



**João Filipe Inácio Lourenço**

Licenciado em Ciências da Engenharia

Eletrotécnica de Computadores

## **Previsão de Produção de Energia Elétrica através de Fontes de Energia Renováveis**

Dissertação para obtenção do Grau de Mestre em  
Engenharia Eletrotécnica e de Computadores

Orientador: Prof. Doutor Pedro Miguel Ribeiro Pereira,  
Professor Auxiliar, Universidade Nova de Lisboa

Júri:

Presidente: Doutor João Francisco Alves Martins, Professor Auxiliar – FCT/UNL

Arguente: Doutor Luís Filipe Figueira Brito Palma, Professor Auxiliar – FCT/UNL

Vogal: Doutor Pedro Miguel Ribeiro Pereira, Professor Auxiliar – FCT/UNL

**Novembro, 2015**



FACULDADE DE  
CIÊNCIAS E TECNOLOGIA  
UNIVERSIDADE NOVA DE LISBOA



Copyright © João Filipe Inácio Lourenço, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.



# Agradecimentos

Ao meu orientador, o Professor Doutor Pedro Miguel Ribeiro Pereira pela amizade, orientação e disponibilidade que demonstrou desde o primeiro dia.

Um grande agradecimento à minha família, em especial aos meus pais e ao meu irmão pelo apoio incondicional que sempre prestaram ao longo da minha vida pessoal e académica e não menos obstante nesta última fase de mestrado. Sem o seu apoio, nunca conseguiria atingir esta etapa.

Obrigado aos meus colegas e amigos de faculdade, em especial para o Renato Paulino, Nuno Ramos, João Ralo, Flávio Silva, Gonçalo Oliveira, André Coelho, Joaquim Pereira, João Soares e Pedro Gomes, pelo seu companheirismo e amizade, com quem tive o prazer de conviver e partilhar esta importante fase da minha vida, sem eles teria sido muito mais difícil ultrapassar esta etapa. Um grande obrigado também para a minha colega e amiga Liliana Sequeira, por me ter acolhido e guiado nos meus primeiros tempos de faculdade.

Agradeço aos meus amigos Hélio Palma, Estefânia Cardoso, Joana Doroteia, Luís Cruz e mais uma vez ao Renato Paulino por estarem sempre presentes, quer nos momentos bons quer nos menos bons.

A todos aqueles que direta ou indiretamente estiveram comigo e sempre me ofereceram o seu apoio, obrigado!



## Resumo

A capacidade de prever precisamente a produção de energia renovável é extremamente relevante tanto do ponto de vista económico como para controlo da estabilidade da rede elétrica. Para tal, é necessário realizar uma previsão das condições meteorológicas adjacentes à produção de energia a partir de fontes de energia renovável. Vários modelos de previsão têm sido utilizados para este fim, desde modelos atmosféricos a modelos estatísticos, onde se destacam métodos como Redes Neurais Artificiais ou a Metodologia de *Box & Jenkins*. Lidar com dados meteorológicos pode revelar algumas complicações devido à possível instabilidade das medições, complicando o desenvolvimento de um modelo de previsão adequado. Neste trabalho pretende-se realizar a previsão de produção a partir de uma instalação fotovoltaica e um gerador eólico através do uso da Metodologia de *Box & Jenkins* para desenvolver um modelo capaz de realizar a previsão das condições meteorológicas para diferentes horizontes temporais medidos no topo do edifício do Departamento de Engenharia Eletrotécnica (DEE) da Faculdade de Ciências e Tecnologia (FCT), Universidade Nova de Lisboa (UNL), e usando esses valores para calcular a produção de energia. Os resultados obtidos revelaram um bom desempenho quando comparados os resultados previstos com os resultados reais para o mesmo período de tempo, garantindo que podem ser utilizados para calcular a previsão de potência produzida através das instalações presentes no local e encorajando novos estudos no tema.



## **Abstract**

The capacity to predict accurately the energy production is extremely relevant both from an economical point of view and to control the stability of the energy network. To do this, it is necessary to perform a forecast of weather conditions related to the production of energy from renewable sources. Several models have been used for this purpose, from physical models to statistical models, which features methods such as Artificial Neural Networks or the Box & Jenkins methodology. Dealing with meteorological data can reveal some complications due to possible unstable measurements, complicating the development of a suitable forecast model. In this work it is intended to calculate the power forecast from a photovoltaic system and a wind generator using the Box & Jenkins methodology in order to develop a method to carry out the forecast of weather conditions for different time horizons measured on top of the Departamento de Engenharia Eletrotécnica (DEE) da Faculdade de Ciências e Tecnologia (FCT), Universidade Nova de Lisboa (UNL), and then using those values to calculate the power production. The results show a good performance when comparing the predicted results with actual results for the same period of time, ensuring that the model can be used to calculate the power forecast produced by the on-site facilities and encouraging new studies in the field.



# Índice

<b>LISTA DE FIGURAS</b> .....	<b>XIII</b>
<b>LISTA DE TABELAS</b> .....	<b>XVII</b>
<b>1. INTRODUÇÃO</b> .....	<b>1</b>
1.1. ENQUADRAMENTO .....	1
1.2. MOTIVAÇÃO E OBJETIVOS .....	2
1.3. ORGANIZAÇÃO.....	3
<b>2. ESTADO DA ARTE</b> .....	<b>5</b>
2.1. MODELOS PREDITIVOS .....	5
2.1.1. Modelos Físicos ou Atmosféricos.....	6
2.1.2. Modelos Estatísticos e de Inteligência Computacional.....	6
2.1.3. Modelos Mistos.....	6
2.2. SÉRIES TEMPORAIS .....	7
2.3. MÉTODOS DE PREVISÃO DE SÉRIES TEMPORAIS.....	8
2.3.1. Modelo de Persistência.....	8
2.3.2. Estimativa por Regressão Linear .....	8
2.3.3. Metodologia de Box & Jenkins.....	9
2.3.4. Redes Neurais Artificiais (RNA).....	10
2.3.5. Sistemas de Inferência Difusa (SID).....	11
<b>3. METODOLOGIA</b> .....	<b>13</b>
3.1. METODOLOGIA ADOTADA .....	13
3.2. RECOLHA DE DADOS E SELEÇÃO DE VARIÁVEIS DE ENTRADA.....	14
3.2.1. Recolha de dados .....	14
3.2.2. Seleção de dados e definição de variáveis de entrada do sistema .....	18
3.2.3. Sinopse.....	21
3.3. SELEÇÃO DO MODELO DE PREVISÃO .....	22
3.4. TRATAMENTO DE DADOS.....	23
3.4.1. Redução do número de amostras.....	24
3.4.2. Filtragem.....	25
3.5. METODOLOGIA DE BOX & JENKINS.....	26
3.5.1. Identificação do Modelo .....	29
3.5.2. Estimação do Modelo.....	34

3.5.3. Validação do Modelo.....	34
<b>4. VALIDAÇÃO.....</b>	<b>35</b>
4.1. DESENVOLVIMENTO DO MODELO DE PREVISÃO.....	35
4.1.1. Modelo de previsão para a temperatura.....	36
4.1.2. Modelo de previsão para a irradiância.....	42
4.1.3. Modelo de previsão para a velocidade do vento.....	48
4.2. PREVISÃO DE POTÊNCIA.....	56
4.2.1. Painéis fotovoltaicos.....	56
4.2.2. Gerador Eólico.....	60
<b>5. CONCLUSÕES E TRABALHOS FUTUROS.....</b>	<b>65</b>
5.1. CONCLUSÕES.....	65
5.2. TRABALHO FUTURO.....	67
<b>BIBLIOGRAFIA.....</b>	<b>69</b>

## Lista de figuras

<i>Figura 2.1 - Exemplo de matriz cognitiva e regras associativas</i> .....	12
<i>Figura 3.1 - Metodologia adotada</i> .....	13
<i>Figura 3.2 - Estação meteorológica instalada no topo do edifício do DEE</i> .....	14
<i>Figura 3.3 - Anemómetro Davis Cup instalado no todo do edifício do DEE</i> .....	15
<i>Figura 3.4 - Termómetro ECT Temperature instalado no topo do edifício do DEE</i> .....	16
<i>Figura 3.5 - Piranómetro PYS SP-110 instalado no topo do edifício do DEE</i> .....	17
<i>Figura 3.6 - Contador de Energia UPT210 [13]</i> .....	18
<i>Figura 3.7 - Gerador eólico instalado no topo do edifício do DEE</i> .....	19
<i>Figura 3.8 - Instalação fotovoltaica instalada no topo do edifício do DEE</i> .....	20
<i>Figura 3.9 - Variáveis de entrada e de saída do sistema</i> .....	21
<i>Figura 3.10 - Arquitetura adotada</i> .....	22
<i>Figura 3.11 - Exemplo das variações diárias das variáveis de entrada do sistema</i> .....	24
<i>Figura 3.12 - Exemplo das variações diárias das variáveis de entrada do sistema sem e com redução do número de amostras</i> .....	25
<i>Figura 3.13 - Exemplo das variações diárias das variáveis de entrada do sistema com e sem filtragem para intervalos entre amostras de 1 minuto (a) e de 15 minutos (b)</i> .....	26
<i>Figura 3.14 - Principais etapas da metodologia de Box &amp; Jenkins [10]</i> .....	28
<i>Figura 3.15 - Série temporal com truncamento abrupto[10]</i> .....	30
<i>Figura 3.16 - Série temporal com decaimento rápido[10]</i> .....	30
<i>Figura 3.17 - Série temporal com decaimento lento [10]</i> .....	30
<i>Figura 3.18 - Séries temporais com decaimento exponencial (a), oscilatório (b) e sinusoidal (c) [10]</i> .....	31
<i>Figura 4.1 - Série temporal da temperatura</i> .....	36
<i>Figura 4.2 - Histograma dos dados recolhidos da temperatura</i> .....	37
<i>Figura 4.3 - Q-QPlot dos dados recolhidos da temperatura</i> .....	37
<i>Figura 4.4 - FAC da série temporal da temperatura</i> .....	38
<i>Figura 4.5 - FAC da série temporal da temperatura após uma diferenciação</i> .....	39
<i>Figura 4.6 - 4-plot dos resíduos do modelo ARIMA (1,1,0)</i> .....	40

<i>Figura 4.7 - Previsão da temperatura até 48 horas à frente através do modelo</i>	
<i>ARIMA (1,1,0)</i> .....	41
<i>Figura 4.8 - Previsão da temperatura até 48 horas à frente através do modelo</i>	
<i>ARIMA (1,1,1)</i> .....	42
<i>Figura 4.9 - Série temporal da irradiância</i> .....	43
<i>Figura 4.10 - Histograma dos dados recolhidos da irradiância</i> .....	44
<i>Figura 4.11 - Q-Qplot dos dados recolhidos da irradiância</i> .....	44
<i>Figura 4.12 - FAC da série temporal da irradiância</i> .....	45
<i>Figura 4.13 - 4-plot dos resíduos do modelo ARIMA (1,0,1)</i> .....	46
<i>Figura 4.14 - Previsão da temperatura até 48 horas à frente através do modelo</i>	
<i>ARIMA (1,0,1)</i> .....	47
<i>Figura 4.15 - Série temporal da velocidade do vento (10 meses)</i> .....	48
<i>Figura 4.16 - Série temporal da velocidade do vento (4 meses)</i> .....	49
<i>Figura 4.17- Histograma dos dados recolhidos da velocidade do vento</i> .....	50
<i>Figura 4.18 - Q-QPlot da velocidade do vento</i> .....	50
<i>Figura 4.19 - FAC da série temporal da velocidade do vento (sem diferenciação)</i>	
.....	51
<i>Figura 4.20 - FAC da série temporal da velocidade do vento (após diferenciação)</i>	
.....	52
<i>Figura 4.21 - 4-Plot dos resíduos da velocidade do vento</i> .....	53
<i>Figura 4.22 - Previsão da velocidade do vento até 24 horas à frente através do</i>	
<i>modelo ARIMA (1,1,1)</i> .....	54
<i>Figura 4.23 - Previsão da velocidade do vento até 12 horas à frente através do</i>	
<i>modelo ARIMA (4,1,1)</i> .....	56
<i>Figura 4.24 - Previsão de potência produzida através dos painéis fotovoltaicos e</i>	
<i>previsão de irradiância</i> .....	58
<i>Figura 4.25 - Comparação entre valores previstos e valores reais de produção</i>	
<i>fotovoltaica</i> .....	59
<i>Figura 4.26 - Curva de potência do gerador eólico instalado no todo do edifício do</i>	
<i>DEE</i> .....	60
<i>Figura 4.27 - Resultados obtidos através da ferramenta cftool do MatLAB</i> .....	61
<i>Figura 4.28 - Previsão de potência produzida através do gerador eólico e</i>	
<i>previsão da velocidade do vento</i> .....	62

*Figura 4.29 - Comparação entre potência produzida (real) pelo gerador eólico e velocidade do vento.....63*



## Lista de tabelas

<i>Tabela 2.1 – Exemplos de modelos de previsão implementados.....</i>	<i>5</i>
<i>Tabela 2.2 - Vantagens e desvantagens do uso de RNA.....</i>	<i>11</i>
<i>Tabela 3.1 - Especificações Gerais do anemómetro Davis Cup [13].....</i>	<i>15</i>
<i>Tabela 3.2 - Especificações gerais do Termómetro ECT Temperature [13].....</i>	<i>16</i>
<i>Tabela 3.3 - Especificações gerais do Piranómetro PYR SP-110 [13].....</i>	<i>17</i>
<i>Tabela 3.4 - Especificações gerais do Gerador Eólico [13].....</i>	<i>19</i>
<i>Tabela 3.5 - Características gerais dos painéis fotovoltaicos instalados [13].....</i>	<i>20</i>
<i>Tabela 3.6 - Modelo indicado de acordo com a FAC do processo.....</i>	<i>33</i>
<i>Tabela 4.1 - Estatísticas da série temporal da temperatura.....</i>	<i>36</i>
<i>Tabela 4.2 - Resultados de previsão para diferentes horizontes temporais.....</i>	<i>41</i>
<i>Tabela 4.3 - Estatísticas da série temporal da irradiância.....</i>	<i>43</i>
<i>Tabela 4.4 - Resultados de previsão para diferentes horizontes temporais.....</i>	<i>47</i>
<i>Tabela 4.5 - Estatísticas da série temporal da velocidade do vento.....</i>	<i>49</i>
<i>Tabela 4.6 - Resultados de previsão para diferentes horizontes temporais.....</i>	<i>54</i>
<i>Tabela 4.7 - Características dos painéis para cálculo da potência produzida [13]</i>	<i>58</i>
.....	58



# 1. Introdução

O presente capítulo tem como objetivo demonstrar a importância da previsão da produção da energia elétrica no contexto atual do mercado de energia e de sustentabilidade ambiental. São referidos também quais as motivações e objetivos que levaram ao desenvolvimento desta dissertação, tal como contribuições futuras em que esta possa resultar.

## 1.1. *Enquadramento*

Como consequência do aumento da população mundial, o consumo de energia elétrica sofre um aumento considerável nos dias que correm, resultando num aumento da procura de energia que, sem medidas apropriadas, terá um efeito severo no clima mundial devido às emissões de gases de efeito de estufa que daí advêm. Uma maneira de estabilizar as emissões destes gases é substituir a fonte de energia, isto é, deixar de depender inteiramente de combustíveis fósseis, substituindo-os por fontes de energia renovável, como, por exemplo, a energia solar e a energia eólica [1].

Hoje em dia, as redes elétricas são na sua maioria centralizadas, transferindo potência das centrais elétricas para os utilizadores finais, no entanto, é expectável que unidades de produção descentralizadas surjam de uma maneira significativa. A introdução de energias renováveis na rede pode resultar numa melhor eficiência no controlo da mesma, sendo que o principal desafio para os operadores da rede é sincronizar, continuamente, a procura de energia com o fornecimento desta [1], [2].

Com o crescimento da procura de energia renovável, os problemas a nível técnico da penetração da mesma na rede elétrica devem ser tomados em conta. A variabilidade dos recursos naturais, as variações sazonais na produção e o alto custo associado ao armazenamento de energia levanta dúvidas quanto à confiabilidade destes sistemas [1].

Ao contrário dos sistemas de produção convencionais, a produção de energia a partir de fontes renováveis não pode ser planeada antecipadamente com precisão, pelo facto da energia solar ser altamente dependente de condições meteorológicas tais como o aparecimento de nuvens ou os ciclos de dia e noite. Já a energia eólica possui um comportamento instável, não conseguindo fornecer para a rede de forma constante [3].

Como tal, a integração da produção através destas fontes requer uma avaliação da disponibilidade precisa e como a produção de energia não possui o mesmo comportamento que a sua procura, torna-se necessário possuir informação exata da disponibilidade desta.

Realizar uma previsão da produção em diferentes horizontes temporais tem um papel fundamental em diversas áreas, seja no armazenamento de energia, em sistemas de controlo em edifícios ou no controlo da rede elétrica, permitindo uma adaptação da carga de forma a otimizar o transporte de energia ou planear atividades de manutenção [1], [2].

O problema com a integração de energia renovável na rede deve-se ao facto que a produção desta não é facilmente prevista, uma vez que depende tanto das condições meteorológicas como de condições específicas do local onde as instalações se encontram. Enquanto as grandes organizações podem gastar o tempo necessário para realizar um modelo de previsão para instalações de produção de grande capacidade, desenvolver estes modelos para instalações de menor dimensões como edifícios ou *smart-homes* não é realizável, tendo esta iniciativa de partir do próprio utilizador que possui a instalação fotovoltaica ou eólica [1], [3], [2].

O objetivo desta dissertação não é desenvolver um modelo de forma a competir com os modelos de previsão comerciais, mas sim investigar a confiabilidade de modelos estatísticos de baixo custo que não requerem informação para além de dados históricos relativos às variáveis em estudo.

## 1.2. *Motivação e Objetivos*

Realizar métodos de previsão para a energia solar e eólica resulta numa melhor qualidade de energia entregue à rede e reduz os custos adicionais com a dependência meteorológica. A combinação destes dois fatores tem sido a principal motivação para vários trabalhos de investigação. Com o desenvolvimento desta dissertação pretende-se desenvolver um sistema capaz de prever a produção de energia através de fontes renováveis num contexto pedagógico mas cuja aplicação possa ser útil caso se pretenda realizar a integração com a rede elétrica.

Esta dissertação tem como objetivo criar um algoritmo de previsão para a energia renovável produzida através da instalação fotovoltaica e do gerador eólico instalado no Departamento de Engenharia Eletrotécnica (DEE) da Faculdade de Ciências e Tecnologia (FCT), Universidade Nova de Lisboa (UNL), utilizando apenas registos de fenómenos meteorológicos obtidos durante o ano de 2013.

### 1.3. *Organização*

A presente dissertação encontra-se organizada em 5 capítulos.

O primeiro capítulo apresenta uma introdução ao trabalho, como está organizada a dissertação, qual a motivação e objetivos que levaram à sua realização.

O segundo capítulo, designado por Estado de Arte, oferece uma visão geral dos modelos de previsão atualmente implementados em todo o mundo. Apresenta uma definição de séries temporais e os métodos de previsão que mais se ajustam a estas.

O terceiro capítulo trata da metodologia aplicada nesta dissertação, desde a recolha de dados e definição das variáveis de entrada do sistema, passando pelo tratamento desses mesmos dados e acabando na escolha do modelo de previsão a implementar, seguindo-se uma explicação mais detalhada sobre o mesmo do que aquela que foi apresentada no segundo capítulo, focando no processo de identificação e de estimação do modelo, de forma a levar a que os resultados obtidos sejam o mais satisfatórios possível.

O quarto capítulo é relativo à validação da metodologia proposta. Sendo apresentadas, passo a passo, cada uma das fases de desenvolvimento do modelo de previsão para cada uma das variáveis, apresentando os resultados obtidos e comparando-os com os resultados expectáveis (reais), por forma a verificar a adequabilidade do modelo. Por último, nesta fase são também convertidos os resultados previstos em potência produzida prevista, de forma a alcançar o objetivo principal desta dissertação.

O quinto capítulo apresenta as conclusões do trabalho desenvolvido e sugestões para trabalhos futuros.

O sexto e último capítulo é referente à bibliografia, apresentando todas as fontes externas que permitiram a realização desta dissertação.



## 2. Estado da Arte

Neste capítulo são abordados os diferentes tipos de técnicas e modelos utilizados em sistemas de previsão. É também apresentada uma definição de série temporal, seguido de uma caracterização dos métodos de previsão que melhor se aplicam a estas.

### 2.1. Modelos Preditivos

O uso de metodologias capazes de criar modelos precisos e robustos têm vindo a ser utilizadas em diferentes áreas, desde controlo de sistemas a previsão de indicadores socioeconómicos, evolução de sistemas, problemas de classificação e também na previsão meteorológica. Tendo em conta a aplicação e o horizonte de previsão a serem estudados, os modelos para realizar a previsão de condições meteorológicas são:

- Modelo Físico (ou Atmosférico);
- Modelo Estatístico e Inteligência Computacional Artificial (análise de séries temporais);
- Modelo Híbrido ou Combinado (Estatístico e Físico).

Estes modelos são utilizados a nível mundial em estações meteorológicas para previsões com um curto/médio horizonte temporal, na tabela 2.1 está, a título exemplificativo, uma lista de alguns países europeus e quais os modelos que utilizam para realizar a previsão meteorológica [4], [5].

