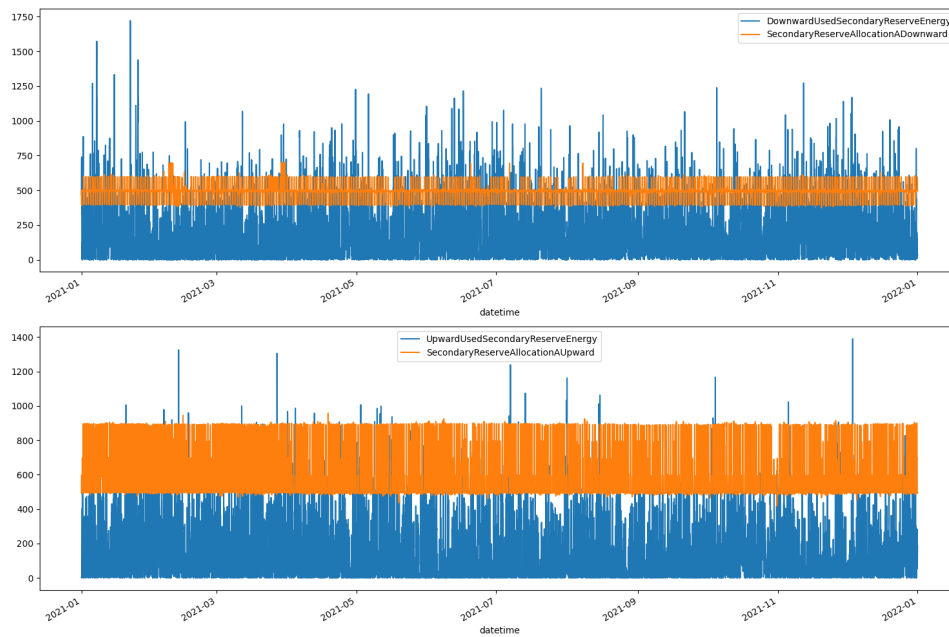


Figure 6.2: Serie Temporal do benchmark 2021

Os metodos em estudo vão ser comparados a esta medida. Sendo que o principal é baixar tanto a alocação perdida, como a alocação a mais. Que se traduzem no erro absoluto.

rmse	abs erro	erro comp	r2 score	mape score	alloc missing	alloc surplus
488.55368458838365	3898946.9000000004	False	-6.008942605247387	85.98375795472323	57046.4	3841900.5

## 6.2 Modelos estatísticos

Antes de entrar para o desenvolvimento de modelos vamos usar métodos e modelos abertos para usar comparativamente. Para as equações apresentadas temos que  $Y_t$  é a variável alvo, no tempo  $t$

Os modelos estatísticos recorrentes em previsões são AR, MA, ARMA, ARIMA, SARIMA para previsões so com um atributo, e para múltiplos atributos VAR.

O modelo AR não teve resultados aplicáveis, logo foi desconsiderado.

