



**Tomás Alexandre Bravo Abreu Cavaco**

Licenciado em Ciências da Engenharia Mecânica

## **Modelação da curva de potência no contexto da detecção de anomalias em turbinas eólicas**

Dissertação para obtenção do Grau de Mestre em  
**Engenharia Mecânica**

Orientador: Doutor Tiago Alexandre Narciso da Silva,  
Professor Auxiliar, Faculdade de Ciências e  
Tecnologia da Universidade Nova de Lisboa

Júri

Presidente: Doutor António Paulo Vale Urgueira, Professor  
Associado, Faculdade de Ciências e Tecnologia  
da Universidade Nova de Lisboa

Vogais: Doutora Susana Maria S. Nascimento M.  
Almeida, Professora Auxiliar, Faculdade de  
Ciências e Tecnologia da Universidade Nova de  
Lisboa;  
Doutor Tiago Alexandre Narciso da Silva,  
Professor Auxiliar, Faculdade de Ciências e  
Tecnologia da Universidade Nova de Lisboa



FACULDADE DE  
CIÊNCIAS E TECNOLOGIA  
UNIVERSIDADE NOVA DE LISBOA

**Fevereiro, 2020**



# Resumo

Atualmente é muito importante saber o estado de condição de uma máquina. Com o desenvolvimento da tecnologia e da complexidade desta, a dificuldade na verificação do estado de funcionamento também tem crescido. Consequentemente, a busca de novas ideias para resolver este problema tem aumentado. As turbinas eólicas são exemplos destas máquinas. Com o crescimento da preocupação com as alterações climáticas, o seu uso tem aumentado consideravelmente. Em busca de novas maneiras de determinar a condição de funcionamento de uma turbina eólica sem interromper o seu funcionamento surgiu o tema deste trabalho que tem como objetivo a monitorização do estado de condição de uma turbina através da sua curva de potência. Este trabalho tem base em técnicas que se enquadram no âmbito da inteligência computacional. Para tal, será feita uma análise semanal onde a curva de potência de uma dada semana será comparada com a curva de potência estimada por um modelo de regressão criado ou ajustado com os dados previamente observados em períodos classificados como sendo representativos de bom estado de funcionamento. Algumas tentativas serão realizadas para tentar aperfeiçoar o modelo com o intuito de obter melhores resultados no final. Serão ainda testados diferentes blocos de dados de treino para verificar qual o melhor conjunto de variáveis de treino para ser usado pelo programa com o intuito de avaliar o estado de condição do equipamento. No final, foi possível identificar anomalias a partir da diferença das curvas de potência e justificar as diferenças com base na forma da curva ou com o número de observações nos blocos de dados usados na previsão.

**Palavras-chave:**

Turbinas eólicas; Detecção de anomalias; Curvas de potência; Processo gaussiano; Modelos de regressão; Energia eólica



# Abstract

Nowadays it is very important to know the condition of a machine. With the development of technology and its complexity, the difficulty in verifying the state of operation has also grown. Consequently, the search for new ideas to solve this problem has increased. Wind turbines are examples of these machines. With the growing concern about global warming, its use has increased considerably. In search of new ways to determine the condition of a wind turbine without interrupting its operation emerged the topic of this work that aims to classify the state of a turbine through its power curve. This work is based on techniques that fall within the scope of computational intelligence. For this purpose, a weekly analysis will be made in which the power curve of a given week will be compared with a power curve estimated by a regression model created and adjusted with data classified as being representative of the good state of functioning. Some attempts are made to improve the regression model in order to obtain better results in the end. Different blocks of training data will also be tested to verify which is the best set of training variables to be used by the program in order to assess the condition of the equipment. In the end, it was possible to identify anomalies from the power difference and justify this differences based on the shape of the curve or the number of observations in the data blocks used in the prediction.

**Keywords:** Wind turbines; Failure detection; Power curves; Gaussian process; Regression models; Wind energy



# Agradecimentos

Depois da conclusão deste trabalho gostava de agradecer a algumas pessoas porque sem elas isto não seria uma realidade.

Queria agradecer ao professor Tiago Silva por todo o tempo que disponibilizou e pelo trabalho que teve para me ajudar.

Queria agradecer aos meus amigos, André Laranjeiro, Márcia Cerquido e Pedro Lopes por todos os momentos passados juntos e por tudo o que fizeram ao longo destes anos.

Um obrigado muito especial à minha família por tudo, mesmo por tudo, o que fizeram, o que passaram e como agiram depois do que aconteceu ao longo desta jornada.

Por último e não menos que os outros. Queria agradecer à minha namorada, Catarina, por ser quem é, por tudo o que fez por mim.

Muito obrigado a toda a gente que me ajudou a chegar onde cheguei hoje.



# Acknowledgements

After completing this work I would like to thank a some people because without them this wouldn't be a reality.

I would like to thank Professor Tiago Silva for all the time he made available and all the work he had helping me.

I would like to thank my friends, André Laranjeiro, Márcia Cerquido and Pedro lopes for every moment spent together and for everything the have done over these years.

A very special thank you to my family for everything, really everything, for what they have done, what they have been through and for how they react after what happended through this journey.

Last but not least. I want to thank my girlfriend, Catarina, for being who she is and for eeverything she done for me.

Thank you to all the people that helped me reach where I am today.



# Índice

<b>Resumo</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>Agradecimentos</b>	<b>vii</b>
<b>Acknowledgements</b>	<b>ix</b>
<b>Índice</b>	<b>xi</b>
<b>Lista de Figuras</b>	<b>xiii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	1
1.2 Objetivo . . . . .	2
1.3 Estrutura do trabalho . . . . .	2
<b>2 Fundamentos Teóricos</b>	<b>3</b>
2.1 Energia eólica . . . . .	3
2.2 SCADA . . . . .	6
2.3 Modelo de regressão Gaussiano . . . . .	7
<b>3 Caso de estudo</b>	<b>9</b>
3.1 Procedimento . . . . .	14
3.2 Determinação do melhor conjunto de variáveis para construção de um modelo de regressão . . . . .	14
3.2.1 Modelo de regressão contruído com dados de velocidade do vento . . . . .	14
3.2.2 Modelo construído com dados de velocidade do vento condicionada à velocidade de rotação do rotor . . . . .	16
3.2.3 Modelo de regressão construído com dados de velocidade de vento e velocidade de rotação do rotor . . . . .	17
3.2.4 Comparação dos resultados obtidos . . . . .	19
3.3 Análise detalhada do estudo 2 . . . . .	21
3.3.1 Análise do ano 2011 . . . . .	21
3.3.2 Tentativa de previsão das anomalias no ano 2012 . . . . .	31

3.3.3	Tentativa de previsão das anomalias no ano 2013 . . . . .	38
3.3.4	Conclusão . . . . .	44
3.4	Tentativa de previsão de uma turbina com o modelo de outra . . . . .	44
<b>4</b>	<b>Conclusão</b>	<b>47</b>
4.1	Trabalhos futuros . . . . .	47
	<b>Bibliografia</b>	<b>49</b>
	<b>Apêndice A Anexos</b>	<b>51</b>
A.1	Deteção de dados omissos . . . . .	51
A.2	Filtragem dos dados . . . . .	51
A.3	Definição dos blocos de dados usados para o treino do modelo . . . . .	52
A.4	Treino . . . . .	54
A.5	Verificação . . . . .	55
A.6	Previsão . . . . .	57

# Lista de Figuras

2.1	Turbina eólica . . . . .	4
2.2	Partes principais de uma turbina eólica . . . . .	4
2.3	Parque eólico <i>offshore</i> (fonte: The Pembrokeshire Herald, Julho 2020) . . . . .	5
3.1	Zona de declive da curva de potência de uma turbina eólica . . . . .	10
3.2	Curva de potência teórica . . . . .	11
3.3	Previsão do modelo para a primeira semana do ano 2011 na turbina 1 . . . . .	13
3.4	Erros na verificação do modelo de regressão construído com dados de velocidade de vento . . . . .	15
3.5	Erro semanal turbina 1 na definição do primeiro modelo construído com dados de velocidade de vento para o ano 2011 . . . . .	15
3.6	Erros na verificação do modelo de regressão construído com dados de velocidade de vento condicionados pela velocidade de rotação do rotor . . . . .	16
3.7	Erros semanal da turbina 1 com modelo de regressão construído com dados de velocidade de vento condicionada pela velocidade de rotação do rotor para o ano de 2011 . . . . .	17
3.8	Erros na verificação do modelo de regressão construído com dados de velocidade de vento e velocidade de rotação do rotor . . . . .	18
3.9	Erro semanal do modelo de regressão construído com dados de velocidade do vento e velocidade de rotação do rotor(turbina 1, 2011) . . . . .	18
3.10	Comparação das verificações do estudo 1 (secção 3.2.1), estudo 2 (secção 3.2.2) e estudo 3 (secção 3.2.3) . . . . .	19
3.11	Comparação dos erros semanais do estudo 1 (secção 3.2.1), estudo 2 (secção 3.2.2) e estudo 3 (secção 3.2.3) . . . . .	20
3.12	Comparação dos erros semanais do estudo 2 (secção 3.2.2) e estudo 3 (secção 3.2.3) . . . . .	20
3.13	Erro semanal para a turbina 1 com modelo de regressão construído com dados da turbina 1 (2011) . . . . .	22
3.14	Curvas de potência da turbina 1 em semanas específicas (2011) . . . . .	22
3.15	Erro semanal para a turbina 2 com modelo de regressão construído com dados da turbina 2 (2011) . . . . .	23
3.16	Curvas de potência da turbina 2 em semanas específicas (2011) . . . . .	23
3.17	Erro semanal para a turbina 3 com modelo de regressão construído com dados da turbina 3 (2011) . . . . .	24

3.18	Curvas de potência da turbina 3 em semanas específicas (2011) . . . . .	24
3.19	Erro semanal para a turbina 4 com modelo de regressão construído com dados da turbina 4 (2011) . . . . .	25
3.20	Curvas de potência da turbina 4 em semanas específicas (2011) . . . . .	26
3.21	Erro semanal para a turbina 5 com modelo de regressão construído com dados da turbina 5 (2011) . . . . .	26
3.22	Curvas de potência da turbina 5 em semanas específicas (2011) . . . . .	27
3.23	Erro semanal para a turbina 6 com modelo de regressão construído com dados da turbina 6 (2011) . . . . .	28
3.24	Curvas de potência da turbina 6 em semanas específicas (2011) . . . . .	28
3.25	Erro semanal para a turbina 7 com modelo de regressão construído com dados da turbina 7 (2011) . . . . .	29
3.26	Curvas de potência da turbina 7 em semanas específicas (2011) . . . . .	29
3.27	Erro semanal para a turbina 8 com modelo de regressão construído com dados da turbina 8 (2011) . . . . .	30
3.28	Curvas de potência da turbina 8 em semanas específicas (2011) . . . . .	30
3.29	Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 1 . . . . .	31
3.30	Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 2 . . . . .	32
3.31	Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 3 . . . . .	33
3.32	Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 4 . . . . .	34
3.33	Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 5 . . . . .	34
3.34	Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 6 . . . . .	35
3.35	Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 7 . . . . .	36
3.36	Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 8 . . . . .	37
3.37	Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 1 . . . . .	38
3.38	Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 2 . . . . .	39
3.39	Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 3 . . . . .	40
3.40	Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 4 . . . . .	41
3.41	Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 5 . . . . .	41

3.42	Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 6 . . . . .	42
3.43	Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 7 . . . . .	43
3.44	Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 8 . . . . .	43
3.45	Comparação dos erros de previsão para as turbinas 3 a 6 considerando o modelo de regressão construído para os dados de cada uma destas turbinas e o construído paraa turbina 1 (2011) . . . . .	46



# 1

## Introdução

### 1.1 Motivação

---

No mundo da engenharia o controlo do estado de uma máquina é algo extremamente necessário, visto que a deterioração desta pode conduzir ao seu mau funcionamento, não sendo possível a realização da sua função. De forma a prevenir possíveis anomalias sem ser preciso interromper o funcionamento da máquina têm-se procurado formas alternativas de verificar a sua condição.

Com o objetivo de reduzir a utilização de combustíveis fósseis tem-se procurado novas formas de obter energia. O conceito de energia renovável vem da utilização de recursos ilimitados, como o sol, vento, água, entre outros. As turbinas eólicas têm a importante função de transformar a energia do vento em energia elétrica sendo portanto uma alternativa aos combustíveis fósseis. Esta não produz gases de efeito estufa e o seu impacto ambiental é geralmente menos problemático do que o de outras fontes de energia. As turbinas eólicas são cada vez mais populares, e por isso, a sua manutenção é de grande interesse.

Como qualquer máquina, estas estão constantemente em evolução. Cada vez maiores, mais complexas mas também mais eficientes. Com o aumento desta complexidade a sua manutenção também fica cada vez mais exigente. Algumas ações de manutenção implicam que a turbina eólica esteja parada, no entanto, deve-se evitar a sua paragem para manutenções preventivas sistemáticas. Isto implicaria custos associados, visto que, não estão a produzir energia quando não trabalham.

Com isto, surgem ideias para tentar controlar a condição de uma turbina enquanto esta está em funcionamento.

## 1.2 Objetivo

---

O objetivo principal deste trabalho é, através da curva de potência de uma turbina eólica, tentar prever o seu estado de funcionamento através da criação de um modelo a partir de um processo de regressão. Para este estudo vai ser explorada a aplicação de técnicas computacionais na deteção de anomalias no funcionamento. Será usada informação de um parque eólico relativa aos anos 2011, 2012 e 2013. Um dos primeiros objetivos que teve de ser alcançado foi perceber como funcionava o algoritmo de regressão e tentar aplicá-lo ao nosso problema. Depois comparar valores obtidos pelo modelo de regressão e em conjunto com o registo de manutenção perceber em que semanas foram registadas anomalias e quando os erros elevados não coincidirem com nenhuma manutenção ou anomalia tentar perceber a origem desse erro, se é causado pela falta de observações nos blocos de dados usados ou por questões relacionadas com o vento mas não consideradas pelo modelo de regressão desenvolvido. Por último, foi verificado se seria possível com o modelo desenvolvido para uma turbina tentar prever o comportamento das restantes turbinas no parque.

## 1.3 Estrutura do trabalho

---

No primeiro capítulo será feita uma introdução do trabalho onde são definidos os respetivos objetivos.

O segundo capítulo contém os fundamentos teóricos para uma melhor compreensão do que será feito ao longo deste trabalho. Os temas abordados serão a energia eólica, SCADA (Supervisory control and data acquisition) e o modelo de regressão.

No terceiro capítulo é onde está explicado o que foi implementado no trabalho, as decisões tomadas e os resultados obtidos.

No capítulo quatro será feita a conclusão, o comentário aos resultados obtidos e ideias para melhorar os mesmos e a sugestão de trabalhos futuros.

## 2

## Fundamentos Teóricos

### 2.1 Energia eólica

---

A energia eólica é uma energia renovável que constitui uma boa alternativa aos combustíveis fósseis e que está sempre em primeira análise disponível em qualquer lugar. A energia eólica pode ser transformada em energia mecânica e só depois em energia elétrica.

A exploração intensiva da energia eólica surgiu nos anos 70 durante a crise do petróleo. Com o medo que este recurso acabasse houve uma busca a diferentes métodos de se obter energia. No entanto, a sua origem é muito mais remota. Diz-se que a utilização de moinhos de vento tem pelo menos três mil anos. Hoje em dia, é usada principalmente para produzir energia elétrica através de aerogeradores ou turbinas eólicas (figura 2.1). As turbinas produzem energia a partir de um gerador que está ligado a um transmissor mecânico. A energia cinética do vento é transformada em energia mecânica através da rotação das pás da turbina. De seguida essa energia é transformada em energia elétrica a partir do gerador. Isto cria uma dependência da velocidade do vento para a produção de energia. Na figura 2.2 estão representados alguns dos principais componentes da turbina eólica.

As pás giram com a ação do vento. Estas estão ligadas ao rotor que por sua vez está ligado à caixa de velocidades. A caixa de velocidades permite transformar as baixas velocidades de rotação do rotor (19-30 rpm) em velocidades que permitem o funcionamento do gerador (1500 rpm). O gerador transforma a energia mecânica em energia elétrica.



Figura 2.1: Turbina eólica

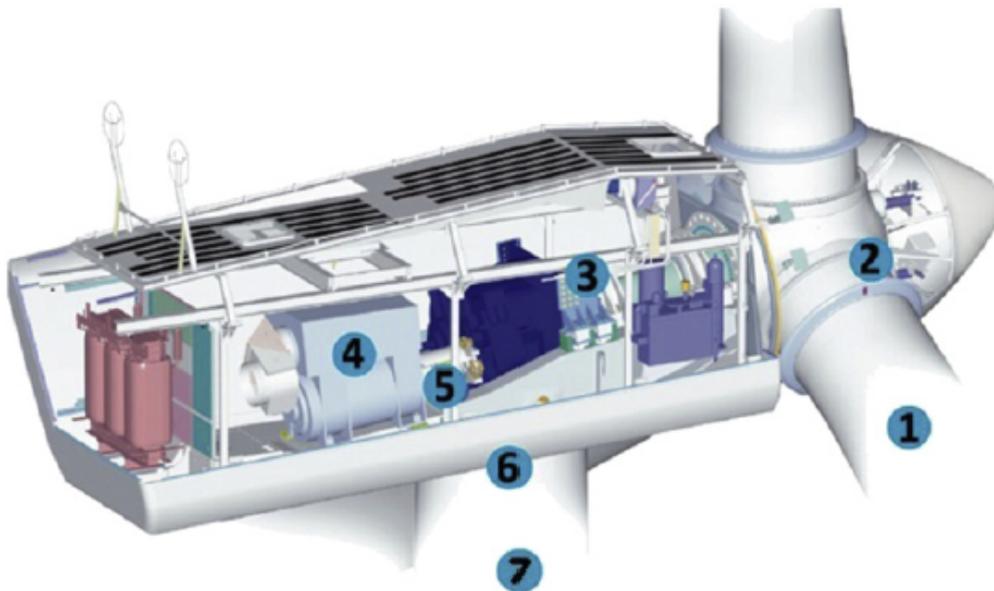


Figura 2.2: Partes principais de uma turbina eólica: (1) Pás, (2) rotor, (3) caixa de velocidades, (4) gerador, (5) rolamento, (6) sistema de Yam e (7) torre

Para a criação de um parque eólico ou para a instalação de uma turbina é muito im-

portante fazer-se um estudo para saber a velocidade do vento e a sua direção porque esta informação é essencial para a maximização do rendimento da mesma. No entanto, todas as turbinas têm uma produção máxima de energia para uma determinada velocidade do vento, que depende do fabricante e do seu modelo. Existem dois tipos de parques eólicos, *onshore* e *offshore*. Os parques *offshore* (figura 2.3) são construídos no mar, ao contrário dos *onshore* que são construídos em terra.



Figura 2.3: Parque eólico *offshore* (fonte: The Pembrokeshire Herald, Julho 2020)

Os benefícios deste tipo de energia são a não libertação de dióxido de carbono para a atmosfera, a redução da utilização de combustíveis fósseis, a melhoria da economia e a criação de emprego. Por outro lado, estas vantagens vêm acompanhadas com alguns impactos negativos, tais como poluição visual, impacto na fauna e poluição sonora. A produção de energia elétrica é limpa mas a produção dos aerogeradores não é, implica um processo produtivo e a respetiva distribuição. De acordo com LIFE-BRIO project, já começa a haver uma preocupação com a reciclagem dos componentes quando chega o final da sua vida útil, como por exemplos as pás que existem estudos para se conseguir reciclar as pás na criação de novas e aumentar a sustentabilidade do processo de produção. As turbinas têm grandes dimensões com cerca de 120 metros de altura e pás com 45 metros de comprimento, existindo modelos com diferentes tamanhos. Por esta razão, o impacto visual causado pelas turbinas pode causar descontentamento na população. A fauna é afetada visto que por um lado para criação de um parque eólico altera o ecossistema local e por outro existe a possibilidade de os pássaros chocarem com as pás da turbina quando esta está em funcionamento.

A potência aerodinâmica disponível do vento afetada pelo coeficiente de potência pode

ser calculada pela equação 2.1 de acordo com Burton et al. (2011).

$$P = \frac{1}{2} C_p \rho A U^3 \quad (2.1)$$

onde  $C_p$  é o coeficiente de potência,  $\rho$  a densidade do ar,  $A$  a área de rotação do rotor e  $U$  a velocidade do vento.

Como ar tem uma densidade muito inferior à da água (cerca de oitocentas vezes menor) a turbina eólica produz muito menos energia do que as centrais hídricas logo têm dimensões maiores. O coeficiente de potência tem um máximo teórico nos 0.593, chamado limite de Betz, que indica o máximo da potência que pode ser extraída do vento independentemente do tipo de turbina em fluxo aberto.

## 2.2 SCADA

---

De acordo com Wang et al. (2014), SCADA é constituído por sensores que são colocados em componentes da turbina e que enviam a informação para uma base de dados onde esta pode ser analisada e processada.

Uma das vantagens do SCADA é a possibilidade de armazenar e controlar todo o tipo de informação num único sítio. Outros benefícios são:

- Melhoria da qualidade do sistema
- Redução dos custos durante a operação
- Melhor desempenho

Esta ferramenta é muito útil na gestão de turbinas eólicas. Estas possuem um número muito elevado de componentes. SCADA permite controlar e armazenar toda a informação sobre a turbina. Algumas das variáveis são: velocidade do vento, velocidade de rotação do rotor, temperatura do ambiente, direção do vento, ângulo das pás, temperaturas de diversos componentes, entre outras. Por norma, este registo é feito de dez em dez minutos sendo registados valor médio, desvio padrão, máximo, mínimo do referido período temporal.

