

4

APLICAÇÃO DO MODELO STAR-TREE

4.1

Introdução

Uma vez conhecida a estrutura de uma árvore STAR, é possível estimá-la através do processo detalhado no capítulo anterior. Porém, esta estrutura é desconhecida de antemão, e a estimação dos parâmetros não-lineares dos nós de decisão e dos parâmetros lineares dos modelos dos nós terminais se torna apenas parte de um ciclo de detecção e estimação de árvores. Não há na literatura forma de estimar de uma só vez o tamanho da árvore. O ciclo de estimação pode ser resumido pela Figura a seguir:

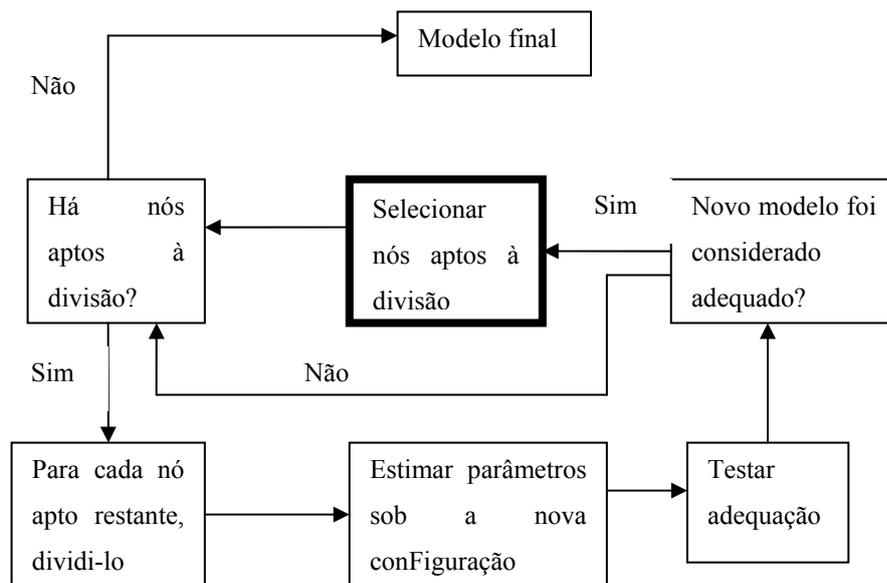


Figura 4.1: Ciclo de detecção de estrutura da árvore STAR

Dois métodos para detecção de árvores serão apresentados neste capítulo, o CART e o de testes de hipótese na divisão de nós. Enquanto que no primeiro, avalia-se a redução do erro, o aumento da complexidade do modelo e a significância amostral dos nós remanescentes a cada iteração de crescimento da árvore, o segundo avalia cada divisão de nós através de sucessivos testes de

hipótese a respeito dos novos nós introduzidos e posteriormente aos parâmetros atribuídos a eles. Ao fim do capítulo, o modelo será testado em uma série de dados reais, e comparado com um modelo linear.

4.2

Método CART

O método CART, ou *Classification and Regression Trees*, envolve a identificação e construção de uma árvore de decisão binária utilizando dados disponíveis na amostra para treinamento. Ele pode ser utilizado tanto para classificação de dados quanto para regressão. Quando a variável dependente é categórica, seu fim é de classificação. Quando é discreta ou contínua, seu fim é de regressão. Para identificar o modelo STAR-Tree, estaremos interessados em sua aplicação em regressão.

O algoritmo CART é uma forma de particionamento recursivo binário, o que implica que cada nó terminal de uma árvore pode ser dividido em dois grupos. O nó que anteriormente era uma folha passa a se tornar um nó de decisão na nova estrutura. O termo recursivo é utilizado, pois a árvore pode ir se ramificando indefinidamente até que algum critério de parada seja atingido. O termo particionamento se refere ao fato de que o conjunto de dados fornecido ao modelo é dividido em diferentes regiões, atendendo ao critério imposto pelo nó de decisão.

O método apresenta fácil interpretação dos resultados, pois o critério de divisão dos nós é bastante simples. Além disso, não faz nenhum tipo de premissa a respeito da distribuição das variáveis envolvidas no modelo, o que apesar de torná-lo menos robusto, pode ser útil em muitos casos, pois independe da necessidade de um analista experiente para parametrizá-lo. Pode-se dizer que é um método automático de aprendizado por máquina. Não se pode dizer o mesmo de métodos mais complexos, que, constantemente, necessitam de estudo dos resultados e reparametrização por parte do analista.