### 6.2.1 MA

AR é blabla

$$Y_t = c + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (6.1)$$

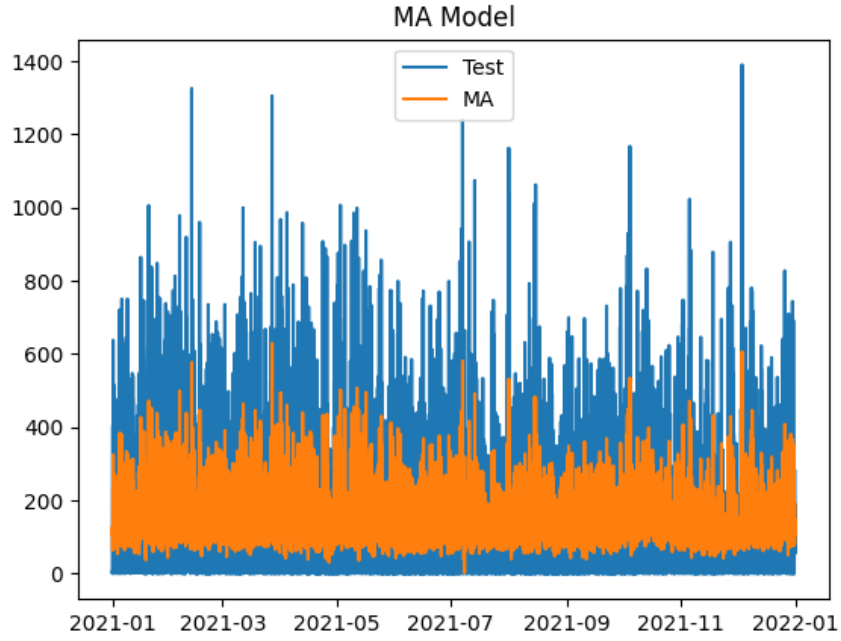


Figure 6.3: Previsões 2021 com modelo MA

### 6.2.2 ARMA

AR é blabla

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (6.2)$$

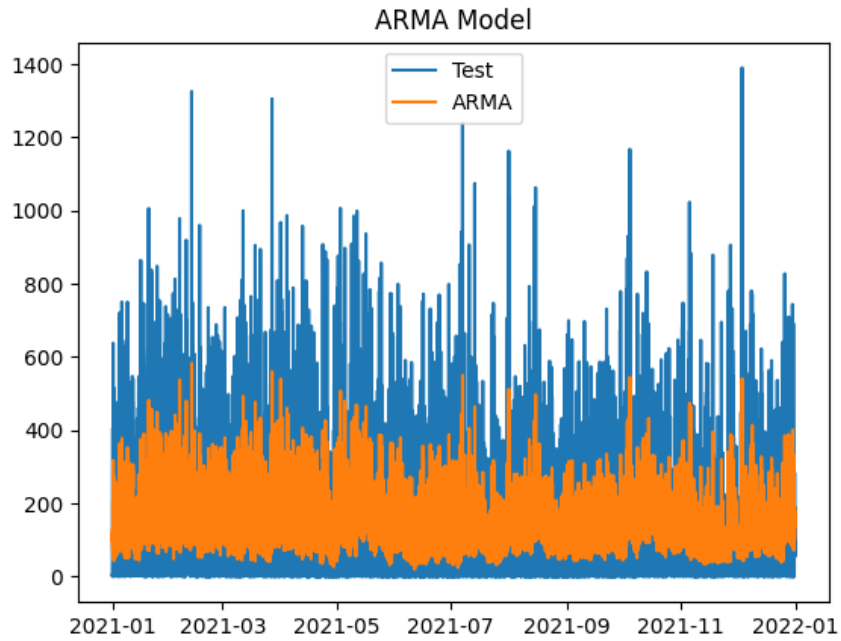


Figure 6.4: Previsões 2021 com modelo ARMA

### 6.2.3 ARIMA

AR eé blabla

$$Y'_t = c + \phi_1 Y'_t - 1 + \phi_2 Y'_t - 2 + \cdots + \phi_p Y'_t - p + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (6.3)$$