## 2.3 Modelo de regressão Gaussiano

---

De acordo com Rasmussen e Williams (2004) para definir uma variável aleatória com base num processo de regressão Gaussiano (equação 2.2) é preciso escolher uma função de média (equação 2.3), e uma função de covariância (equação 2.4). Cada uma delas é definida por parâmetros que a caracterizam. O  $E$  representa o valor esperado. Onde  $x$  (variável aleatória multivariada) para a qual a covariância é calculada para cada variável que compõe  $x$ .

$$f(x_p) \sim GP(m(x_p), k(x_p, x_q)) \quad (2.2)$$

$$m(x_p) = E[f(x_p)] \quad (2.3)$$

A equação 2.4 representa a função de covariância.

$$k(x_p, x_q) = E[(f(x_p) - m(x_p))(f(x_q) - m(x_q))] \quad (2.4)$$

Neste trabalho será usado o algoritmo de Rasmussen e Williams (2004) que contém um conjunto de funções média e de funções de covariância.

A função de média  $m_\phi$  (com parâmetros  $\phi$ ) é uma função escalar definida em todo o domínio que determina o valor expectável (equação 2.3) para um determinada variável  $x$ .

Existem dois tipos de funções média, simples e compostas. Exemplos de funções média simples são, lineares, constantes, zero. As compostas definidas por um conjunto de funções simples. Pode ser através da soma de funções simples ou através do produto das mesmas, entre outras operações.

Escolheu-se a função *zero mean function* (Equação 2.5). Esta função de média não tem qualquer parâmetro para se definir.

$$m(x_p) = 0 \quad (2.5)$$

A função de covariância que será usada neste trabalho é a exponencial quadrática. Esta função especifica a covariância entre dois pares de variáveis.

$$cov(f(x_p), f(x_q)) = k(x_p, x_q) = \exp\left(-\frac{1}{2} |x_p - x_q|^2\right) \quad (2.6)$$

Estas escolhas foram feitas com base nas opções tomadas na bibliografia de referência Papatheou et al. (2017) e, visto que estas apresentavam melhores resultados que outras experimentadas no trabalho.

A distribuição conjunta dos resultados do treino,  $f$ , e o teste,  $f^*$ , de acordo com a prior é:

$$\begin{bmatrix} f \\ f^* \end{bmatrix} \sim N \left( 0, \begin{bmatrix} k(x, x) & k(x, x^*) \\ k(x^*, x) & k(x^*, x^*) \end{bmatrix} \right) \quad (2.7)$$

Se existirem  $m$  pontos de treino e  $n$  pontos de teste,  $k(x, x^*)$  é uma matriz com  $m \times n$  das covariâncias avaliadas para todos os pares de pontos de treino e teste. O mesmo acontecerá para as restantes matrizes da equação 2.7.

O valor das funções  $f^*$  pode ser obtido pela junção de funções de distribuição posteriores. Normalmente não é possível obter estes valores, apenas versões com ruído:  $y = f(x) + \epsilon$  onde  $\epsilon$  tem variância  $\sigma_n^2$

$$f^* | x, y, x^* \sim N(m(f^*), cov(f^*)) \quad (2.8)$$

$$m(f^*) = k(x^*, x) [k(x, x) + \sigma_n^2 l]^{-1} y \quad (2.9)$$

$$cov(f^*) = k(x^*, x^*) - k(x^*, x) [k(x, x) + \sigma_n^2 l]^{-1} k(x, x^*) \quad (2.10)$$

Para o caso deste trabalho a função de covariância da equação 2.6 é adaptada, tomando a forma,

$$k_y(x_p, x_q) = \sigma_f^2 \exp \left( \frac{-(x_p - x_q)^2}{2l^2} \right) + \sigma_n^2 \delta_{pq} \quad (2.11)$$

onde  $\sigma_f^2$  é a variância do bloco de dados,  $l$  fator de escala,  $\sigma_n^2$  a variância do ruído,  $x_p$  os pontos do modelo e  $x_q$  os pontos para previsão.

O modelo de regressão criado pelo processo Gaussiano otimiza os parâmetros  $\sigma_n^2$  e  $l$  que caracterizam o modelo. O modelo de regressão  $f(x)$  estima a curva de potência semanal da turbina.

# 3

## Caso de estudo

Este estudo é feito para um parque eólico com oito turbinas do mesmo modelo estão potencialmente sujeitas a condições de funcionamento diferentes. Foi feita uma análise individual a cada turbina com o objetivo de determinar o estado de funcionamento da mesma com o intuito de detetar anomalias. Para tal, recorreu-se aos blocos de dados fornecidos pelo sistema SCADA. Este bloco de dados contém informação sobre a potência ativa (kW), a velocidade do vento (m/s), direção do vento, velocidade do rotor em RPM, temperatura exterior e a temperatura em alguns componentes. Para este trabalho foram usadas a potência ativa cujos valores variam entre 0 e 2500 kW, velocidade do vento que pode variar entre 0 e 25 m/s e velocidade do rotor cujo valor nominal deve estar contido no intervalo de 9.6 e 16.9 RPM.

O trabalho começou com uma análise dos blocos de dados fornecidos. Em cada ano existem cerca de 52561 registos por turbina. Para algumas turbinas não foi registada pelo SCADA qualquer informação durante certos períodos de tempo. Foi assumido que por alguma razão o sistema terá sido desligado para detetar este tipo de fenómenos foram comparados as datas de cada registo com um bloco de dados completo com o código no anexo A.1. Para facilitar a programação e a coerência do estudo foram complementados os blocos de valores omissos com o objetivo de completar os blocos de dados para que a fase de filtragem dos dados para os intervalos nominais indicados fosse executada de forma correta.

Foram usados blocos de dados relativos aos anos de 2011 até 2013, por serem os anos a que tínhamos acesso aos dados. Isto será importante para uma melhor aproximação à realidade do modelo de regressão obtido.

Numa primeira fase, serão contruídos três modelos considerando blocos de dados diferentes para se avaliar o modelo de regressão mais promissor. O primeiro terá em conta apenas a velocidade do vento como variável. O segundo será na mesma uma variável (a velocidade do vento), mas desta vez condicionada à velocidade nominal de rotação do rotor. O último será treinado com duas variáveis, a velocidade do vento e a velocidade de rotação do rotor. Estas variáveis serão filtradas porque apenas são importantes pontos que estejam

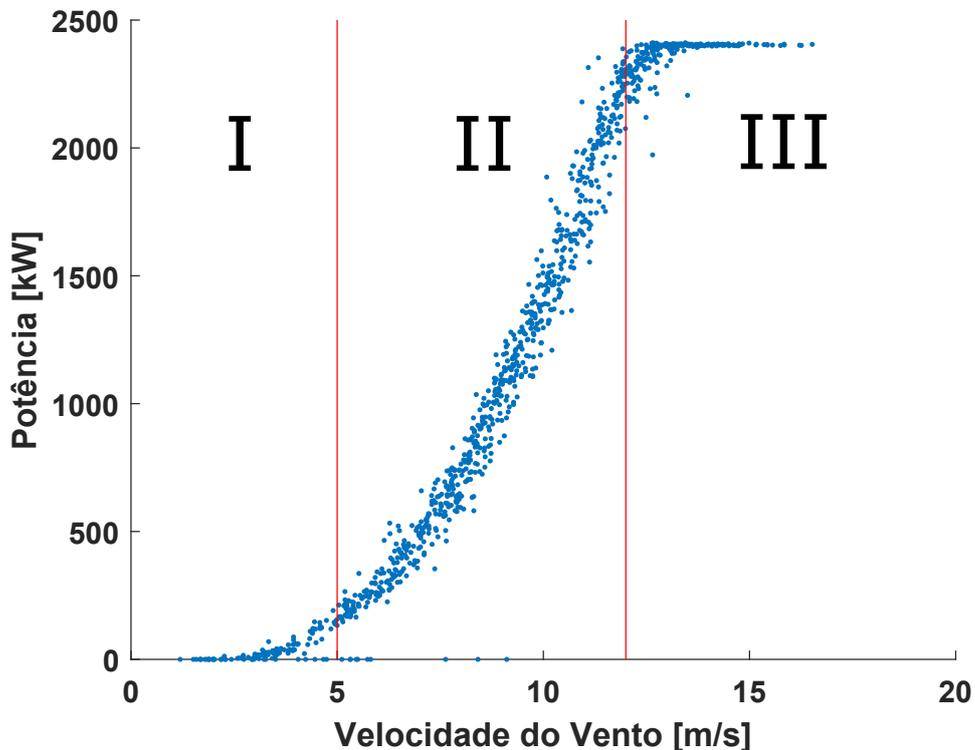


Figura 3.1: Zona de declive da curva de potência de uma turbina eólica

compreendidos no intervalo de valores definido pelo fabricante do equipamento. Será apenas estudada a zona de declive da curva de potência. A partir do conhecimento da curva de potência, sabemos que o declive desta é afetado pela condição da turbina, portanto serão apenas estudados pontos cujos valores estão compreendidos entre 5 m/s e 12 m/s como está representado na figura 3.1. Esta decisão foi tomada após a análise de algumas curvas de potência onde se verificou que, para algumas semanas, a turbina não chegava a produzir a potência máxima. Também foi feito um modelo de teste para a curva de potência completa, no entanto devido ao número reduzido de pontos no patamar onde a potência é máxima (zona III na figura 3.1) o modelo de regressão ficava menos preciso na zona de declive da curva de potência (zona II na figura 3.1). Com isto, foi tomada a decisão de se estudar apenas a zona de declive da curva de potência. A última restrição foi que a velocidade do rotor terá de estar compreendida entre 9.6 e 16.9 RPM porque é só neste intervalo nominal é que a turbina está a produzir potência elétrica. O código usado na filtragem dos dados está presente no anexo A.2.

É importante perceber as diferentes fases deste estudo. Na fase de treino é dada informação ao algoritmo para que os parâmetros destes sejam ajustados e assim criar um modelo de regressão cada vez mais correto. Nesta fase são consideradas apenas curvas de potência normais, ou seja, curvas de potência semelhantes a curva de potência fornecida pelo fabricante

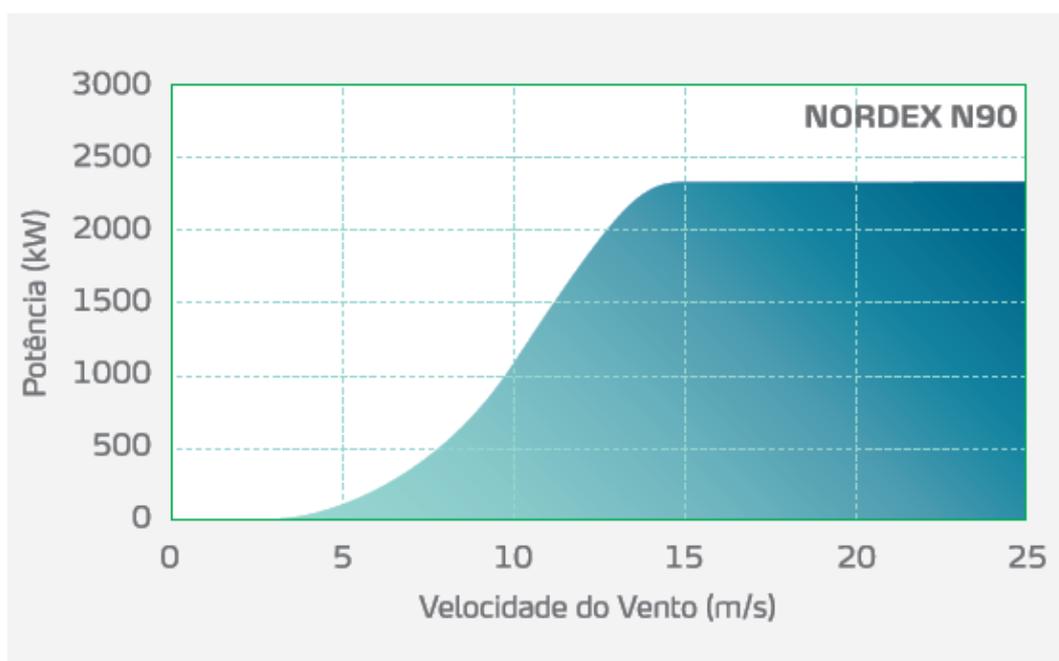


Figura 3.2: Curva de potência teórica

e sem registos de anomalias. A segunda fase é a validação. Esta parte é um teste do modelo criado, onde serão usados um quinto dos pontos utilizados na fase anterior. Por último, a fase da previsão ou a fase de teste, são fornecidos ao programa todos os pontos para se conseguir classificar as curvas de potência. Com este processo o algoritmo é capaz de prever a potência que a turbina deve estar a produzir para uma determinada velocidade do vento a partir de toda a informação fornecida na fase de treino.

A realização destes testes é feita com uma turbina escolhida sem nenhum critério objetivo. Foi escolhida a turbina 1. Através da representação das curvas de potência para o ano de 2011 foram escolhidas seis curvas que podem ser consideradas como "boas", curvas equivalentes à curva teórica representada na figura 3.2. Esta classificação é atribuída através da experiência visto que não temos informação sobre se realmente a turbina está com um funcionamento anormal, no entanto temos sempre a hipótese de confirmar se existe algum evento próximo de manutenção. Estas seis curvas de potência serão usadas para treinar o modelo de regressão em todos os treinos. No final, com base no modelo de regressão que apresentar melhores resultados será feito o estudo das restantes turbinas.

Para o treino foi preciso definir em primeiro lugar o tipo de funções média, covariância e densidade probabilidade que teriam de ser utilizadas. Foram escolhidas as mesmas funções de média, covariância e densidade probabilidade usadas no artigo Papatheou et al. (2017). Para além destas funções também é preciso definir os parâmetros do modelo. O processo de ajuste dos parâmetros do modelo de regressão são o alvo do processo de aprendizagem com base nos dados e este é executado com uma função contida no próprio algoritmo do processo

Gaussiano. Para o mesmo bloco de dados foram executadas várias iterações do processo de otimização de parâmetros e verificou-se que os parâmetros obtidos eram sempre iguais. Com as funções e com os parâmetros é possível ajustar o modelo para o bloco de dados (anexo A.4).

Depois dos modelos de regressão terem sido treinados para os três casos de treino vai ser feita uma verificação para se ter uma ideia se o modelo ficou bem definido. Esta verificação consiste em escolher cerca de um quinto dos pontos utilizados para treino, aleatoriamente, para fazer a previsão dos mesmos como pode ser comprovado no em anexo A.5. Ou seja, quando se fez o treino do modelo, o que essencialmente se está a fazer é dar informação ao programa para ele saber que para uma determinada velocidade de vento deve estar a produzir uma determinada potência. Quanto mais pontos forem utilizados no treino mais informação terá o modelo de regressão. O que se espera nesta verificação é que utilizando os pontos que foram usados para o treino do modelo, o erro, a diferença entre a potência que a turbina está a produzir e a prevista pelo modelo (figura 3.3) seja baixo. Para o cálculo deste erro foi utilizada a equação:

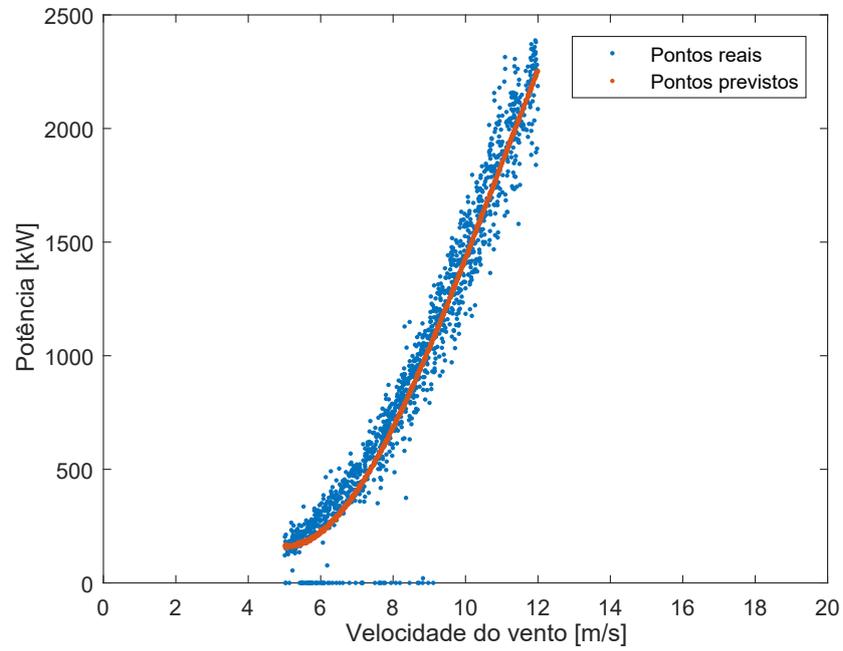
$$Erro = \frac{100}{N\sigma_y^2} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.1)$$

onde  $N$  é o número de pontos utilizados na previsão,  $\sigma_y^2$  é a variância,  $y_i$  o valor da potência que está a ser produzida e  $\hat{y}_i$  a potência que o programa prevê que esteja a ser produzida no período temporal  $i$ .

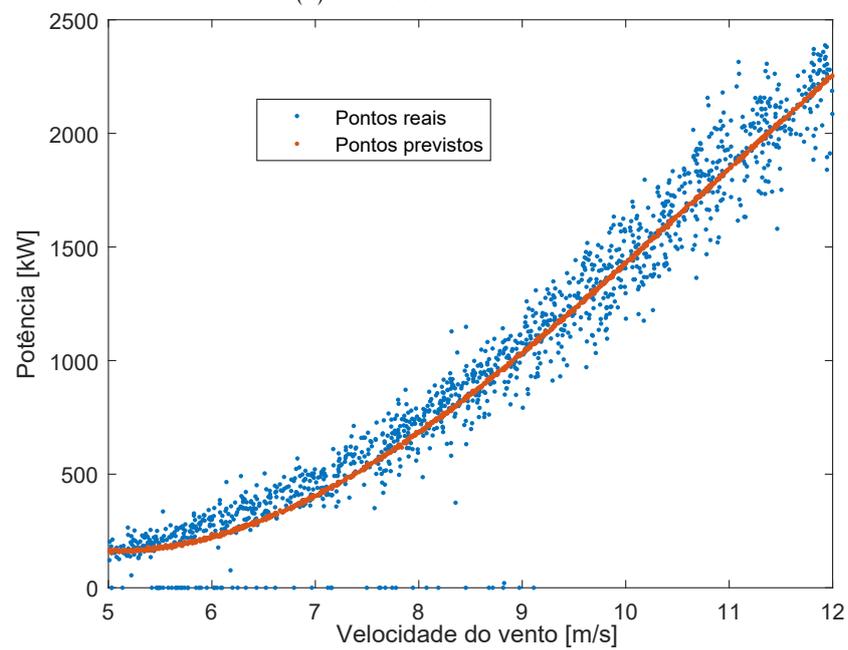
Depois também será feita uma análise anual para cada um dos treinos que consiste em fazer o mesmo raciocínio que foi feito na verificação explicado anteriormente com os dados de cada semana. Através dos dados, o programa faz a previsão da potência e depois é calculado o erro para a respetiva semana.

Este vai ser o princípio utilizado para o objetivo deste trabalho que é a deteção de anomalias. Quando existe uma anomalia é expectável que a potência que a turbina está a produzir varie. É importante perceber que existem muitos outros fatores que podem influenciar flutuações na potência e consequentemente no erro.

Com a capacidade de previsão dos três modelos de regressão contruídos escolhe-se o que apresenta melhores resultados. Com essa escolha feita, faz-se a análise para o ano 2011. Esta tem como objetivo a previsão das anomalias, no entanto tem uma particularidade em comparação com a mesma análise que será feita para os anos 2012 e 2013. Para o ano 2011 serão usados os dados de 2011 para treino e para a previsão, ao contrário dos outros dois anos onde serão utilizados os dados do ano anterior no caso de 2012 e dados do ano 2011 para prever 2013. Isto para tentar aproximar o estudo mais da realidade tendo em conta que estes dados de histórico estão disponíveis e poderão ser considerados para análise.



(a) Previsão do modelo



(b) Previsão do modelo ampliada

Figura 3.3: Previsão do modelo para a primeira semana do ano 2011 na turbina 1

## 3.1 Procedimento

---

Com o objetivo de prever o comportamento da turbina, a criação do respetivo modelo de regressão é conseguida com base no seguinte procedimento:

1. Filtrar os blocos de dados de acordo com as restrições.
2. Identificar as curvas de potência que serão utilizadas para o treino do modelo.
3. Treinar um modelo para cada turbina com as suas respetivas curvas classificadas como boas.
4. Com o modelo definido pelos respetivos parâmetros, selecionar uma percentagem de pontos do conjunto de dados utilizados no ponto 2 para ser feita uma previsão dos mesmos e comparar essa previsão com o modelo treinado. Este passo é uma verificação para ver se o modelo está bem definido.
5. Com o modelo bem definido, para cada turbina, fazer a previsão de cada semana e a respetiva comparação.

## 3.2 Determinação do melhor conjunto de variáveis para construção de um modelo de regressão

---

### 3.2.1 Modelo de regressão contruído com dados de velocidade do vento

Como explicado anteriormente, esta parte do trabalho foi feita apenas para a turbina 1. Depois de serem escolhidas as seis curvas de potência representativas da turbina em bom funcionamento foi construído o modelo de regressão. A figura 3.4 apresenta o erro nas cinquenta verificações que foram feitas. De acordo com o trabalho da referência Papatheou et al. (2017) erros da ordem dos erros da figura 3.4 são considerados bons, ou seja, o modelo está bem definido e pode-se fazer o previsão semanal para o ano de 2011 e comparar as semanas com o erro mais elevado com o registo de manutenção para descobrir se existe alguma situação relatada sobre uma anomalia na respetiva semana. O erro semanal da turbina 1 está representado na figura 3.5.