Tabela 2.1 – Exemplos de modelos de previsão implementados

Modelo de previsão	Desenvolvido por:	Método	Localização
WPPT	Universidade de Copenhaga	Estatístico	Dinamarca
Zephyr	Riso & IMM	Combinado	Dinamarca
Previento	Universidade de Oldenburg	Estatístico	Alemanha
AWPPS	Armines	Estatístico	Irlanda, Creta
RAL	RAL	Estatístico	Irlanda
SIPREÓLICO	University Carlos III	Estatístico	Espanha
Prediktor	Risco	Físico	Espanha e Alemanha

### 2.1.1. Modelos Físicos ou Atmosféricos

O tipo de modelo mais utilizado pelos investigadores da área meteorológica é conhecido por *Numeric Weather Prediction* (NWP). Este tipo de modelo fornece a previsão meteorológica para horizontes de previsão até 8 horas e para o seu uso é necessário definir condições iniciais e de contorno, bem como o mapa digital do terreno e a rugosidade do mesmo.

O domínio do modelo é referente à área coberta pelo modelo de previsão. Associado ao tipo de domínio existem diversos tipos de fronteiras, em modelos de área limitada existem fronteiras laterais, superiores e inferiores, enquanto nos modelos globais (cobrem toda a superfície terrestre) existem apenas fronteiras superiores.

O princípio básico de um modelo NWP é calcular numericamente a evolução meteorológica através da integração de diversas equações associadas à circulação meteorológica (como por exemplo, as equações da termodinâmica) com uma condição inicial obtida através de observações meteorológicas [5].

### 2.1.2. Modelos Estatísticos e de Inteligência Computacional

Este tipo de modelos conta com uma grande variedade de métodos, de onde se destacam os modelos de inteligência computacional (Redes Neurais, Sistemas de Inferência Difusa, etc.) e os modelos estatísticos (Modelo ARIMA, Regressão linear, etc.). Esta abordagem, devido ao facto da sua precisão ser inversamente proporcional com o aumento do horizonte de previsão, é apenas adequada para previsões a curto prazo [5].

### 2.1.3. Modelos Mistos

Esta abordagem combina o modelo numérico de previsão (NWP) com os métodos estatísticos, com as redes neurais e a lógica difusa, de forma a aproveitar as vantagens de cada um no desenvolvimento do modelo de previsão.

Estes modelos são, geralmente, aplicados em locais onde já está implementado um modelo NWP, mas que não tenha uma resolução adequada para o horizonte de previsão desejado, sendo que a saída do modelo NWP será utilizada como entrada para os modelos estatísticos ou de inteligência computacional [5].

## 2.2. Séries Temporais

Uma série temporal é um conjunto de observações ordenadas no tempo, não necessariamente igualmente espaçadas, que apresentam dependência serial, isto é, dependência entre instantes de tempo, por exemplo: os dados de uma variável aleatória  $z$  (consumo de energia) no instante  $t$ , com  $t$  variando de 1 até  $N$  (número total de observações), devem, de certa maneira, conter informações necessárias para que seja determinado o valor dessa variável no instante  $t + 1$ . As séries temporais podem ser classificadas como discretas, contínuas, determinísticas, estocásticas, multivariadas e multidimensionais. A maneira tradicional de analisar uma série temporal é através da sua decomposição nas componentes de tendência, cíclica e de sazonalidade [6].

As séries temporais podem ser representadas como a combinação de três componentes:

- **Tendência:** É um movimento que traduz a influência de fatores que fazem com que a série aumente ou diminua o seu valor com o passar do tempo. Esta componente caracteriza-se como um movimento ascendente ou descendente de longa duração. Quando a série temporal não apresenta qualquer tendência ascendente ou descendente ela é denominada de série estacionária.
- **Cíclica:** são caracterizados pelas oscilações de subida e de queda nas séries, de forma suave e repetida, ao longo da componente de tendência. Por exemplo, ciclos meteorológicos.
- **Sazonal:** corresponde às oscilações de subida e de descida que ocorrem num determinado período do ano, do mês, da semana ou do dia. A diferença entre as componentes sazonais e cíclica é que a componente sazonal possui movimentos facilmente previsíveis, ocorrendo em intervalos regulares de tempo, enquanto movimentos cíclicos tendem a ser irregulares.

Existe também uma quarta componente, chamada de componente de erro. Esta ocorre quando, após um efeito de tendência, cíclico ou sazonal, surgem flutuações de período curto, com deslocamento inexplicável.

Em geral, ao estudar uma série temporal, pretende-se descrever a mesma, verificar as suas características mais relevantes e as suas possíveis relações com outras séries, ou realizar previsões de curto prazo (*forecast*) a partir de valores históricos da série. O número de instantes para o qual é realizada a previsão é chamado de horizonte de previsão.

## 2.3. Métodos de Previsão de Séries Temporais

Aqui são abordados alguns modelos de previsão de séries temporais, dos quais será escolhido um para desenvolver para a previsão da produção de energia a partir da instalação fotovoltaica e do gerador eólico instalado no topo do edifício do DEE.

### 2.3.1. Modelo de Persistência

Um dos modelos mais utilizados na previsão de curto-prazo da velocidade do vento é o método de persistência. Este método corresponde ao método da média móvel simples em que a previsão é a média das  $N$  observações mais recentes de uma série  $x$ , como pode ser observado na equação 2.1.

$$X_t = \frac{1}{N} \sum_{i=1}^N X_{t-i} \quad (2.1)$$

Este método é considerado o método de previsão mais simples, uma vez que a previsão que realiza é apenas baseada nos últimos valores da série. O método de previsão é utilizado no caso de os dados disponíveis quanto á previsão do vento se rem insuficientes e deve ser utilizado apenas para previsões de curto-prazo (horizonte temporal de poucas horas) [5], [7].

### 2.3.2. Estimativa por Regressão Linear

Um dos métodos para previsão de qualquer série temporal é a descrição da série como uma função cúbica da sua variável, através da aplicação da teoria da regressão. Esta teoria permite que se estabeleçam relações entre variáveis que se relacionam e cujas informações estão disponíveis (através de dados históricos), relações às quais se associam modelos de regressão.

Uma vez estabelecida a relação pelo modelo de regressão, é necessário avaliar a confiança que nela se pode colocar, realizando testes estatísticos. Entre os vários tipos básicos de informação a considerar, insere-se a informação que descreve as mudanças assumidas por uma variável ao longo do tempo (dados de séries temporais). Para este tipo de informação é possível estabelecer relações que descrevem as situações observadas por meio de modelos de regressão.

Um modelo de regressão de uma equação serve como base para obter uma revisão pontual, à qual se associa um intervalo de confiança, dando origem a previsões de intervalos de confiança da previsão, construídos para que se obtenha uma margem de erro em torno da previsão pontual, definindo então bandas de confiança de previsão [5], [8].

### 2.3.3. Metodologia de *Box & Jenkins*

Segundo *Box & Jenkins* (1994), a análise de séries temporais tem como objetivo principal a realização de previsões. Esta metodologia permite realizar a previsão de valores futuros de uma série temporal tendo por base apenas os seus valores presentes e passados. O processo realiza-se através da correlação temporal existente entre os valores. De acordo com os autores, a realização desta metodologia é representada por um conjunto de processos estocásticos denominados modelos ARIMA (*AutoRegressive Integrated Moving Average*) onde, em cada instante de tempo, existe um conjunto de valores que a série pode assumir, aos quais estão associadas diferentes probabilidades de ocorrência [9].

A metodologia de *Box & Jenkins* consiste num processo constituído por 5 etapas:

1. Estabelecer a estacionaridade da série temporal. Caso a série não seja estacionária, é necessário proceder a sucessivas diferenciações da série até que esta obtenha essa característica.
2. Identificar um modelo que se ajuste aos dados em estudo através de análise da série temporal, mais propriamente das suas funções de autocorrelação e autocorrelação parcial.
3. Especificar o modelo e estimar os parâmetros do mesmo.
4. Realizar testes de validação de forma a garantir que o modelo escolhido descreve os dados em estudo de forma adequada. Deve ser feita uma análise aos resíduos, sendo que estes não devem apresentar correlações entre eles.
5. Após a verificação da adequabilidade do modelo, este estará pronto a ser utilizado para realizar a previsão num horizonte de tempo pré-definido.

A metodologia de *Box & Jenkins* apresenta como principal vantagem a sua metodologia sistemática, e o facto de permitir estimações tanto pontuais como durante um intervalo de tempo pré-definido e baseado numa série de hipóteses e testes estatísticos para validação do modelo [5], [10].

### 2.3.4. Redes Neurais Artificiais (RNA)

As redes neurais fazem parte de um campo de estudo da Inteligência artificial, que é uma área que desenvolve técnicas que permitem incorporar num ambiente computacional, mesmo que de forma limitada, algumas funcionalidades dos sistemas biológicos no que se refere aos seus aspetos de inteligência. Por este motivo oferecem uma alternativa na resolução de problemas complexos, são capazes de aprender a partir de exemplos e de lidar com dados incompletos e uma vez treinadas podem realizar previsões instantaneamente [1], [11].

As unidades de processamento (neurónios) constituem os principais elementos de uma rede neuronal. A organização das unidades de processamento define a arquitetura ou topologia de uma rede neuronal, que pode ser de uma só camada ou multicamada.

Um dos métodos de treino mais utilizados é o algoritmo de retro propagação do erro. Este algoritmo tem como base matemática o método da descida mais acentuada (*steepest descent*).

O treino da rede neuronal com o algoritmo de retro propagação do erro é realizado de forma supervisionada, onde são apresentados um a um exemplos de um conjunto de treino. Durante a fase de treino, uma passagem completa pelo conjunto de treino é designado por época. Uma prática comum é realizar o ajuste dos pesos e polarizações da rede durante várias épocas até que um determinado critério de convergência seja atingido. Este procedimento deve ser executado com extremo cuidado para evitar um sobre ajustamento da rede neuronal aos dados de treino, sendo necessário garantir a capacidade de generalização da rede neuronal.

Um método muito comum para evitar o sobre ajustamento consiste em dividir os dados em treino e validação, este último conjunto é utilizado para aferir a qualidade do treino e a capacidade de generalização da rede durante a fase de treino, em cada época é calculado o erro no conjunto de validação parando-se quando ao fim de um número definido de iterações o erro aumentar sucessivamente, sendo a melhor solução aquela que dá o erro menor no processo de teste.

O treino de retro propagação do erro pode ser realizado de três formas distintas:

- Modo *Offline*: os parâmetros são ajustados somente ao final de cada época (processamento de todo o conjunto de observações), ou seja, os parâmetros da rede são ajustados somente ao final de cada ciclo. Desta forma, o treino é menos influenciado pela ordem de apresentação dos padrões, é menos suscetível a oscilações, porém a velocidade de aprendizagem geralmente é mais baixa. Esta solução é a mais adequada quando os dados estão afetados por ruído.

- Modo Sequencial ou estocástico: os parâmetros são ajustados ao final do processamento de cada observação (padrão), isto é, os pesos da rede são ajustados no final do processamento de cada observação.
- Modo *Online*: cada novo exemplo de treino disponível é propagado desde a entrada da rede até à saída, sendo depois o erro à saída retro propagado apenas uma vez, atualizando de imediato os pesos da rede. Poderá ser conveniente previamente realizar um treino *offline* com os dados disponíveis (se existirem), para determinar os pesos iniciais do treino *online*.

Na tabela 2.2 estão resumidas algumas das vantagens e as desvantagens do uso de técnicas de redes neuronais.

Tabela 2.2 - Vantagens e desvantagens do uso de RNA

<b>Vantagens</b>	<b>Desvantagens</b>
Qualidade superior comparada com outras técnicas estatísticas.	As redes podem permitir chegar a conclusões que contrariam a teoria.
Autoaprendizagem, apenas se baseiam em dados históricos.	Treino demorado, o treino da rede em função da aplicação pode demorar horas.
Baixo custo de implementação.	Necessário grande volume de dados para teste da rede.
Capacidade de generalização, mesmo com lacunas nos dados (ou dados incompletos) as redes podem preencher essas lacunas sem degradação.	

### 2.3.5. Sistemas de Inferência Difusa (SID)

Os sistemas de inferência difusa (SID) utilizam princípios de lógica difusa, fornecendo respostas ponderadas pelos graus de pertença das variáveis independentes de entrada do sistema.

Esses sistemas são modelos empíricos, nos quais o conhecimento é adquirido através de experiências passadas. O conhecimento de um sistema de inferência difusa pode ser armazenado em forma matricial (figura 2.1) chamada de matriz cognitiva. Essa matriz contém uma coleção de regras que indicam qual das respostas é mais significativa para um dado vetor de entrada.

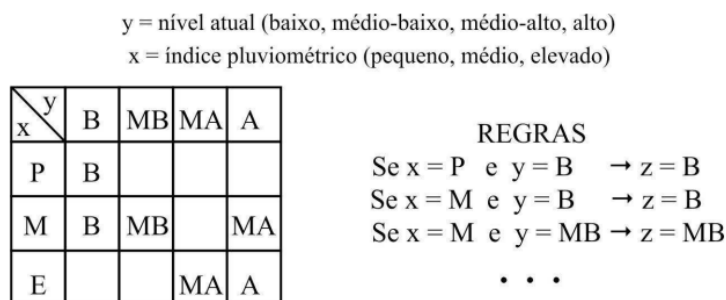


Figura 2.1 - Exemplo de matriz cognitiva e regras associativas

A lógica difusa distingue-se da lógica clássica (Booleana) por permitir graus de pertença diferentes de 0 e 1. Para um conjunto difuso existem graus intermédios de pertença, uma vez que elementos podem pertencer a mais do que um conjunto. O conceito de pertença é diferente do de probabilidade, uma vez que, enquanto a probabilidade indica a possibilidade de ocorrência de um evento, o grau de pertença indica o quanto o evento pertence a uma classe considerada.

A aprendizagem de um sistema de inferência difusa passa pelo estabelecimento de regras difusas que associam a cada configuração de conjuntos de entrada, uma saída representada pelo conjunto difuso correspondente. Métodos de aprendizagem automáticos, que utilizam dados amostrais no estabelecimento de regras associativas, são normalmente utilizados quando o grau de complexidade e o número de dimensões ultrapassam a capacidade humana de relacionar causas e efeitos. O método de aprendizagem organizativo consiste, basicamente, numa pesquisa dos dados amostrais, com contagem das ocorrências de combinações dos agrupamentos de entrada aos de saída. As regras são armazenadas em uma matriz cognitiva onde cada dimensão é uma variável do sistema, e o valor armazenado é o peso da conexão. A utilização de um número elevado de conjuntos e de variáveis de entrada resulta em sistemas com dimensão excessiva. Se a amostra de dados para a aprendizagem não possuir situações suficientes para a geração das regras correspondentes, pode ocorrer falta de conhecimento disponível para a generalização do modelo [12].

## 3. Metodologia

Neste capítulo é abordado todo o processo realizado ao longo da dissertação até desenvolvimento do modelo de previsão, isto é, desde a aquisição de dados, análise dos mesmos, seleção do tipo de modelo que melhor se ajusta para a previsão de produção de energia eólica através das fontes de energia renovável presentes no departamento e explicação do mesmo. Todos os cálculos e gráficos apresentados nas páginas seguintes foram realizados através da utilização do *software MatLAB*.

### 3.1. Metodologia Adotada



Figura 3.1 - Metodologia adotada

Como é possível observar na figura 3.1, a metodologia adotada para esta dissertação passa por quatro fases principais:

1. Recolha de dados – fase em que foram recolhidos dados históricos sobre diversas variáveis que podem influenciar a produção de energia elétrica, nesta fase foi também decidido quais as variáveis que faz sentido considerar como variáveis de entrada no sistema;
2. Seleção do Modelo de Previsão – fase em que é identificado o modelo de previsão mais adequado para as variáveis de entrada;
3. Tratamento de dados – fase em que as variáveis de entrada do sistema sofreram todos os ajustes necessários para serem introduzidas num modelo de previsão;
4. Desenvolvimento e validação – fase em que é desenvolvido o modelo de previsão e que os resultados obtidos através do mesmo são comparados com valores reais por forma a averiguar o desempenho do modelo.

## 3.2. *Recolha de dados e seleção de variáveis de entrada*

Neste capítulo será abordado o método de recolha de dados, desde a especificação da instrumentação utilizada para a recolha dos dados meteorológicos à seleção dos mesmos como variáveis de entrada do sistema a desenvolver.

### 3.2.1. Recolha de dados

O primeiro passo para o desenvolvimento de um modelo capaz de prever a produção a partir de fontes de energia renovável passa por uma recolha de dados meteorológicos durante um longo período de tempo, para o trabalho em questão estes dados foram recolhidos através de uma estação meteorológica presente no local onde estão instalados os sistemas de produção em estudo, entre Janeiro e Dezembro de 2013, com um intervalo de tempo de um minuto. Estes dados contemplam registos relativos a temperatura, irradiância, direção e velocidade do vento.

#### 3.2.1.1. *Estação Meteorológica*



Figura 3.2 - Estação meteorológica instalada no topo do edifício do DEE

A estação meteorológica em questão (figura 3,2) encontra-se instalada no terraço do Departamento de Engenharia Eletrotécnica (DEE) da Faculdade de Ciências e Tecnologias da Universidade Nova de Lisboa (FCT-UNL) e consiste num anemómetro Davis Cup, num termómetro ECT Temperature e num piranómetro PYR, modelo SP-110 [13].

### 3.2.1.2. Anemómetro Davis Cup



Figura 3.3 - Anemómetro Davis Cup instalado no todo do edifício do DEE

Na figura 3.3 encontra-se uma fotografia do anemómetro instalado na central meteorológica do DEE, sendo que as suas especificações gerais podem ser consultadas na tabela 3.1.

Tabela 3.1 - Especificações Gerais do anemómetro Davis Cup [13]

<b>Especificações Gerais</b>	
Sensor: Direção do vento	Estabilizador vertical e potenciómetro
Sensor: Velocidade do vento	Copos/pás rotacionais
<b>Alcance</b>	
Direção do vento	0° - 360°
Velocidade do vento	0 - 58 m/s
<b>Precisão</b>	
Direção do vento	± 7%
Velocidade do vento	± 5%
<b>Resolução</b>	
Direção do vento	1°
Velocidade do vento	0,45 m/s

### 3.2.1.3. Termómetro ECT Temperature

Da estação meteorológica faz também parte um termómetro ECT Temperature constituído por um termístor e protegido por um escudo de irradiação, como é possível observar na figura 3.4 [13].



Figura 3.4 - Termómetro ECT Temperature instalado no topo do edifício do DEE

As especificações gerais do termómetro ECT Temperature encontram-se descritos na tabela 3.2.

Tabela 3.2 - Especificações gerais do Termómetro ECT Temperature [13]

<b>Tipo de Sensor</b>	Termístor
<b>Alcance</b>	-40°C a 50°
<b>Precisão</b>	De -20°C a 5°C: $\pm 1^\circ\text{C}$ De 40°C a 50°C: $\pm 1^\circ\text{C}$ De -40°C a -20°C: $\pm 1^\circ\text{C}$ De 5°C a 40°C: 0,5°C
<b>Resolução</b>	Inferior ou igual a 0°C: 0,015°C Superior a 0°C: 0,016°C
<b>Velocidade de medição</b>	10 ms

<b>Condições de operação</b>	-40°C a 50°C
<b>Comprimento do cabo</b>	3 m

#### 3.2.1.4. Piranómetro PYR SP-110

Por forma a medir a irradiância, foi instalado um piranómetro (sensor de célula de silício) PYR SP-110, da empresa Apogee Instruments Inc., como representado na figura 3.5 [13].



Figura 3.5 - Piranómetro PYS SP-110 instalado no topo do edifício do DEE

Na tabela 3.3 encontram-se as especificações gerais do instrumento em questão.

Tabela 3.3 - Especificações gerais do Piranómetro PYR SP-110 [13]

<b>Alcance</b>	0 – 1750 $W/m^2$
<b>Precisão</b>	± 5%
<b>Condições de operação</b>	-40°C a 55°C 0% a 100% de humidade relativa À prova de água

### 3.2.1.5. Contadores de energia

Para aquisição dos dados de produção e de consumo do departamento, são utilizados três contadores, todos eles da marca Algodue Eletronica, modelo UPT210 (figura 3.6) [13].



Figura 3.6 - Contador de Energia UPT210 [13]

Um dos contadores encontra-se ligado entre o consumo e a rede de forma a saber os consumos internos do edifício do DEE, enquanto os outros dois se encontram ligados ao gerador eólico e à instalação fotovoltaica presentes no topo do edifício por forma a medirem a produção de energia destes sistemas.

### 3.2.2. Seleção de dados e definição de variáveis de entrada do sistema

Por forma a garantir a melhor escolha possível das variáveis de entrada do modelo de previsão, é necessário realizar uma análise do sistema de produção instalado. Este sistema é constituído por um gerador eólico e por um conjunto de painéis fotovoltaicos, cujas características serão detalhadas neste subcapítulo. Este sistema encontra-se instalado no mesmo local da estação meteorológica, pelo que os dados recolhidos podem ser introduzidos diretamente no modelo de previsão sem sofrerem qualquer tipo de transformação (ex: caso o gerador eólico estivesse a uma altura superior à do anemómetro seria necessário calcular a velocidade do vento tendo em conta essa diferença de altitude).

### 3.2.2.1. Gerador Eólico



Figura 3.7 - Gerador eólico instalado no topo do edifício do DEE

O gerador eólico em estudo (figura 3.7) é um modelo FD3.6-2000-10L da empresa Yangzhou Shenzhou Wind-Driven Generator, na tabela 3.4 é possível consultar as características gerais do gerador segundo o fabricante [13].