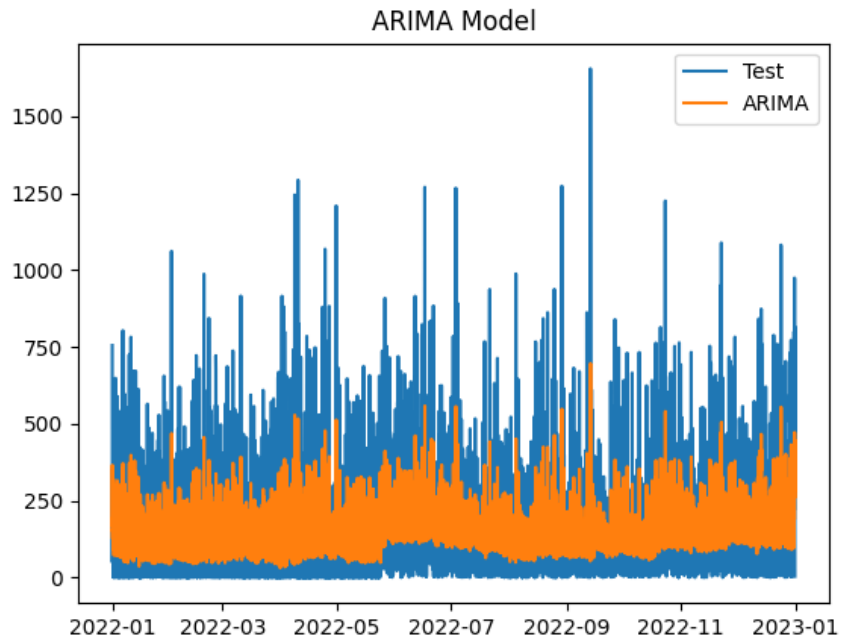


Figure 6.5: Previsões 2021 com modelo ARIMA

### 6.2.4 SARIMA

AR eé blabla

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (6.4)$$



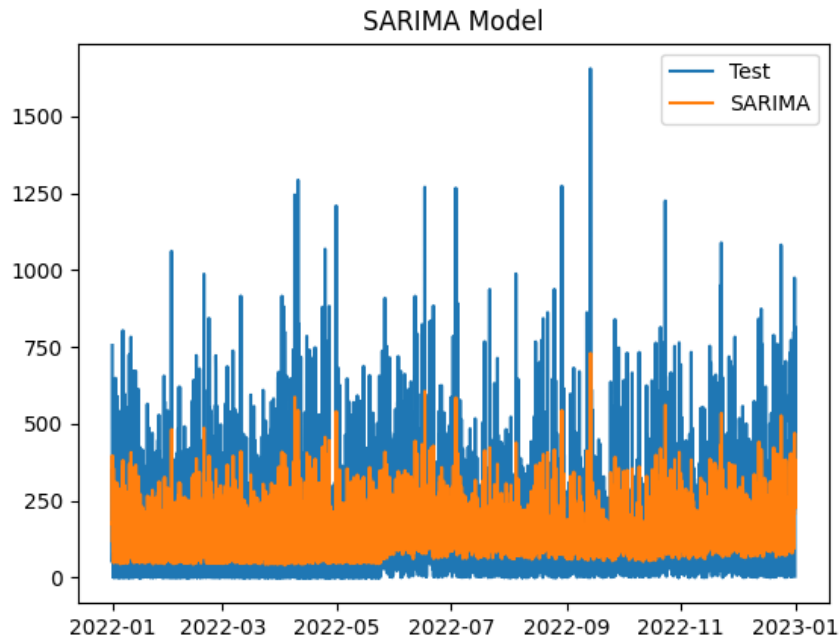


Figure 6.6: Previsões 2021 com modelo SARIMA

TODO: criar csv com resultados disto

### 6.3 Forecat

TODO! Com o propósito de desenvolver este estudo, e deixar ferramentas para a replicação do mesmo, foi criado uma biblioteca em python para desenhar as arquitecturas em estudo.

#### 6.3.1 Construtor de modelos

#### 6.3.2 subsection Gerador de dados (depth 2)

distribuição

clustering

### 6.4 Treino e Resultados

Realizaram-se várias experiências, onde em cada um se ia eliminando alguns dos objectos em estudo. Em casa experiências toda a parametrização era igual, à excepção do objecto de estudo.

#### 6.4.1 Arquiteturas e numeros de epocas

Nesta experiência foi testado o resultado das várias arquiteturas em estudo, como também o impacto do numero de epocas na qualidade dos modelos- As arquiteturas estudadas foram:

- VanillaDense
- VanillaCNN

- VanillaLSTM
- StackedCNN
- StackedLSTM
- EncoderDecoder
- UNET

O modelos foram treinados em 200 epocas, sendo que foram salvos a cada 15 epocas, de forma conseguirmos perceber os contextos nos saltos de epocas.

As parametrizações usadas:

- loss : mean squared error
- Metodo activação no meio : relu
- Metodo activação no fim : relu
- optimizador : Adam
- Janela temporal em X : 168 horas (1 semana)
- Janela temporal em Y : 24 horas (1 semana)
- Fracção de treino : 95

### 6.4.2 Funções de Perda (Loss)

TODO: o que é a loss function?