Como desvantagem, pode-se citar a mesma característica levantada como vantagem anteriormente: o método não possui embasamento estatístico. Ele simplesmente busca um equilíbrio entre redução do erro da estimação e o nível de complexidade do modelo.

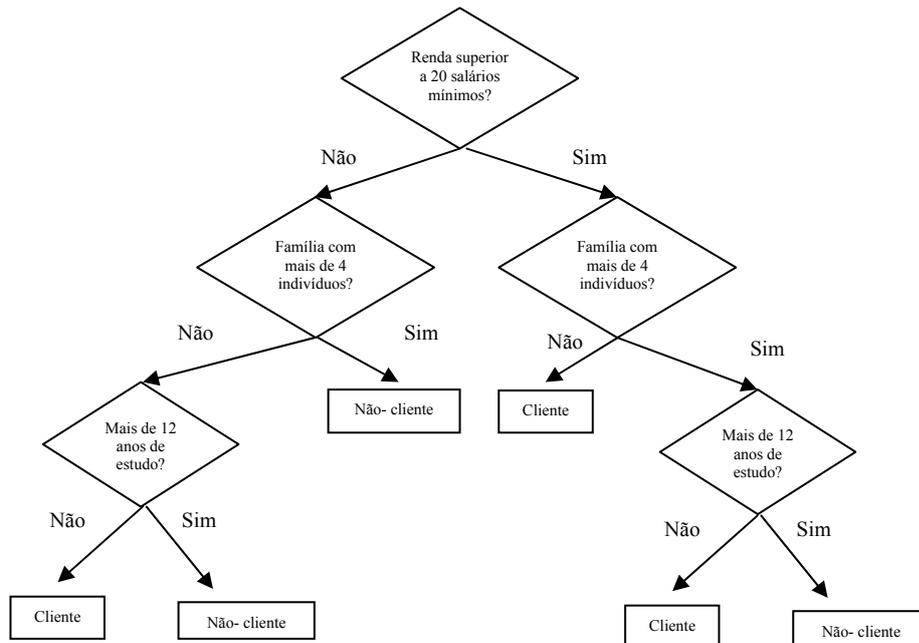


Figura 4.2: Exemplo de aplicação do CART para classificação

Nesta implementação, a metodologia CART foi alterada para aceitar transições suaves entre nós. A decisão deixa portanto de ser binária, e passa a ser regida pelas funções de pertinência do modelo STAR a ser estimado. A cada iteração, o algoritmo avalia a soma dos quadrados dos resíduos (*sum of squared errors* – *SSE*) como critério de ganho e a quantidade de parâmetros como custo de complexidade do modelo. O SSE da árvore em uma dada iteração é dado por:

$$SSE_{Arv(i)} = \sum_{t=1}^T \left(y_t - \hat{y}_t^{Arv(i)} \right)^2$$

(4-1)

Onde $Arv(i)$ representa a estrutura da árvore na iteração de estimação i .

Para avaliar a qualidade da estrutura detectada, é utilizada uma estatística que pondera custo e complexidade. A medida escolhida para executar tal tarefa foi o AICc (Akaike Information Corrected Criterion, ver[1]).

$$AICc_{Arv(i)} = \frac{\ln(SSE_{Arv(i)})}{n} + \frac{n+k}{n-k-2} \quad (4-1)$$

Onde n é o comprimento da série e k é a quantidade de parâmetros envolvidos no modelo.

Se $AICc_{Arv(i+1)} < AICc_{Arv(i)}$, o crescimento da árvore se torna justificado.

Outro critério de parada adotado em conjunto com o AICc foi a pertinência remanescente em qualquer um dos novos regimes após a divisão. Se não for significativa, a divisão é descartada. Este limiar para o descarte é definido através de parametrização.

Algoritmo CART

```

modelo_atual = GeraAR(p)
aicc_0 = ∞
Estima(Y, modelo_atual)
aicc_1 = CalculaAICc(Y, modelo_atual)
Enquanto (aicc_1 < aicc_0 )
{
  aicc_0 = aicc_1

  Para cada folha n em modelo_atual
  {

    Se (não (n é terminal) )
    {
      AdicionaFilhos(n)

      tempAicc = aicc_1

      Estima(Y, modelo_atual)

      aicc_1 = CalculaAICc(Y, modelo_atual)

      Se (não (aicc_1 < aicc_0 e DivisãoSignificante(n)))
      {
        RemoveFilhos(n)
        MarcaComoTerminal(n)
        aicc_1 = tempAicc
      }
    }
  }
}

}

}

```

Detalhamento do Algoritmo CART

4.3

Estimação por teste de hipótese

Na metodologia CART, a busca por variáveis de divisão é exaustiva, e o critério de divisão de nós é baseado apenas em uma função de custo. Porém, os modelos em si não são avaliados, seja em sua formação, seja na existência de dois modelos derivados de uma divisão de nós.