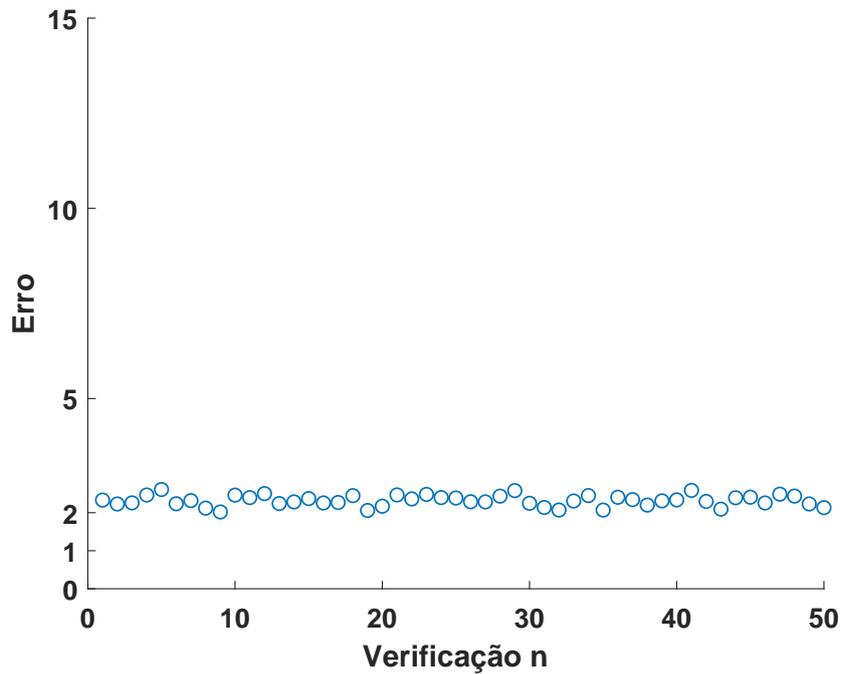


Figura 3.4: Erros na verificação do modelo de regressão construído com dados de velocidade de vento

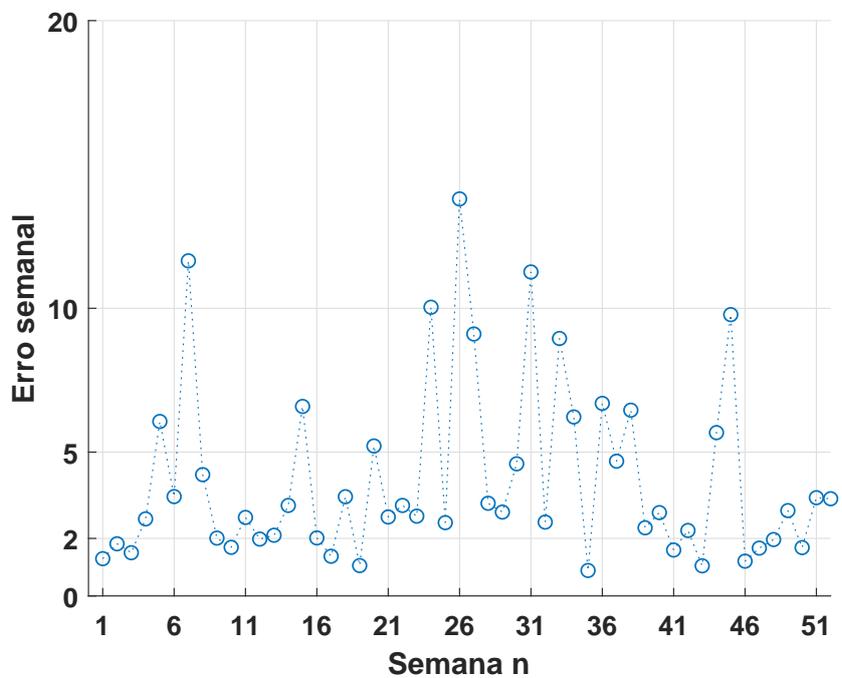


Figura 3.5: Erro semanal turbina 1 na definição do primeiro modelo construído com dados de velocidade de vento para o ano 2011

### 3.2.2 Modelo construído com dados de velocidade do vento condicionada à velocidade de rotação do rotor

Semelhante ao que foi feito anteriormente mas com uma restrição maior dos pontos, para além de estarem restritos à velocidade do vento para apenas se trabalhar na zona com declive da curva de potência, haverá no entanto uma outra restrição, a velocidade do rotor de cada ponto tem de estar compreendida entre 9.6 e 16.9 RPM. Por velocidade do vento condicionada à velocidade de rotação do rotor entende-se que para um determinado ponto no tempo onde seja observada velocidade do vento e potência produzida estando a velocidade de rotação do rotor compreendida no intervalo nominal. Se a velocidade de rotação do rotor para esse mesmo instante não estiver compreendida entre o intervalo definido o ponto não será admitido para treino, visto que, de acordo com os dados fornecidos pelo fabricante não há produção de energia acontece nesse intervalo.

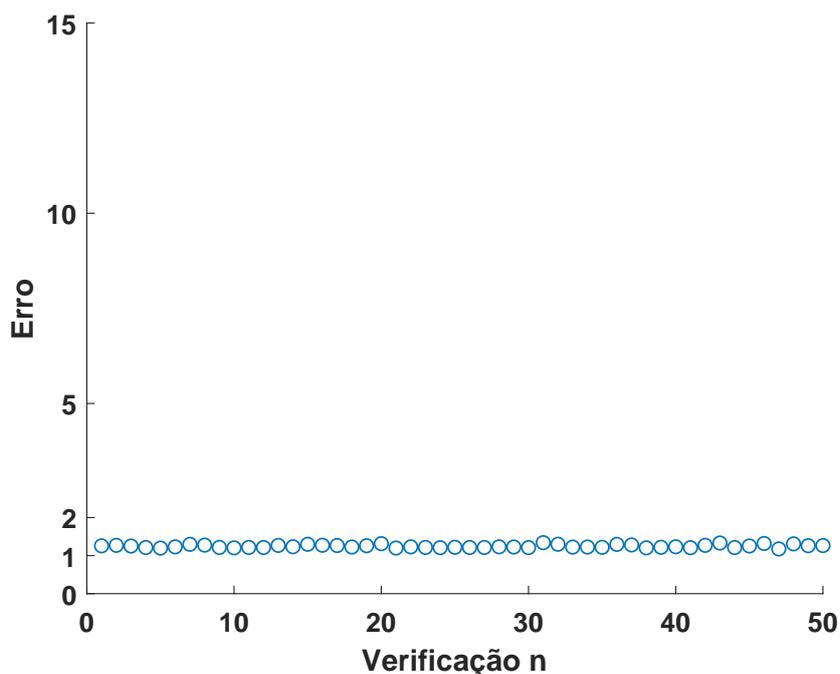


Figura 3.6: Erros na verificação do modelo de regressão construído com dados de velocidade de vento condicionados pela velocidade de rotação do rotor

Em comparação com a verificação do ponto anterior (figura 3.4), nota-se que, na figura 3.6 o erro é mais baixo, podendo ser concluído que o modelo foi bem definido. É importante também salientar que o erro da figura 3.6 é calculado para o mesmo bloco de dados utilizado no caso anterior.

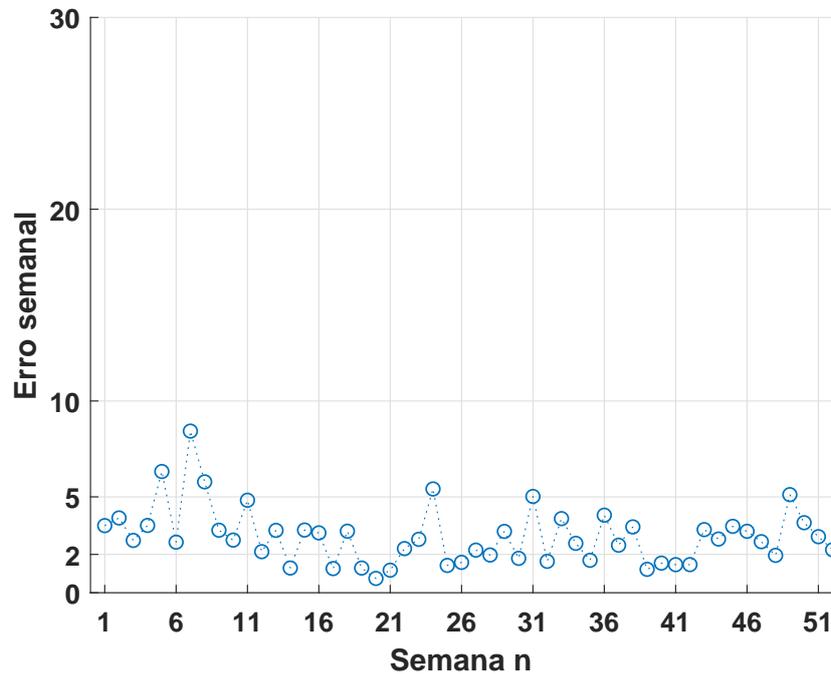


Figura 3.7: Erros semanal da turbina 1 com modelo de regressão construído com dados de velocidade de vento condicionada pela velocidade de rotação do rotor para o ano de 2011

### 3.2.3 Modelo de regressão construído com dados de velocidade de vento e velocidade de rotação do rotor

Por último foi feito um teste com o modelo de regressão construído com dados de duas variáveis, ou seja, usar a velocidade do vento e a velocidade de rotação da turbina. Embora os erros na figura 3.8 sejam menores que os apresentados na figura 3.4 e na figura 3.6, na figura 3.9, os erros semanais (figura 3.9 são muito superiores aos erros da figura 3.5 e 3.7, isto será comentado na secção seguinte.

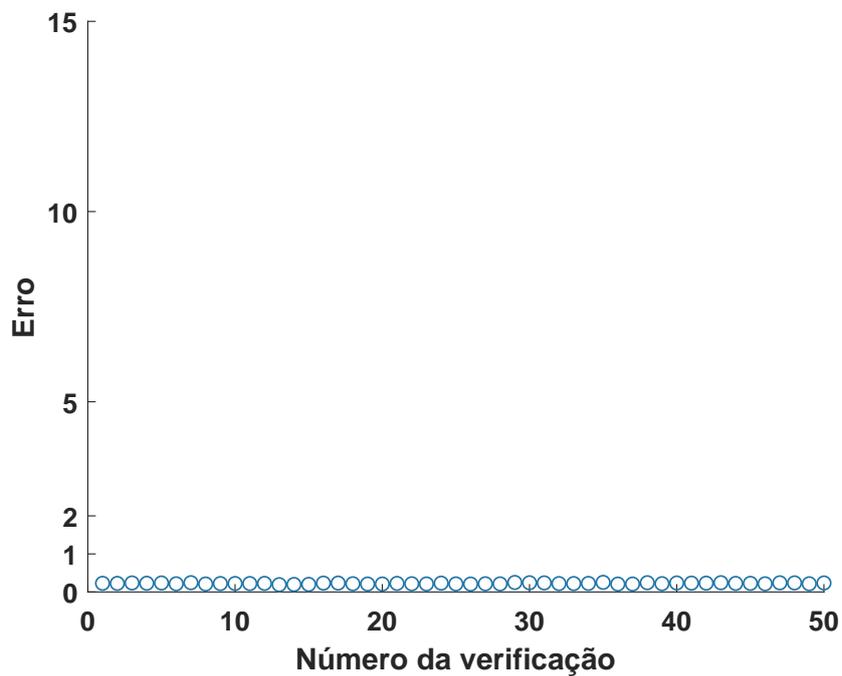


Figura 3.8: Erros na verificação do modelo de regressão construído com dados de velocidade de vento e velocidade de rotação do rotor

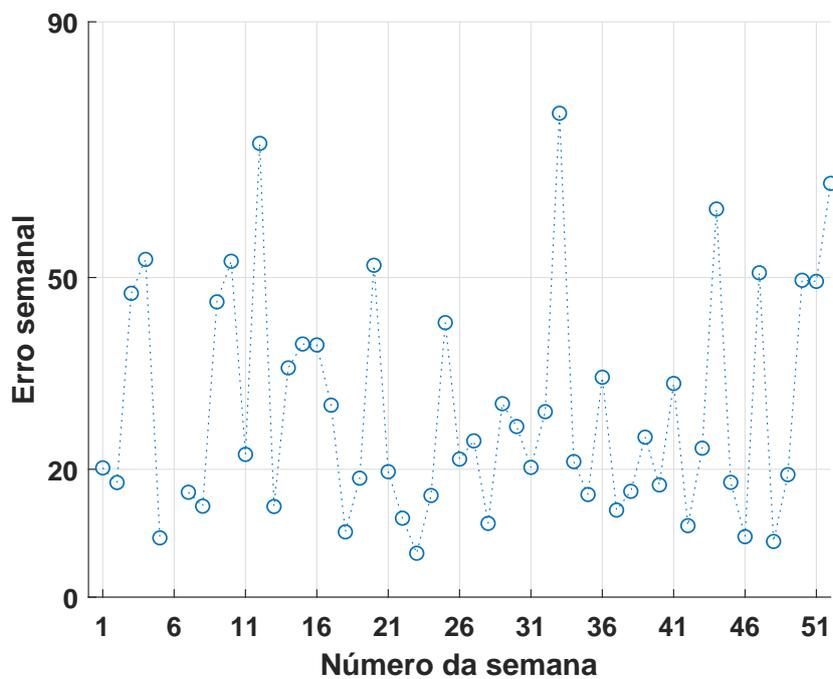


Figura 3.9: Erro semanal do modelo de regressão construído com dados de velocidade do vento e velocidade de rotação do rotor(turbina 1, 2011)

### 3.2.4 Comparação dos resultados obtidos

Para a realização do estudo mais rigoroso, vai ser escolhido um dos três modelos definidos anteriormente. Esta escolha será feita a partir da análise e comparação dos erros obtidos. A partir da figura 3.10 podemos concluir que a modelo da secção 3.2.3 apresenta melhores resultados na verificação seguido da secção 3.2.2 e por último da secção 3.2.1. No entanto, na análise semanal, da secção 3.2.3 apresenta erros elevados em comparação com os outros dois como é visível na figura 3.11. Com a restrição da informação, a dimensão dos blocos de dados pode ser demasiado reduzida, especialmente na análise semanal. A verificação como é feita dentro do mesmo bloco de dados, não é afetada, pois foi escolhido um bloco que continha um número suficiente de observações. No entanto, não é possível garantir que aconteça o mesmo na análise semanal. A figura 3.12 representa com melhor detalhe apenas os erros da análise semanal da secção 3.2.2 e 3.2.1. Apesar de os resultados de ambos os estudos serem muito próximos, é possível verificar que, no geral, o secção 3.2.2 apresenta erros semanais inferiores, logo, será usado a velocidade do vento condicionada a velocidade de rotação do rotor para construir o modelo de regressão para as restantes turbinas.

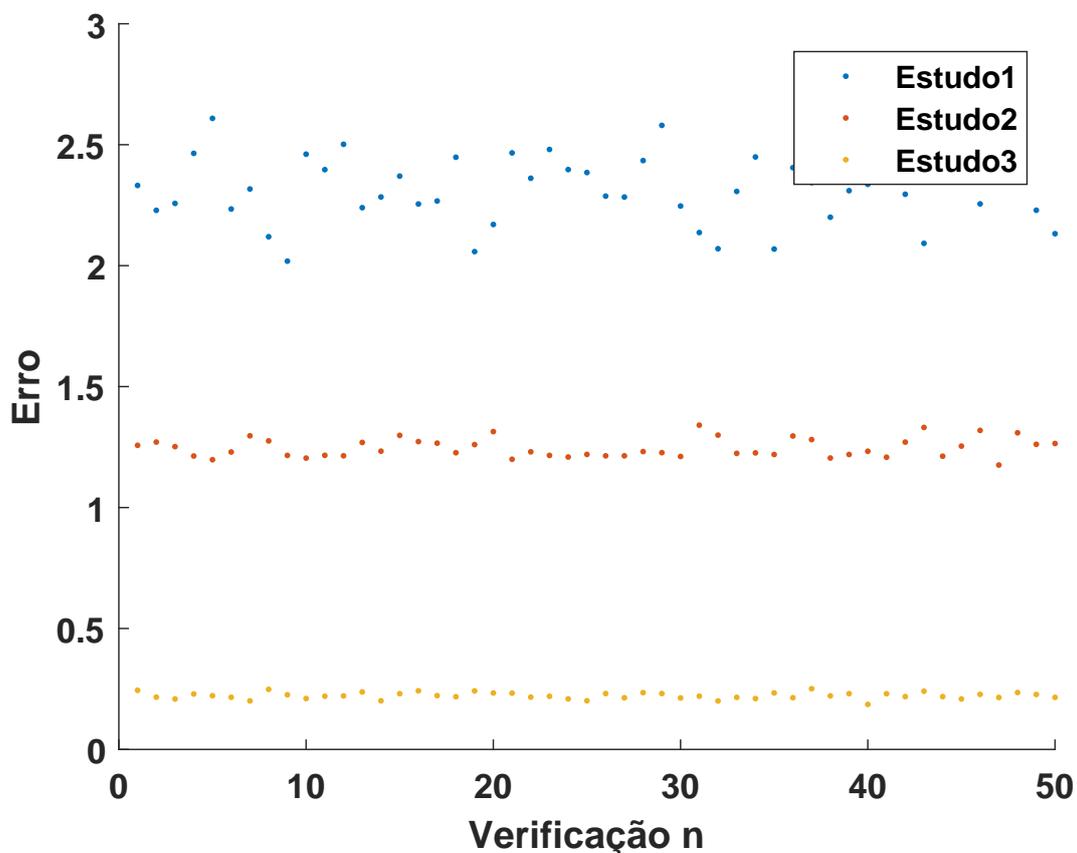


Figura 3.10: Comparação das verificações do estudo 1 (secção 3.2.1), estudo 2 (secção 3.2.2) e estudo 3 (secção 3.2.3)

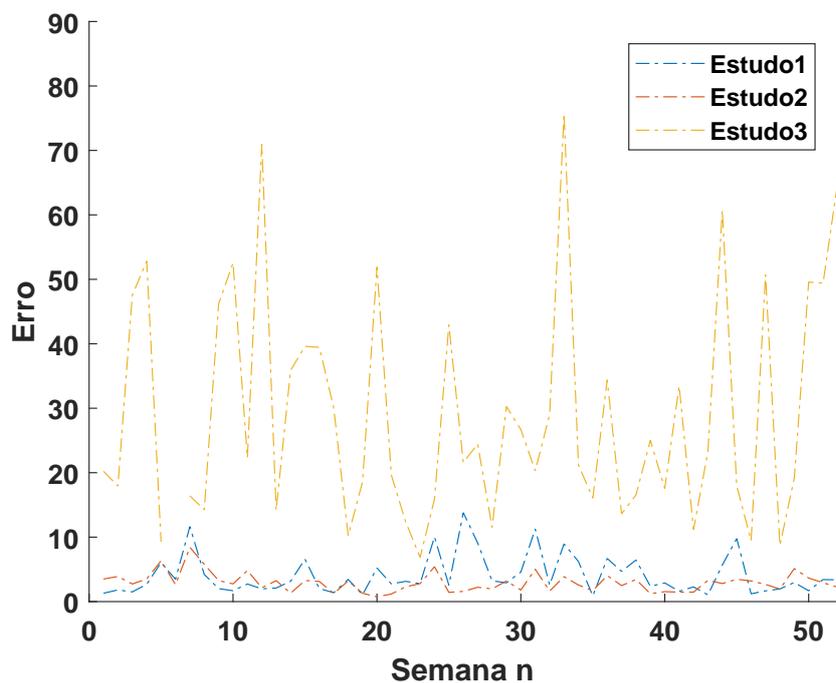


Figura 3.11: Comparação dos erros semanais do estudo 1 (secção 3.2.1), estudo 2 (secção 3.2.2) e estudo 3 (secção 3.2.3)

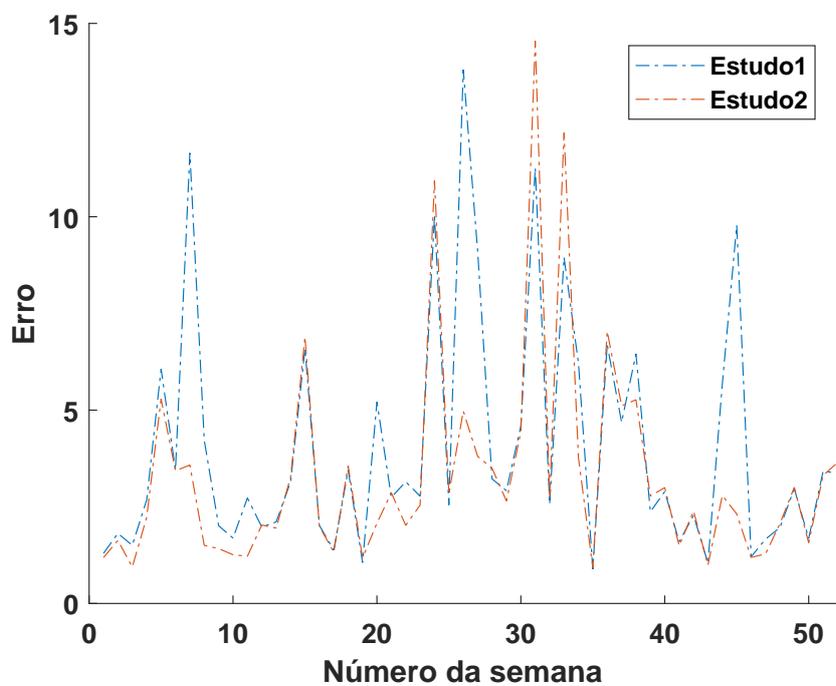


Figura 3.12: Comparação dos erros semanais do estudo 2 (secção 3.2.2) e estudo 3 (secção 3.2.3)

### 3.3 Análise detalhada do estudo 2

---

Na secção 3.3 será feita uma análise para todas as turbinas com a velocidade do vento condicionada a velocidade de rotação do rotor. Numa primeira parte será criado um modelo de regressão para cada turbina apenas com observações no ano de 2011. Em seguida, verificar-se-á se os erros mais elevados de cada modelo para a própria turbina coincidem com semanas onde a esta esteve a funcionar de forma anómala de acordo com o registo de eventos de manutenção. Nos gráficos, como o gráfico da figura 3.13, pode-se detetar três tipos de fenómenos que causam o possível aumento do erro. A vermelho estão representadas as reparações registadas no registo de manutenção. A amarelo as manutenções que podem causar a interrupção no funcionamento da turbina que pode gerar um aumento no erro para essa semana. Por último a verde estão representadas as semanas com a curva de potência fora do normal. Isto pode ser causado por duas situações ou houve poucos pontos registados pelo SCADA ou simplesmente não houve vento e consequentemente foi observada uma potência baixa.