Esta experiência consiste em rever que função de perda é melhor aplicavel ao problema. Sendo um problema de regressao linear, de valores bastante oscilatórios e com uma distribuição exponencial, temos algumas loss functions que já são reconhecidas para o problema.

- mean absolute error
- mean squared error
- loss : mean absolute error

### 6.4.3 Hiperparametrização

#### 6.4.3.1 Activação

#### 6.4.3.2 Optimizadores

### 6.4.4 Janelas Temporais

Um dos pontos deste trabalho é perceber a fesiabilidade de usar dados de previsão do dia anterior (DA) para estes atributos energéticos. Algo que pode ser também aplicado no futuro a outros dados que não DA, mas sim a 3 horas, ou a 8 horas. Para perceber esta flexibilidade, mas especialmente para escolher as melhores janelas temporais a usar neste modelos, vamos testar várias combinações. Mantendo em mente que o objectivo é prever 24 horas, para os casos onde o alvo não dá um previsão de

## 6. MÉTODOS

24 horas, é necessário criar um número de modelos para fazer as 24 horas. Para validação apenas é usado o espaço temporal previsto, e não múltiplos modelos.

Dado as análises de autocorrelação iremos usar como janelas para treino o conjunto [24, 48, 96, 168] para prever o conjunto [1, 4, 8, 12, 24]

Para além destes foram também testadas combinações com janelas de treino 8 e 12 horas. Estas mostraram rapidamente que janelas de treino menores que as de previsão funcionam muito mal.

### 6.4.5 Classificação

Como descrito em (ref)... existe também o uso de tanto classes como valores lineares para resolução de problemas de regressão, também chamado *cluster-wise regression*. Para este teste mudamos um pouco o modelo em uso. Ao invés de apenas uma camada interpretativa, fazemos duas, em paralelo, sendo que uma resolve a regressão e a outra a classificação. Outro caso, proposto aqui, é usar uma nova camada interpretativa, que combina as duas saídas anteriores (linear e classificação), e resolve novamente para os valores lineares.

Estes modelos não terão apenas uma saída, mas várias, como as arquiteturas MultiTail, mas neste caso cada uma resolve para um problema diferente, com funções de perda, e ativações diferentes.

TODO: desenho destas duas camadas interpretativas

### 6.4.6 Pesos

Por último foi testado o impacto do uso de pesos nos modelos. Estes pesos são o peso que aquele alvo TODO: explicar pesos

#### 6.4.6.1 Modelos lineares

Para os modelos lineares o peso que é adiciona ao modelo é a distância à média. Este peso serve para dar mais importância a valores facilmente considerados outliers.

#### 6.4.6.2 Modelos Lineares e de Classificação

Aqui o peso é dado por saída. Para as saídas lineares o peso dados é o mesmo que apresentado anteriormente, para a saída de classificação, o peso é o inverso da frequência da classe. Distribuindo assim a importância de treino pela frequência das classes. Sendo uma prática comum especialmente quando as distribuições são muito desiguais, como o caso em estudo. É aqui estudada a aplicação destes pesos individualmente, e em conjunto. Os pesos aqui são também normalizados de modo a que o maior peso em cada um deles seja 1, e logo a multiplicação dos dois esteja dentro das mesmas dimensões de relevância.

$$P_m = |y - \text{mean}| \quad (6.5)$$

## 6.5 Considerações adicionais

Foram realizados testes adicionais que não obtiveram resultados passivos de boa interpretação, e foram imediatamente descartados, como:

- Janela temporal em X : 96, 48, 24

- otimizador : todos os otimizadores disponiveis na biblioteca keras
- loss : todas as outra loss functions de regressão disponiveis.
- epocas : influência do numero de epocas nos modelos, foram treinados até 20000 epocas alguns modelos mas à medida que a perda ia estagnando na assintomta, o modelo ia apenas piorando.