Para um dado modelo linear:

$$y_t = (B_0 \times X) + \varepsilon_t, \text{ onde } B_0 = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \quad (4-2)$$

Temos como possível alternativa, um modelo LSTAR:

$$y_t = f(x_{0,t}; \Phi) + \varepsilon_t \quad (4-3)$$

$$= (B_1 \times X) \times G[x_{0,t}; \gamma_0; c_0] + (B_2 \times X) \times (1 - G[x_{0,t}; \gamma_0; c_0]) + \varepsilon_t$$

$$\text{Onde } B_i = \begin{bmatrix} \beta_{i,0} \\ \beta_{i,1} \\ \vdots \\ \beta_{i,p} \end{bmatrix}, i=1,2$$

Consideremos agora que o modelo (4-2) representa uma árvore de uma profundidade apenas, e o modelo (4-3), uma árvore de dois níveis, com um nó de decisão e dois nós terminais. A evolução de (4-2) para (4-3) representa, portanto uma decisão de divisão de nó, ampliando o modelo STAR-Tree. Para modelar um teste de hipótese que avalia a importância desta divisão, pode-se considerar como hipótese nula que $H_0: \gamma_0 = 0$, e como alternativa que $H_a: \gamma_0 > 0$. Isto porque quando $\gamma_0 = 0$, o modelo (4-3) equivale a:

$$y_t = (B_1 \times X) \times \frac{e^0}{1+e^0} + (B_2 \times X) \times \left(1 - \frac{e^0}{1+e^0}\right) + \varepsilon_t$$

$$y_t = \frac{(B_1 + B_2)}{2} \times X + \varepsilon_t$$

Ou seja, um modelo onde os parâmetros lineares decorrem da ponderação constante dos vetores B_1 e B_2 equivale ao modelo de ordem inferior descrito em

(4-1). Quando assume-se que $\gamma_0 = 0$, temos um problema de identificação em torno da função de transição G , pois o parâmetro c_0 pode assumir qualquer valor sem que a função de verossimilhança, utilizada na estimação dos parâmetros do modelo, se altere.

A solução proposta por [7] envolve aproximar a função G original para uma expansão de Taylor de 3ª ordem em torno de $\gamma = 0$. Temos que :

$$y_t = A^T \times \begin{bmatrix} 1 \\ x_{0,t} \\ x_{0,t}^2 \\ x_{0,t}^3 \end{bmatrix} + e_t \quad (4-4)$$

$$G[x_{0,t}; \gamma_0; c_0] = \frac{e^{-\gamma_0(x_{0,t} - c_0)}}{1 + e^{-\gamma_0(x_{0,t} - c_0)}} \quad (4-4)$$

Sendo assim:

$$G[x_{0,t}; \gamma_0; c_0]_{\gamma=0} \approx \quad (4-5)$$

$$G[x_{0,t}; 0; c] + \frac{dG}{d\gamma_0}[x_{0,t}; 0; c] \times \gamma_0 + \frac{dG}{dc_0}[x_{0,t}; 0; c](c_0 - c) + \frac{1}{2} \left(\frac{d^2G}{d\gamma_0^2}[x_{0,t}; 0; c](\gamma_0)^2 + 2 \frac{d^2G}{d\gamma_0 dc_0}[x_{0,t}; 0; c](\gamma_0)(c_0 - c) + \frac{d^2G}{dc_0^2}[x_{0,t}; 0; c](c_0 - c)^2 \right) + \frac{1}{6} \left(\frac{d^3G}{d\gamma_0^3}[x_{0,t}; 0; c](\gamma_0)^3 + 2 \frac{d^3G}{d\gamma_0^2 dc_0}[x_{0,t}; 0; c](\gamma_0)^2(c_0 - c) + 2 \frac{d^3G}{d\gamma_0 dc_0^2}[x_{0,t}; 0; c](\gamma_0)(c_0 - c)^2 + \frac{d^3G}{dc_0^3}[x_{0,t}; 0; c](c_0 - c)^3 \right)$$