Tabela 3.4 - Especificações gerais do Gerador Eólico [13]

<b>Potência Nominal (W)</b>	2000
<b>Tensão Nominal (V)</b>	120
<b>Diâmetro das pás do rotor (m)</b>	3,2
<b>Velocidade de arranque (m/s)</b>	3
<b>Velocidade nominal do vento (m/s)</b>	9
<b>Velocidade de corte (m/s)</b>	16
<b>Rotação nominal das pás (rpm)</b>	400
<b>Material das pás</b>	Fibra de vidro
<b>Número de pás</b>	3
<b>Gerador</b>	Alternador de Magnetos Permanentes

Este modelo de gerador eólico não é um sistema fixo, isto é, o gerador alinha-se com a direção do vento de forma a garantir o máximo aproveitamento possível a todos os instantes, como tal, tendo em conta os dados relativos ao vento que a estação meteorológica são apenas a velocidade e a direção do vento, será tomada como variável de entrada para o modelo apenas a velocidade do vento.

### 3.2.2.2. Instalação Fotovoltaica



Figura 3.8 - Instalação fotovoltaica instalada no topo do edifício do DEE

A instalação fotovoltaica presente no DEE consiste num conjunto de 7 painéis fotovoltaicos (figura 3.8), cinco deles da empresa *Bangkok Solar*, do modelo *BS 40* e dois da empresa *SWEA* com o modelo *SLW 130*. Na tabela 3.5 encontram-se as características gerais dos dos tipos de painéis, em condições STC [13].

Tabela 3.5 - Características gerais dos painéis fotovoltaicos instalados [13]

	<i>Bangkok Solar – BS40</i>	<i>SWEA – SWL 130</i>
<b>Tipo de célula</b>	Silício amorfo	Silício policristalino
<b>Potência Máxima, <math>P_{max}</math> (W)</b>	40	130
<b>Tensão à potência máxima, <math>V_{max}</math> (V)</b>	44,8	40,8

<b>Corrente à potência máxima,</b> $I_{max}$ (A)	0,93	3,2
<b>Tensão de circuito aberto, <math>V_{oc}</math></b> (V)	62,2	50,1
<b>Corrente de curto-circuito,</b> $I_{sc}$ (A)	1,14	3,21
<b>Área total (<math>m^2</math>)</b>	$0,635 \times 1,245 = 0,8$	$1,2 \times 0,99 = 1,2$

A produção de energia fotovoltaica depende diretamente da irradiância e da temperatura, sendo que estas entram numa série de equações relativas ao cálculo da potência produzida por um painel fotovoltaico, como será detalhado no capítulo 4 desta dissertação, como tal, os dados provenientes da estação meteorológica relativos a irradiância e temperatura terão que ser tomados como variáveis de entrada do modelo de previsão.

### 3.2.3. Sinopse

Após a análise dos sistemas de produção de energia renovável presentes no DEE, bem como dos dados recolhidos através da estação meteorológica, concluiu-se que as variáveis a considerar como variáveis de entrada no modelo de previsão devem ser: a velocidade do vento, a temperatura e a irradiância. Por outro lado, a variável de saída do modelo deve ser a produção total de energia renovável, como se encontra ilustrado na figura 3.8.

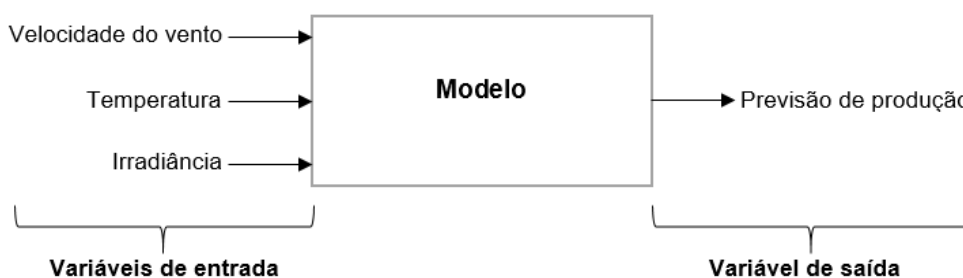


Figura 3.9 - Variáveis de entrada e de saída do sistema

Poder-se-ia desenvolver um método de previsão baseado na produção da instalação fotovoltaica e da instalação eólica medida pelos contadores descritos em 3.2.1.5, mas acontece que os dados recolhidos apresentam várias lacunas entre eles, desde horas a semanas, o que indica que existiu algum problema com a contagem de energia, o que impossibilitou a utilização desses dados para o desenvolvimento do modelo.

### 3.3. Seleção do Modelo de Previsão

Na figura 3.9 está representado o sistema adotado, este sistema foi desenvolvido tendo em conta a definição das variáveis de entrada do sistema e o estudo realizado previamente no capítulo do estado da arte e os parágrafos seguintes são correspondentes à sua explicação.

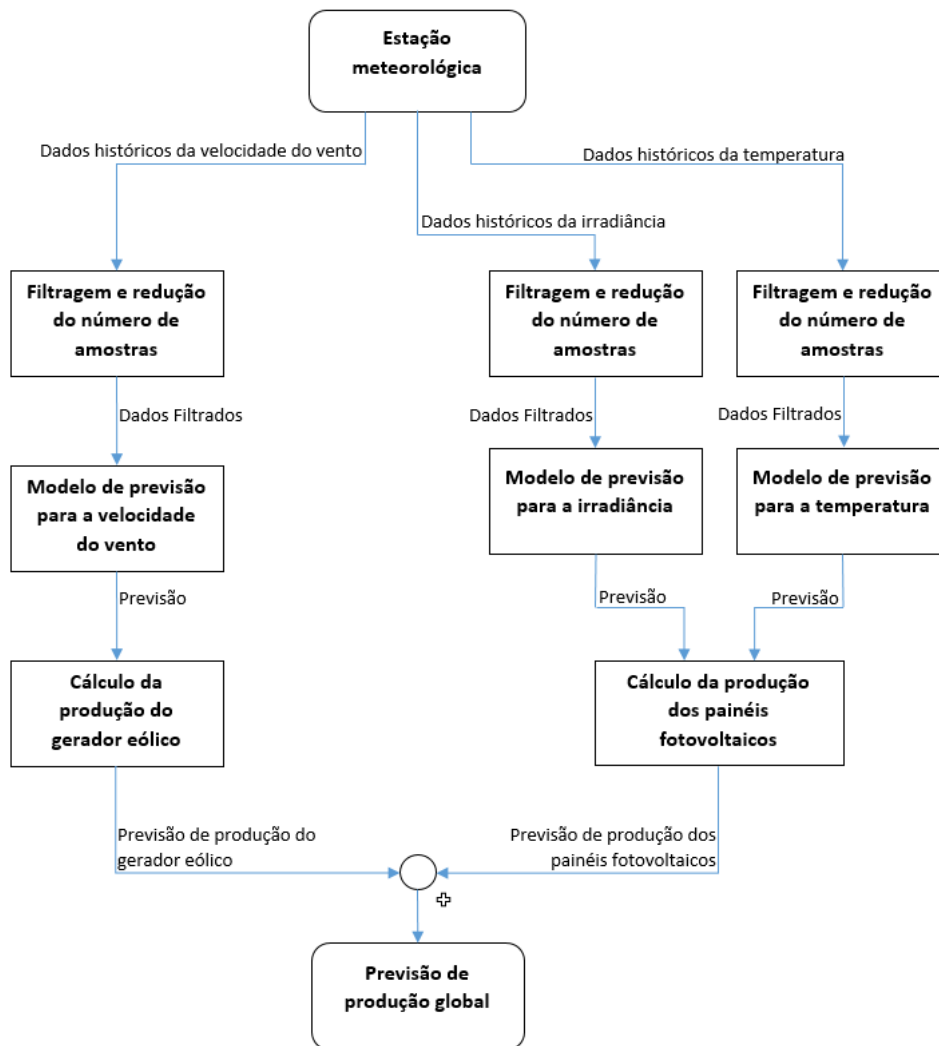


Figura 3.10 - Arquitetura adotada

Uma vez definidas as variáveis de entrada do modelo de previsão, é necessário proceder para a seleção do tipo de modelo a desenvolver. Esta escolha recaiu sobre os modelos estatísticos em detrimento dos modelos de inteligência computacional artificial, uma vez que existem dados históricos suficientes para garantir uma boa aprendizagem do modelo e também devido ao elevado processamento que seria necessário realizar para aplicar um modelo de inteligência artificial. Dos modelos estatísticos possíveis, a escolha recaiu sobre uma modelização de *Box & Jenkins*, uma vez que, normalmente, as previsões obtidas através deste tipo de modelo, para um curto horizonte temporal, apresentam bons resultados.

Uma vez que os dois sistemas de produção são independentes, tal como todas as variáveis de entrada, será realizada a previsão de cada uma das variáveis individualmente. Para que os modelos recebam os dados nas melhores condições possíveis, estes são filtrados e o seu número de amostras é reduzido de forma a suavizar o sinal para uma melhor aprendizagem do modelo.

Obtidos os resultados de previsão, a produção de cada um dos sistemas de produção é calculada através de uma estimativa baseada na curva de potência, para o caso do gerador eólico, e de acordo com as equações que descrevem o comportamento dos painéis fotovoltaicos, para a produção de energia solar. Estas duas produções são então somadas por forma a obter a previsão de produção global, que é a variável de saída do sistema.

No capítulo 3.4 é explicada a forma como foi realizado o tratamento de dados, enquanto o capítulo 3.5 retrata, de uma forma mais detalhada do que foi previamente retratado no capítulo do estado da arte, a metodologia de *Box & Jenkins*.

### 3.4. *Tratamento de dados*

Uma vez definidas as variáveis de entrada e saída do sistema, é necessário efetuar um tratamento dos dados que, posteriormente, serão utilizados para treinar o modelo ARIMA, como tal, foi desenvolvido um excerto de código que permite apresentar, num gráfico, as variações de cada uma das séries temporais ao longo de um dia (figura 3.10), por forma a verificar se os dados poderiam ou não ser introduzidos no modelo sem qualquer tipo de tratamento.

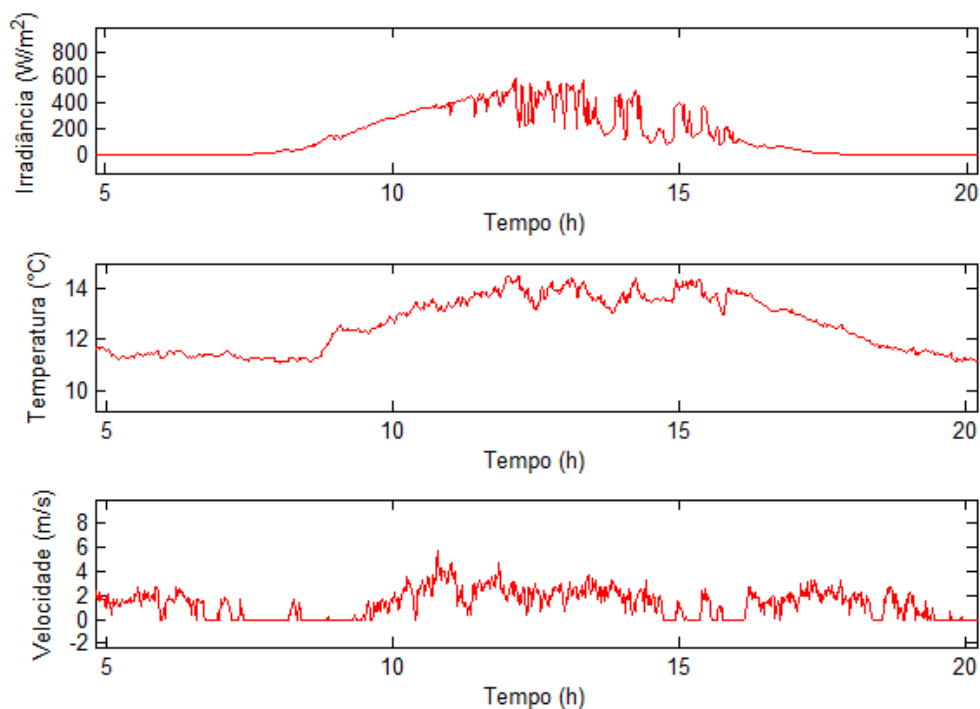


Figura 3.11 - Exemplo das variações diárias das variáveis de entrada do sistema

É visível, pela figura 3.10, que as três séries temporais apresentam variações bruscas de comportamento, este fenómeno pode ocorrer devido a influências exteriores que influenciam as medições efetuadas pela estação meteorológica (passagem de nuvens, ventos frios, etc.) ou até mesmo devido a possíveis falhas/avarias durante essas medições. Acontece que estas variações bruscas têm um impacto negativo no desenvolvimento do modelo, uma vez que quanto mais instável for o padrão da série temporal mais dificuldade existirá durante a fase de treino do modelo, dificultando o processo de aprendizagem do mesmo.

### 3.4.1. Redução do número de amostras

Por forma a estabilizar o comportamento das séries temporais, reduzindo as variações no comportamento referidas anteriormente, desenvolveu-se um código que permite diminuir a frequência do sinal, evitando ao máximo a perda de informação. Este código permite definir qual o intervalo de tempo entre amostras e realiza um cálculo da média das amostras originais durante esse período, por exemplo, se for considerado um intervalo entre amostras,  $z$ , de 15 minutos, o valor no instante  $x(15)$  será dado pela média dos 15 valores anteriores, como é possível observar na equação 3.1.

$$x(j) = \frac{\sum_{i=z}^{i=0} x(j-i)}{z} \quad (3.1)$$

Como tal, ao aplicar a equação 3.1 aos dados originais é reduzido o número de amostras, suavizando o comportamento da série mas sem perder a sua informação, como pode ser observado na figura 3.11.

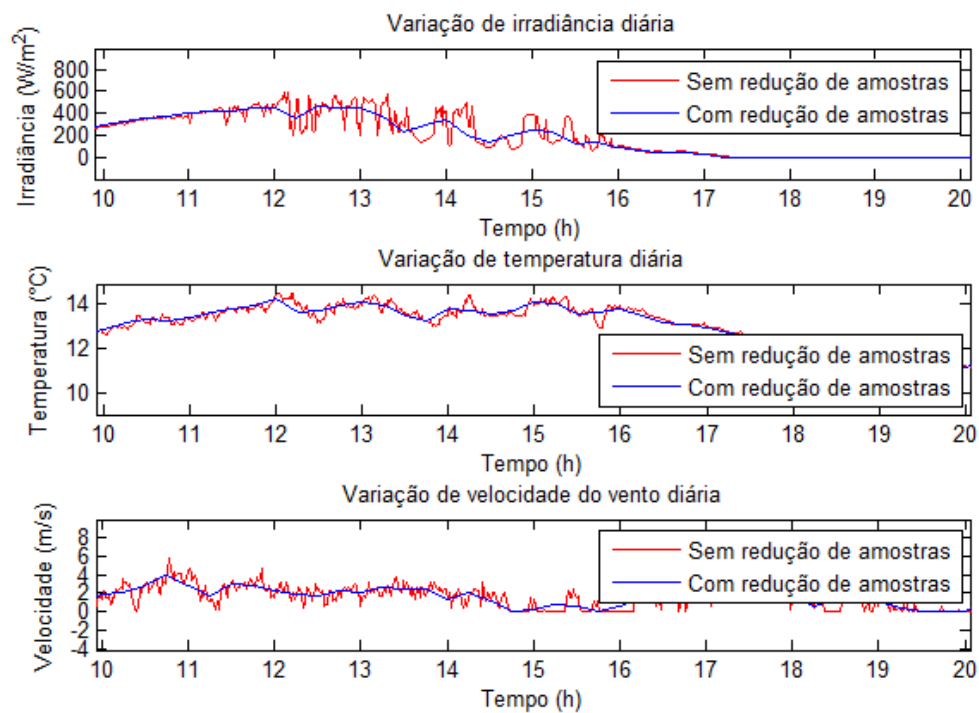


Figura 3.12 - Exemplo das variações diárias das variáveis de entrada do sistema sem e com redução do número de amostras

### 3.4.2. Filtragem

Apesar de o processo de redução do número de amostras servir, de certa forma, como um primeiro filtro para os dados, estes continuam a poder exibir picos e variações bruscas, como tal, um processo de suavização pode ser necessário, principalmente caso não se realize a redução do número de amostras ou o intervalo entre estas escolhido seja pequeno. Para tal efeito, foi utilizado um filtro de média móvel que suaviza os dados substituindo cada ponto pela média dos valores vizinhos numa dada janela de amostragem, sem perder o formato original dos dados de entrada, eliminando apenas o ruído existente.

Os resultados obtidos após a utilização deste filtro podem ser observados, a título exemplificativo, na figura 3.12, onde os dados são filtrados sem a existência de uma redução do número de amostras (a) e onde o tempo entre amostras é de 5 minutos (b).

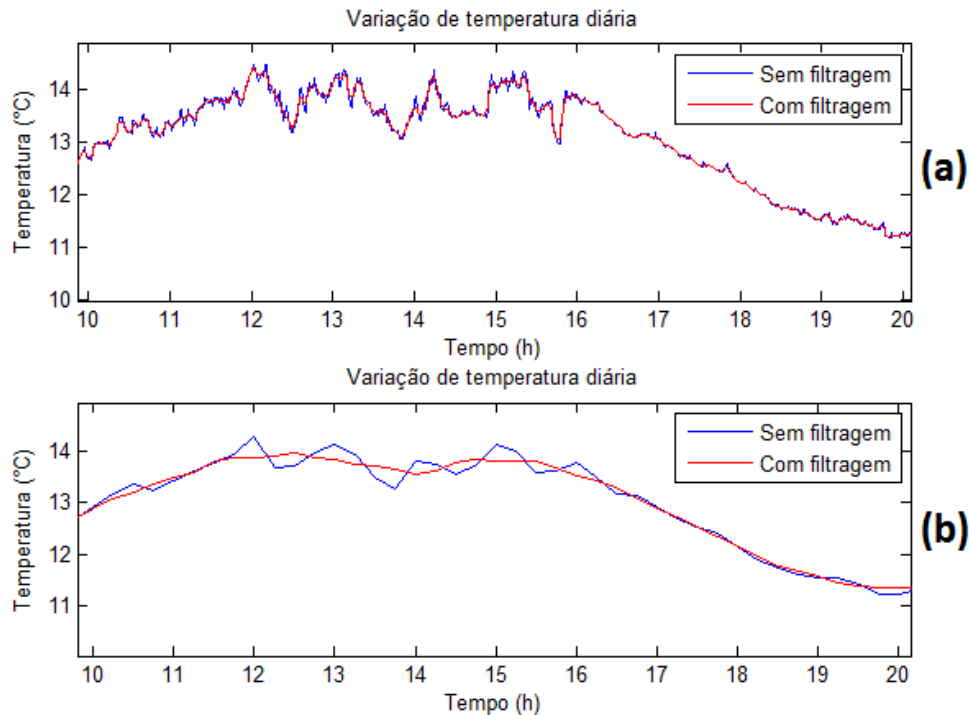


Figura 3.13 - Exemplo das variações diárias das variáveis de entrada do sistema com e sem filtragem para intervalos entre amostras de 1 minuto (a) e de 15 minutos (b)

### 3.5. Metodologia de Box & Jenkins

Como foi referido anteriormente, uma abordagem bastante utilizada para a análise de modelos paramétricos é conhecida por metodologia de *Box & Jenkins* (1970). Esta metodologia consiste no ajuste de modelos Autorregressivos Integrados e de Médias Móveis, ARIMA, a séries temporais de valores observados para que a diferença entre os valores gerados pelos modelos e os valores observados resulte em séries de resíduos de comportamento aleatório em torno de zero [10].

Nestes modelos, ao contrário dos modelos de regressão clássicos, não se pode assumir independência entre observações, pelo contrário, estes vão modelar o grau de autocorrelação entre desvios e observações desfasadas.

De forma geral, quando se faz referência a modelos ARIMA, considera-se que esses modelos são ajustados à série original. Por sua vez, para os modelos ARMA, considera-se que a série é uma série diferenciada. Considerando  $Y_t$  uma série já diferenciada, os modelos ARMA podem ser definidos, de forma geral, pela equação 3.2.

$$Y_t = \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (3.2)$$

Os parâmetros  $p$  e  $q$  representam o número de parâmetros relativos aos comprimentos de defasagem em que se observam valores significativos de autocorrelação e que correspondem a particularidades da série que devem ser explicadas pelo modelo.

Estes modelos podem ser separados em dois modelos complementares, os modelos de média móvel e os modelos autorregressivos. Os primeiros correspondem a processos de médias móveis de ordem  $q$  em que cada observação  $Y_t$  é gerada por uma média ponderada dos erros aleatórios  $q$  períodos no passado. Estes modelos são definidos através da equação 3.3:

$$Y_t = \alpha + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (3.3)$$

onde os parâmetros  $\theta_1, \dots, \theta_q$  podem ser positivos ou negativos. O sinal negativo no terceiro termo em diante corresponde a uma convenção.

Por outro lado, o modelo autorregressivo genérico é dado pela equação 3.4:

$$Y_t = \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (3.4)$$

Este processo modela uma autorregressão da variável  $Y_t$  com uma defasagem dela própria, para os  $p$  períodos de defasagem em que a autocorrelação parcial entre as variáveis é significativa.