Todos os metodos foram realizados utilizando código em python, que está aberto em <https://github.com/JotaFan/renewable-generation-into-reserve-markets>

# Resultados e discussão

Os resultados das experiências são apresentados por experiência, sendo que cada um vai testando e eliminando parametros na modulação. Após análise inicial na modelação foi concluido que usar tentar modelar apenas um dos atributos de cada vez leva a melhor resultados que tentar modelar os dois no mesmo modelo. Foi também visto que os dois atributos em causa são analogos, logo os sistemas que melhor represemtam um dos atributos tambem são semelhantemente eficazes no outro. Assim todas as experiências foram realizadas usando apenas um dos atributos alvo, sendo este o "UpwardUsedSecondaryReserveEnergy", e em que o atributo de alocação comparativo é "SecondaryReserveAllocationAUpward".

## 7.1 Métricas

A nível de comparação efectiva de modelos iremos usar as seguintes metricas: RMSE, erro absoluto, r2 score, percentagem optima, alocação em falta, alocação a mais. Sendo que as métricas de decisão final são as que representam melhor o objectivo: baixar os custo de alocação das reservas secundárias. logo serão as metricas de alocção em falta e em demasia, sendo que a soma das mesmas é o erro absoluto.

$$RMSE(y, \hat{y}) = \sqrt{\frac{\sum_{i=0}^{N-1} (y_i - \hat{y}_i)^2}{N}} \quad (7.1)$$

$$\text{Absolute Error} = \sum_{i=0}^{N-1} |y_i - \hat{y}_i| \quad (7.2)$$

$$R^2 \text{ score} = 1 - \frac{\sum_{i=0}^{N-1} (y_i - \hat{y}_i)^2}{\sum_{i=0}^{N-1} (y_i - \bar{y})^2} \quad (7.3)$$

$$\text{alocação em falta} = \begin{cases} 0 & , \text{if } \hat{y} \geq y \\ y - \hat{y} & , \text{otherwise} \end{cases} \quad (7.4)$$

$$\text{optimal percentage} = \frac{1}{N} \sum_{i=1}^n 1[\hat{y}_i \geq y_i \& \hat{y}_i \leq alloc](\text{not: [13]}) \quad (7.5)$$

Onde  $\hat{y}$  são as previsões dos modelos,  $y$  são os valores reais utilizados, "UpwardUsedSecondaryReserveEnergy", e  $alloc$  são os valores alocados "SecondaryReserveAllocationAUpward"

## 7.2 Experiências

Para comparação o erro absoluto para o ano de 2021 na alocação feita é de 3889367.4

## 7.2.1 Arquiteturas e numeros de epocas

### 7.2.1.1 Arquiteturas

name	rmse	abs erro	erro comp	r2 score	mape score	alloc missing	alloc surplus	optimal percentage	better allocation	beter percentage
VanillaDense	1200.20	10073188.28	False	-41.27	192.50	487.24	10072701.04	1.13	0.77	1.17
StackedCNN	453.12	3717806.99	True	-5.03	82.67	52038.16	3665768.83	35.04	34.48	38.34
UNET	183.58	1166186.91	True	0.01	17.02	687071.63	479115.27	60.39	60.39	85.47
VanillaCNN	174.96	1124205.72	True	0.10	16.45	614355.90	509849.82	61.79	61.79	86.18

### 7.2.1.2 Epocas

## 7.2.2 Funções de Perda (Loss)

As funções de perda foram estudas com duas arquitecturas diferentes, de modo a conseguir ter uma melhor noção do impacto das mesmas.

name	rmse	abs erro	erro comp	r2 score	mape score	alloc missing	alloc surplus	optimal percentage	better allocation	beter percentage
UNETmse	180.84	1269085.40	True	0.04	23.32	521000.39	748085.00	68.78	68.78	87.31
StackedCNNmae	198.29	1083872.29	True	-0.15	5.93	937786.28	146086.01	43.65	43.65	83.32
StackedCNNmape	232.60	1257578.80	True	-0.59	1.01	1252220.35	5358.46	19.15	19.15	80.60
StackedCNNmse	176.71	1103152.68	True	0.08	14.81	673158.53	429994.14	58.75	58.75	85.68
StackedCNNwl	2002.81	17092114.23	False	-116.72	320.21	0.00	17092114.23	0.05	0.00	0.05
UNETmae	195.49	1083463.60	True	-0.12	7.14	887646.74	195816.87	45.95	45.95	83.70
StackedCNNsle	225.42	1761315.40	True	-0.49	40.41	320813.27	1440502.13	79.46	79.46	89.93
StackedCNNsde	612.31	4978280.05	False	-10.00	81.86	3325165.78	1653114.27	18.27	18.06	30.14
UNETsle	211.40	1142731.21	True	-0.31	4.70	1048097.40	94633.80	39.64	39.64	82.11