A função logística descrita em (4-4) apresenta as seguintes derivadas parciais:

$$\frac{dG}{d\gamma_0} = (c_0 - x_{0,t})G(1 - G) \quad (4-6)$$

$$\frac{dG}{dc_0} = (\gamma_0)G(1 - G)$$

$$\frac{d^2G}{d\gamma_0^2} = (c_0 - x_{0,t})^2 (G(1-G)^2 - G^2(1-G))$$

$$\frac{d^2G}{dc_0^2} = (\gamma_0)^2 (G(1-G)^2 - G^2(1-G))$$

$$\frac{d^2G}{d\gamma_0 dc_0} = (c_0 - x_{0,t}) (\gamma_0) (G(1-G)^2 - G^2(1-G))$$

$$\frac{d^3G}{d\gamma_0^3} = (c_0 - x_{0,t})^3 (G(1-G)^3 - 4G^2(1-G)^2 + G^3(1-G))$$

$$\frac{d^3G}{dc_0^3} = (\gamma_0)^3 (G(1-G)^3 - 4G^2(1-G)^2 + G^3(1-G))$$

$$\frac{d^3G}{d\gamma_0^2 dc_0} = (c_0 - x_{0,t})^2 (\gamma_0) (G(1-G)^3 - 4G^2(1-G)^2 + G^3(1-G))$$

$$\frac{d^3G}{d\gamma_0 dc_0^2} = (c_0 - x_{0,t}) (\gamma_0)^2 (G(1-G)^3 - 4G^2(1-G)^2 + G^3(1-G))$$

Com a hipótese nula de que $\gamma_0 = 0$, todos os termos das derivadas parciais que envolvem multiplicação desta variável são anulados, e função de pertinência G assume o valor de 0.5. A aproximação de Taylor para a função de pertinência se reduz a :

$$G[x_{0,t}; \gamma_0; c_0]_{\gamma_0=0} \approx \tag{4-7}$$

$$\frac{1}{2} + \frac{1}{4} \gamma_0 (c - x_{0,t}) + 0 \times (c - x_{0,t})^2 (\gamma_0)^2 - \frac{1}{48} (c - x_{0,t})^3 (\gamma_0)^3$$

Voltando à equação (4-3), ela pode ser reescrita como:

$$y_t = \phi_0 + \lambda_0 \times G[x_{0,t}; \gamma_0; c_0] \tag{4-8}$$

Onde

$$\phi_0 = B_2$$

$$\lambda_0 = B_1 - B_2$$

Sendo assim, unindo (4-7) e (4-8), após alguma manipulação chega-se a :

$$y_t = [\alpha_0 \quad \alpha_1 \quad \alpha_2 \quad \alpha_3] \times \begin{bmatrix} 1 \\ x_{0,t} \\ x_{0,t}^2 \\ x_{0,t}^3 \end{bmatrix} + e_t \quad (4-9)$$

Os termos α_i estão em função de $\phi_0, \lambda_0, \gamma_0$ e c_0 , e $e_t = \varepsilon_t + R$, onde R é o restante da expansão de Taylor.

Sendo assim, a hipótese nula pode ser reescrita como:

$$H_0 : \alpha_i = 0, i = 1, 2, 3 \quad (4-10)$$

Conforme o modelo STAR-Tree cresce, a mesma lógica apresentada acima continua a ser aplicada. A relação entre parâmetros e variáveis é linearizada através de uma expansão de Taylor de 3ª ordem, e a hipótese nula sobre o nó a ser dividido passa a ser equivalente a anular parte dos coeficientes resultantes dessa expansão. O termo eliminado sob a hipótese nula sempre inclui informação relativa ao modelo derivado da divisão de nós.

A partir destes resultados, a seguinte estratégia de teste LM para divisão de nós é apresentada por [7]:

1. Calcular as estatísticas auxiliares h_t e v_t , onde:

$$h_t = \frac{df(x_{0,t}; \Phi)}{d\Phi} \Big|_{H_0} \quad (4-11)$$

e

$$v_t = \left[\left(\prod_{i=0}^{n-1} G[x_{i,t}; \gamma_i; c_i] \right) x_{n,t} \quad \left(\prod_{i=0}^{n-1} G[x_{i,t}; \gamma_i; c_i] \right) x_{n,t}^2 \quad \left(\prod_{i=0}^{n-1} G[x_{i,t}; \gamma_i; c_i] \right) x_{n,t}^3 \right]$$

A estatística h_t é calculada a partir da expansão de Taylor, sob a hipótese nula. Isto significa que o vetor Φ terá como itens os parâmetros α que não foram eliminados, e os parâmetros não lineares das funções G de nós de nível superior. No caso mais simples, estes parâmetros não lineares não existem, $\Phi = [\alpha_0]$, e $h_t = 1$.