#### 3.3.1 Análise do ano 2011

##### Modelo para turbina 1

Este modelo já ficou definido anteriormente. Na figura 3.13 a semana 7, 45 e 48 correspondem a reparações da caixa de velocidades, caixa de yam (sistema que permite a rotação da turbina) e sistema de refrigeração respetivamente. O erro na semana 24 corresponde a uma manutenção elétrica da turbina. Os erros nas curvas de potência das semanas 31, 33 e 37 são resultado de vento fraco nessa altura do ano visto que em comparação com as outras semanas o bloco de dados tem mais ou menos a mesma dimensão. Na semana 7 também existe uma curva de potência fora do normal porque durante essa semana houve ventos muito elevados e poucos registos com vento baixo.

As figuras 3.14a e 3.14b representam as curvas de potência da turbina 1 na semana 24 e 45 respetivamente. A figura 3.14a representa uma semana onde existiu uma manutenção na turbina e a semana na figura 3.14b representa uma semana onde existiu uma reparação.

##### Modelo da turbina 2

Com um novo modelo criado para a turbina 2 e em conjunto com o histórico da manutenção da mesma, podemos concluir que na semana 24 e 48 houve manutenções na turbina. Semelhante ao que aconteceu na turbina 1 as semanas 31 e 33 são resultados de ventos fracos e não por falta de pontos. Houve duas avarias, reparação das pás na semana 26 e nas semanas 43, 44 e 45 uma reparação da caixa multiplicadora. O erro na semana 44 foi eliminado para

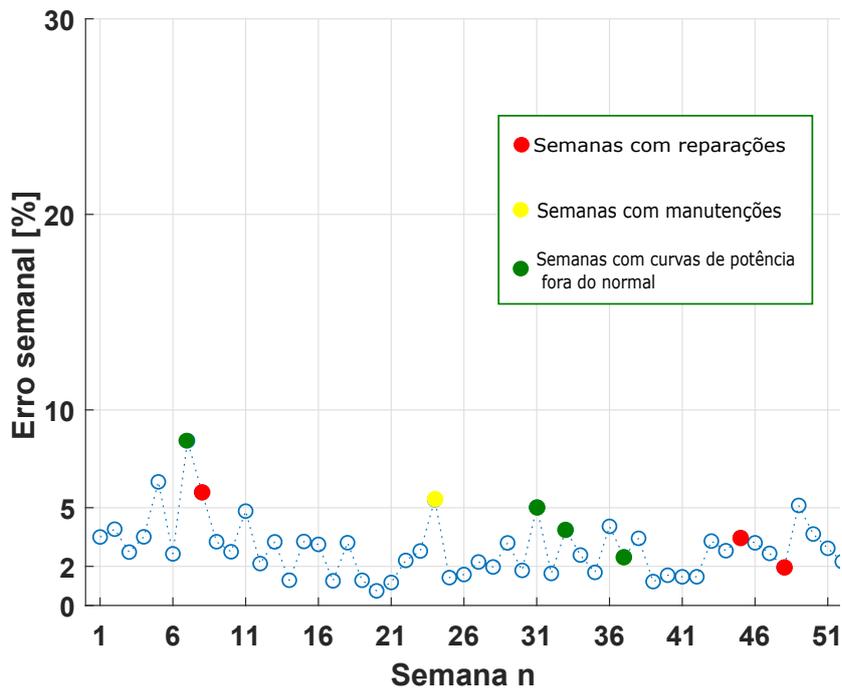
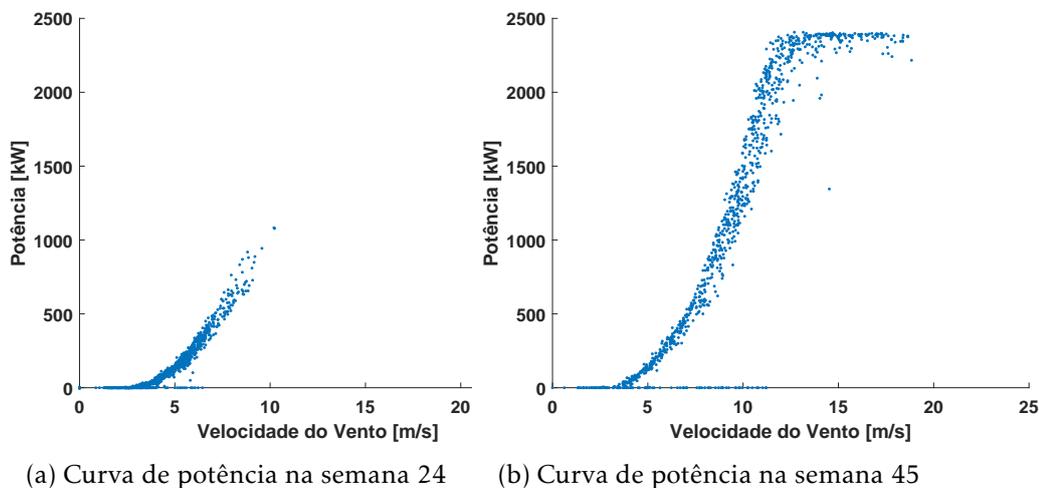


Figura 3.13: Erro semanal para a turbina 1 com modelo de regressão construído com dados da turbina 1 (2011)

simplificação do gráfico que era muito elevado relativamente aos restantes, nesta semana a potência máxima da turbina foi reduzida para 30% da potência máxima que causou o erro elevado. Na semana seguinte a potência máxima passou para 70% melhorando assim o erro. Este aumento gradual é efetuado para repôr a turbina em funcionamento após a substituição da caixa multiplicadora. Na semana 7 voltou a acontecer o mesmo que aconteceu na turbina



(a) Curva de potência na semana 24

(b) Curva de potência na semana 45

Figura 3.14: Curvas de potência da turbina 1 em semanas específicas (2011)

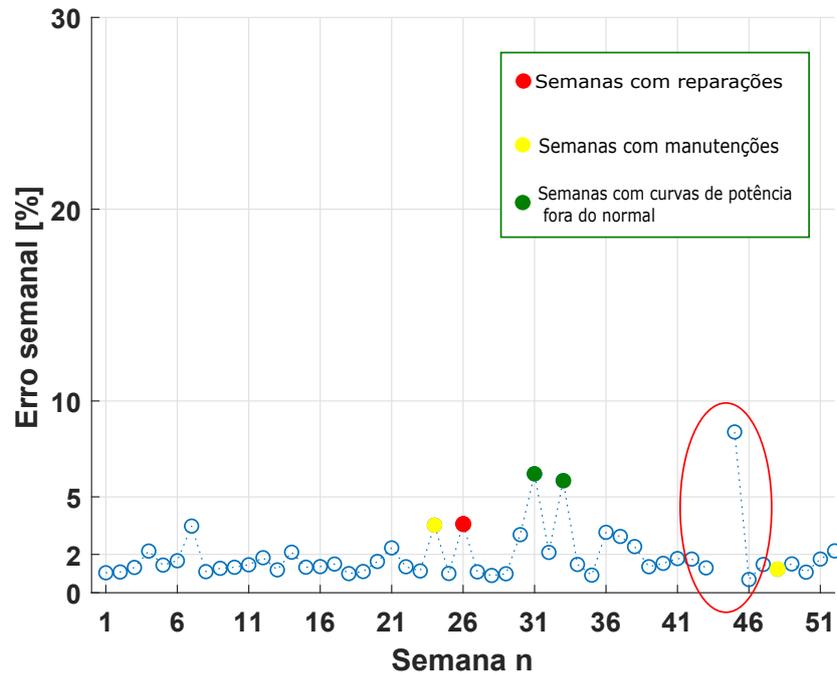
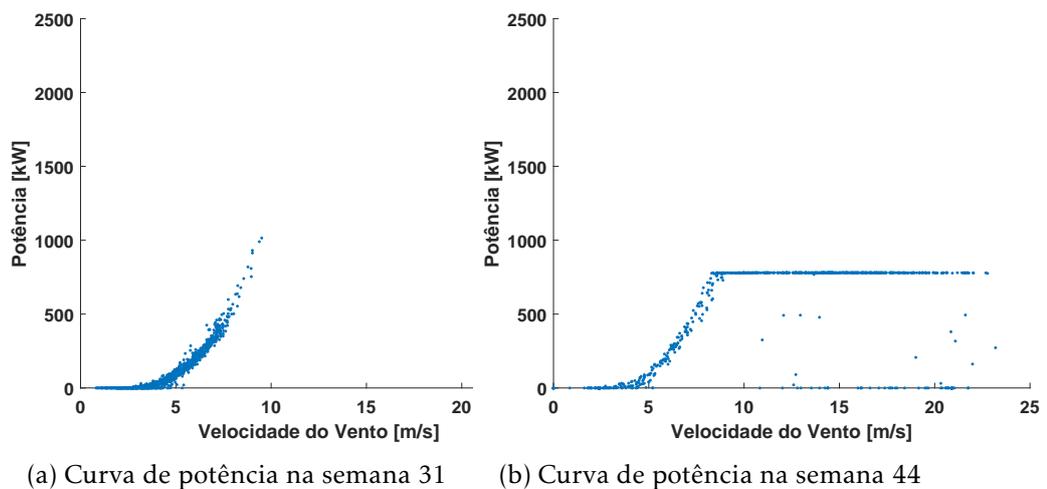


Figura 3.15: Erro semanal para a turbina 2 com modelo de regressão construído com dados da turbina 2 (2011)

1, com o vento fraco durante esta semana o erro aumentou.

As figuras 3.16a e 3.16b representam as curvas de potência da turbina 2 na semana 31 (semana com falta de vento) e 44 (semana com substituição da caixa multiplicadora) respectivamente.



(a) Curva de potência na semana 31

(b) Curva de potência na semana 44

Figura 3.16: Curvas de potência da turbina 2 em semanas específicas (2011)

### Modelo da turbina 3

Depois de comparar a figura 3.17 com o registo de manutenção não houve nenhuma reparação importante nesta turbina para o ano 2011. Os pontos com erros mais elevados foram causados pela curva de potência fora do normal tal como nas turbinas anteriores visto que as condições atmosféricas afetam o parque todo.

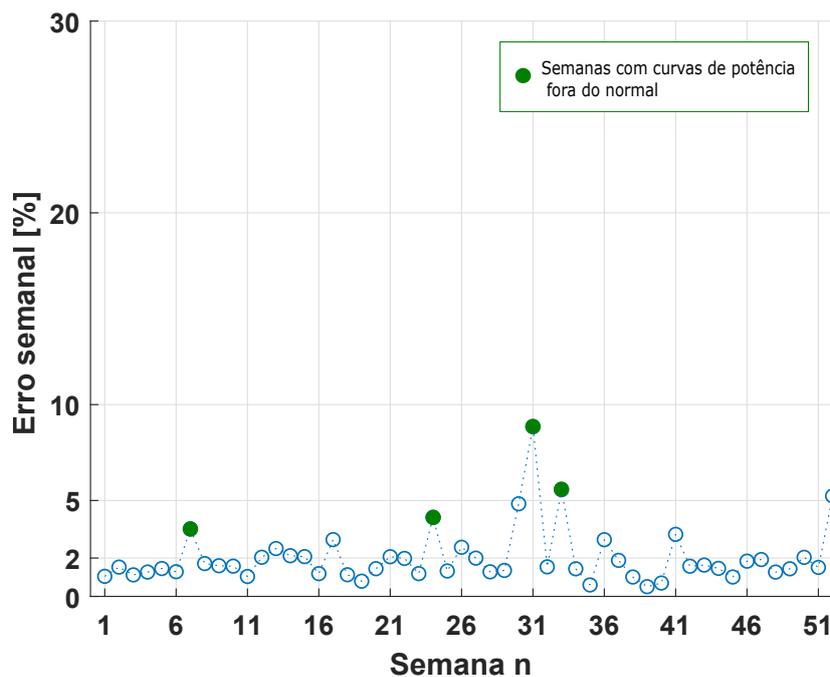


Figura 3.17: Erro semanal para a turbina 3 com modelo de regressão construído com dados da turbina 3 (2011)

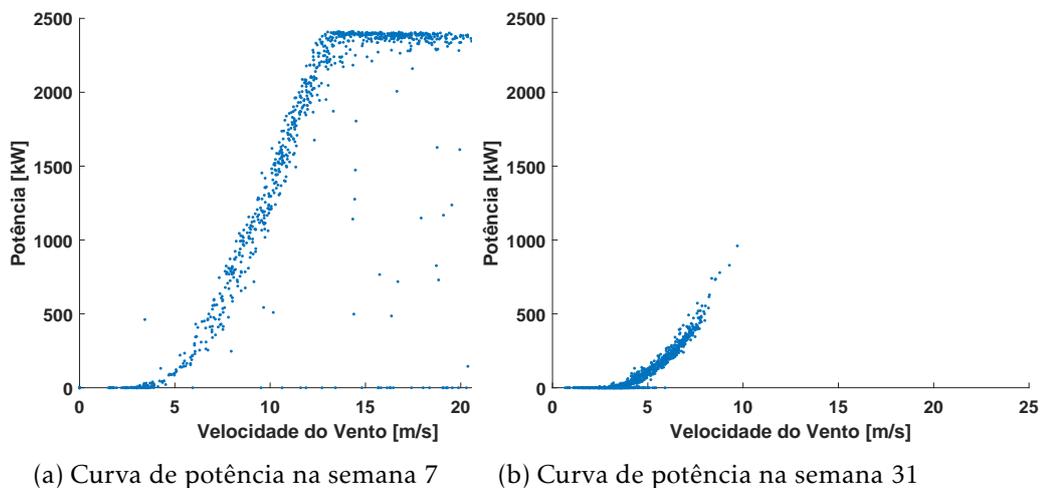


Figura 3.18: Curvas de potência da turbina 3 em semanas específicas (2011)

As figuras 3.18a e 3.18b representam as curvas de potência da turbina 3 na semana 7 e 31 respectivamente. A semana 7 apresenta poucas observações e a semana 31 apresenta pouco vento registrado.

#### Modelo da turbina 4

Na figura 3.19, a anomalia registrada na turbina 4 corresponde à semana 23. Nesta semana também houve uma manutenção elétrica na turbina. Os restantes pontos já foram explicados nas turbinas anteriores.

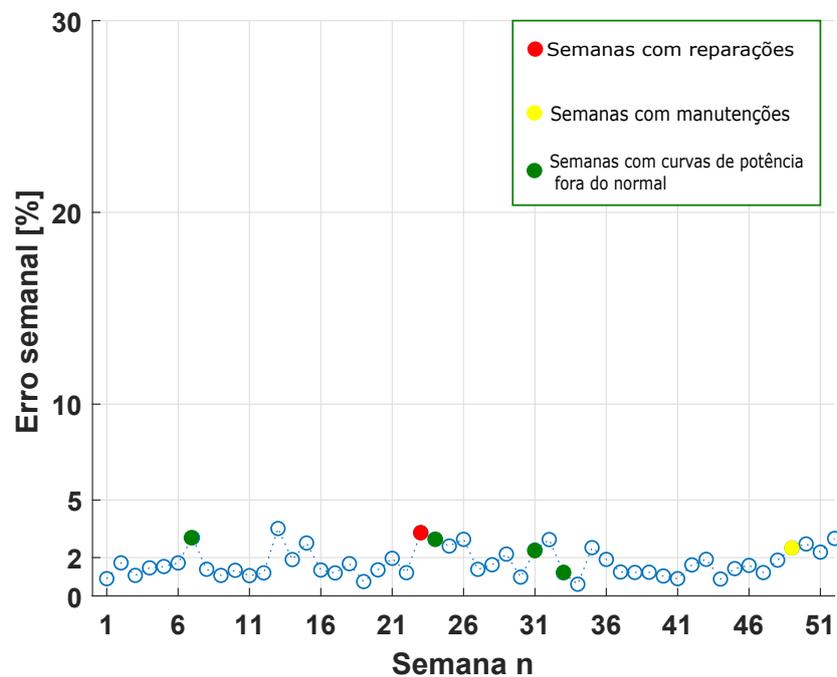


Figura 3.19: Erro semanal para a turbina 4 com modelo de regressão construído com dados da turbina 4 (2011)

As figuras 3.20a e 3.20b representam as curvas de potência da turbina 4 na semana 23 (semana com anomalia registrada) e 49 (semana com manutenção registrada) respectivamente.

#### Modelo da turbina 5

Para esta turbina as anomalias não coincidem com as semanas com maiores erros, estes são causados pelas condições atmosféricas.

As figuras 3.22a e 3.22b representam as curvas de potência da turbina 5 na semana 24 (semana com pouco vento) e 36 (semana com poucas observações) respectivamente.

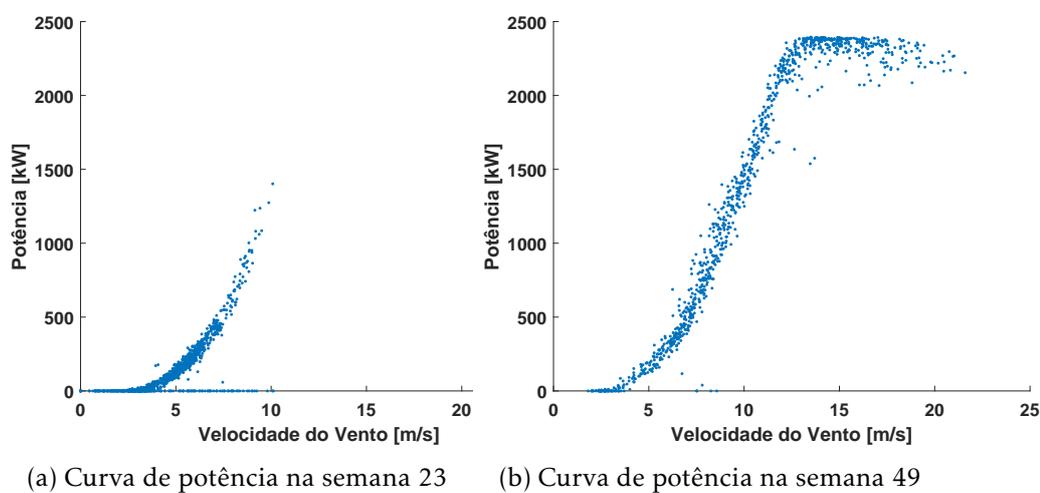


Figura 3.20: Curvas de potência da turbina 4 em semanas específicas (2011)

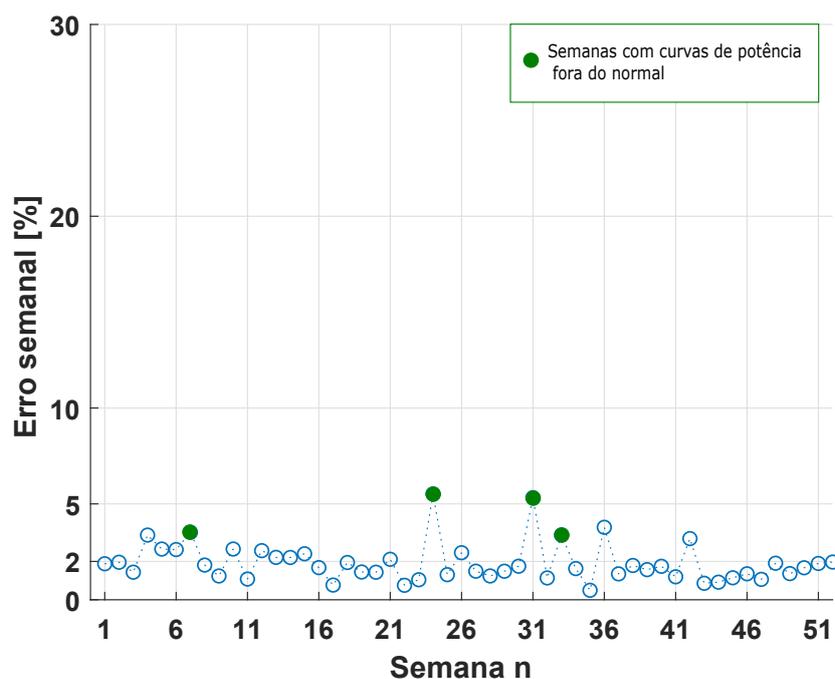


Figura 3.21: Erro semanal para a turbina 5 com modelo de regressão construído com dados da turbina 5 (2011)

### Modelo da turbina 6

De acordo com a figura 3.23 a turbina 6 existem anomalias nas semanas 5 e 12 no sistema de pitch. Nas semanas 15 e 25 foram feitas duas manutenções. A partir do registo da manutenção não se consegue concluir porque é que o erro a partir da semana 21 é elevado. No entanto, depois da análise das curvas de potência é possível verificar que na semana 21

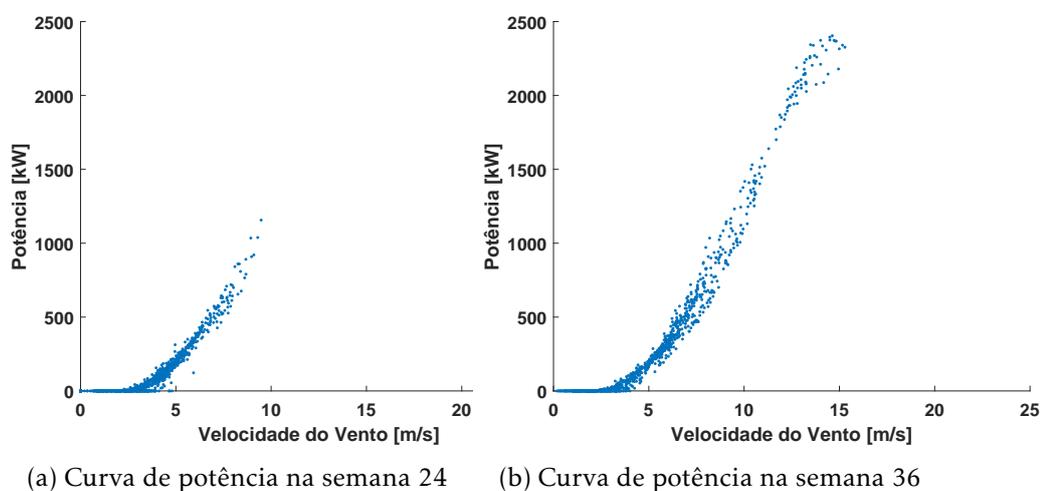


Figura 3.22: Curvas de potência da turbina 5 em semanas específicas (2011)

existem poucos pontos comparando com as restantes o que pode explicar este aumento do erro. A temperatura podia ser uma explicação para o aumento do erro das semanas 21 à 31, visto que a potência que a turbina produz está dependente da temperatura exterior devido à densidade do ar, no entanto com a informação obtida por SCADA relativamente à temperatura exterior nota-se que não existe uma grande diferença de temperatura em relação às outras semanas. No entanto, nota-se que a velocidade de rotação do rotor é muito reduzida para estas semanas.