O modelo de Box-Jenkins assume que a série temporal é estacionária. Box e Jenkins recomendaram a diferenciação de séries não estacionárias em uma ou mais vezes até atingir a estacionaridade, é esta característica que difere o modelo ARIMA de um modelo ARMA.

Os modelos ARIMA são construídos através de um ciclo iterativo, onde a escolha da estrutura do modelo é baseada nos seus próprios dados, como tal, estes modelos são capazes de descrever os processos de geração de séries temporais para os previsores. Na figura 3.13 estão representadas as principais etapas deste ciclo.

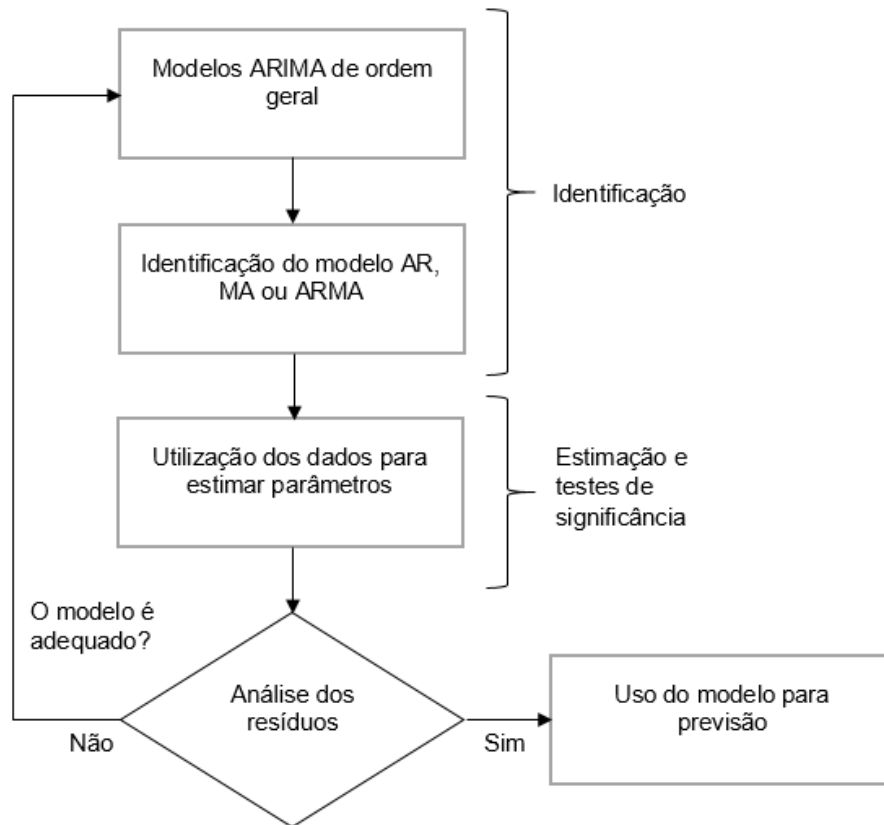


Figura 3.14 - Principais etapas da metodologia de *Box & Jenkins* [10]

A metodologia de *Box & Jenkins* é então constituída por três fases principais, identificação do modelo, estimação do modelo e, por último, validação (teste de adequação) do modelo.

### 3.5.1. Identificação do Modelo

Nesta fase são abordados as diferentes etapas para a identificação do modelo que melhor se ajusta à série temporal [10], [14].

#### 3.5.1.1. Estacionaridade e Sazonalidade

O primeiro passo no desenvolvimento do modelo é determinar se a série temporal em estudo é ou não estacionária e se existe algum tipo de sazonalidade.

Muitos processos temporais apresentam sazonalidade, isto é, apresentam flutuações periódicas. Caso a sazonalidade esteja presente no comportamento da série temporal, esta deve ser incorporada no modelo da série temporal. Caso a sazonalidade seja significativa, o gráfico de autocorrelação deverá mostrar alta correlação nos *lags* iguais ao período, por exemplo, para dados diários com efeito sazonal, será expectável que existam picos nas *lags* 24, 48, 72, *etc.* Caso os dados sejam mensais, é esperado que a sazonalidade seja observável através das *lags* 12, 24, 36, *etc.*

Nesta etapa da metodologia de *Box & Jenkins*, o objetivo passa por detetar a sazonalidade, caso exista, e identificar a ordem dos termos autorregressivos sazonais e dos termos de média móvel sazonal, a sazonalidade não deve ser removida antes de ajustar o modelo, deve sim ser incluída a ordem dos termos sazonais na especificação do mesmo. Quando o período é conhecido, os valores destes termos são idênticos ao período da série, isto é, para dados mensais com sazonalidade, tipicamente, seria incluído um termo AR sazonal ou um termo MA sazonal igual a 12.

Ao contrário da sazonalidade, que a sua existência não é uma obrigatoriedade para o desenvolvimento do modelo, a estacionaridade do processo é fundamental pois esta garante a possibilidade de fixar os parâmetros do modelo válidos para realizar a revisão com base nos dados do passado.

Um processo estacionário tem a propriedade de que a sua média, variância e covariância não mudam no decorrer do tempo ou quando a probabilidade de ocorrerem flutuações no processo em torno da média é a mesma em qualquer momento do processo.

Caso o processo não seja estacionário, é necessário realizar uma transformação por forma a remover a não-estacionaridade através da diferenciação dos dados, isto é, dada uma série  $X_t$ , é criada uma nova série  $Y_t = X_t - X_{t-1}$ . Os dados diferenciados vão então conter um ponto a menos que os dados originais.

Por forma a verificar a estacionaridade de uma s3rie temporal, 3 3 poss3vel recorrer 3 an3lise dos gr3ficos de autocorrela3o do processo:

- Uma s3rie pode ser considerada estacion3ria se existir um truncamento abrupto (figura 3.14) ou um padr3o de decaimento r3pido (figura 3.15);
- Caso exista um padr3o de decaimento lento para zero entre amostras, a s3rie 3 considerada n3o-estacion3ria (figura 3.16)

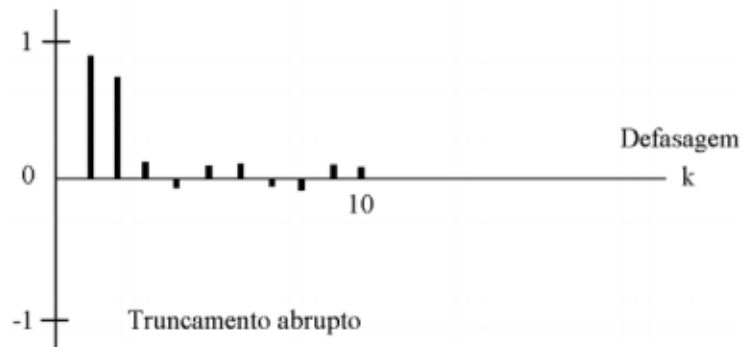


Figura 3.15 - S3rie temporal com truncamento abrupto[10]

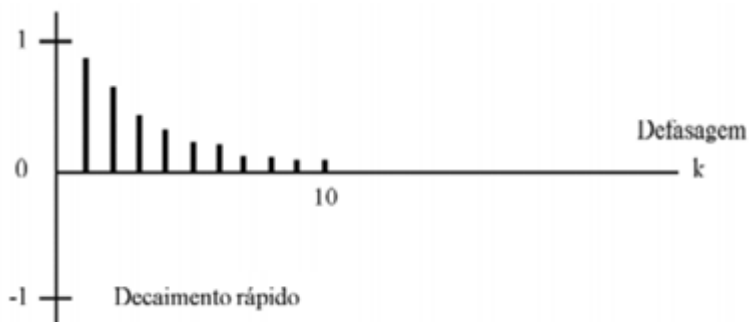


Figura 3.16 - S3rie temporal com decaimento r3pido[10]

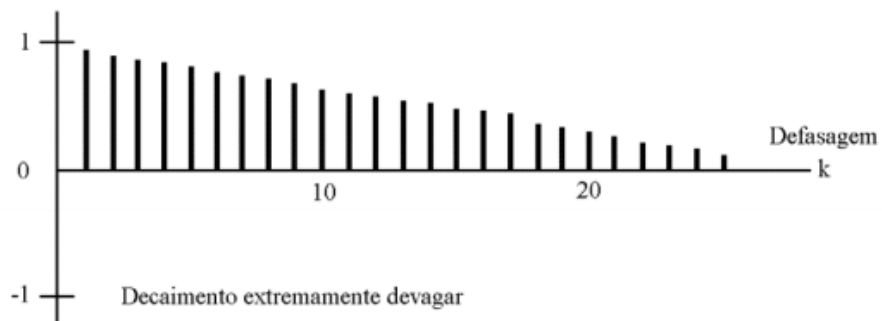


Figura 3.17 - S3rie temporal com decaimento lento [10]

É de ressaltar que este decaimento exponencial pode ou não ter oscilações, ou apresentar a forma de uma onda sinusoidal, como está exemplificado na figura 3.17.

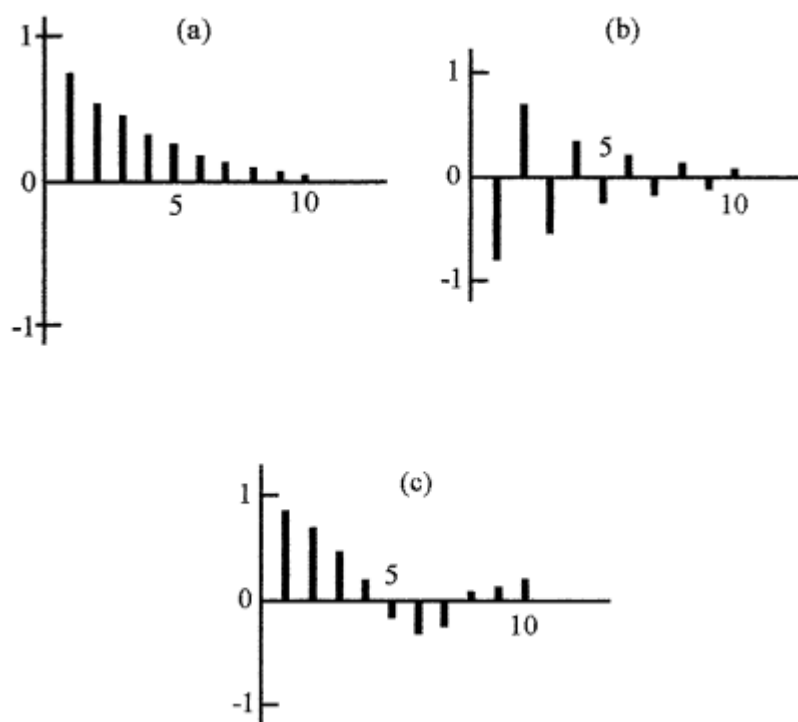


Figura 3.18 - Séries temporais com decaimento exponencial (a), oscilatório (b) e sinusoidal (c) [10]

Alternativamente à análise dos gráficos de autocorrelação, é possível aplicar o Teste da Raiz Unitária de *Dickey-Fuller* [10].

### 3.5.1.2. Teste da Raiz Unitária de *Dickey-Fuller*

Considerando um processo estocástico (uma coleção de variáveis aleatórias indexadas por um conjunto de índices, que representa, por exemplo, a evolução temporal de um sistema) de raiz unitária:

$$Y_t = \rho Y_{t-1} + \varepsilon_t, -1 \leq \rho \leq 1, \quad (3.2)$$

onde  $\varepsilon_t$  é um termo de erro de ruído branco.

Se  $\rho = 1$  (raiz unitária), o processo gerador da série  $Y_t$  é o processo aleatório e a série é não-estacionária.

Subtraindo  $Y_{t-1}$  nos dois lados da igualdade, obtém-se:

$$Y_t - Y_{t-1} = \rho Y_{t-1} - Y_{t-1} + \varepsilon_t \quad (3.3)$$

$$\therefore \Delta Y_{t-1} = \delta Y_{t-1} + \varepsilon_t \quad (3.4)$$

, onde  $\delta = \rho - 1$ .

Um procedimento de teste da raiz unitária pode ser aplicado a esta equação. Dessa forma, estima-se os parâmetros da equação e testa-se  $\delta = 0$ .

Se  $\delta = 0, \rho = 1 \Rightarrow$  raiz unitária e a série original é não estacionária.

No entanto, a distribuição *t de student* não se aplica nestes casos, como tal, é utilizado o teste de *Dickey-Fuller*, que verifica se o valor da estatística *t* estimado para o parâmetro de  $Y_{t-1}$  segue a distribuição da estatística  $\tau$ . Dessa forma, o teste de *Dickey-Fuller* verifica se a hipótese nula é rejeitada ou não em determinados níveis de significância estatística, conforme valores tabelados. Se o valor em módulo de  $t < |\text{estatística } \tau|$ , não é possível rejeitar a hipótese nula, ou seja, a hipótese de não-estacionaridade.

Caso a hipótese nula seja rejeitada, pode existir um de três casos:

- $Y_t$  é uma série temporal estacionária com média zero;
- $Y_t$  é uma série temporal estacionária com média diferente de zero;
- $Y_t$  é uma série estacionária em torno de tendência determinística.

Similar ao teste de *Dickey-Fuller*, existe o teste de *Dickey-Fuller* aumentado. A principal diferença entre estes dois é que o teste de *Dickey-Fuller* aumentado é utilizado para modelos de series temporais de maior complexidade que o teste simples. O resultado obtido no teste aumentado é uma constante que, caso seja negativa, quanto maior for o seu valor absoluto mais forte é a rejeição da hipótese de raiz unitária, caso seja positiva a hipótese nula não pode ser rejeitada.

### 3.5.1.3. Identificação dos Parâmetros p e q

Uma vez garantida a estacionaridade da série temporal e confirmada a existência ou não de sazonalidade, o passo seguinte é identificar os termos autorregressivos e de média móvel, p e q, respetivamente [6], [10], [14]

A principal ferramenta para se fazer isto são os gráficos de autocorrelação e o gráfico de autocorrelação parcial. A amostra do gráfico de autocorrelação e a amostra do gráfico de autocorrelação parcial são comparados ao comportamento teórico destes gráficos quando a ordem é conhecida.

Para um processo autorregressivo de ordem 1, AR(1), o gráfico de autocorrelação deve apresentar um decréscimo exponencial. Para processos autorregressivos de ordem superior, é expectável que o gráfico apresente uma mistura de componentes de decrescimento exponencial e/ou sinusoidal amortecido.

Para os processos autorregressivos de ordem superior a 1, o gráfico de autocorrelação necessita ser suplementada com um gráfico de autocorrelação parcial. A autocorrelação parcial de um processo AR(p) torna-se nula na *lag* p+1 e superiores.

Um processo de média móvel de ordem q, MA(q), torna-se nulo na *lag* q+1 e superiores, sendo esta a principal ferramenta para determinar a ordem do processo. Por outro lado, a função de autocorrelação parcial não apresenta utilidade na identificação deste parâmetro.

Caso um processo possua tanto uma componente autorregressiva como de média móvel, é normal que as funções de autocorrelação e de autocorrelação parcial não possuam o mesmo comportamento que as funções teóricas. Isto faz com que as ordens dos modelos mistos sejam particularmente difíceis de se identificar. Desenvolver bons modelos através da utilização e análise destes gráficos passa muito por um processo de tentativa e erro.

A tabela seguinte resume então qual o modelo indicado, dependendo da forma da função de autocorrelação do processo.

Tabela 3.6 - Modelo indicado de acordo com a FAC do processo

<b>Forma</b>	<b>Modelo Indicado</b>
Exponencial, decaindo para zero.	Modelo autorregressivo, necessário utilizar o gráfico de autocorrelação parcial para identificar a ordem do modelo regressivo.
Alternando entre valores positivos e negativos, decaindo para zero.	Modelo autorregressivo, necessário utilizar o gráfico de autocorrelação parcial para identificar a ordem do modelo regressivo.
Um ou dois picos, valores restantes praticamente nulos.	Modelo de média móvel, ordem é determinada através da <i>lag</i> em que o gráfico se torna nulo.
Decai após poucas <i>lags</i> .	Modelo misto.

Nula ou próxima de zero.	Os dados são essencialmente aleatórios.
Valores altos nos intervalos fixados.	Inclui o termo autorregressivo sazonal.
Não decai para zero.	A série não apresenta estacionaridade.

### 3.5.2. Estimação do Modelo

Uma vez que a estimação dos parâmetros do modelo é um problema não linear, esta etapa deve ser desenvolvida através da utilização de *software* que ajusta os modelos de *Box & Jenkins* [14].

### 3.5.3. Validação do Modelo

O diagnóstico dos modelos de *Box & Jenkins* é feito através da análise dos resíduos da série temporal. Caso o modelo escolhido seja adequado, os resíduos devem ser independentes, caso as suas distribuições sejam normais, ou ruído branco com uma distribuição fixa com média e variância constante [14]. Caso estas suposições não sejam verificadas é necessário reajustar o modelo

Caso existam vários modelos cuja análise dos resíduos garanta a sua adequabilidade, a seleção de um modelo em detrimento dos outros deve ser realizada através da comparação entre os resultados obtidos e os valores expectáveis (reais). Esta comparação pode ser realizada através de testes estatísticos como o MAPE (*Mean Absolute Percent Error*), que mede a dimensão do erro de previsão em termos percentuais, e o RMSE (*Root Mean Square Deviation*), que representa o desvio padrão médio entre os valores previstos e os valores observados. Estes dois testes estão representados nas equações 3.5 e 3.6 [15], [16].

$$MAPE = \left( \frac{1}{n} \sum \frac{|Real - Previsão|}{|Real|} \right) \times 100 \quad (3.5)$$

$$RMSE = \frac{1}{n} \sum (Real - Previsão)^2 \quad (3.6)$$

Estes dois testes serão os testes estatísticos utilizados para validação dos resultados obtidos para a previsão de cada uma das variáveis temporais, realizando-os para diferentes períodos dentro do horizonte temporal.

## 4. Validação

Neste capítulo da dissertação é realizado o desenvolvimento de cada um dos modelos de previsão, isto é, para a velocidade do vento, para a irradiância e para a temperatura, realizando uma análise estatística de cada série temporal, identificando quais os modelos que mais se aplicam através da análise das funções de autocorrelação e de autocorrelação parcial das séries temporais e realizando os ajustes necessários nos parâmetros dos modelos para obter o melhor desempenho possível. Os resultados são então comparados com valores reais por forma a verificar o seu desempenho. Numa última fase, os valores previstos são convertidos para potência produzida.

### 4.1. *Desenvolvimento do Modelo de Previsão*

Os dados utilizados no desenvolvimento dos modelos de previsão correspondem ao período de 10 meses, de Janeiro de 2013 a Outubro de 2013, uma vez que foi verificado que existem dados em falta no mês de Novembro e a inclusão desses dados prejudicaria o treino do modelo. Este período foi também escolhido de forma a ser possível comparar os resultados de previsão finais com um período em que os dados recolhidos estejam em boas condições. No caso da série temporal da velocidade do vento foi necessário reduzir ainda mais o período de teste uma vez que foram detetados mais dados em falta, como vai ser possível observar adiante. Para diminuir a carga computacional necessária para realizar o treino dos modelos foi realizada uma diminuição do número de amostras das séries temporais, passando estas a serem intervaladas por um período de tempo de 15 minutos.

Este capítulo é constituído por três subcapítulos, cada um deles respetivo ao desenvolvimento do modelo de previsão para cada uma das variáveis de entrada no sistema. As previsões realizadas são posteriormente utilizadas para converter os valores da velocidade do vento, da temperatura e da irradiância em potência gerada pelo gerador eólico e pelos painéis fotovoltaicos, respetivamente.

#### 4.1.1. Modelo de previsão para a temperatura

Neste subcapítulo trata-se todo o desenvolvimento do modelo de previsão para a temperatura, bem como uma análise exploratória da série temporal.

##### 4.1.1.1. Análise exploratória da série temporal

A partir dos dados históricos previamente referidos foi construída a série de médias da temperatura ilustrada na figura 4.1.

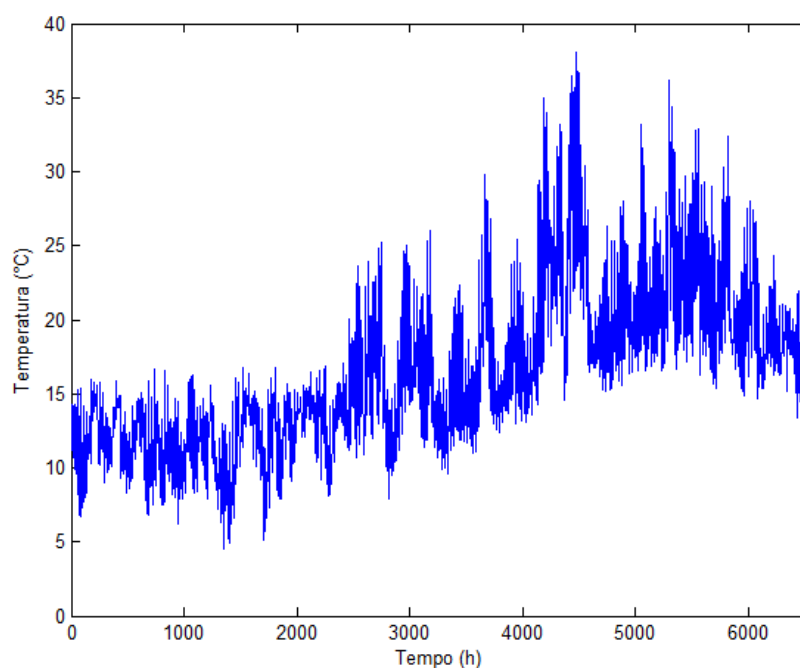


Figura 4.1 - Série temporal da temperatura

De seguida, na tabela 4.1 são apresentadas algumas estatísticas descritivas dos registos da temperatura medida no local em estudo.