A nível de percentagem de modelo melhor, elas estão todas bastante renhidas, mas ha uma clara vantagem quando vemos a percentagem de melhor alocação ou de alocação optima. Sendo que a perda que avança é a de MSLE (Mean Square Log Error). Esta perda é tendicionalmente usada para distribuições exponencias, e onde temos bastantes outliers, que devem ser considerados. O que é o caso no nosso problema. Esta experiência valida a observação feito na análise estatística sobre o mesmo.

## 7.2.3 Funções de Perda (Loss)

Aqui devem constar gráficos e sua análise crítica e ligação com a secção 2 da revisão bibliográfica no sentido de comparar valores e discutir diferenças, por exemplo. A legenda dos gráficos deve seguir a das figuras, isto é, porque também são figuras.

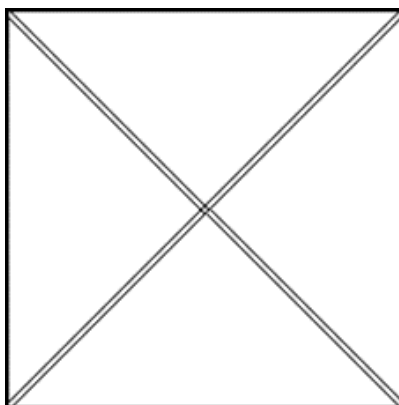


Figure 7.1: Exemplo de como considerar um gráfico

## **Conclusões e sugestões futuras**

Aqui são dadas as respostas às perguntas de investigação formuladas na secção 1.2. Não fazer aqui a discussão dos resultados. Essa discussão deve ser feita no capítulo 7.2.3. Não esquecer de indicar sugestões futuras para que um colega possa dar continuidade ao trabalho desenvolvido.