As figuras 3.24a e 3.24b representam as curvas de potência da turbina 6 na semana 15 (semana com manutenção) e 21 (semana com falta de vento) respetivamente.

#### Modelo da turbina 7

Para a turbina 7, houve uma reparação do sistema de pitch na semana 46 e uma outra reparação na semana 51, como pode ser verificado na figura 3.25. A temperatura também poderia ser uma justificação para os erros elevados para esta turbina. No entanto, como para a turbina 6, não houve uma grande diferença da temperatura exterior nestas semanas.

As figuras 3.26a e 3.26b representam as curvas de potência da turbina 7 na semana 33 (semana com pouco vento) e 51 (semana com uma anomalia registada) respetivamente.

#### Modelo da turbina 8

Na turbina 8 houve uma reparação na caixa multiplicadora na semana 46 e duas manutenções na semana 21 e 39 como pode ser verificado na figura 3.27.

As figuras 3.28a e 3.28b representam as curvas de potência da turbina 8 na semana 21 e 46 respetivamente.

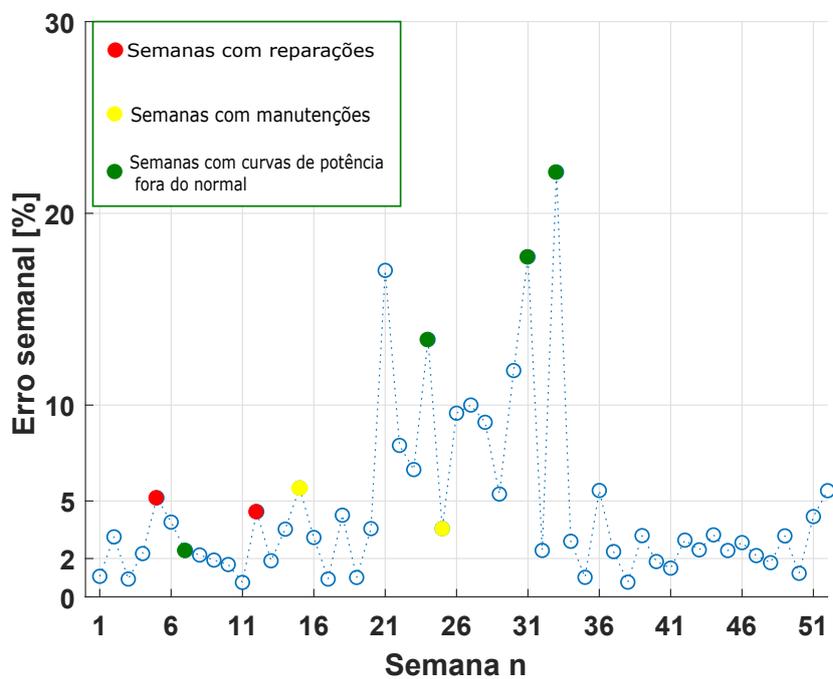


Figura 3.23: Erro semanal para a turbina 6 com modelo de regressão construído com dados da turbina 6 (2011)

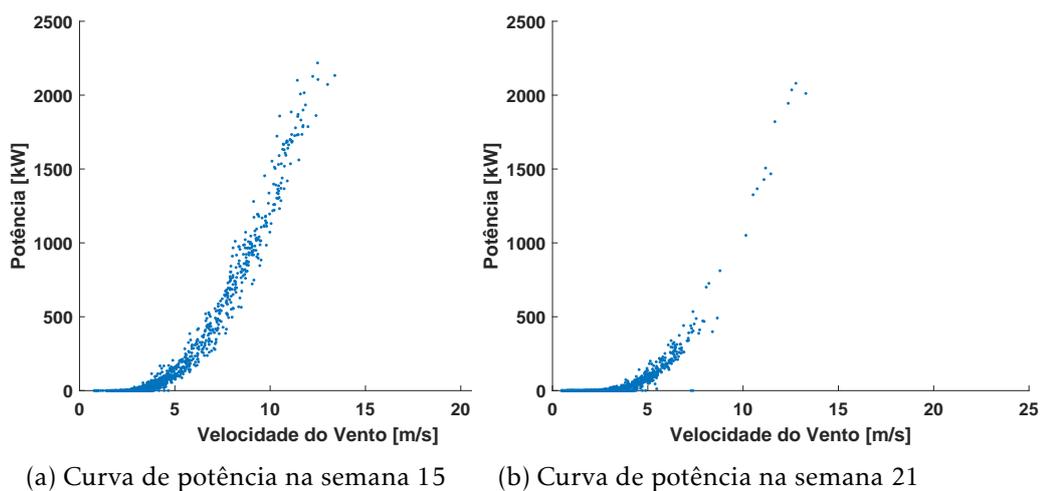


Figura 3.24: Curvas de potência da turbina 6 em semanas específicas (2011)

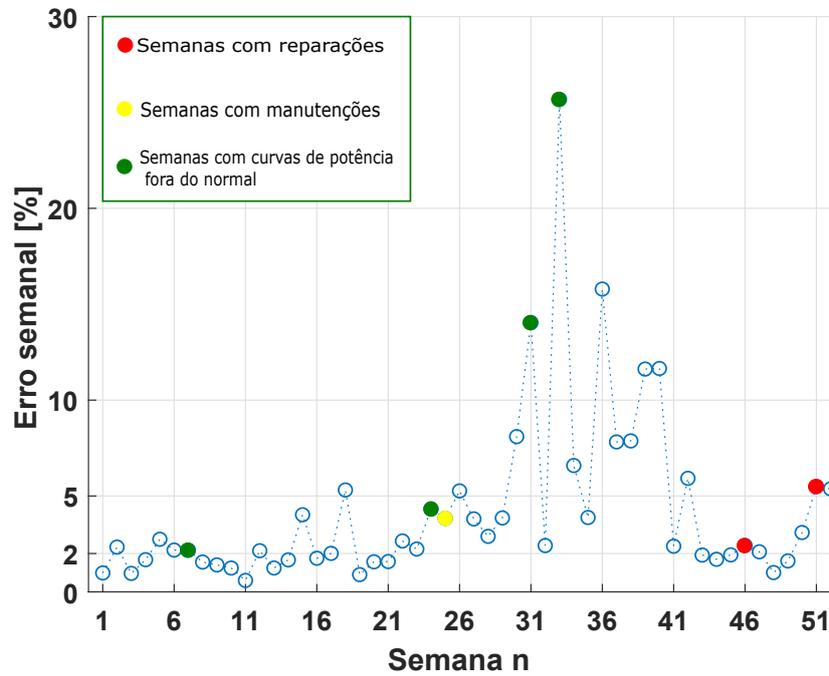
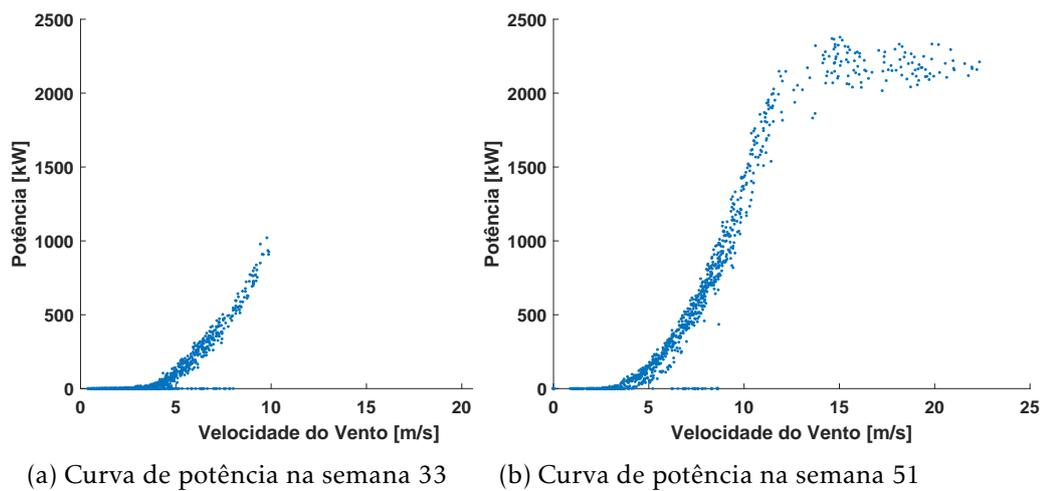


Figura 3.25: Erro semanal para a turbina 7 com modelo de regressão construído com dados da turbina 7 (2011)



(a) Curva de potência na semana 33

(b) Curva de potência na semana 51

Figura 3.26: Curvas de potência da turbina 7 em semanas específicas (2011)

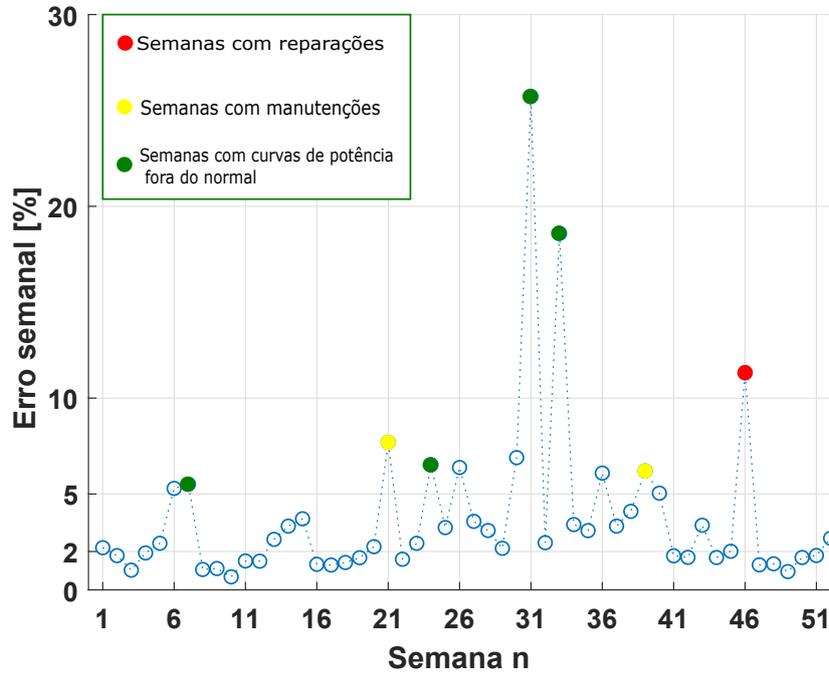
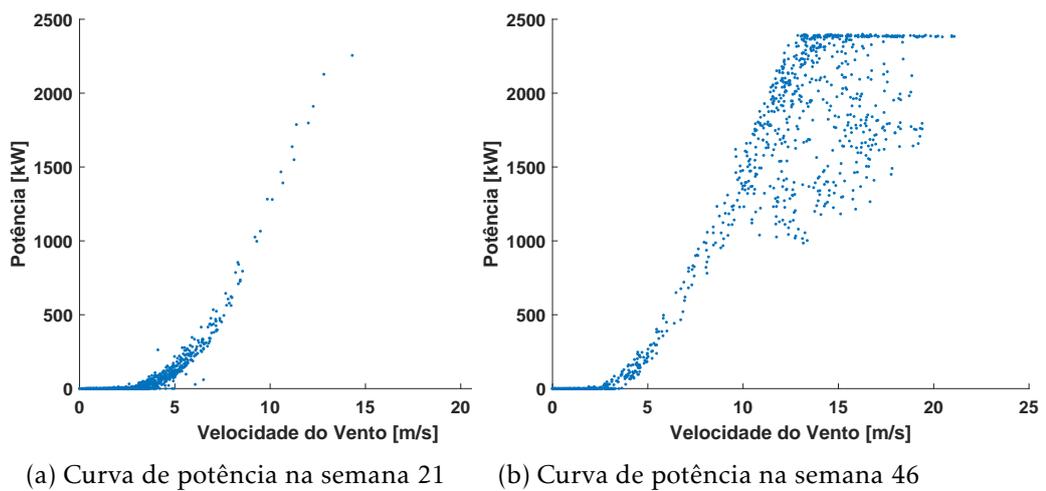


Figura 3.27: Erro semanal para a turbina 8 com modelo de regressão construído com dados da turbina 8 (2011)



(a) Curva de potência na semana 21

(b) Curva de potência na semana 46

Figura 3.28: Curvas de potência da turbina 8 em semanas específicas (2011)

### 3.3.2 Tentativa de previsão das anomalias no ano 2012

Os resultados anteriores foram satisfatórios, no entanto, o ideal seria adicionar o maior volume de informação ao modelo para este ser cada vez mais completo. Por questões computacionais e pelo caminho escolhido ao longo deste trabalho, haverá uma tentativa de previsão do ano de 2012 apenas com informação do ano 2011. Isto para tentar aproximar o estudo da realidade. Os pontos utilizados para a construção do modelo de regressão serão os mesmos dos da análise do ano 2011.

#### Modelo da turbina 1

A partir do registo de manutenção é possível verificar que para o ano 2012 não existiram praticamente nenhuma anomalia que pudesse interferir com a curva de potência. Apenas foram registadas uma manutenção na semana 22 e uma substituição das baterias de pitch na semana 46. Depois da comparação com a figura 3.29 na semana 14 a curva de potência apresenta um número reduzido de observações registadas. Nas semanas 29, 32 e 48 a curva de potência têm um número suficiente de observações no entanto o erro continua elevado o que nos leva a pensar que seja causado por condições atmosféricas no parque.

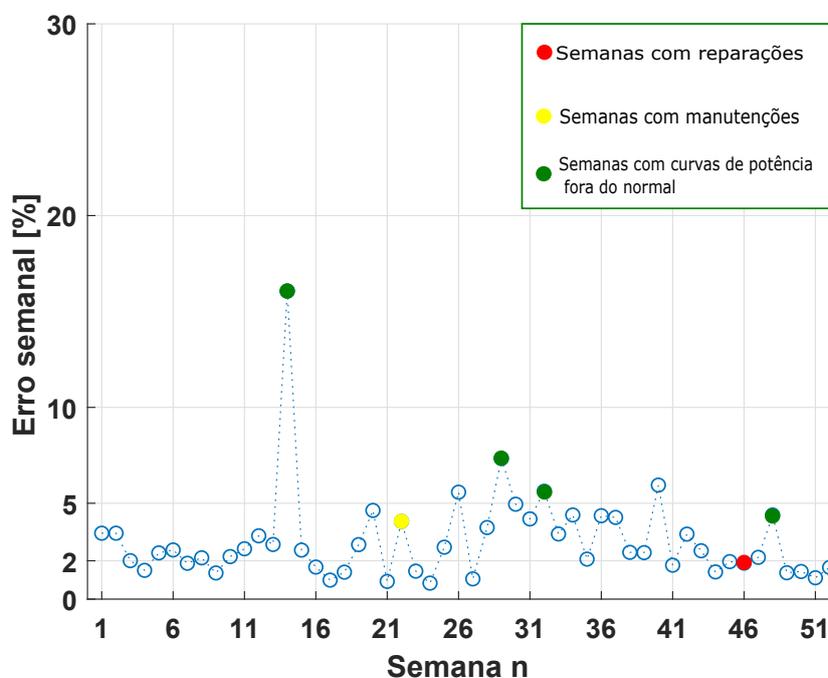


Figura 3.29: Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 1

### Modelo da turbina 2

Na figura 3.30 a semana 14 apresenta poucas observações registadas. Na semana 29, 32 e 48 o erro é justificado por condições atmosféricas no parque. Na semana 25 substituição do óleo nas caixas de *gam* e *pitch* da turbina, na semana 36 substituição de baterias e na semana 44 substituição da caixa de *gam*.

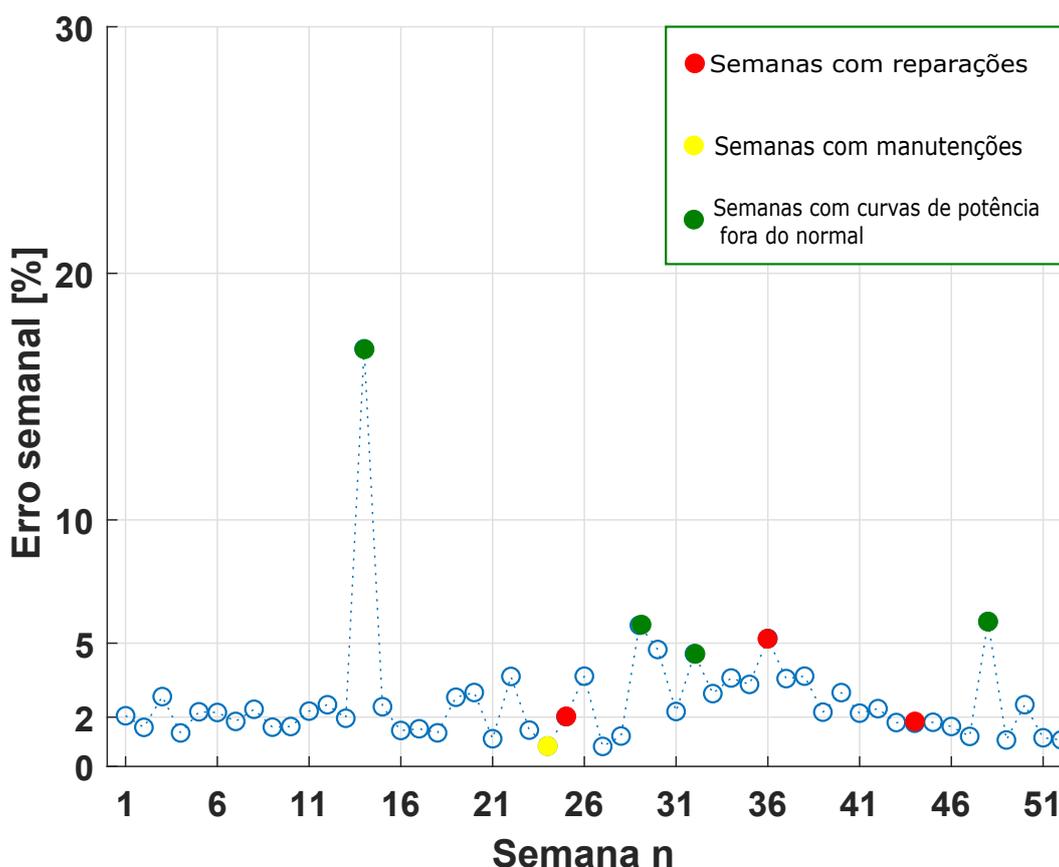


Figura 3.30: Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 2

### Modelo da turbina 3

A partir da figura 3.31 é possível verificar que os erros das semanas 14, 29, 32 e 48 continuam na turbina 3. No entanto, esta apresenta mais três semanas com a curva de potência com falta de observações registadas. O único fenómeno presente no registo de manutenção da turbina 3 é uma manutenção com substituição de caixas *gam* e *pitch* na semana 25.