2. Estimar o modelo mais simples, sob H_0 , e calcular a estatística

$$SSR_0 = \sum_{t=1}^T \left(y_t - \hat{y}_t \right)^2 = \sum_{t=1}^T \left(\hat{\varepsilon}_t \right)^2$$

3. Calcular o mesmo para a regressão de $\hat{\varepsilon}_t$ em h_t e v_t . Computar a soma dos resíduos como SSR_1 .
4. Calcular a estatística de teste:

$$\frac{(SSR_0 - SSR_1) / 3}{SSR_1 / (T - (p + 3))} \quad (4-12)$$

Onde p é a quantidade de elementos no vetor Φ e T é o tamanho da série. A estatística possui distribuição F com 3 e $T-(p+3)$ graus de liberdade.

A modelagem de divisão de nós é recursiva e pode ser melhor descrita a partir do seguinte algoritmo¹:

DetectaArvore(X,Y,N[0...n],a)

Entrada:

X : matriz com candidatas a variáveis de transição

Y: série temporal

N[0...n]: vetor de nós sujeitos a divisão

a: nível de significância do teste

Variáveis globais:

NProx[0...n]: Nós candidatos a divisões na próxima profundidade

ctProx=-1 : contador de nós candidatos a divisões na próxima profundidade

Para cada nó n em N

```
{
  nCandidatos[0...n]
  rankingVars[0...n]
  pValores[0...n]
  i=-1
  Para cada série x em X
  {
    pValor= TesteLM(n,Y,x)
    Se (pValor<a)
    {
      i=i+1
      nCandidatos[i]=n
      rankingVars[i]=x
      pValores[i]=pValor
    }
  }
}
```

¹ No sistema a ser descrito no capítulo 5, o algoritmo foi escrito em uma modelagem orientada a objeto. Esta é uma representação estruturada do mesmo.

```

}
//Ordena-se os vetores nCandidatos e rankingVars em função dos pValores
//encontrados
nCandidatos= Ordena(nCandidatos,pValores)
rankingVars = Ordena(rankingVars,pValores)

Para cada nó n em nCandidatos
{
  Se (ContaElementos(rankingVars)>0)
  {
    //Aplica-se temporariamente a divisão de nós
    AdicionaFilhos(n)
    Para cada vetor v em rankingVars
    {
      //Estima os parâmetros para a árvore do nó n, com a variável de transição v
      // aplicada ao nó de decisão n e retorna um vetor de parâmetros lineares
      B[0...n]= EstimaParametros(n,v)

      // Testa os parâmetros do vetor B condicionalmente aos parâmetros não
lineares
      // encontrados com nível de significância 'a'.
      Se (TestaHipotese(B)>=a) //teste rejeita hipótese
        RemoveFilhos(n)
      Senão //teste não rejeita hipótese
      {
        Para cada nó f em Filhos(n)
        {
          ctProx = ctProx +1
          NProx[ctProx]=f
        }
      }
    }
  }
  Se(ContaElementos(NProx)>0)
  {
    //Passa para a próxima profundidade
    DetectaArvore(X,Y, NProx,a)
  }
}
}
}

```

4.4

Aplicação a dados reais

Nesta seção será feita uma comparação de desempenho na estimação de séries reais entre dois modelos: o modelo multivariado autorregressivo com variáveis exógenas (ARX) e o modelo STAR-Tree estimado por testes de hipótese.

A série a ser estudada representa o consumo de combustível de caminhões fora de estrada na área de mineração. O consumo de combustível representa um dos maiores custos na operação destas máquinas. Os equipamentos de mineração mais modernos registram eletronicamente além do próprio consumo, uma série de variáveis medidas na mesma unidade de tempo. A frequência de medições é de aproximadamente duas a cada hora, e é feita a cada ciclo de operação do veículo.

Um ciclo de operação representa as atividades de carga, movimentação e descarga do caminhão. Nele, são medidas grandezas como a carga movimentada (em toneladas), as distâncias percorridas (carregado e descarregado, em km), o total de combustível consumido (em litros), e os tempos gastos em cada atividade (tempos de ociosidade, carga, descarga, de movimentação carregado e descarregado, todos em segundos). A estimação de um modelo que permita entender o comportamento do consumo (em função ou não das variáveis apresentadas) e fazer previsões de curto prazo é importante para o planejamento da operação.