Tabela 4.1 - Estatísticas da série temporal da temperatura

<b>Número de amostras</b>	26179
<b>Média (m/s)</b>	16,4239
<b>Mediana (m/s)</b>	15,3222
<b>Desvio-padrão (m/s)</b>	5,3489
<b>Mínima (m/s)</b>	4,7267
<b>Máxima (m/s)</b>	37,2932

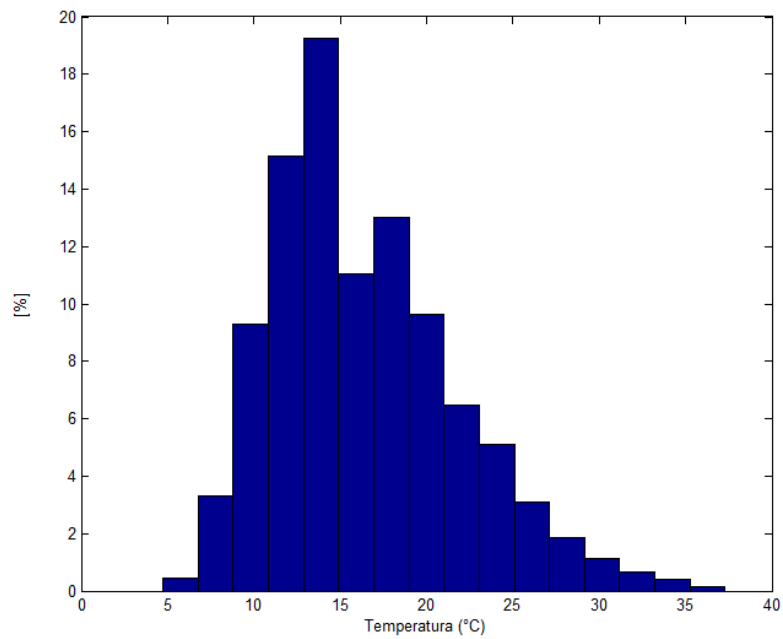


Figura 4.2 - Histograma dos dados recolhidos da temperatura

Na figura 4.2, o histograma da distribuição de frequências da temperatura no DEE revela que na maior parte do tempo a temperatura não ultrapassa os 25 °C. Adicionalmente, o Q-QPlot na figura 4.3 sugere que os registos da temperatura não seguem uma distribuição normal, resultado este confirmado por meio do *MatLAB* através do teste de *Kolmogorov-Smirnov*, uma vez que este rejeita a hipótese nula de que os dados em estudo tenham distribuição normal.

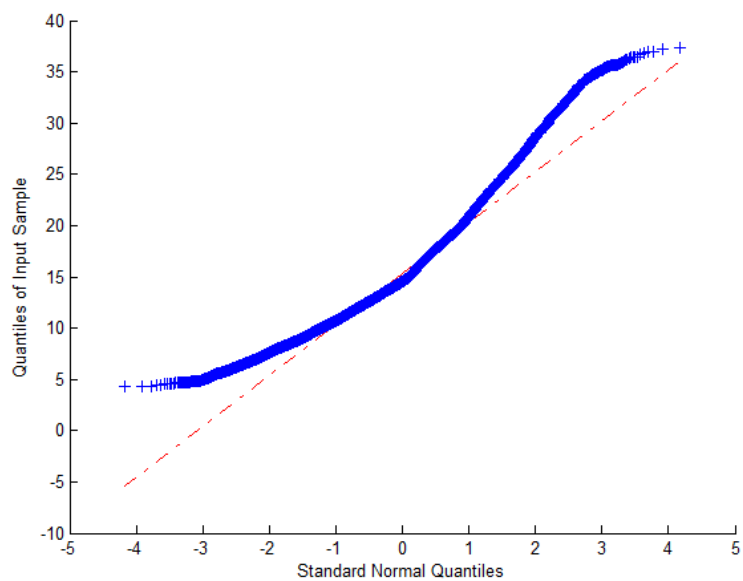


Figura 4.3 - Q-QPlot dos dados recolhidos da temperatura

#### 4.1.1.2. Identificação do modelo

Na figura 4.4 está apresentada a função de autocorrelação (FAC) dos registos da temperatura, como é possível observar esta exibe picos localizados em intervalos de 96 amostras (um dia, visto que o intervalo entre amostras é de 15 minutos) que decaem lentamente, portanto, a série temporal não é estacionária e possui sazonalidade.

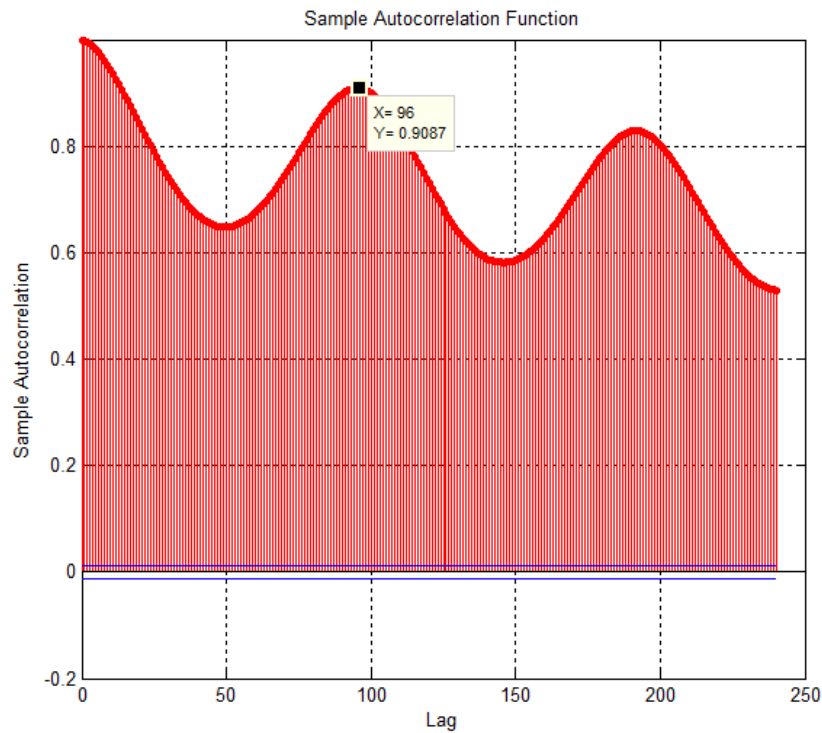


Figura 4.4 - FAC da série temporal da temperatura

Como foi referido anteriormente, para poder ser possível aplicar a metodologia de *Box & Jenkins*, a série temporal em estudo tem obrigatoriamente de apresentar estacionaridade, logo, é necessário realizar uma transformação no processo para lhe conceder essa característica. Para tal, os dados recolhidos foram alvo de uma diferenciação e de seguida foi verificado que se a série temporal apresentava estacionaridade através da aplicação do teste de *Dickey-Fuller*, que rejeitou a hipótese nula de a série possuir uma raiz unitária. Posto isto, é necessário voltar a analisar o gráfico relativo à FAC da série temporal, agora após uma diferenciação (figura 4.5).

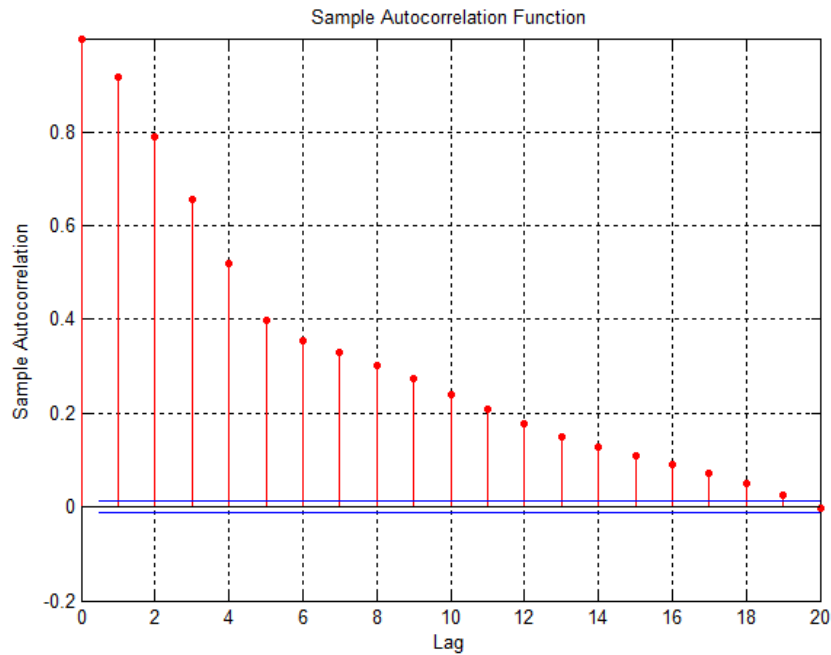


Figura 4.5 - FAC da série temporal da temperatura após uma diferenciação

Como é possível observar, o gráfico da função de autocorrelação obtido depois da série original ser diferenciada apresenta um decrescimento exponencial, esta característica é própria de um processo autorregressivo de ordem igual a 1. É de ressaltar que embora na figura 4.5 o número de *lags* tenha sido reduzido por forma a observar mais detalhadamente o comportamento da função de autocorrelação, o termo autorregressivo sazonal permanece após a diferenciação da série, tendo de ser introduzido no modelo a desenvolver.

#### 4.1.1.3. Desenvolvimento do modelo

Em função das características da série relatadas anteriormente, propõe-se uma modelização estatística baseada num modelo ARIMA com componente sazonal de 24 horas de ordem (1,1,0). Após a estimação do modelo, é necessário fazer um diagnóstico do mesmo através da análise dos seus resíduos, como pode ser observado na figura 4.6.

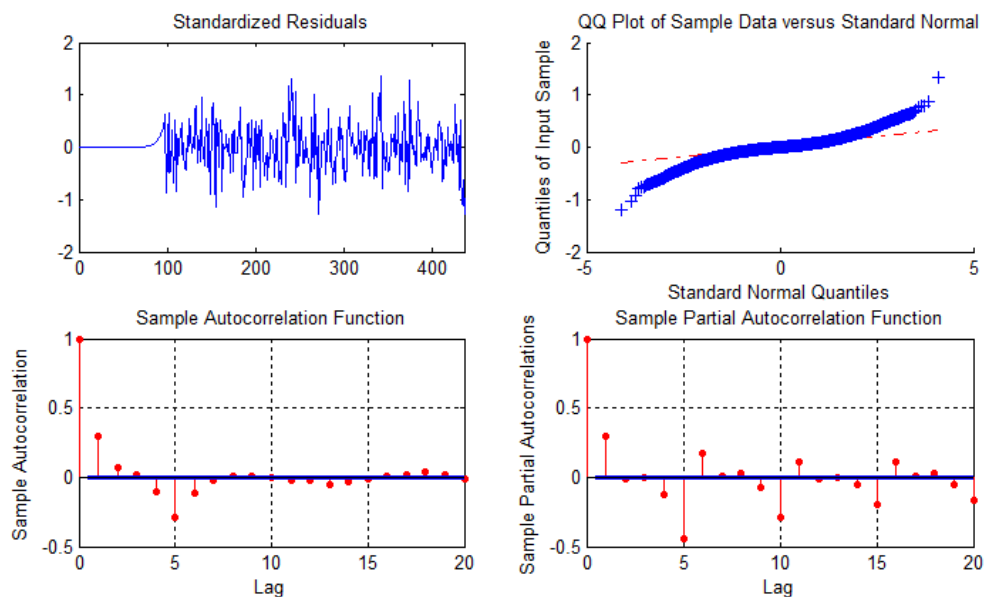


Figura 4.6 - 4-plot dos resíduos do modelo ARIMA (1,1,0)

Através da análise do 4-plot dos resíduos obtidos utilizando o modelo ARIMA (1,1,0), é possível concluir que o modelo proposto é adequado para os dados introduzidos, uma vez que os resíduos não apresentam correlações significativas entre amostras e possuem uma média nula, o facto de os resíduos não estarem normalmente distribuídos, como foi referido anteriormente, não se apresenta como uma característica de rejeição do modelo.

Uma vez definido o modelo, o passo seguinte passa por gerar a previsão através da utilização do mesmo. A previsão é realizada para um horizonte temporal de 48 horas e recai sobre os últimos 2 dias do mês de Outubro, sendo os resultados obtidos comparados com os dados recolhidos nesse mesmo período (figura 4.7).

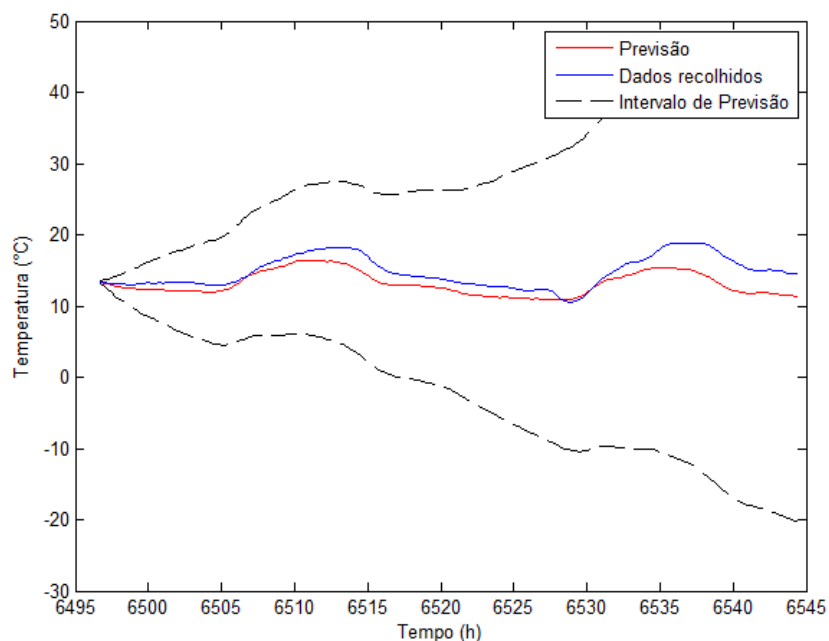


Figura 4.7 - Previsão da temperatura até 48 horas à frente através do modelo ARIMA (1,1,0)

Apesar dos resultados obtidos com o modelo (1,1,0) serem satisfatórios, outros modelos foram testados, aumentando a ordem do processo autorregressivo, concluindo-se que, quanto maior a ordem do processo, maior correlação existe entre os resíduos e, conseqüentemente, resultando numa pior adequabilidade do modelo.

Como tal, foi então colocada a hipótese de o modelo ideal ser um modelo misto, visto que este tipo de modelos não é reconhecido através da observação da série temporal ou das suas funções de correlação, foram realizados testes estatísticos por forma a verificar o desempenho do modelo misto na previsão da temperatura, como pode ser observado na tabela 4.2 pelas estatísticas de erro médio absoluto (MAPE) e pela raiz do erro quadrático médio (RMSE).

Tabela 4.2 - Resultados de previsão para diferentes horizontes temporais

Modelo	Teste	Até 3h	Até 6h	Até 12h	Até 24h	Até 48h
ARIMA (1,1,0) com sazonalidade	RMSE (°C)	0,21	0,68	0,66	1,83	4,09
	MAPE (%)	3,1	5,6	5,4	7,8	10,8
ARIMA (1,1,1) com sazonalidade	RMSE (°C)	<u>0,21</u>	<u>0,65</u>	<u>0,59</u>	<u>1,77</u>	<u>3,93</u>
	MAPE (%)	<u>3,1</u>	<u>5,4</u>	<u>5,1</u>	<u>7,6</u>	<u>10,5</u>

ARIMA (1,1,2) com sazonalidade	RMSE (°C)	0,22	0,68	0,65	1,82	4,07
	MAPE (%)	3,2	5,5	5,4	7,8	10,7

Através dos resultados da análise estatística realizada às previsões obtidas pelos diferentes modelos, é possível afirmar que o modelo que mais se ajusta aos dados em estudo é um modelo ARIMA (1,1,1) com sazonalidade de 24 horas. Este modelo apresenta um desempenho positivo na previsão a curto tempo da temperatura, os resultados obtidos para o MAPE e para a RMSE revelam que o modelo apresenta um melhor resultado quanto menor for o horizonte de previsão, sendo que para valores superiores a 24 horas este começa a ser pouco fiável. Na figura 4.8 está então representada a previsão da temperatura para 48 horas em intervalos de 15 minutos através da utilização de um modelo ARIMA (1,1,1).

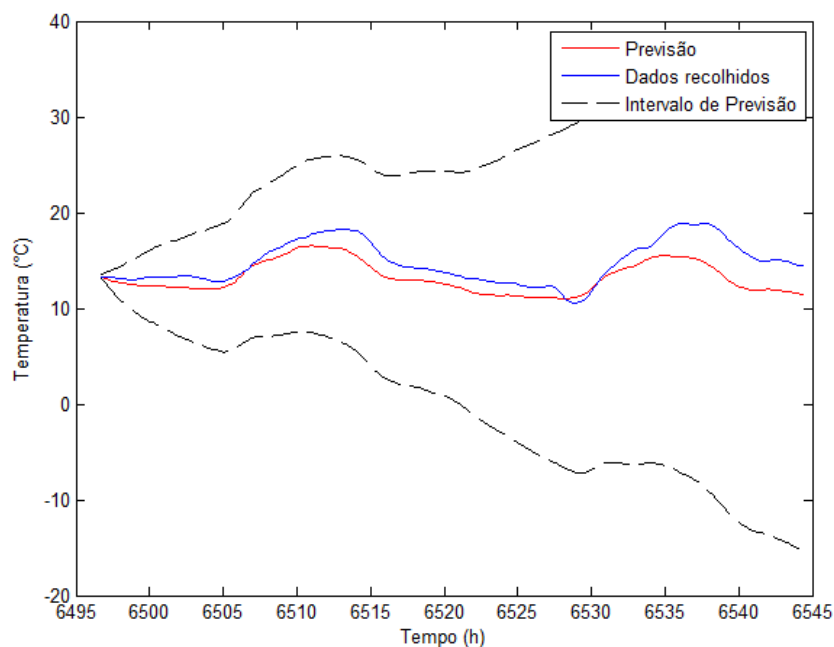


Figura 4.8 - Previsão da temperatura até 48 horas à frente através do modelo ARIMA (1,1,1)

#### 4.1.2. Modelo de previsão para a irradiância

Neste subcapítulo trata-se todo o desenvolvimento do modelo de previsão para a irradiância, bem como uma análise exploratória da série temporal.

#### 4.1.2.1. Análise exploratória da série temporal

A partir dos dados históricos previamente referidos foi construída a série de médias da irradiância ilustrada na figura 4.9.

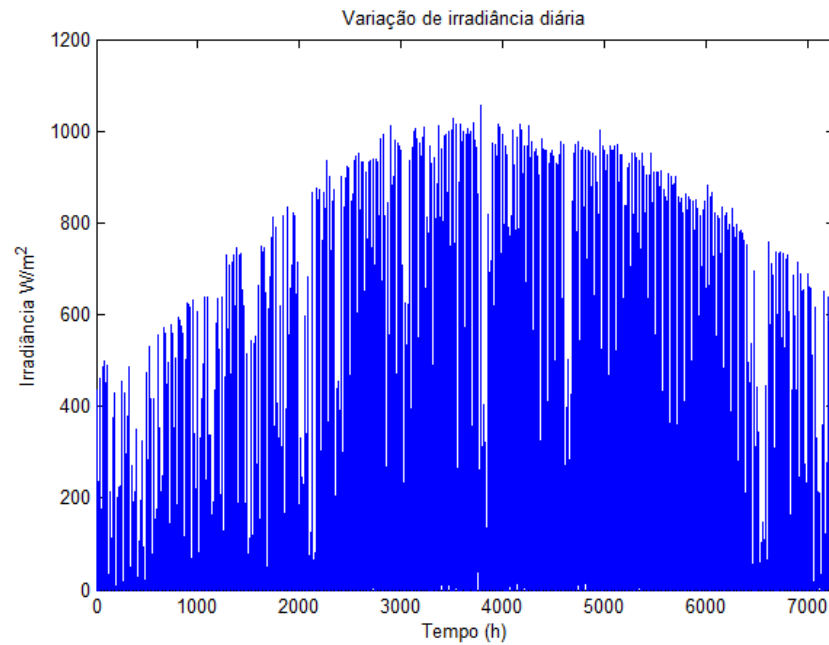


Figura 4.9 - Série temporal da irradiância

Na tabela 4.3, estão presentes as estatísticas descritivas da série temporal da irradiância.