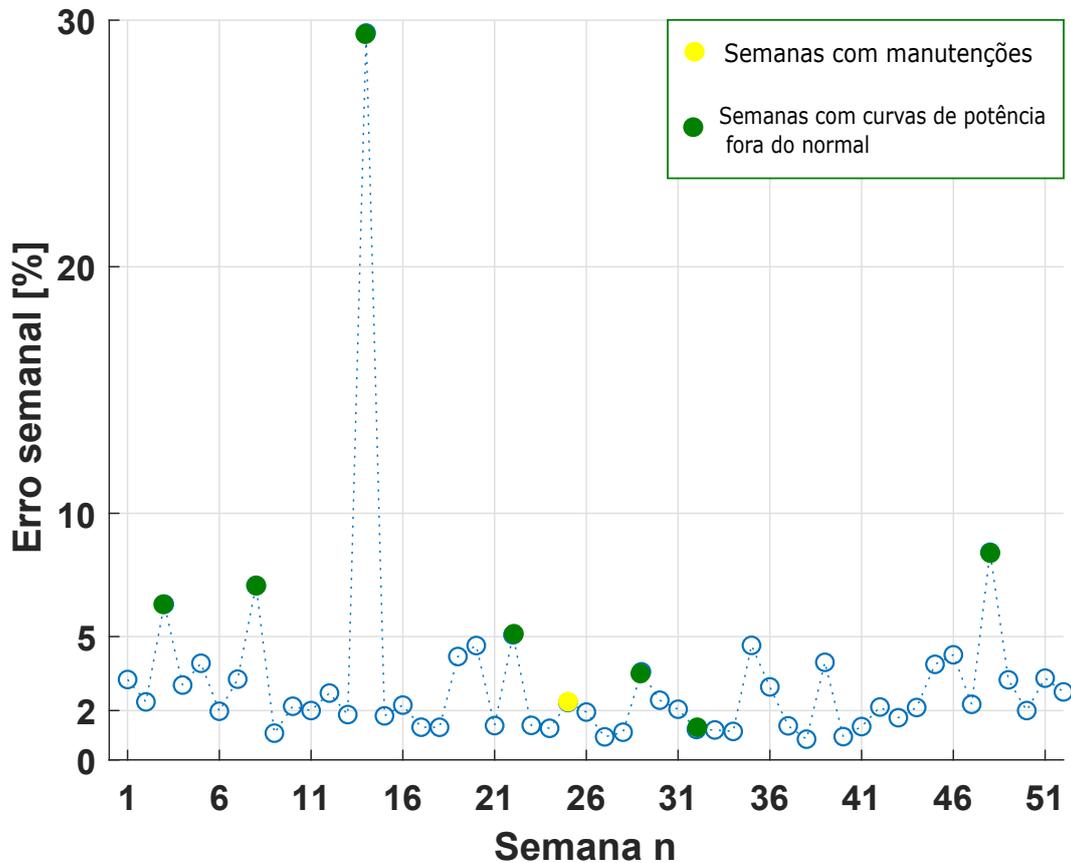


Figura 3.31: Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 3

#### Modelo da turbina 4

Com o registo de manutenção é possível verificar que houve uma manutenção na semana 26. Da semana 5 à semana 11 foram registadas várias paragens devido a um erro na turbina. Isto pode justificar os erros destas semanas na figura 3.32.

#### Modelo da turbina 5

Na turbina 5 não houve qualquer tipo de anomalia registada no ano de 2012 que afetasse a curva de potência, apenas tem uma manutenção registada na semana 26. Tirando os pontos que têm origem nas condições atmosféricas, é possível verificar que na figura 3.33 não existe erro para a semana 14 isto é porque o erro é muito superior aos restantes e por isso foi retirado do gráfico para simplificar a sua leitura. O erro da semana 22 deve-se ao número reduzido de pontos registados nessa semana.

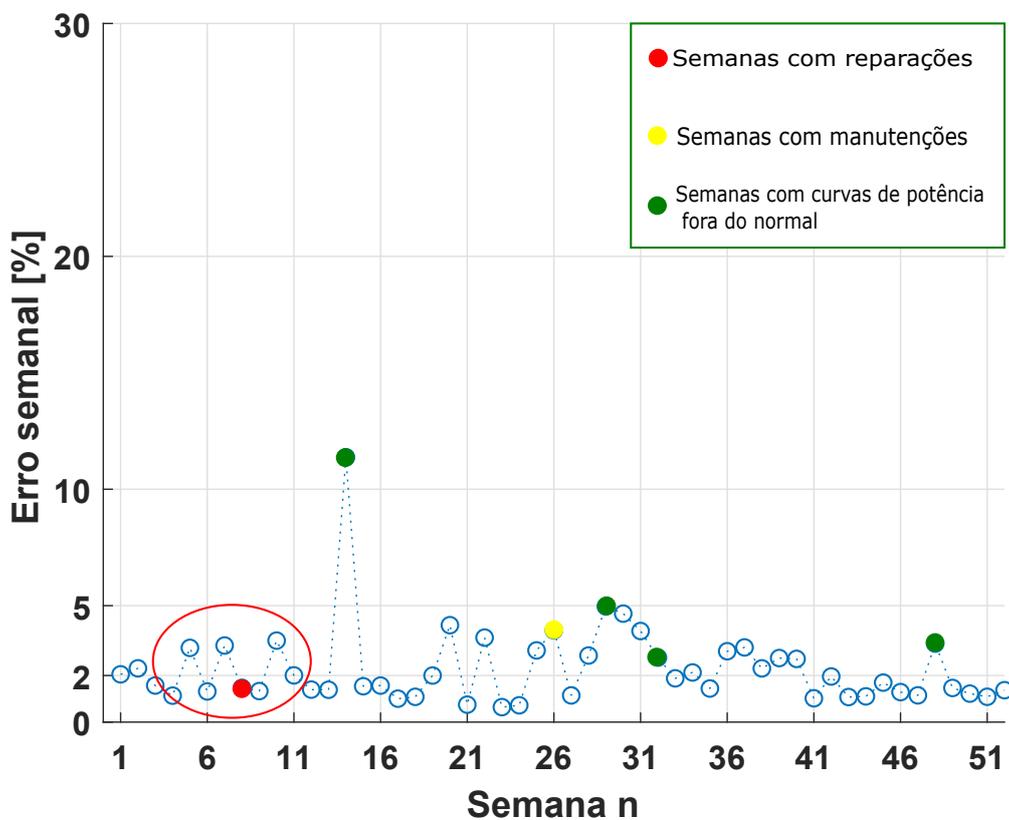


Figura 3.32: Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 4

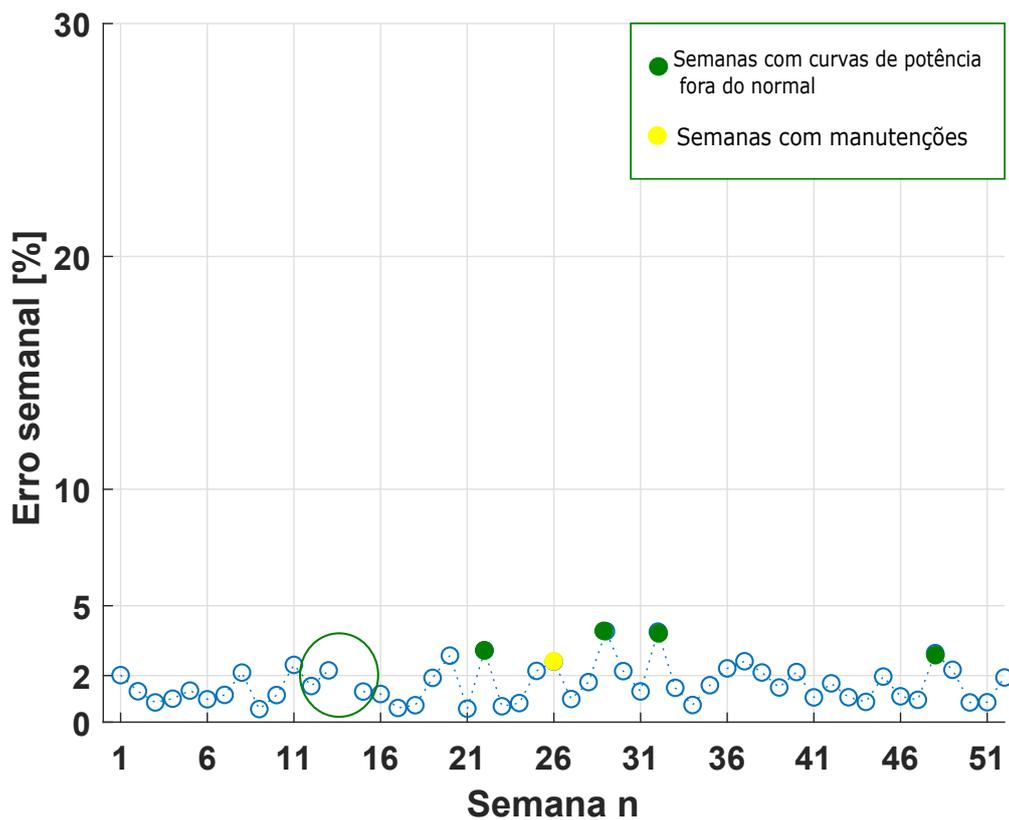


Figura 3.33: Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 5

### Modelo da turbina 6

É curioso que para o ano 2011 a turbina 6 foi uma das que apresentou maiores erros. Para o ano de 2012 esta teve o mesmo comportamento sendo que é uma das turbinas com os erros mais elevados como pode ser verificado na figura 3.34. Na semana 8 esta apresentou um erro mais elevado devido à falta de pontos na curva de potência. Na semana 22 o erro foi retirado do gráfico por ser demasiado elevado, no entanto este erro também é justificado pela falta de dados nessa semana. Da semana 25 à 40 o erro é elevado mas não foi possível determinar a razão para este acontecimento.

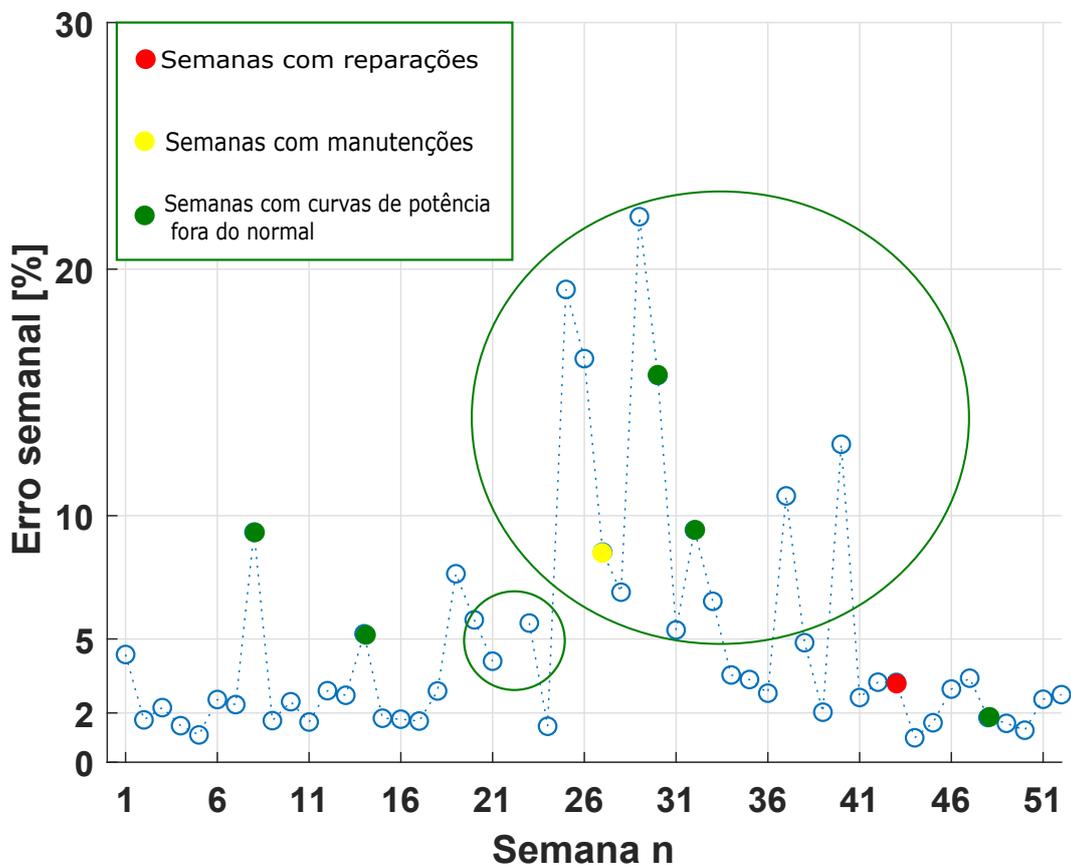


Figura 3.34: Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 6

### Modelo da turbina 7

Com recurso à figura 3.35 é importante salientar que os erros das semanas 3, 5 e 19 são erros causados por falta de observações nas respetivas . Na semana 8 houve uma substituição de rolamentos e na semana 15 problemas com *pitch*.

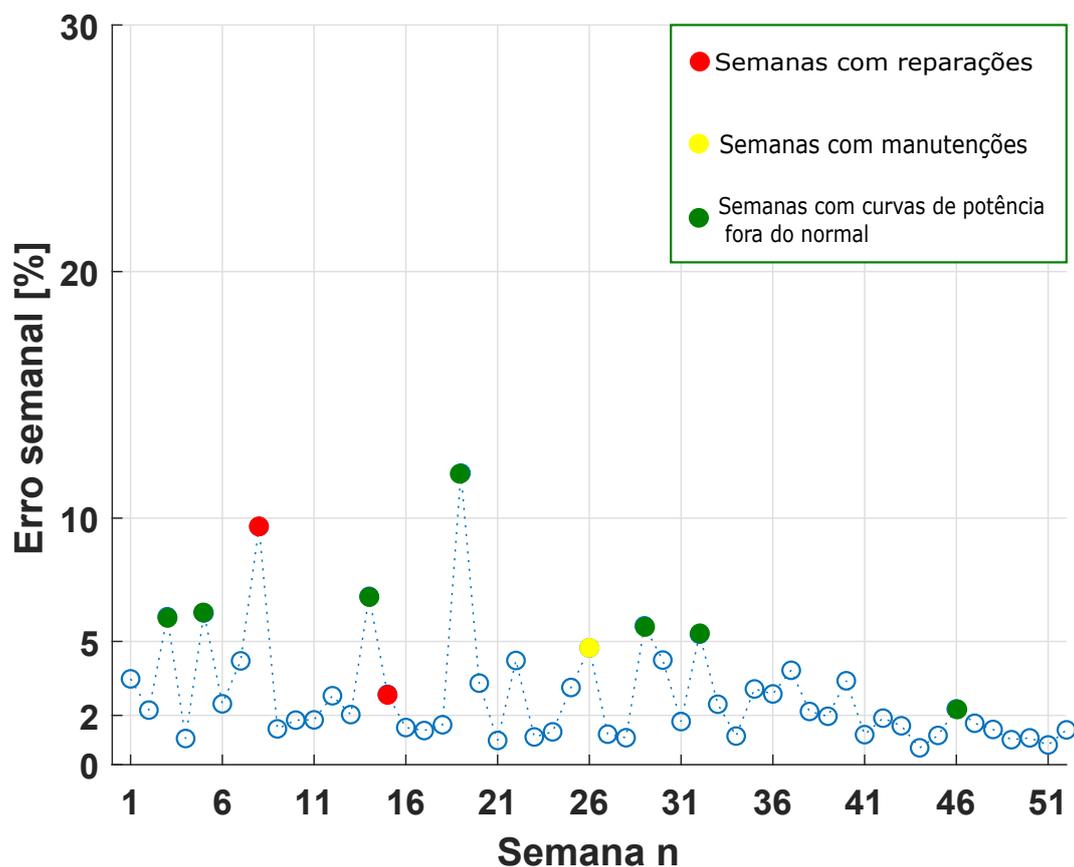


Figura 3.35: Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 7

### Modelo da turbina 8

Embora não haja informação sobre qualquer informação no registo de manutenção, a turbina 8 tem as curvas de potência das semanas posteriores à semana 21 com falta de vento o que explica os erros apresentados nestas semanas. Não há qualquer tipo de informação sobre o que possa causar isto visto que há pontos suficientes nos blocos de dados utilizados.

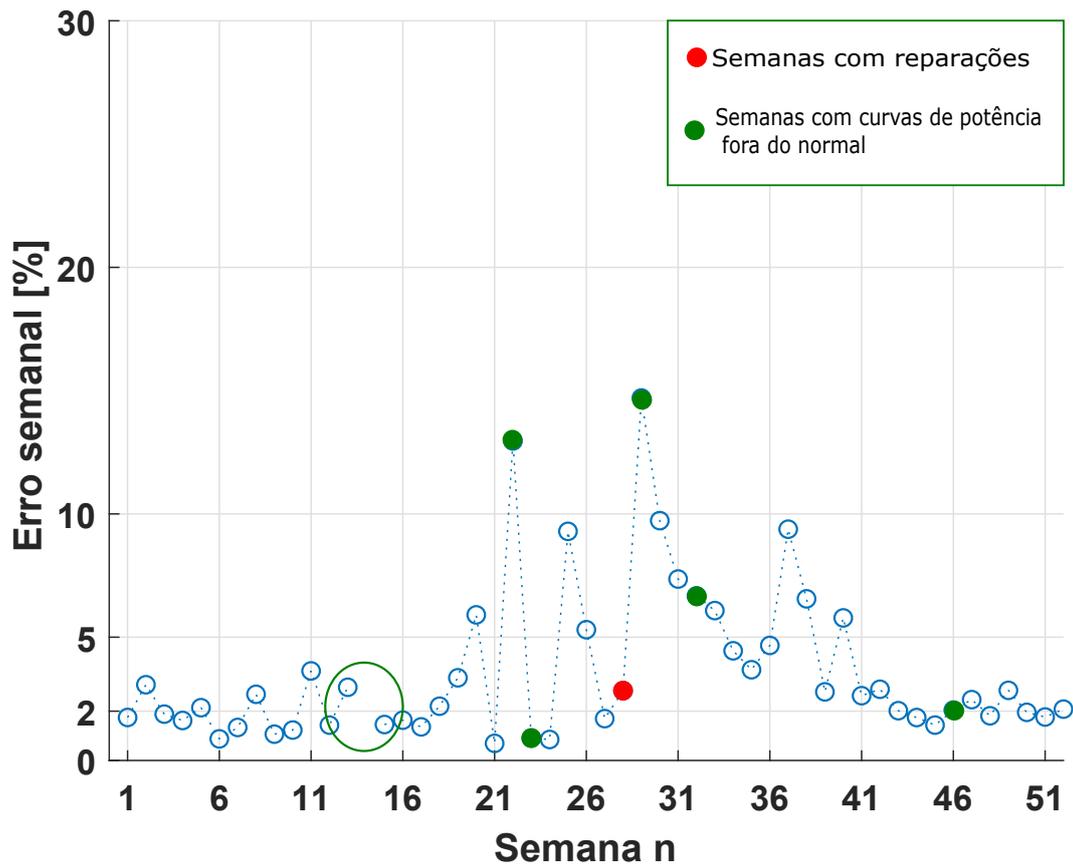


Figura 3.36: Previsão do ano 2012 com modelo de regressão construído com dados de 2011 da turbina 8

### 3.3.3 Tentativa de previsão das anomalias no ano 2013

Na secção 3.3.3 vai ser seguido o mesmo raciocínio da secção 3.3.2, ou seja, com a informação do ano 2011 haverá uma tentativa de previsão do ano de 2013. Para o ano de 2013 apenas temos informação sobre 49 semanas.

#### Modelo da turbina 1

Recorrendo ao registo de manutenção e ao gráfico da figura 3.37 foram encontradas reparações nas pás na semana 31 e 33. Uma manutenção na semana 29 e nas semanas 3, 10, 16, 21, 27 e 34 havia um número reduzido de observações.