# Referências

- [1] Fernando Lopes. “From wholesale energy markets to local flexibility markets: structure, models and operation”. In: *Local Electricity Markets* (Jan. 2021), pp. 37–61. DOI: [10.1016/B978-0-12-820074-2.00009-5](https://doi.org/10.1016/B978-0-12-820074-2.00009-5).
- [2] William F. Watson, Paul L. Joskow, and Richard Schmalensee. “Markets for Power: An Analysis of Electrical Utility Deregulation”. In: *Southern Economic Journal* 51 (2 Oct. 1984), p. 640. ISSN: 00384038. DOI: [10.2307/1057863](https://doi.org/10.2307/1057863).
- [3] Fred C. Schweppe et al. “Spot Pricing of Electricity”. In: (1988), p. 384. URL: [https://books.google.com/books/about/Spot\\_Pricing\\_of\\_Electricity.html?hl=pt-PT&id=QZ8ACAAAQBAJ](https://books.google.com/books/about/Spot_Pricing_of_Electricity.html?hl=pt-PT&id=QZ8ACAAAQBAJ).
- [4] Ricardo J. Bessa et al. “Reserve Setting and Steady-State Security Assessment Using Wind Power Uncertainty Forecast: A Case Study”. In: *IEEE Transactions on Sustainable Energy* 3 (4 Oct. 2012), pp. 827–836. ISSN: 1949-3029. DOI: [10.1109/TSTE.2012.2199340](https://doi.org/10.1109/TSTE.2012.2199340).
- [5] Célia Carneiro. “Mecanismos de funcionamento do Mercado de Serviços de Sistema - aplicação ao MIBEL”. In: (Nov. 2016). URL: <https://core.ac.uk/download/pdf/302870251.pdf>.
- [6] Camila Fernandes, Pablo Frías, and Javier Reneses. “Participation of intermittent renewable generators in balancing mechanisms: A closer look into the Spanish market design”. In: *Renewable Energy* 89 (Apr. 2016), pp. 305–316. ISSN: 0960-1481. DOI: [10.1016/J.RENENE.2015.12.037](https://doi.org/10.1016/J.RENENE.2015.12.037). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0960148115305437>.
- [7] Marco Agostini et al. “The participation of small-scale variable distributed renewable energy sources to the balancing services market”. In: *Energy Economics* 97 (May 2021), p. 105208. ISSN: 0140-9883. DOI: [10.1016/J.ENERCO.2021.105208](https://doi.org/10.1016/J.ENERCO.2021.105208).
- [8] Stef van Buuren and Karin Groothuis-Oudshoorn. “mice: Multivariate Imputation by Chained Equations in R”. In: *Journal of Statistical Software* 45 (3 Dec. 2011), pp. 1–67. ISSN: 1548-7660. DOI: [10.18637/JSS.V045.I03](https://doi.org/10.18637/JSS.V045.I03). URL: <https://www.jstatsoft.org/index.php/jss/article/view/v045i03>.
- [9] S F Buck. “A Method of Estimation of Missing Values in Multivariate Data Suitable for use with an Electronic Computer”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 22 (2 1960), pp. 302–306. ISSN: 00359246. URL: <http://www.jstor.org/stable/2984099>.
- [10] V L Helen Josephine, A.P. Nirmala, and Vijaya Lakshmi Alluri. “Impact of Hidden Dense Layers in Convolutional Neural Network to enhance Performance of Classification Model”. In: *IOP Conference Series: Materials Science and Engineering* 1131 (1 Apr. 2021), p. 012007. ISSN: 1757-8981. DOI: [10.1088/1757-899X/1131/1/012007](https://doi.org/10.1088/1757-899X/1131/1/012007).
- [11] Hansika Hewamalage, Christoph Bergmeir, and Kasun Bandara. “Recurrent Neural Networks for Time Series Forecasting: Current status and future directions”. In: *International Journal of Forecasting* 37 (1 Jan. 2021), pp. 388–427. ISSN: 01692070. DOI: [10.1016/j.ijforecast.2020.06.008](https://doi.org/10.1016/j.ijforecast.2020.06.008).



## REFERÊNCIAS

- [12] Rogério Luís de C. Costa. “Convolutional-LSTM networks and generalization in forecasting of household photovoltaic generation”. In: *Engineering Applications of Artificial Intelligence* 116 (Nov. 2022), p. 105458. ISSN: 0952-1976. DOI: [10.1016/J.ENGAPPAI.2022.105458](https://doi.org/10.1016/J.ENGAPPAI.2022.105458).
- [13] Kenneth E. Iverson. *A programming language*. 1962, p. 11. ISBN: 0471430145. URL: <https://www.jsoftware.com/papers/APL.htm>.

# Anexos

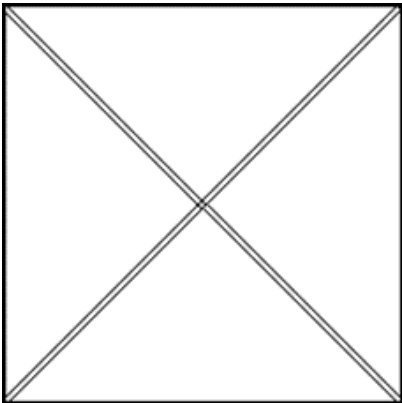


Figure A.1: Exemplo de como considerar um gráfico nos anexos.

Table A.1: Isto é um exemplo de uma tabela. Se fôr igual(copiada) a outro autor deve ser pedido autorização para reproduzir.

Title 1	Title 2	Title 3	Title 4
entry 1	data	data	data
	data	data	data
	data	data	data
entry 2	data	data	data
	data	data	data
entry 3	data	data	data
	data	data	data
	data	data	data
	data	data	data
entry 4	data	data	data