Tabela 4.3 - Estatísticas da série temporal da irradiância

<b>Número de amostras</b>	26179
<b>Média (m/s)</b>	221,1530
<b>Mediana (m/s)</b>	27,9297
<b>Desvio-padrão (m/s)</b>	300,8523
<b>Mínima (m/s)</b>	0
<b>Máxima (m/s)</b>	1055,1210

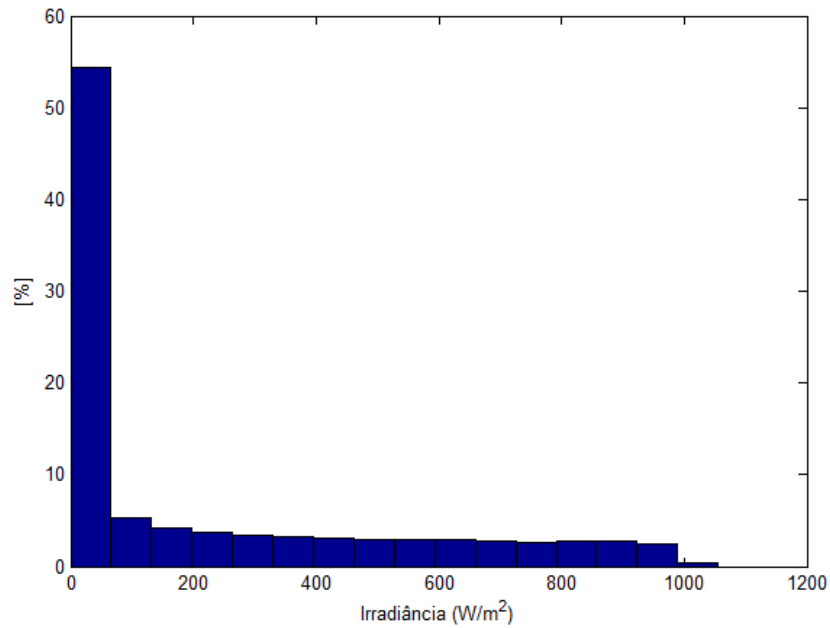


Figura 4.10 - Histograma dos dados recolhidos da irradiância

Como seria expectável, o histograma da distribuição de frequência da irradiância (figura 4.10) concentra-se mais quanto menor o seu valor, visto que a irradiância é menor quanto mais próxima do período noturno e nula durante este. O Q-QPlot da série (figura 4.11) demonstra claramente que esta série não possui uma distribuição probabilística normal.

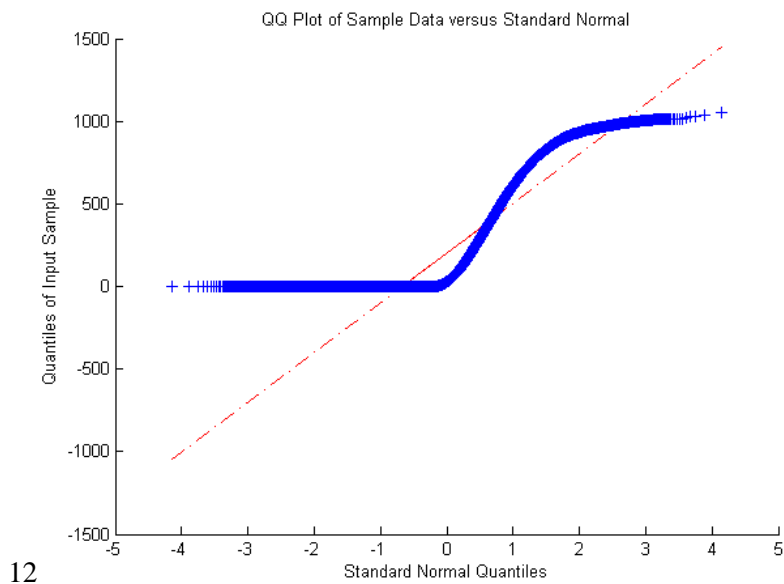


Figura 4.11 - Q-Qplot dos dados recolhidos da irradiância

#### 4.1.2.2. Identificação do modelo

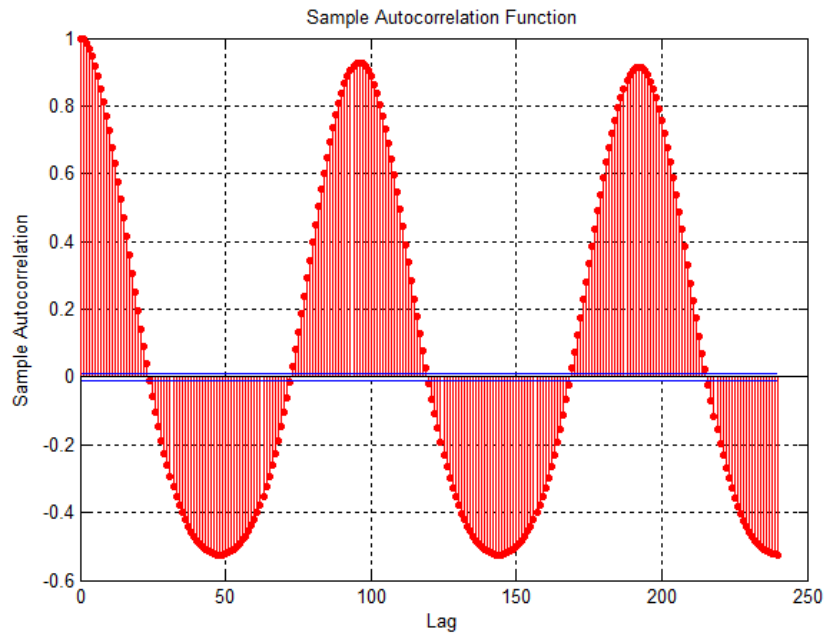


Figura 4.12 - FAC da série temporal da irradiância

A função de autocorrelação (FAC) dos registos da irradiância (figura 4.12) exhibe picos localizados em intervalos de 96 amostras (um dia, visto que o intervalo entre amostras é de 15 minutos) que decaem rapidamente, indicando que a série é estacionária, o comportamento sinusoidal que se pode observar é devido ao termo sazonal autorregressivo, fazendo com que existam correlações de elevado valor a cada 96 amostras (24 horas), de qualquer forma, para confirmar a estacionaridade da série foi realizado o teste de Dickey Fuller aumentado, sendo que este indicou que se deve rejeitar a hipótese nula de o processo possuir uma raiz unitária, concluindo-se então que a série é estacionaria e, como tal, não é necessário realizar qualquer diferenciação, implicando que o parâmetro  $d$  do modelo a desenvolver será igual a 0.

Através da observação do comportamento da função de autocorrelação da série temporal, é possível concluir que a seleção do modelo não recairá sobre um processo autorregressivo (AR), uma vez que a série não apresenta um decaimento exponencial para zero nem alterna entre valores positivos e negativos, decaindo para zero também. A hipótese do modelo ideal ser um processo de média móvel (MA) também é excluído visto que o comportamento da FAC não apresenta apenas alguns picos com valores elevados e os restantes pontos praticamente nulos. Como tal, conclui-se que o modelo a aplicar para a previsão da irradiância deve ser um modelo misto, o facto da FAC da série decair rapidamente após poucas *lags* também apoia esta decisão.

#### 4.1.2.3. Desenvolvimento do modelo

Como não é possível definir os parâmetros  $p$  e  $q$  de um modelo misto através da observação das suas funções de autocorrelação e autocorrelação parcial, estes terão de ser estimados através de tentativa-erro, isto é, é definido um modelo misto, neste caso um modelo ARIMA (1,0,1) com sazonalidade de 24 horas, realizando uma análise aos resíduos e dos erros associados à previsão realizada através desse modelo e posteriormente ajustando os valores dos parâmetros caso isso resulte numa melhor adequabilidade do modelo. Na figura 4.13 está então representada o 4-plot dos resíduos obtidos utilizando este modelo.

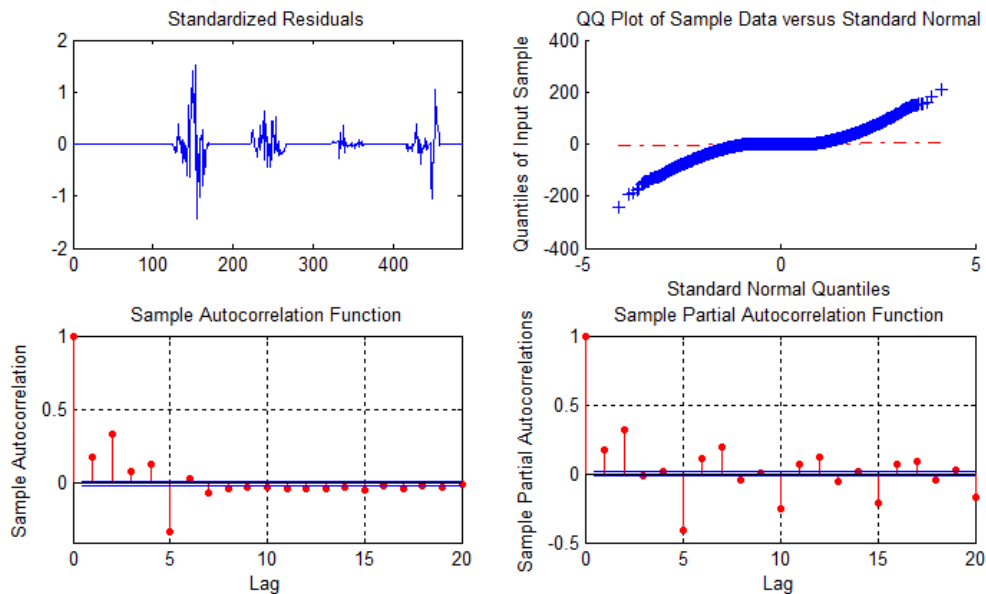


Figura 4.13 - 4-plot dos resíduos do modelo ARIMA (1,0,1)

Através da análise do 4-plot dos resíduos obtidos utilizando o modelo misto proposto, é possível concluir que o mesmo é adequado para os dados introduzidos, uma vez que os resíduos não apresentam correlações significativas entre amostras e possuem uma média nula, o facto de não existir uma distribuição normal dos resíduos não se apresenta como uma característica de rejeição do modelo.

Após concluir que um modelo misto se apresenta adequado para a previsão da série temporal em estudo, é necessário realizar a previsão utilizando o mesmo e fazer o cálculo dos erros associados a essa previsão para que se possa comparar os resultados obtidos através da utilização do modelo ARIMA (1,0,1) com outros modelos, variando apenas os parâmetros  $p$  e  $q$ . Na figura 4.14 está então representada a previsão da irradiância, para um horizonte temporal de 48 horas.

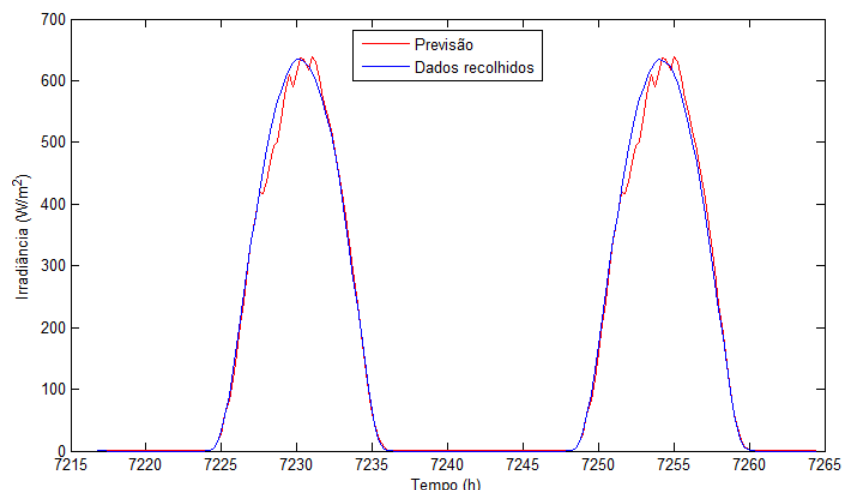


Figura 4.14 - Previsão da temperatura até 48 horas à frente através do modelo ARIMA (1,0,1)

Similarmente com o que ocorreu com o modelo de previsão da temperatura, ao aumentar o termo autorregressivo (AR) do modelo de previsão da irradiação a autocorrelação entre os seus resíduos aumenta significativamente, excluindo dessa forma qualquer processo AR superior a 1. Foram então aplicados testes estatísticos aos resultados previstos em comparação com os dados reais (recolhidos) para o modelo acima proposto, tal como para modelos com ajustes no parâmetro  $q$  por forma a determinar qual o melhor modelo para a previsão da irradiação. Na tabela 4.4 estão presentes esses resultados através das estatísticas de erro médio absoluto (MAPE) e de raiz do erro quadrático médio (RMSE).

Tabela 4.4 - Resultados de previsão para diferentes horizontes temporais

<b>Modelo</b>	<b>Teste</b>	<b>Até 12h</b>	<b>Até 24h</b>	<b>Até 48h</b>
<u>ARIMA (1,0,1) com sazonalidade</u>	<u>RMSE (°C)</u>	<u>211,23</u>	<u>226,35</u>	<u>257,84</u>
	<u>MAPE (%)</u>	<u>7,5</u>	<u>7,1</u>	<u>8,1</u>
ARIMA (1,0,2) com sazonalidade	RMSE (°C)	211,28	226,38	257,88
	MAPE (%)	10,7	8,6	8,7
ARIMA (1,0,3) com sazonalidade	RMSE (°C)	211,43	226,47	257,96
	MAPE (%)	8,7	7,5	8,6

Como pode ser observado, variações no processo de média móvel não só têm um pequeno impacto no desempenho do modelo de previsão, como diminuem a adequabilidade do mesmo. Os

elevados valores de RMSE variam dependendo dos dados históricos, que apresentam grandes variações nos valores da irradiação (possivelmente devido a nuvens ou a mau funcionamento do sensor) mas permanecem semelhantes independentemente do ajuste do termo de média móvel. Conclui-se então que o modelo ARIMA (1,0,1) com sazonalidade de 24 horas é o modelo ideal para previsão a curto tempo da irradiância.

#### 4.1.3. Modelo de previsão para a velocidade do vento

Neste subcapítulo trata-se todo o desenvolvimento do modelo de previsão para a velocidade do vento, bem como uma análise exploratória da série temporal.

##### 4.1.3.1. Análise exploratória da série temporal

A figura 4.15 representa a série de médias da velocidade do vento construída a partir dos dados recolhidos.

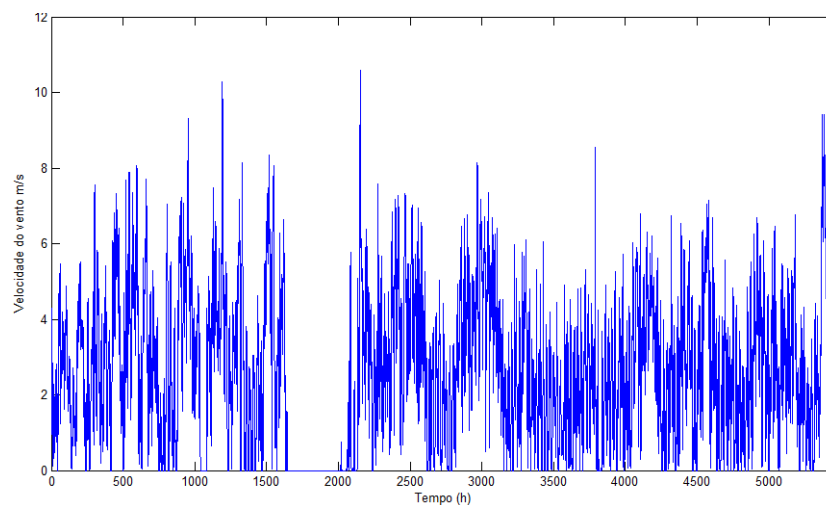


Figura 4.15 - Série temporal da velocidade do vento (10 meses)

Como é possível observar pela figura, todas as amostras entre a hora 1644 até à hora 2008 apresentam valor nulo, significa isto que existiu algum problema no anemómetro da estação meteorológica e a inclusão destes valores no modelo seria prejudicial para o mesmo, como tal, para o modelo de previsão da velocidade do vento, não é possível considerar o mesmo período de treino que foi considerado para a irradiância e para a temperatura, sendo utilizado então os dados

recolhidos desde junho e setembro de 2013, obtendo uma nova série de médias da velocidade do vento, que pode ser observada na figura 4.16.

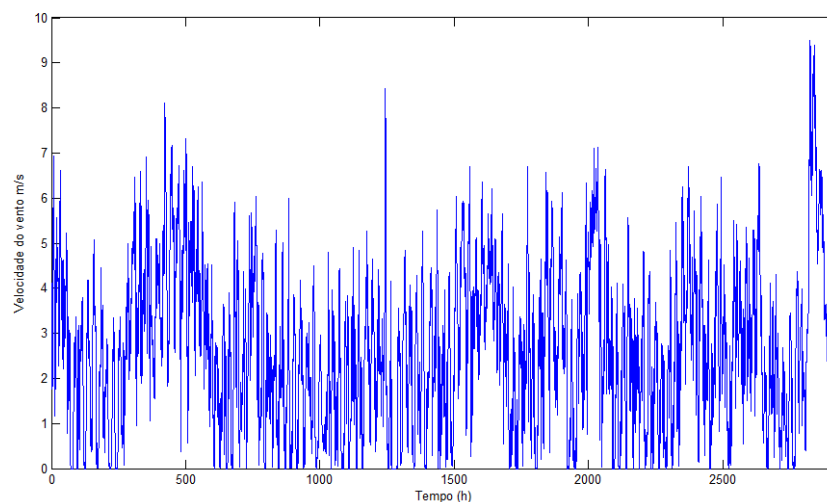


Figura 4.16 - Série temporal da velocidade do vento (4 meses)

Na tabela 4.5 são apresentadas algumas estatísticas descritivas dos registros da temperatura medida no local em estudo.

Tabela 4.5 - Estatísticas da série temporal da velocidade do vento

<b>Número de amostras</b>	11715
<b>Média (m/s)</b>	2,7016
<b>Mediana (m/s)</b>	2,6045
<b>Desvio-padrão (m/s)</b>	1,7490
<b>Mínima (m/s)</b>	0
<b>Máxima (m/s)</b>	9,4887

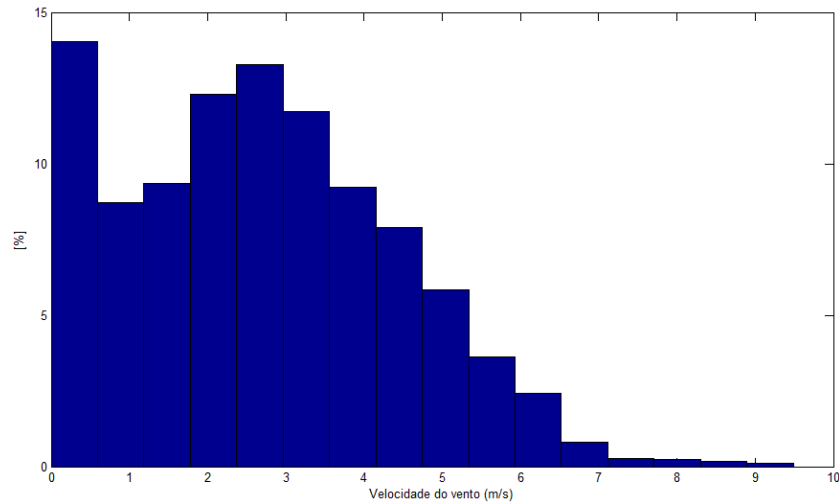


Figura 4.17- Histograma dos dados recolhidos da velocidade do vento

Na figura 4.17, o histograma da distribuição de frequências da temperatura no DEE revela que na maior parte do tempo a velocidade não ultrapassa os 3 m/s, este fator revela que na maior parte do tempo o aerogerador não irá produzir qualquer energia uma vez que, como foi referido anteriormente, este só começa a produzir para velocidades iguais ou superiores a 3 m/s. Adicionalmente, o Q-QPlot na figura 4.18 sugere que os registos da temperatura não seguem uma distribuição normal, resultado este confirmado por meio do *MatLAB* através do teste de *Kolmogorov-Smirnov*, uma vez que este rejeita a hipótese de que os dados apresentem uma distribuição normal.