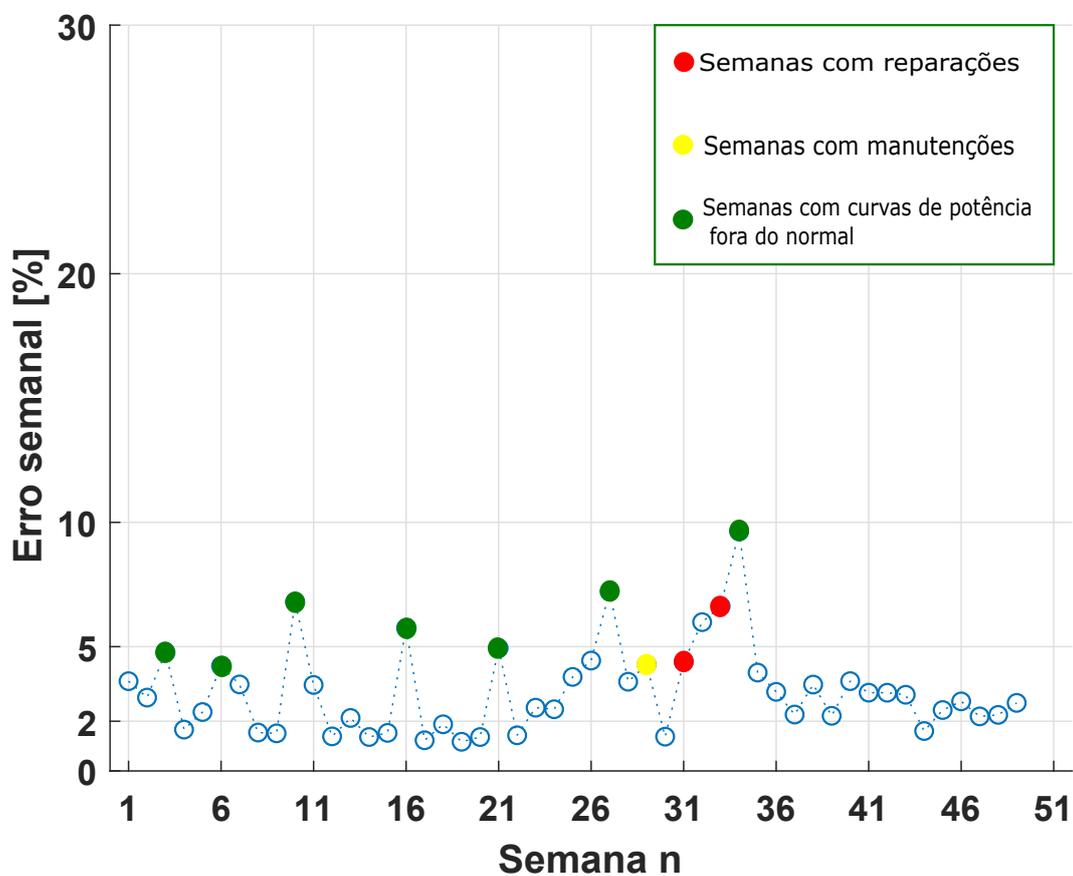


Figura 3.37: Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 1

### Modelo da turbina 2

A partir da figura 3.38 e do registo de manutenção é possível verificar que para a turbina 2 não existe nenhuma anomalia que afetasse a curva de potência, no entanto é possível detetar que nas semanas 3, 6 e 26 o número de observações foi reduzido o que aumenta o erro para essa semana. Para os erros elevados da semana 21 à 31 não existe qualquer justificação.

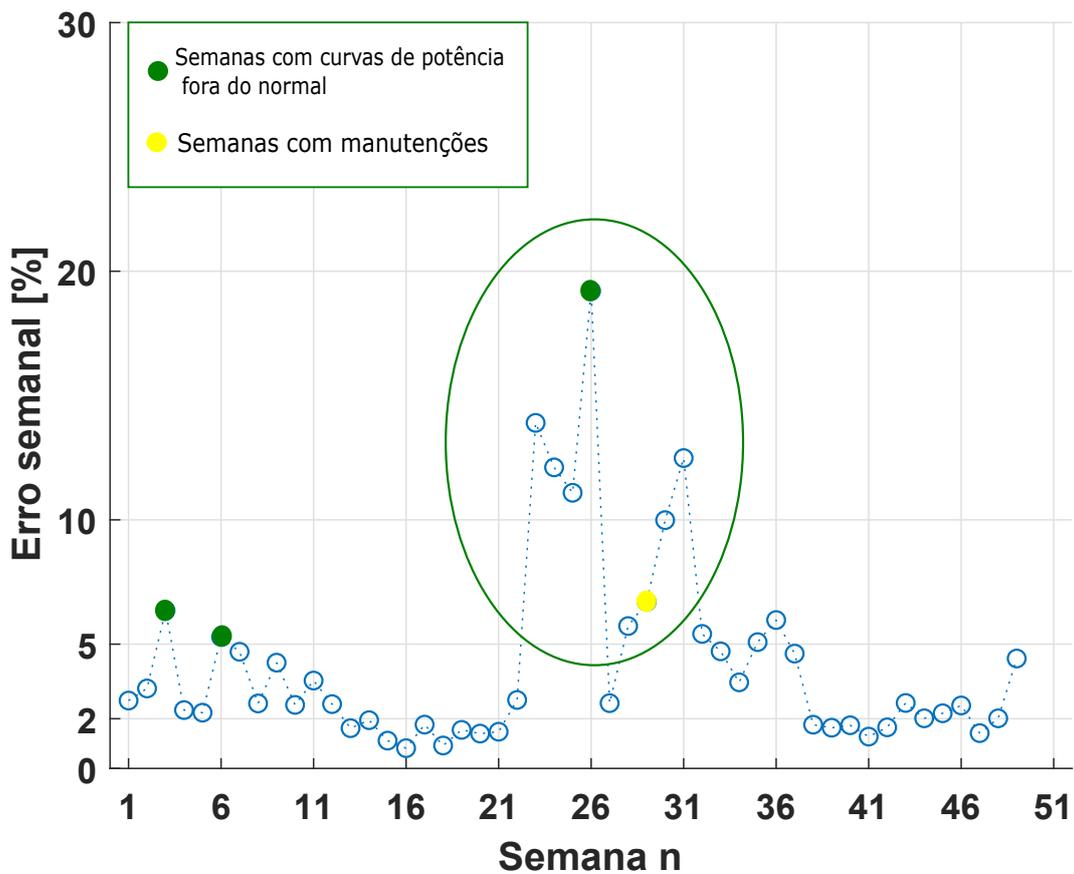


Figura 3.38: Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 2

### Modelo da turbina 3

Na figura 3.39 os erros elevados das semanas 3, 10, 26 e 29 são causados pela falta de observações nessas semanas. No entanto, existe uma manutenção na semana 28 e uma anomalia no sistema hidráulico na semana 39.

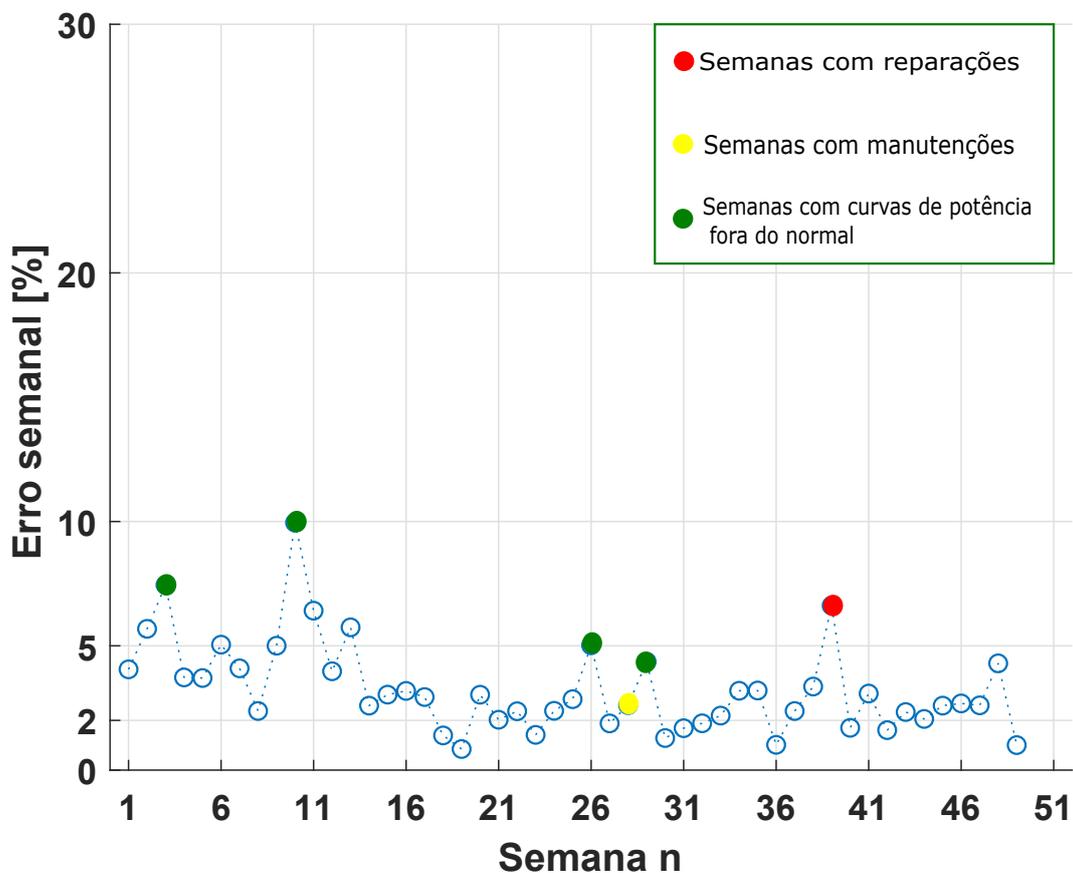


Figura 3.39: Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 3

#### Modelo da turbina 4

Na figura 3.40, os erros da semana 3, 6, 10, 16 e 28 deve-se a uma curva de potência com poucas observações. Na semana 21 uma anomalia com um relé de proteção.

#### Modelo da turbina 5

Para o ano 2013, a turbina 5 apresenta erros elevados no geral, como pode ser observado na figura 3.41. Há três anomalias a registrar para a turbina 5: na semana 16 foi detetado um ruído enquanto a turbina estava em funcionamento, na semana 34 uma anomalias das pás.

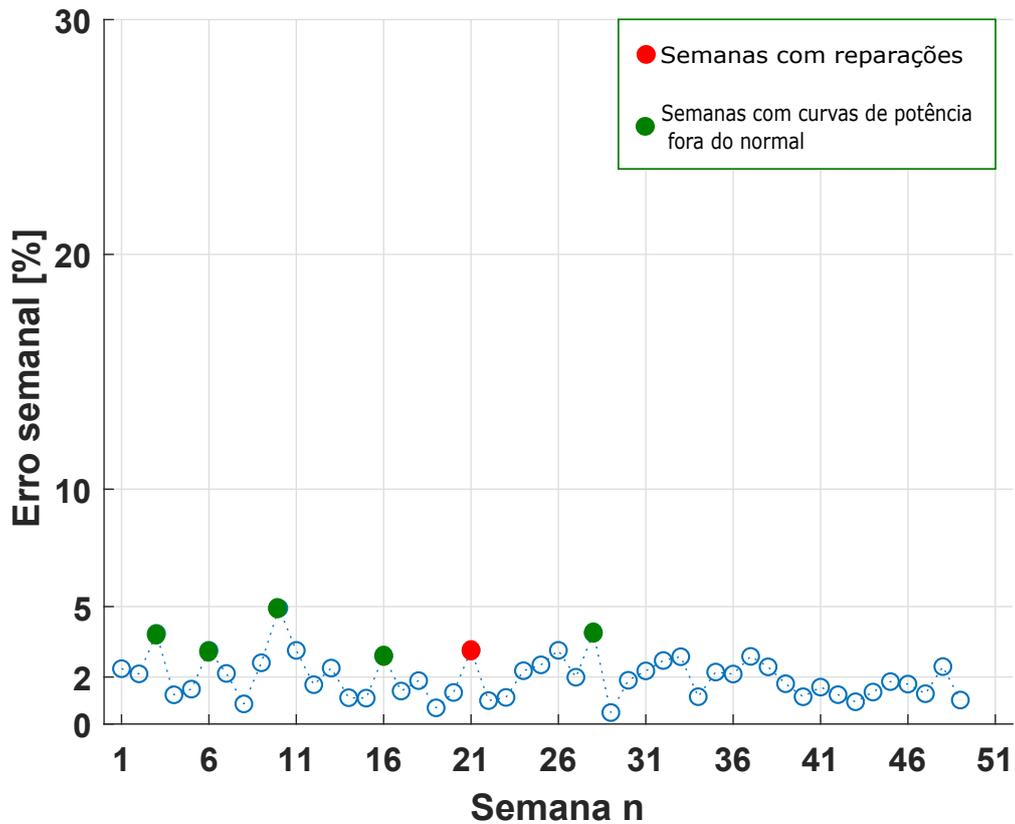


Figura 3.40: Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 4

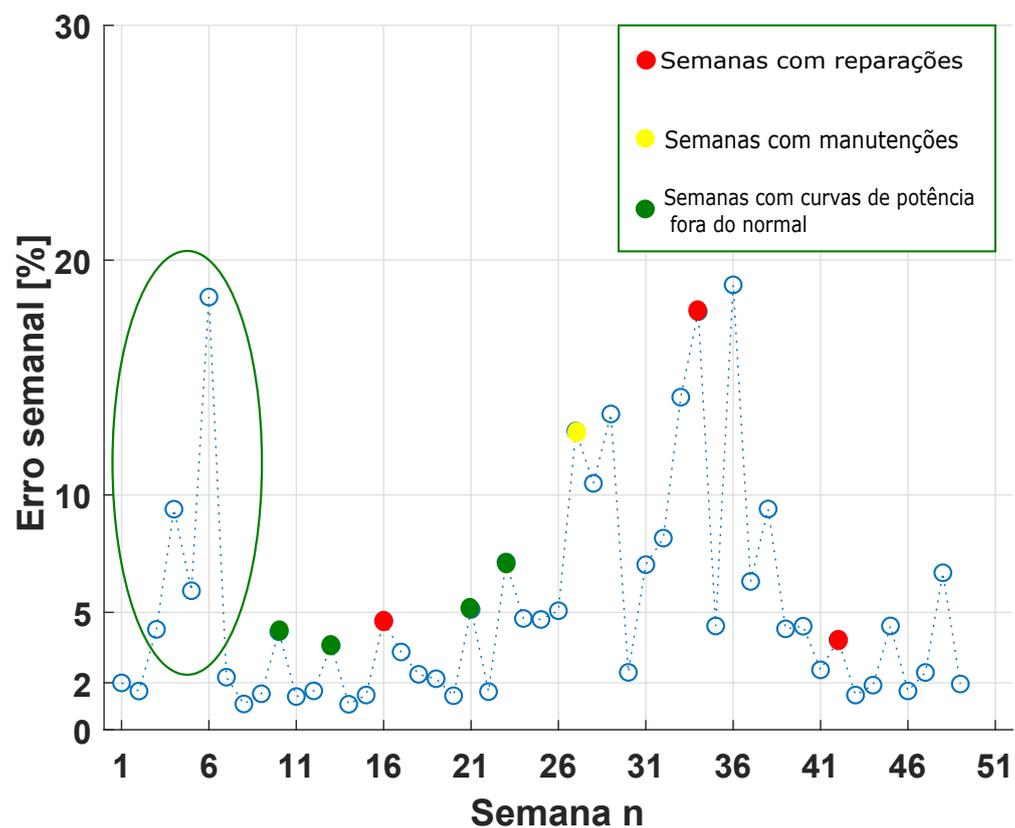


Figura 3.41: Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 5

### Modelo da turbina 6

Para a turbina 6, no ano de 2013, a turbina apresenta os erros elevados, no entanto no registo de manutenção não se encontra o motivo. Na figura 3.42 estão representados os erros semanais.

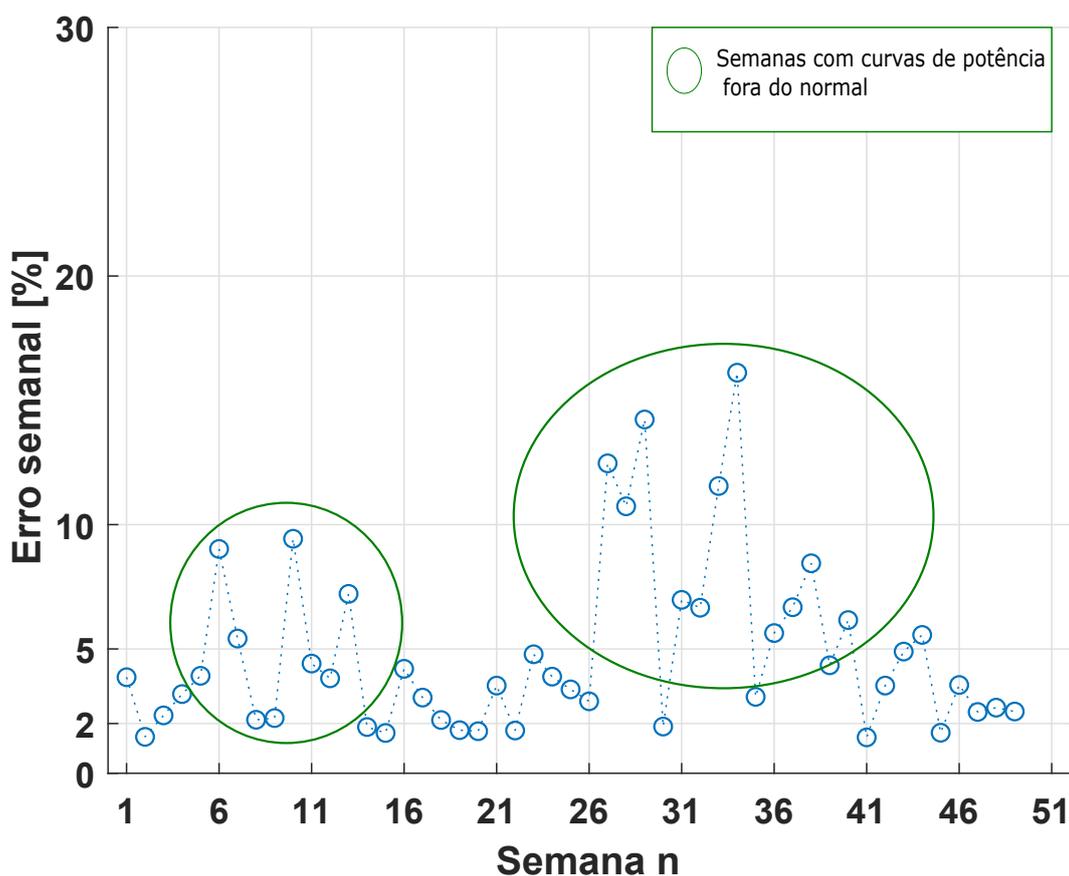


Figura 3.42: Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 6

### Modelo da turbina 7

Na figura 3.43 foi possível detetar uma manutenção na semana 27 e uma reparação na caixa de velocidades na semana 34. Esta anomalia é muito interessante porque percebe-se a evolução do erro até à semana da intervenção.

### Modelo da turbina 8

A turbina 8 não apresenta qualquer tipo de anomalia que possa afetar a curva de potência no ano de 2013. No entanto, existem umas semanas, representadas na figura 3.44, onde a curva de potência apresenta erros elevados.

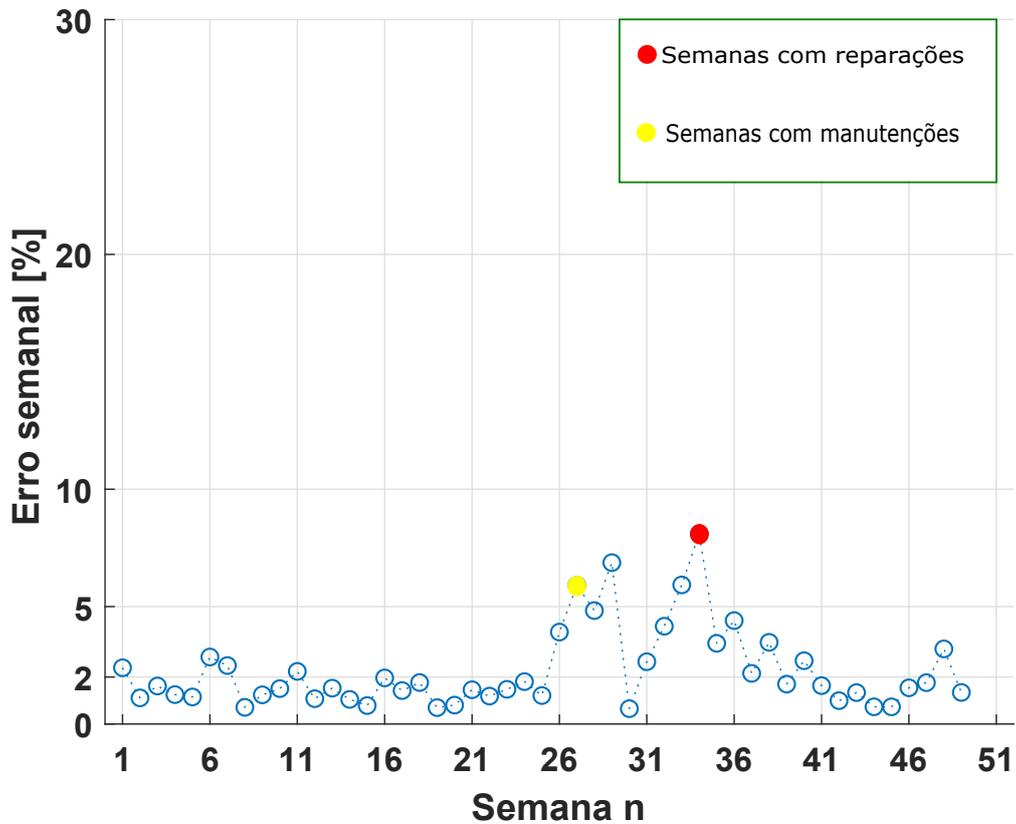


Figura 3.43: Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 7

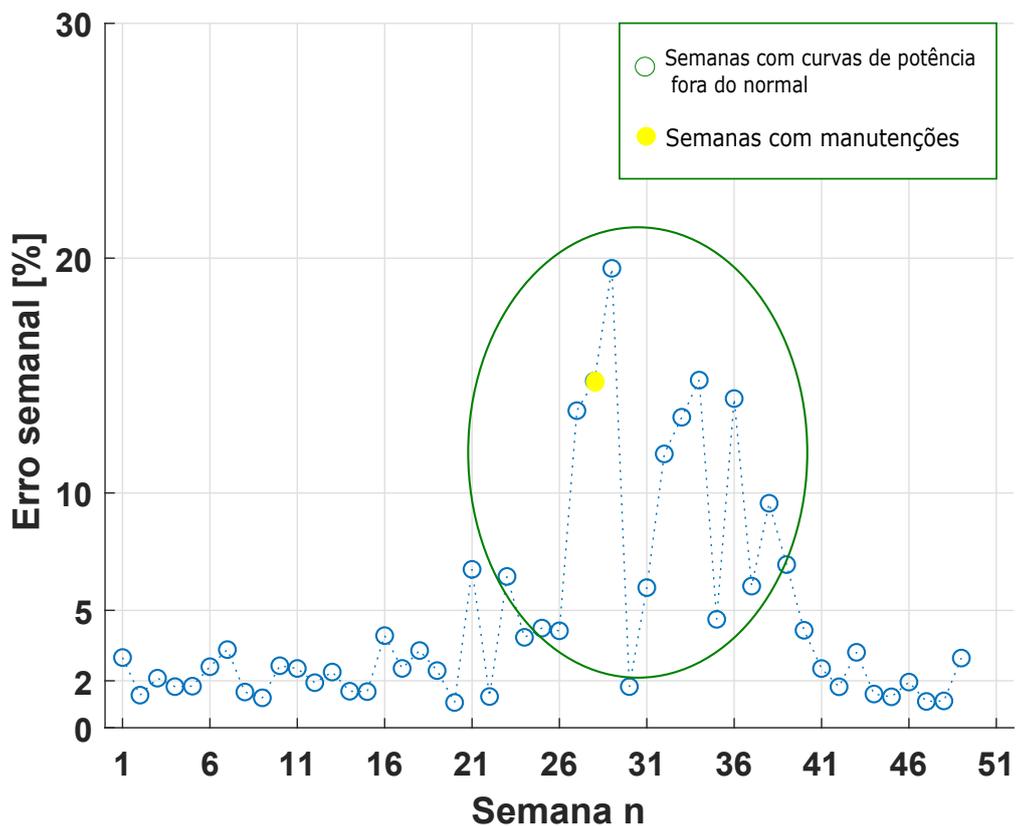


Figura 3.44: Previsão do ano 2013 com modelo de regressão construído com dados de 2011 da turbina 8

### 3.3.4 Conclusão

No geral, os resultados obtidos foram satisfatórios, permitindo a identificação de anomalias. É preciso ter em conta que todas as turbinas estão sujeitas a condições diferentes. Dada a orografia do terreno e a sua disposição no parque eólico é adequado que se faça este estudo individualmente para cada turbina. Quanto aos resultados, embora seja possível identificar anomalias pelo erro, outras não se conseguem. Os erros elevados podem ser resultados de duas situações, ou a curva de potência da turbina está longe da criada pelo modelo de regressão ou, para a dada semana, existirem poucos pontos, por exemplo quando a turbina está parada. Também é preciso perceber que o facto de se estar a trabalhar à semana pode acontecer que a anomalia ocorra nos primeiros dias da semana mas o erro semanal mantém-se baixo porque o resto da semana a turbina está a funcionar corretamente. Um resultado que também é importante é o facto de existirem erros elevados em todas as turbinas para as mesmas semanas por não haver vento na zona onde o parque está situado como por exemplo as semanas 31 e 33 para 2011 como pode ser verificado na figura 3.13 e nas figuras equivalentes das restantes turbinas.

## 3.4 Tentativa de previsão de uma turbina com o modelo de outra

---

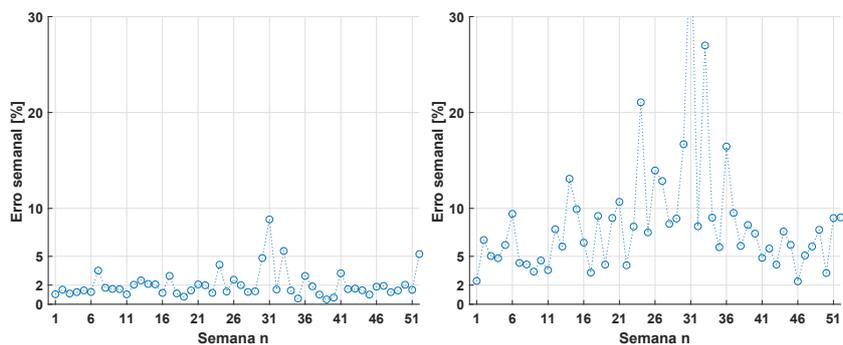
Na secção 3.4 o objetivo é tentar verificar se é possível com um modelo de uma turbina prever o estado de funcionamento de outra. O intuito deste "cruzamento de modelos" tem a haver com a possibilidade de termos um novo equipamento para o qual não temos histórico e, desta forma, ter uma indicação da condição desse equipamento. A figura 3.45 mostra a comparação entre os erros de previsão dos modelos de regressão para as turbinas 3 a 6 considerando os dados observados para cada um destes equipamentos e os mesmos considerando o modelo de regressão construído para os dados da turbina 1 apenas. A comparação é efectuada para o ano de 2011. É importante verificar que como seria expectável, a previsão não é igual porque todas as turbinas estão sujeitas a condições diferentes de funcionamento. Isto confirma a decisão de se estudar cada turbina individualmente a partir da criação de um modelo de regressão dedicado para cada.

A partir das figuras 3.45a e 3.45b é possível verificar que o erro subiu no geral quando usado o modelo de regressão com dados de outra turbina. No entanto, os erros mais elevados coincidem com as mesmas semanas, como por exemplo as semanas 31 e 24.

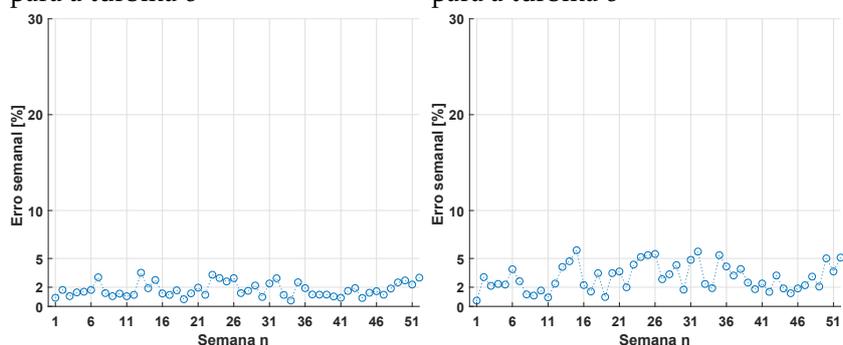
Nas figuras 3.45c a 3.45f o erro não varia com a mesma magnitude que o exemplo anterior o que nos leva a pensar que talvez fosse possível determinar o funcionamento das turbinas

4 e 5 a partir do modelo da turbina 1.

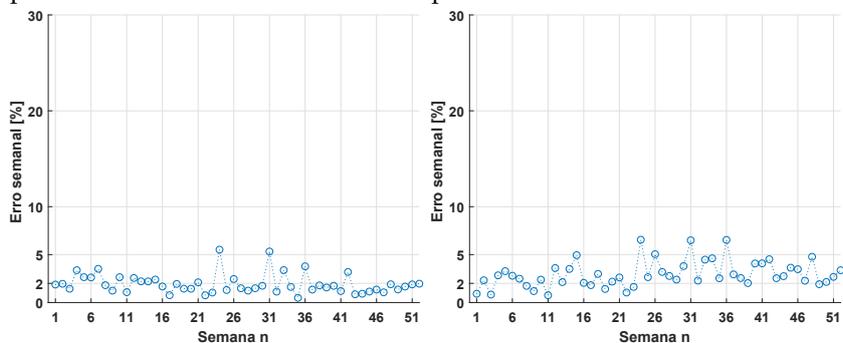
Por último, nas figuras 3.45g e 3.45h o erro é menor quando se usa o modelo de regressão construído com dados da turbina 1 para prever a turbina 6. Isto não significa que o modelo da turbina 6 esteja mal definido, o que significa é que o comportamento desta turbina neste ano é parecido ao comportamento normal da turbina 1.



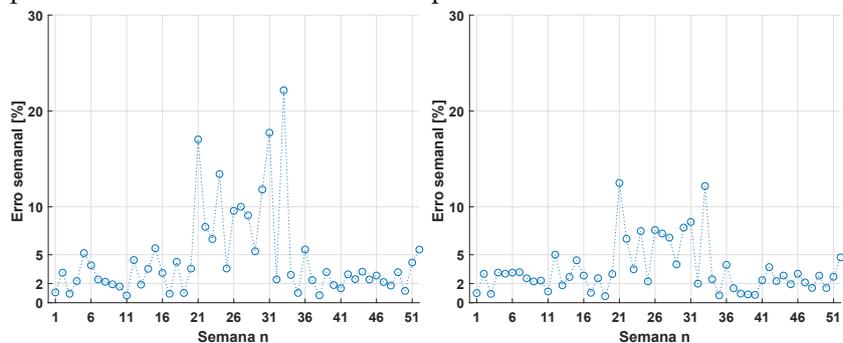
(a) Erro da previsão do modelo 3 para a turbina 3 (b) Erro da previsão do modelo 1 para a turbina 3



(c) Erro da previsão do modelo 4 para a turbina 4 (d) Erro da previsão do modelo 1 para a turbina 4



(e) Erro da previsão do modelo 5 para a turbina 5 (f) Erro da previsão do modelo 1 para a turbina 5



(g) Erro da previsão do modelo 6 para a turbina 6 (h) Erro da previsão do modelo 1 para a turbina 6

Figura 3.45: Comparação dos erros de previsão para as turbinas 3 a 6 considerando o modelo de regressão construído para os dados de cada uma destas turbinas e o construído para a turbina 1 (2011)

# 4

## Conclusão

No presente existe um aumento na procura de novas técnicas de manutenção relacionadas com as turbinas eólicas. Este trabalho tem como objetivo melhorar a manutenção de um parque eólico, com a criação de modelos de regressão para prever o comportamento de uma turbina, e assim, tentar detectar o seu funcionamento anómalo.

Os resultados finais deste estudo foram satisfatórios. É preciso ter em conta que este tema ainda é bastante recente e para uma fase inicial era importante conseguir perceber o como funcionava e como se aplicava o processo de regressão. Foram tomadas decisões para tentar criar um modelo mais correto possível. Como foi referido ao longo deste trabalho quanto mais informação o modelo tiver melhor será a sua previsão, no entanto, houve um limite máximo de informação fornecida ao programa devido à carga computacional limitada. Mas tendo em conta todas as limitações os resultados foram positivos e acredito que será possível abordar este tema por outros caminhos. Em termos da detecção de anomalias também é preciso ter em conta que não se pode garantir que todas as anomalias sejam detetadas. No entanto, foi possível justificar grande parte dos erros mais elevados da previsão do processo de regressão, seja por falta de elementos nos blocos de dados, curva de potência fora do normal, manutenções feitas pela empresa ou anomalias.

### 4.1 Trabalhos futuros

---

Como trabalhos futuros proponho uma análise feita com mais capacidade computacional para ser possível adicionar mais elementos de treino ao modelo de regressão e para o processo de treino ser mais rápido. Também pode ser interessante, treinar o modelo com

informação adicional que influencie o funcionamento da turbina, como por exemplo, a temperatura ambiente, entre outras. As funções de média e covariância também podem ser alteradas para tentar perceber se é possível uma melhor aproximação.

Este trabalho foi realizado numa perspetiva da engenharia mecânica, podendo gerar um novo conjunto de ideias e conceitos ligados à aprendizagem computacional.

# Bibliografia

- Burton, T., Jenkins, N., Sharpe, D., Bossanyi, E., 2011. Wind Energy Handbook. John Wiley & Sons.
- Dervilis, N., Papatheou, E., Antoniadou, I., Cross, E.J., Worden, K., 2016. On the usage of active learning for SHM. ISMA 16 , 14–17.
- IEC 61400-12-1, 2005. Wind energy generation systems-Part 12-1: Power performance measurements of electricity producing wind turbines. International Electrotechnical Commission .
- LIFE-BRIO project, . Demonstration of wind turbine rotor blade recycling into the coal clough wind farm decommissioning opportunity. URL: <http://www.lifebrio.eu/index.php/the-life-brio-project-2/>.
- Papatheou, E., Dervilis, N., Maguire, A.E., Campos, C., Antoniadou, I., Worden, K., 2017. Performance monitoring of a wind turbine using extreme function theory. Renewable Energy 113, 1490–1502. doi:10.1016/j.renene.2017.07.013.
- Rasmussen e Williams, C., 2004. Gaussian processes for machine learning. International journal of neural systems 14, 69–106. doi:10.1142/S0129065704001899.
- Simões, J., 2016. Controlo de Condição de Componentes de Turbinas Eólicas. Tese de Mestrado. FCT, Universidade Nova de Lisboa .
- Wang, K.S., Sharma, V.S., Zhang, Z.Y., 2014. (scada) data based condition monitoring of wind turbines. Advances in Manufacturing 2, 61–69. doi:10.1007/s40436-014-0067-0.



## Apêndice A

### Anexos

#### A.1 Detecção de dados omissos

---

```
% Carrega o bloco sem dados omitidos %
dadoscorretos=xlsread('frt_11.xls',1,'F:F');
% Carrega o bloco para verificar se está completo %
dadoserrados=xlsread('frt_13.xls','F:F');
% Ciclo para detetar falhas no bloco de dados %
for i=1:52560
if dadoscorretos(i)~=dadoserrados(i);
disp(i)
return
end
end
```

#### A.2 Filtragem dos dados

---

```
% Iniciar ciclo para as oito turbinas e para os três anos %

for turbina=1:8
for ano=2011:2013
```

```
for semana=1:52
cd E:\TESE\MAIN\Estudo\Estudo2\dados_semanais
clear dados

% Carregar os blocos de dados completos %

filename=strcat('Turbina',num2str(turbina),'(',num2str(semana),'-',num2str(ano),')');
load(filename)

% Filtragem dos dados filtrados %

s=size(teste_semanal);s=s(1,1);cnt=0;
for i=1:s
if teste_semanal(i,2)>=5 && teste_semanal(i,2)<=12 && teste_semanal(i,4)>=9.6 && teste_se
cnt=cnt+1;
dados(cnt,1)=teste_semanal(i,1);
dados(cnt,2)=teste_semanal(i,2);
end
end

% Guardar bloco de dados filtrado %

cd E:\TESE\MAIN\Estudo\Estudo2\dados_filtrados
save(filename,'dados')
end
end
end

cd E:\TESE\MAIN\Estudo\Estudo2
```

### A.3 Definição dos blocos de dados usados para o treino do modelo

---

```
function definirPtrreino = definirPtrreino(turbinas)
```

### A.3 - Definição dos blocos de dados usados para o treino do modelo

---

```
% Definição das semanas que serão utilizadas %

if turbinas==1
curvas=[semana ano ; semana ano ; semana ano];
elseif turbinas==2
curvas=[semana ano ; semana ano ; semana ano];
elseif turbinas==3
curvas=[semana ano ; semana ano ; semana ano];
elseif turbinas==4
curvas=[semana ano ; semana ano ; semana ano];
elseif turbinas==5
curvas=[semana ano ; semana ano ; semana ano];
elseif turbinas==6
curvas=[semana ano ; semana ano ; semana ano];
elseif turbinas==7
curvas=[semana ano ; semana ano ; semana ano];
elseif turbinas==8
curvas=[semana ano ; semana ano ; semana ano];
end
s=size(curvas);s=s(1,1);

% Carrega o bloco de dados da respectiva semana %

for i=1:s
semana=curvas(i,1);
ano=curvas(i,2);
filename=strcat('Turbina',num2str(turbinas),'(',num2str(semana),'-',num2str(ano),')');
cd E:\TESE\MAIN\Estudo\Estudo2\dados_filtrados
load(filename)

% Criação de uma nova variável se for a primeira semana %

if i==1
pontostreino=dados;
else

% Juntar o blocos de dados à variável previamente criada %

pontostreino=[pontostreino;dados];
```

54

```
end
end
cd E:\TESE\MAIN\Estudo\Estudo2

% Guarda os blocos de dados que serão utilizados para o treino %

save('pontosparatreino','pontostreino')
```

## A.4 Treino

---

```
% Carregar os pontos para o treino do modelo %

load pontosparatreino
x=pontostreino(:,2);
y=pontostreino(:,1);

% Definição das funções de média e covariância %

meanfunc=@meanZero; covfunc=@covSEiso; likfunc = @likGauss;

% Definição aleatória dos parâmetros para serem otimizados %

hyp = struct('mean', [], 'cov', [0 0], 'lik', -1);

% Otimização dos parâmetros da regressão %

hyp2 = minimize(hyp, @gp, -2000, @infGaussLik, meanfunc, covfunc, likfunc, x, y);

% Valores máximos e mínimos da velocidade do vento e potência produzida %

mx=max(x(:));
xm=min(x(:));
my=max(y(:));
ym=min(y(:));
```

```

% Verificação dos blocos de dados para detetar erros %

exp(hyp2.lik);
K = feval(covfunc, hyp2.cov, x);

% Divisão do espaço a definição do modelo %

z = linspace(xm, mx, 401)';

% Criação do modelo %

n1m1 = gp(hyp2, @infGaussLik, [], covfunc, likfunc, x, y)
[m s2] = gp(hyp2, @infGaussLik, [], covfunc, likfunc, x, y, z);

% Guardar as variáveis importantes %

save('treino', 'hyp2', 'm', 's2', 'z', 'meanfunc', 'covfunc', 'likfunc', 'x', 'y')

```

## A.5 Verificação

---

```

function verificacao = verificacao(turbinas)

%Início do ciclo para as cinquenta verificações %

for i=1:50
clearvars -except i erroverificacao turbinas

% Carregar os pontos e resultados do treino da regressão %

load pontosparatreino
load treino

```

## A.5 - Verificação

---

```
% Selecionar os pontos aleatórios do bloco de dados utilizados no treino %

erroverificacao(i,1)=i
tamanho=size(pontostreino);
tamanho=tamanho(1,1);
pteste=tamanho*0.5;
pteste=round(pteste);
k=randperm(tamanho,pteste);
teste1=pontostreino(k,:);
r=true(1,tamanho);
r(k)=false;
nteste=pontostreino(r,:);

% Previsão do modelo para o bloco de dados selecionado anteriormente %

xs=teste1(:,2);
ys=teste1(:,1);
[y_mu ys2] = gp(hyp2, @infGaussLik, [], covfunc, likfunc, x, y, xs);

% Cálculo do erro %

vy=var(ys,1);
c=(ys-y_mu);d=c.^2;N=size(d);s=sum(d);
erroverificacao(i,2)=(50*sum(d))/(N(1,1)*vy);
end

% Criação do registo dos erros %

t=table(erroverificacao);
writetable(t,'erroverificação')
save('erroverificacao','erroverificacao')

% Representação %

figure
axes('Parent',gcf,'FontWeight','bold','FontSize',12), hold on
plot(erroverificacao(:,1),erroverificacao(:,2),'o')
axis([0 50 0 15])
xlabel('Verificação n')
```

```

ylabel('Erro')
yticks([0 1 2 5 10 15])
gravar=strcat('verificacao',num2str(turbinas),'(2)');
gravafig(gcf,gravar)
cd E:\TESE\MAIN\Estudo\Estudo2

```

## A.6 Previsão

---

```

function previsao = previsao(turbinas)
close all

% Seleciona a turbina para testar o modelo da variavel turbinas %

for turbina=1:8

% Ciclo da previsão para cada semana %

for i=1:52
clearvars -except i erroteste turbina turbinas
ano_falha=2011;
semana_falha=0+i;
erroteste(i,1)=semana_falha;
filename=strcat('Turbina',num2str(turbina),'(',num2str(semana_falha),'-',num2str(ano_falha)

% Carregar o bloco de dados para previsão %

cd E:\TESE\MAIN\Estudo\Estudo2\dados_filtrados
load(filename)
cd E:\TESE\MAIN\Estudo\Estudo2

% Carregar os resultados do treino da regressão %

load treino

```

## A.6 - Previsão

---

```
tamanho=size(dados);
tamanho=tamanho(1,1);
pteste=tamanho;
if tamanho<=pteste
pteste=tamanho;
end

% Selecionar os pontos do bloco de dados, todos foram utilizados mas esta fórmula permite

k=randperm(tamanho,pteste);
teste1=dados(k,:);
r=true(1,tamanho);
r(k)=false;
nteste=dados(r,:);

xs=teste1(:,2);
ys=teste1(:,1);
pontos=size(xs);pontos=pontos(1,1);

% Previsão %

[ymu ys2] = gp(hyp2, @infGaussLik, [], covfunc, likfunc, x, y, xs);

% Cálculo do erro %

vy=var(ys,1);
c=(ys-ymu);d=c.^2;N=size(d);s=sum(d);
erroteste(i,2)=(50*sum(d))/(N(1,1)*vy);
erroteste(i,3)=vy
erroteste(i,4)=pontos

% Retirar erros demasiado elevados %

if erroteste(i,2)>=35
erroteste(i,2)=NaN
end
end

% Guardar informação sobre o erro %
```

```
t=table(erroteste);
writetable(t,'erroteste')
filename=strcat(num2str(turbinas),num2str(turbina),'erro(2)');
cd E:\TESE\MAIN\Estudo\Estudo2\erros
save(filename,'erroteste')
cd E:\TESE\MAIN\Estudo\Estudo2

% Representação do erro semanal %

figure
axes('Parent',gcf,'FontWeight','bold','FontSize',12), hold on
plot(erroteste(:,1),erroteste(:,2),'o')
axis([0 52 0 30]) %Normalmente limite 30=20
xlabel('Semana n')
ylabel('Erro semanal')
grid on
xticks(1:5:52)
yticks([0 2 5 10 20 30])
figura=strcat(num2str(turbinas),num2str(turbina),'errosemanal',num2str(ano_falha),'(2)');
gravafig(turbina,figura)
cd E:\TESE\MAIN\Estudo\Estudo2
end
close all
```

## A.6 - Previsão

---