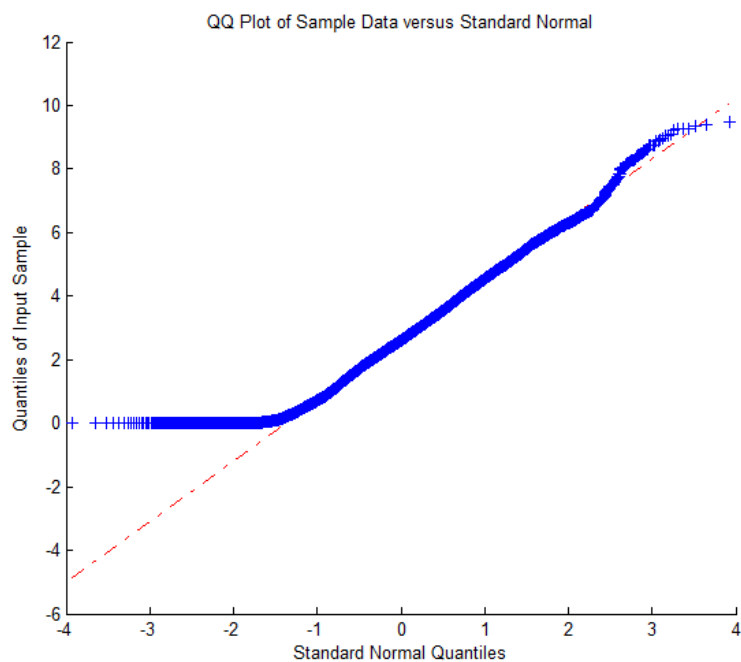


Figura 4.18 - Q-QPlot da velocidade do vento

#### 4.1.3.2. Identificação do modelo

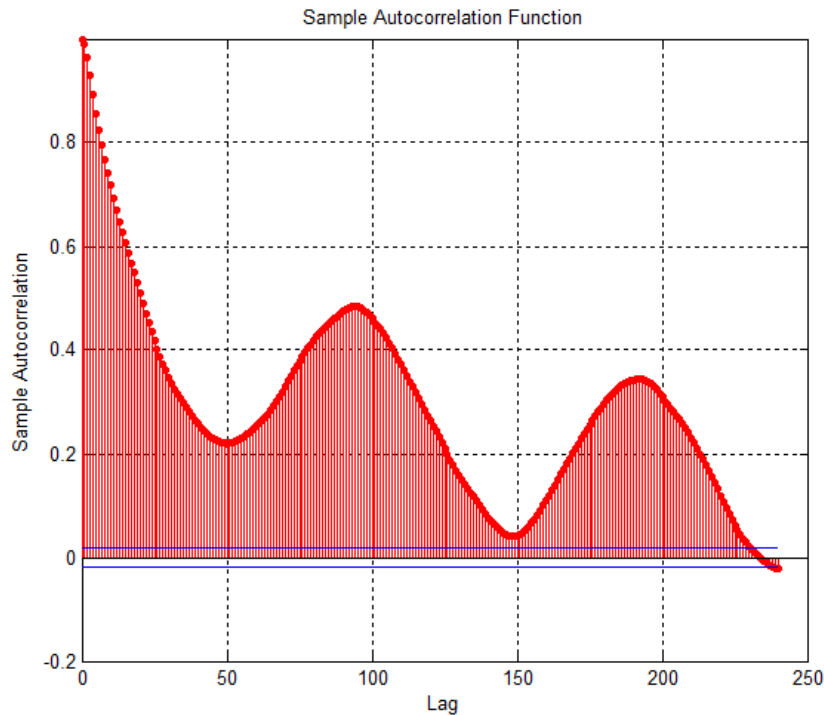


Figura 4.19 - FAC da série temporal da velocidade do vento (sem diferenciação)

Como é possível observar na figura 4.19, a função de autocorrelação apresenta um decaimento lento para zero, o que sugere que a série não apresenta estacionaridade. Uma vez que um modelo estatístico apenas pode ser aplicado quando na presença de uma série estacionária, os dados recolhidos são diferenciados.

Analisando a figura 4.20, é possível observar que a série diferenciada, através da sua FAC, já aparenta possuir estacionaridade, uma vez que esta possui um padrão de decaimento rápido. A sua estacionaridade foi confirmada através do teste de *Dickey-Fuller*, que rejeitou a hipótese da série possuir uma raiz unitária, desta forma, como a estacionaridade é alcançada após uma diferenciação, o parâmetro  $d$ , do modelo a desenvolver, terá o valor 1.

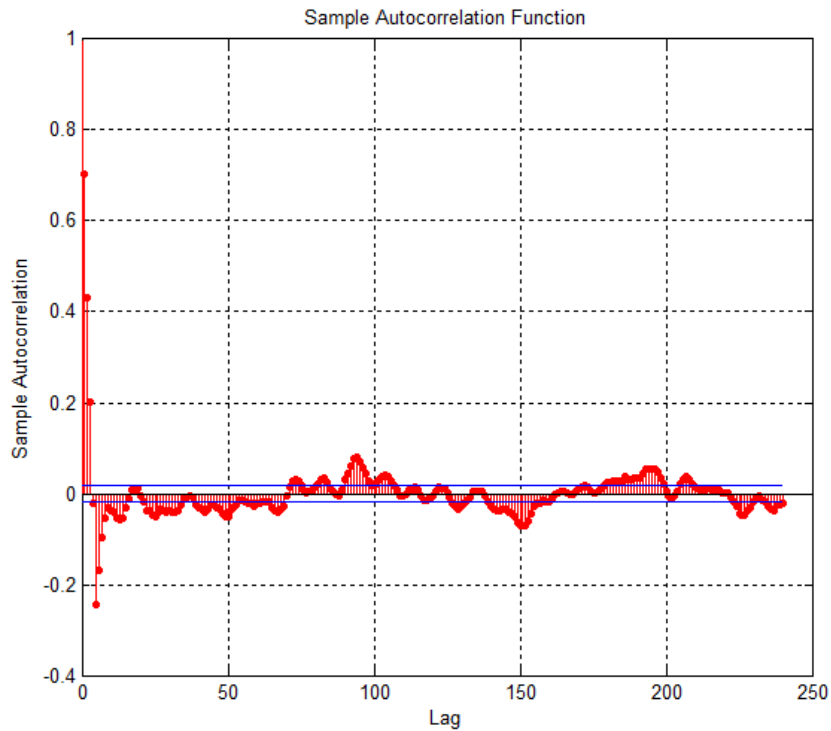


Figura 4.20 - FAC da série temporal da velocidade do vento (após diferenciação)

A função de autocorrelação do processo apresenta uma queda rápida para zero após poucas *lags*, o que indica que provavelmente o modelo que melhor se aplica aos dados da velocidade do vento será um modelo misto.

#### 4.1.3.3. Desenvolvimento do modelo

Tal como foi referido anteriormente, não é possível definir os parâmetros  $p$  e  $q$  de um misto através da observação das suas funções de autocorrelação e autocorrelação parcial. Como tal, é tomado como base de testes um modelo ARIMA (1,1,1), com sazonalidade de 24 horas, e realizando uma análise aos seus resíduos (figura 4.21).

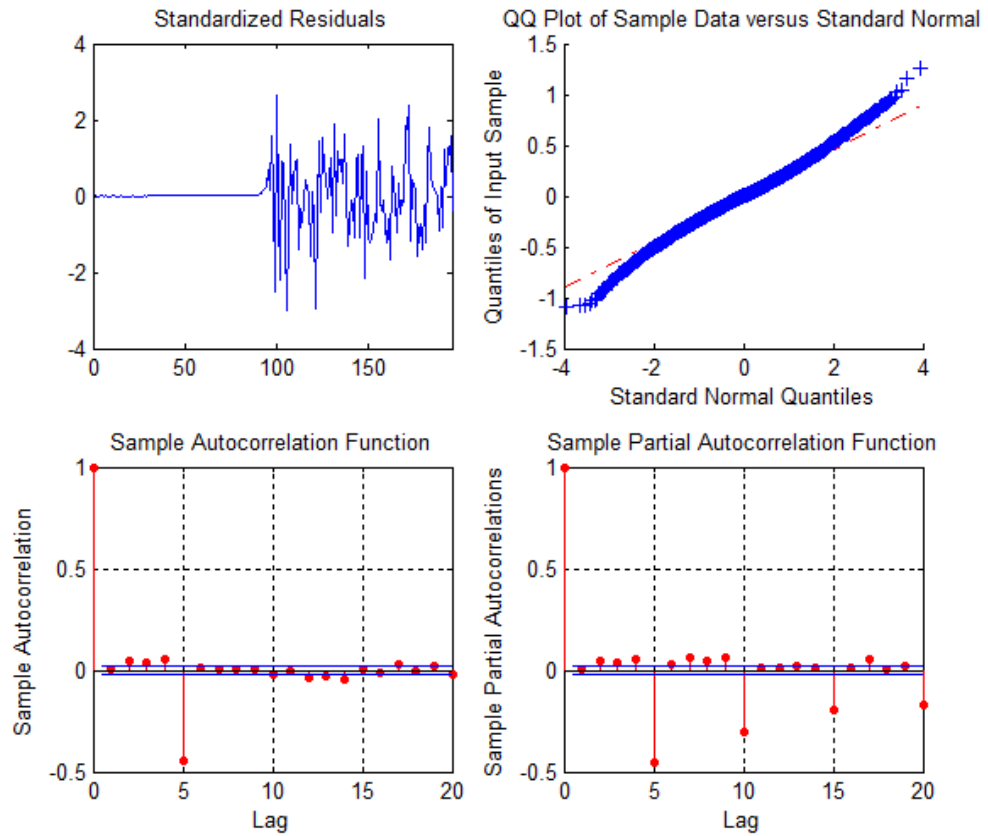


Figura 4.21 - 4-Plot dos resíduos da velocidade do vento

Como é possível observar na figura 4.21, o modelo proposto apresenta uma elevada adequabilidade nas primeiras 100 *lags*, com os valores dos resíduos muito próximos de zero. O modelo não apresenta também nenhuma correlação significativa entre amostras e possui uma distribuição normal, o que garante que o modelo proposto possui um elevado desempenho e está pronto para a fase de previsão.

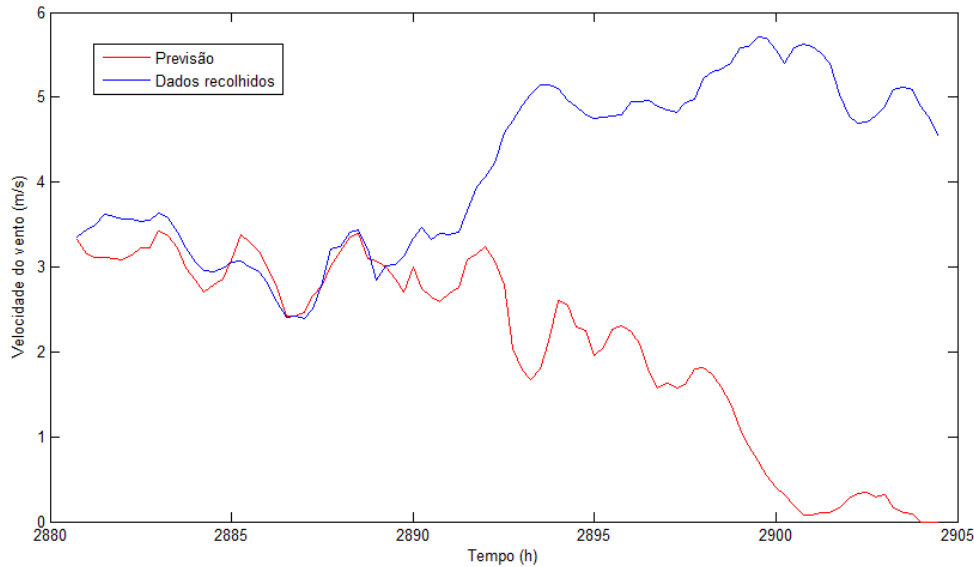


Figura 4.22 - Previsão da velocidade do vento até 24 horas à frente através do modelo ARIMA (1,1,1)

Na figura 4.22 está representada a previsão da velocidade do vento num horizonte temporal de 24 horas. Ao contrário do que foi observado nas séries temporais da irradiância e da temperatura, a velocidade do vento apresenta um comportamento muito mais irregular, tornando-se difícil realizar previsões com horizontes temporais superiores a 24 horas. Como é possível observar, o modelo proposto apresenta um elevado desempenho nas primeiras 10 horas de previsão quando comparado com os dados reais, após este período, o comportamento da previsão toma valores que não demonstram qualquer relação com os valores reais, uma vez que decaem para zero. Apesar deste fator, tendo em conta a elevada imprevisibilidade do vento, considera-se que os resultados obtidos são satisfatórios, ainda assim, foram testados outros modelos através do ajuste dos parâmetros  $p$  e  $q$ , e comparados com os resultados obtidos com o modelo ARIMA (1,1,1). Os resultados obtidos estão presentes na tabela 4.6.

Tabela 4.6 - Resultados de previsão para diferentes horizontes temporais

Modelo	Teste	Até 6h	Até 12h	Até 24h
ARIMA (1,1,1) com sazonalidade	RMSE (m/s)	0,0824	0,2316	8,4116
	MAPE (%)	7,6	9,8	43,5
ARIMA (2,1,1) com sazonalidade	RMSE (m/s)	0,1297	0,3261	9,0795
	MAPE (%)	9,3	12,4	46,3

ARIMA (3,1,1) com sazonalidade	RMSE (m/s)	0,0889	0,1570	7,0833
	MAPE (%)	7,6	9,4	39,9
<u>ARIMA (4,1,1) com sazonalidade</u>	<u>RMSE (m/s)</u>	<u>0,0671</u>	<u>0,1589</u>	<u>7,4625</u>
	<u>MAPE (%)</u>	<u>6,4</u>	<u>8,7</u>	<u>40,61</u>
ARIMA (1,1,2) com sazonalidade	RMSE (m/s)	0,1324	0,3305	9,1024
	MAPE (%)	9,4	12,51	46,41
ARIMA (2,1,2) com sazonalidade	RMSE (m/s)	0,0662	0,1771	7,8220
	MAPE (%)	6,6422	8,8	41,56
ARIMA (3,1,2) com sazonalidade	RMSE (m/s)	0,0683	0,1916	8,02
	MAPE (%)	6,96	9,0	42,1527
ARIMA (3,1,3) com sazonalidade	RMSE (m/s)	0,0759	0,1575	7,34
	MAPE (%)	6,97	9,0	40,4437
ARIMA (3,1,4) com sazonalidade	RMSE (m/s)	0,1112	0,1650	6,86
	MAPE (%)	8,7419	10,17	39,6954

Como se pode observar na tabela 4.6, após sucessivos ajustes nos termos autorregressivos e de média móvel, conclui-se que os modelos mistos que apresentam melhor adequabilidade para a realização da previsão da velocidade do vento é um modelo ARIMA (3,1,1) e o modelo ARIMA (4,1,1), ambos com sazonalidade de 24 horas. O aumento do parâmetro de média móvel não resultou em qualquer melhoria na previsão da velocidade do vento. Conclui-se também que só faz sentido considerar a previsão até 12 horas, uma vez que os valores após esse instante não revelam qualquer proximidade com os valores reais.

Por forma a calcular a previsão da potência gerada através do gerador eólico, a previsão da velocidade do vento será realizada então através do uso do modelo ARIMA (4,1,1) com sazonalidade de 24 horas. Na figura 4.23 está representada a previsão obtida utilizando este modelo.

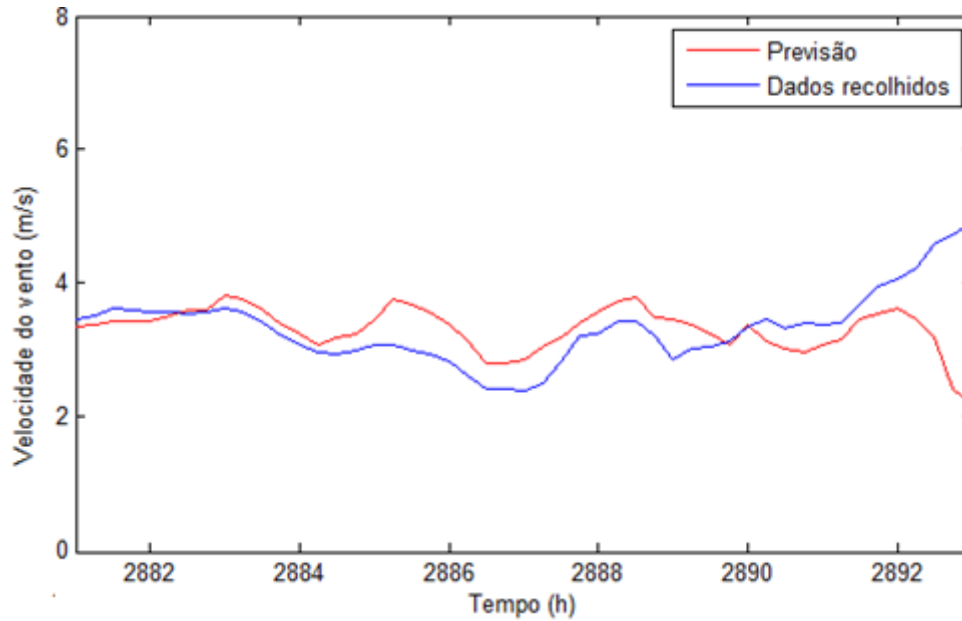


Figura 4.23 - Previsão da velocidade do vento até 12 horas à frente através do modelo ARIMA (4,1,1)

## 4.2. Previsão de Potência

Após a obtenção dos dados relativos à previsão da velocidade do vento, da temperatura e da irradiância é necessário converter esses registos em potência produzida. Para isso são utilizadas duas metodologias diferentes, uma para a produção a partir do gerador eólico e outra para a produção a partir dos painéis fotovoltaicos.

Para o gerador eólico, é realizada uma estimativa baseada na curva de potência do aerogerador. Cada turbina tem uma curva característica associada que descreve o processo ideal de conversão da energia cinética da velocidade do vento para potência elétrica produzida. Esta curva é fornecida pelo fabricante.

Já para os painéis fotovoltaicos, os dados referentes às previsões da temperatura e da irradiância são convertidos para potência elétrica através da aplicação de várias equações que descrevem o comportamento geral de um painel fotovoltaico, estas equações tomam em conta as características gerais da instalação fotovoltaica que são fornecidas pelo fabricante.

### 4.2.1. Painéis fotovoltaicos

Para obter a potência elétrica produzida por um painel fotovoltaico unicamente a partir de registos de temperatura ambiente ( $T_{ambiente}$ ) e irradiância ( $G$ ) é necessário realizar alguns

cálculos intermédios de forma a determinar a temperatura das células do painel e a potência de referência (potência máxima) que um painel pode produzir em determinado instante. Como tal, estas variáveis são calculadas através das equações 4.1, 4.2 e 4.3 [17]:

- Temperatura das células,  $T_{cel}$ :

$$T_{cel} = T_{ambiente} + \left( \frac{NOCT - 20}{0,8} \right) \times G \times 10^{-3} \quad (4.1)$$

- Potência de referência,  $P_{ref}$ :

$$P_{ref} = P_{max} \times \frac{G}{G_{STC}} \quad (4.2)$$

- Potência elétrica produzida (DC),  $P_{elec}$ :

$$P_{elec} = P_{ref} - P_{ref} \times \frac{T_{coef}}{100} \times (T_{cel} - 25^{\circ}C) \quad (4.3)$$

Como é possível observar pelas equações anteriores, para calcular a potência elétrica ideal produzida por um painel fotovoltaico é necessário considerar algumas constantes que são características gerais de cada painel, determinadas após testes em laboratório sob condições STC (*Standard Test Conditions*), isto é, sob uma irradiância de  $1000 \text{ W/m}^2$  e temperatura das células de  $25^{\circ}C$  e que são fornecidas pelo fabricante:

- $NOCT$ , ou *Nominal Operating Cell Temperature* é definida como a temperatura atingida pelas células submetidas a condições STC.
- $T_{coef}$ , ou coeficiente de temperatura do painel, é um número que descreve a forma como o painel fotovoltaico lida com temperaturas superiores a  $25^{\circ}C$ .
- $P_{max}$ , ou potência máxima de pico, é a potência máxima que o painel pode produzir em condições ideais.

Estas características variam entre painéis fotovoltaicos, no caso específico do DEE, como foi referido anteriormente, estão instalados 7 painéis fotovoltaicos de dois tipos diferentes, estando as suas respetivas características definidas na tabela 4.7.

Tabela 4.7 - Características dos painéis para cálculo da potência produzida [13]

	<b>Bangkok Solar – BS40</b>	<b>SWEA – SWL 130</b>
Número de painéis	5	2
<i>NOCT</i>	45 °C	45 °C
<i>T<sub>coef</sub></i>	0,19 %/°C	0,19 %/°C
<i>P<sub>max</sub></i>	40 W	130 W

Uma vez definidos os valores característicos de cada painel, foi calculada a potência elétrica produzida a cada instante, através do cálculo da potência elétrica de cada painel através das equações anteriormente apresentadas e multiplicando esses valores instantâneos pelo número de painéis de cada tipo por forma a determinar a potência elétrica produzida por toda a instalação fotovoltaica (equação 4.4).

$$P_{elec_{instalação}} = (P_{elec_{BS40}} \times 5) + (P_{elec_{SWL130}} \times 2) \quad (4.4)$$

Na figura 4.24 está então representada a previsão de produção da instalação fotovoltaica num horizonte temporal de 48 horas, com intervalos de 15 minutos. Está também representada a curva referente à irradiância de forma a compreender melhor o impacto da mesma na produção de energia.

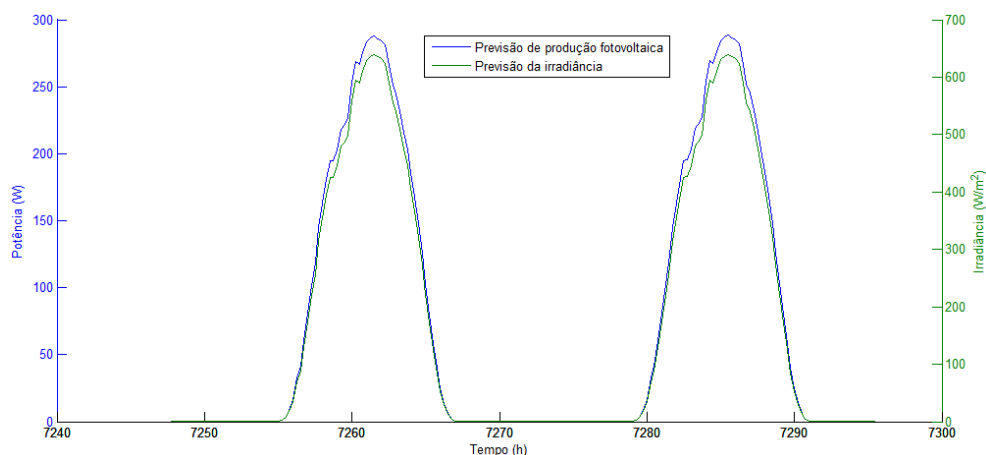


Figura 4.24 - Previsão de potência produzida através dos painéis fotovoltaicos e previsão de irradiância

Como é possível observar na figura 4.24, a produção de potência a partir da instalação fotovoltaica está diretamente relacionada com o valor da irradiância previsto. A influência da temperatura na produção não é igualmente observável uma vez que apenas interfere diretamente com a temperatura das células do painel e para observar a essa influência seria necessário comparar a produção em dois períodos de tempo distintos que possuísem os mesmos valores de irradiação mas valores de temperatura diferentes. Na figura 4.25 encontra-se uma comparação entre os valores de potência previstos pelo modelo e os valores reais recolhidos durante o mesmo período.

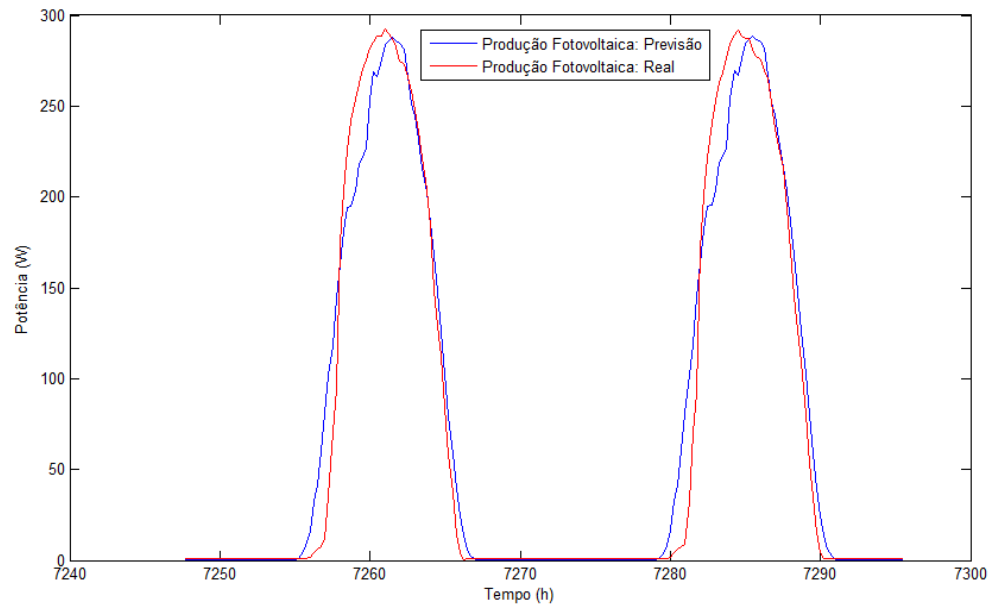


Figura 4.25 - Comparação entre valores previstos e valores reais de produção fotovoltaica

Através da figura é possível constatar que, neste específico período de teste, existe alguma discrepância entre os valores de previsão de produção e os valores reais recolhidos. A diferença entre os períodos iniciais e finais de cada dia não se devem a um mau desempenho do método de previsão (visto que nesses instantes a previsão de irradiância e temperatura apresentou um bom desempenho) mas sim a diferenças entre o desempenho ideal dos painéis fotovoltaicos e o seu desempenho real, possivelmente devido a desgaste das células fotovoltaicas. Apesar da diferença que se pode observar no pico de produção em cada um dos dias, a energia total de ambas as séries temporais apresenta valores muito semelhantes: 14,735 kWh previstos contra 14,365 kWh reais, revelando-se como um fator favorável para a análise do desempenho do modelo desenvolvido.

## 4.2.2. Gerador Eólico

A previsão da produção de potência elétrica a partir do gerador eólico é realizada intersectando os valores previstos de velocidade do vento com a curva de potência fornecida pelo fabricante (figura 4.26).

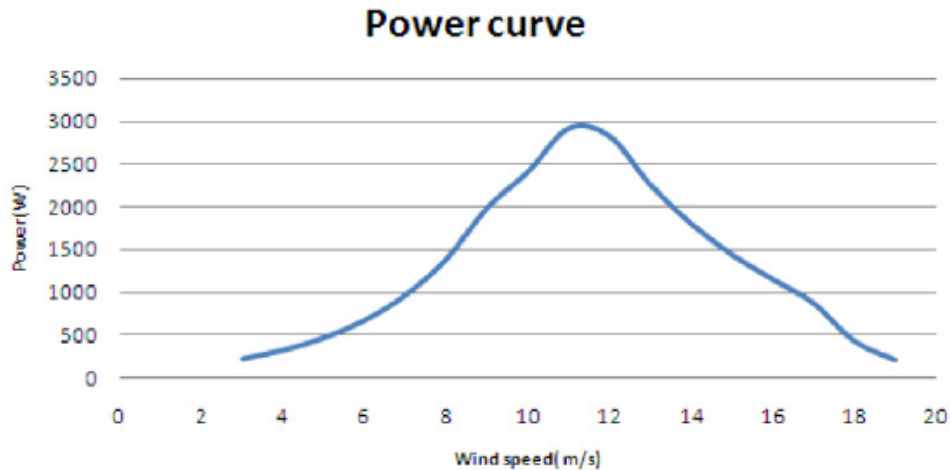


Figura 4.26 - Curva de potência do gerador eólico instalado no todo do edifício do DEE

Para tal, foram retirados alguns pontos da curva de potência do gerador e utilizados na função *cftool* do *MatLAB* de forma a obter a equação do polinómio que melhor represente a curva de potência em questão. Os resultados obtidos indicam que o polinómio que melhor se ajusta possui grau 7, como é possível observar na figura 4.27.

Na figura 4.27 é possível observar os valores que representam a qualidade da aproximação da curva aos pontos introduzidos. O valor de “SSE” corresponde à soma dos quadrados devido a erro, quanto menor o seu valor menor será a componente de erro a si associado, garantindo que existe uma boa aproximação. O valor obtido revela que a curva polinomial tem algum erro a si associada. Por outro lado, os resultados obtidos para o teste *R-Square*, que mede a capacidade da aproximação explicar a variação dos dados, é muito próximo de 1, o que garante que a curva explica cerca de 99% da variação dos dados introduzidos. O resultado obtido para o RMSE indica que o erro médio entre a curva obtida e o valor real é de aproximadamente 114,2W.

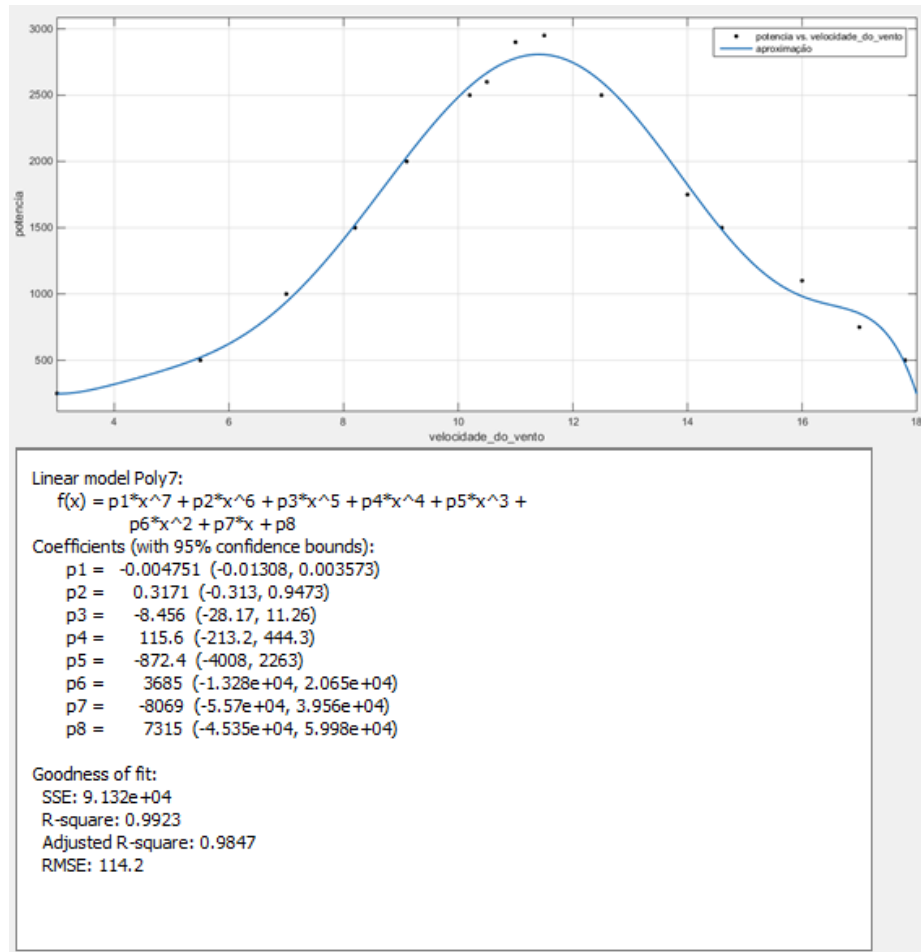


Figura 4.27 - Resultados obtidos através da ferramenta cftool do *MatLAB*

Uma vez obtido o polinómio que melhor representa a curva de potência do gerador, foi calculada a potência gerada tendo em conta os valores da velocidade do vento obtidos através da previsão descrita na secção 4.1.3, não esquecendo as limitações da velocidade de arranque e de corte do gerador enunciados na secção 3.2.2. Na figura 4.28 está então representada a previsão de produção partir do gerador eólico num horizonte temporal de 12 horas com um intervalo entre amostras de 15 minutos.

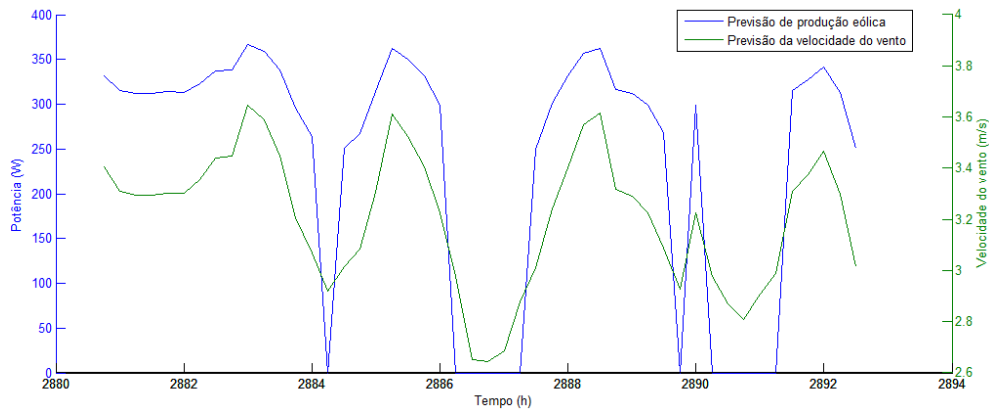


Figura 4.28 - Previsão de potência produzida através do gerador eólico e previsão da velocidade do vento

Como seria expectável a previsão de produção do gerador eólico é proporcional com a velocidade do vento, sendo que, para valores de velocidade inferiores a 3 m/s e produção torna-se inexistente, como se pode constatar na figura 4.28, por exemplo, entre as horas 2886 e 2888. O mesmo ocorre quando a velocidade do vento é superior a 16 m/s, a velocidade de corte do gerador, sendo que não é possível observar este corte graficamente visto que no local em estudo não é previsível obter-se valores de velocidade tão elevados.

Ao contrário do que foi realizado na secção 4.2.1 para a previsão de produção do gerador eólico, não é possível apresentar uma comparação entre os resultados finais da previsão de potência eólica produzida e os valores reais da mesma, uma vez que, como é observável na figura 4.29, a produção do gerador eólico não é de todo coincidente com o que seria expectável ao observar a curva de potência fornecida pelo fabricante.

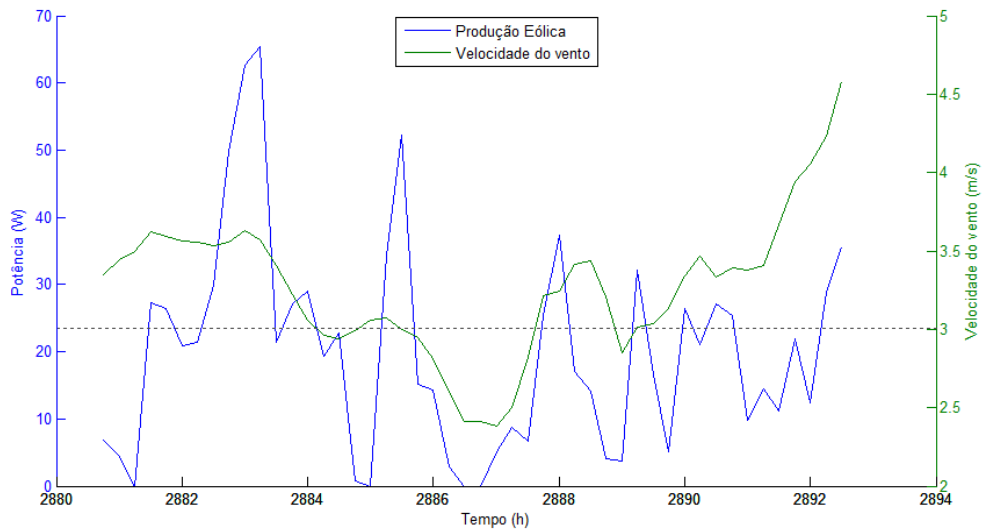


Figura 4.29 - Comparação entre potência produzida (real) pelo gerador eólico e velocidade do vento

Como é possível observar na figura 4.29, a potência produzida pelo gerador eólico não cumpre o apresentado na curva de potência cedida pelo fabricante (figura 4.26). Sendo que a única forma de verificar o desempenho da previsão realizada seria criando, experimentalmente, uma nova curva de potência baseada em dados recolhidos de velocidade do vento e de potência produzida pelo gerador eólico, o que não foi possível realizar durante esta dissertação devido à falta de dados de produção previamente referenciada.



## 5. Conclusões e Trabalhos Futuros

### 5.1. Conclusões

Com o intuito de realizar uma previsão da produção de energia a partir de fontes renováveis instaladas no topo do edifício do Departamento de Engenharia Eletrotécnica (DEE) da Faculdade de Ciências e Tecnologia (FCT), Universidade Nova de Lisboa (UNL), foi desenvolvido um sistema capaz de prever as condições meteorológicas para diferentes horizontes temporais.

O desenvolvimento desta dissertação passou por diversas etapas. Inicialmente foi realizado um estudo prévio dos modelos de previsão existentes e da sua aplicabilidade para o caso específico em estudo, com maior foco nos modelos estatísticos devido ao facto de apenas ser necessário dados históricos recolhidos no local para o desenvolvimento dos modelos. Entre outros, foram estudados modelos como Redes Neuronais Artificiais (RNA), Sistemas de Inferência Difusa (SID) ou a Metodologia de *Box & Jenkins*.

Após o estudo teórico, foi necessário realizar uma análise dos dados recolhidos por forma a perceber quais deles fariam sentido considerar como variáveis de entrada do modelo de previsão a desenvolver, concluindo-se que, para a instalação fotovoltaica faria sentido considerar os dados relativos à temperatura e à irradiância, enquanto no caso do gerador eólico apenas faria sentido utilizar os dados relativos à velocidade do vento, ignorando a direção do vento, uma vez que o aerogerador se alinha com este. Nesta fase foi também decidido que cada uma das variáveis deveria ser tratada de forma individual, isto é, desenvolvendo um modelo de previsão para cada uma das séries temporais a si associadas e, posteriormente, convertendo os valores previstos para potência produzida.

Uma vez realizada a análise dos dados, tornou-se necessário realizar um tratamento dos mesmos de forma a remover os valores extremos e as variações bruscas das séries temporais, uma vez que estas são representativas de aleatoriedade, para tal, foi realizada uma suavização dos dados através de um filtro de média móvel e também uma redução do número de amostras de forma a diminuir o peso computacional do processo de treino dos modelos.

Realizado o tratamento de dados, foram desenvolvidos os modelos de previsão para a temperatura, a irradiância e a velocidade do vento. O processo escolhido foi a metodologia de *Box & Jenkins* dado os bons resultados evidenciados no estudo teórico, a sua simplicidade e não exigirem um grande desempenho computacional no seu processo de treino. Após os ajustes necessários aos modelos obtiveram-se os seguintes resultados:

- Para previsão da temperatura, o modelo que mais se ajustou foi o modelo ARIMA (1,1,1) com sazonalidade de 24 horas. Este modelo demonstrou um elevado desempenho nos testes estatísticos realizados, principalmente durante as primeiras 12 horas. Para horizontes temporais superiores a 24 horas o modelo perde qualidade significativamente;
- Para previsão da irradiância, o modelo com melhor desempenho foi o modelo ARIMA (1,0,1) com sazonalidade de 24 horas. Este modelo demonstrou um desempenho aceitável para horizontes temporais até 48 horas, sendo que por vezes apresenta um erro médio um pouco elevado, muito devido ao facto do possível aparecimento de nebulosidade que afeta drasticamente a irradiância medida e cujo comportamento é difícil de prever;
- Para previsão da velocidade do vento, o modelo que melhor desempenho revelou foi o modelo ARIMA (4,1,1) com sazonalidade de 24 horas. Apesar do comportamento altamente oscilatório da velocidade do vento, o modelo desenvolvido apresenta um elevado desempenho na previsão durante um horizonte temporal de 12 horas, sendo que após este período os resultados previstos perdem qualquer semelhança que tivessem com os resultados reais.

Uma vez obtidos os resultados de previsão para cada uma das variáveis, foi realizado o cálculo da potência produzida utilizando os resultados previstos. No caso da instalação fotovoltaica os valores previstos de irradiância e temperatura ambiente são utilizados nas equações que definem a produção de energia solar a partir dos painéis fotovoltaicos instalados, já a produção a partir do gerador eólico foi calculada cruzando os valores previstos para a velocidade do vento com uma curva polinomial gerada que melhor representa a curva de potência do gerador eólico.

Quando comparada a produção global da instalação fotovoltaica com a produção real obtida no mesmo período de tempo, o modelo apresenta uma previsão de energia produzida muito semelhante com a real apesar de apresentar algumas diferenças no comportamento horário da série temporal, diferenças essas que se devem a diferenças entre o comportamento ideal dos painéis fotovoltaicos e o seu comportamento real. Por outro lado, a comparação entre a produção prevista para o gerador eólico não foi comparada com a produção real uma vez que o equipamento não opera de acordo com a curva de potência especificada pelo fabricante.

O resultado final resume-se então a um sistema em que, introduzidos registos passados relativos a temperatura, irradiância e velocidade do vento, realiza uma previsão da produção de energia a partir de uma instalação fotovoltaica e um gerador eólico instalados no topo do edifício do DEE para um horizonte temporal e um intervalo entre amostras definidos pelo utilizador.

## 5.2. *Trabalho Futuro*

Como trabalho futuro, sugere-se que seja criada uma base de dados onde os dados recolhidos, quer pela estação meteorológica quer pelos contadores, sejam armazenados por forma a serem utilizados para desenvolver melhores métodos de previsão. Sugere-se também que seja realizado o desenvolvimento de um método híbrido, que integre as características dos modelos de *Box & Jenkins* e das redes neuronais artificiais, visto que estas apresentam um melhor desempenho para problemas não lineares. Seria também interessante a integração de um modelo de previsão para a produção de energia em tempo real numa plataforma de monitorização da energia consumida e produzida no edifício do DEE.



## Bibliografia

- [1] A. G. C. de R. Vaz, “Photovoltaic Forecasting with Artificial Neural Networks,” 2014.
- [2] J. D. P. Fernandes, “Previsão de produção de energia elétrica a partir do vento.”
- [3] C. Voyant, M. Muselli, C. Paoli, and M. L. Nivet, “Hybrid methodology for hourly global radiation forecasting in Mediterranean area,” *Renew. Energy*, vol. 53, pp. 1–11, 2013.
- [4] <https://www.ncdc.noaa.gov>, “National Centers For Environmental Information.” .
- [5] I. V. G. da Silva, “Previsão de Vento para Geração de Energia Elétrica.”
- [6] R. J. H. G. Athanasopoulos, “Forecasting: principles and practice.”
- [7] G. Kariniotakis, P. Pinson, N. Siebert, G. Giebel, and R. Barthelmie, “The State of the Art in Short-term Prediction of Wind Power - From an offshore Perspective,” *Wind Energy*, pp. 20–21, 2004.
- [8] M. A. Cavalcanti Neto, “A Previsão com o Modelo de Regressão,” no. 1987, pp. 1–66, 1991.
- [9] Luiza Maria Oliveira da Silva, “Uma Aplicação de Árvores de Decisão, Redes Neurais e KNN para a Identificação de Modelos ARMA Não-Sazonais e Sazonais,” 2005.
- [10] M. A. Cavalcanti Neto, “A Previsão com a Metodologia de Box-Jenkins,” *Mater. Didático*, no. Figura 1.
- [11] C. Xi, “Artificial Neural Network and Fuzzy Logic in forecasting short-term Temperature,” no. 120295, 2013.
- [12] F. Dornelles, O. C. Pedrollo, and J. A. Goldenfum, “Sistemas de Inferência Difusa na Previsão de Nível do Rio,” pp. 1–20.
- [13] J. Cunha, “Sistema de monitorização online,” 2013.
- [14] <http://www.bertolo.pro.br/MetodosQuantitativos/Previsao/pmc4.htm>, “Introdução à Análise de Séries Temporais.” .
- [15] <https://www.kaggle.com>, “Root Mean Squared Error.” .
- [16] <http://www.vanguardsw.com/>, “Mean Absolute Percent Error.” .
- [17] G. M. Masters, *Renewable and efficient electric power systems*. 2004.