Para os modelos que trabalham com variáveis exógenas, foram derivadas as seguintes séries a partir dos dados originais:

- Percentual do tempo carregando (%)
- Percentual do tempo parado e vazio (%)
- Percentual do tempo andando e vazio (%)
- Percentual do tempo parado e carregado (%)
- Percentual do tempo andando carregado (%)
- Velocidade média – equipamento vazio (km/h)
- Velocidade média – equipamento carregado (km/h)
- Carga movimentada (ton)

A série a ser estimada será o consumo médio horário, em litros por hora, medido a cada ciclo de operação. Para a comparação, foi selecionada uma série referente a aproximadamente um mês de operação de um caminhão de mineração.

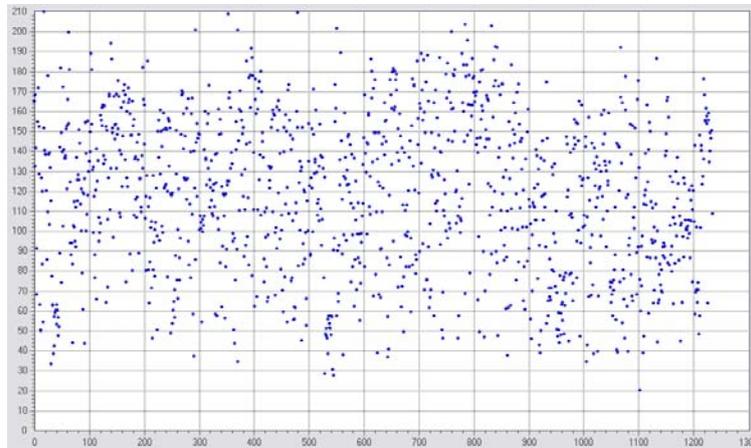


Figura 4.3: Série de consumo médio horário de combustível, de janeiro a fevereiro de 2005

As variáveis explicativas introduzidas não produzem efeitos permanentes sobre o equipamento no curto prazo. Em outras palavras, apesar de ser razoável inferir que há correlação no médio e longo prazo entre, por exemplo, excesso de carga e consumo, o mesmo não se pode dizer do curto prazo. Um ciclo de operação que exige muito da máquina não exerce influência sobre os seguintes.

No entanto, as condições de aplicação do equipamento na mesma unidade de tempo do ciclo exercem influência significativa sobre o consumo de combustível. Um modelo que consiga capturar o efeito dessas variáveis sobre a série em questão teria bastante aplicação no planejamento da produção.

A seguir serão exibidas as estatísticas utilizadas para comparar o desempenho dos métodos utilizados para a estimação:

Tabela 4.1: Descrição das estatísticas de avaliação dos modelos

Estatística	Descrição
Mean Absolute Error (MAE)	$\frac{1}{N} \sum_{t=1}^N y_t - \hat{y}_t $
Mean Absolute Percentage Error (MAPE)	$\frac{100}{N} \sum_{t=1}^N \left \frac{y_t - \hat{y}_t}{y_t} \right $
Standard Error (s.e.)	$\sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2}$

4.4.1

Modelo ARX

Para esta avaliação, foram introduzidas, além das variáveis exógenas já citadas, as dez últimas observações da série de consumo de combustível. A equação geral do modelo é dada por:

$$y_t = \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{i=1}^q \beta_i x_i + e_t \quad (4-13)$$

Onde y_{t-i} são as variáveis autoregressivas, x_i são as variáveis exógenas, e φ_i e β_i seus respectivos parâmetros.

A cada iteração de estimação foram aplicados testes de hipótese sobre os parâmetros das variáveis com a hipótese nula $H_0 : \varphi_i = 0$ e $H_0 : \beta_i = 0$. Dentre os parâmetros em que a hipótese nula não fosse rejeitada com um nível de significância de $\alpha=5\%$, aquela de menor p-valor era eliminada, e a estimação era realizada novamente. A seguir os parâmetros que não foram rejeitados pelos teste de hipótese:

Tabela 4.2: Parâmetros resultantes da estimação

Parâmetro	Estimativa	Intervalo de confiança (95%)
Percentual do tempo carregando	31,67	(25,10; 38,24)
Percentual do tempo andando vazio	144,05	(123,27;164,84)
Percentual do tempo andando carregado	277,63	(229,62; 325,64)

Os parâmetros com maior correlação com a série original foram os percentuais de tempo carregando, andando vazio e carregado. A Figura 4.4 mostra a relação das variáveis selecionadas com a série de consumo:

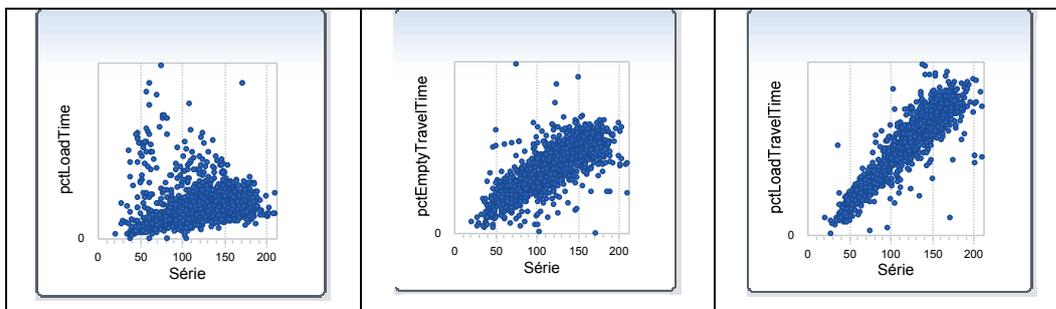


Figura 4.4: Variáveis selecionadas X série de consumo

A seguir foi realizada uma análise sobre os resíduos da estimação. Eles apresentaram média próxima de zero, e o histograma dos mesmos permite visualizar uma falta de simetria na distribuição amostral do erro. Uma das caudas se mostra mais prolongada do que a outra.

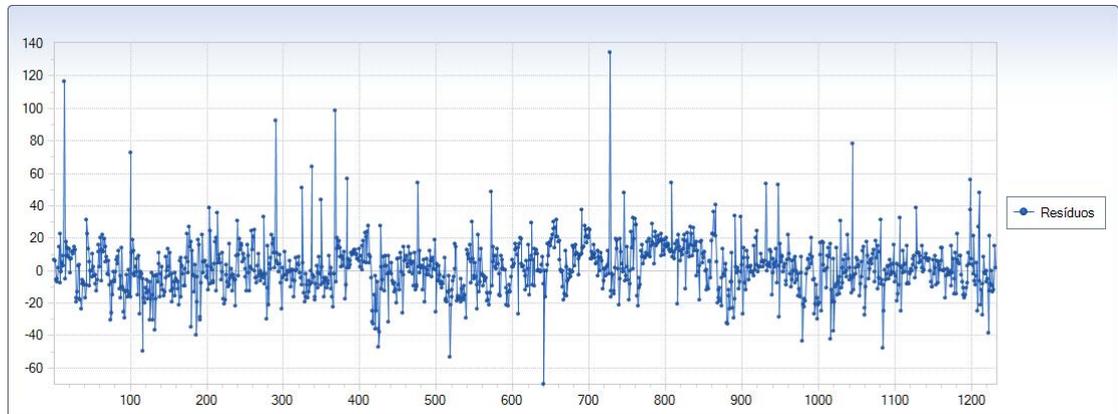


Figura 4.5: Série de resíduos da estimação

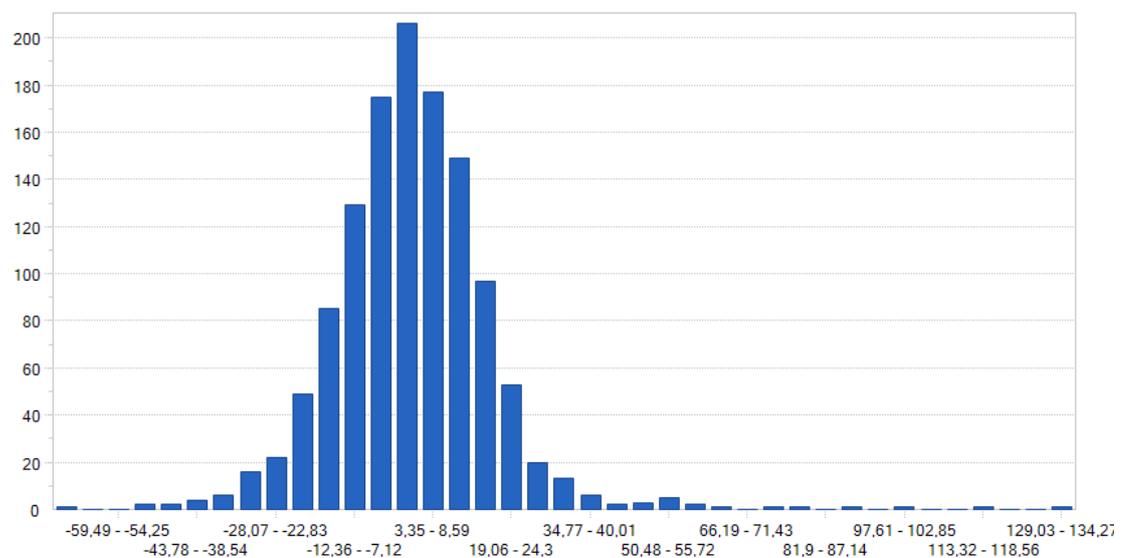


Figura 4.6: Histograma dos resíduos da estimação

É possível ver também através do gráfico da probabilidade empírica acumulada a tendência em subestimar os valores da série:

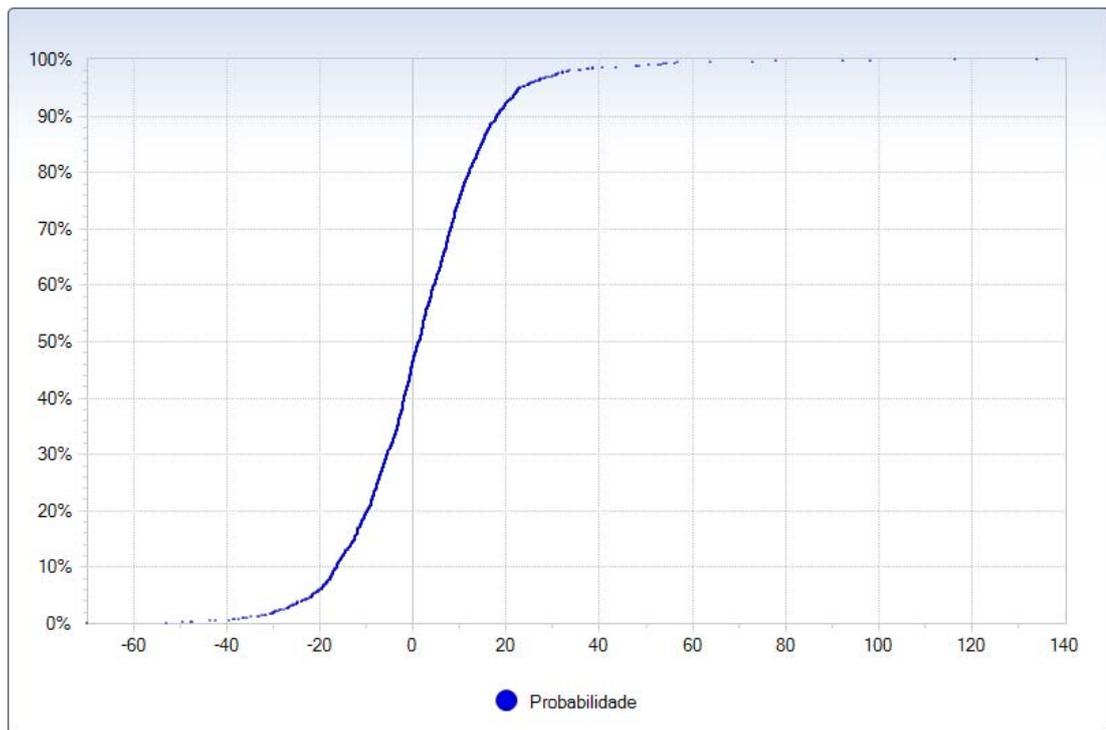


Figura 4.7: Probabilidade empírica acumulada dos resíduos

Para avaliar a hipótese de normalidade dos resíduos, foi realizado um teste de Jarque-Bera com nível de significância de 5%. O teste confirmou os indícios da inspeção visual, rejeitando a hipótese de normalidade.

A seguir as estatísticas resultantes da estimação do modelo:

Tabela 4.3: Estatísticas de avaliação do modelo ARX

Estatística	Valor
MAE	11,407
MAPE	10,908
s.e.	16,130

4.4.2

Modelo STAR-Tree

Para a estimação através do modelo STAR-Tree, as variáveis exógenas também serão utilizadas como candidatas a variáveis de transição, e os modelos dos nós terminais irão conter tanto as variáveis exógenas quanto auto-regressivas até a ordem 10. O crescimento da árvore de decisão se dará através de testes de hipótese, com significância inicial equivalente a 5%. Conforme a árvore for

crecendo, este nível de significância se tornará cada vez mais restritivo, penalizando modelos de ordem muito alta. Em cada modelo terminal também será aplicado um teste de hipótese com o intuito de eliminar-se variáveis irrelevantes. Neste teste, realizado a cada variável testada, a hipótese nula é a de que aquela variável é igual a zero. A significância utilizada neste teste é de 5%. A Figura 4.8 mostra a estrutura final da árvore estimada:

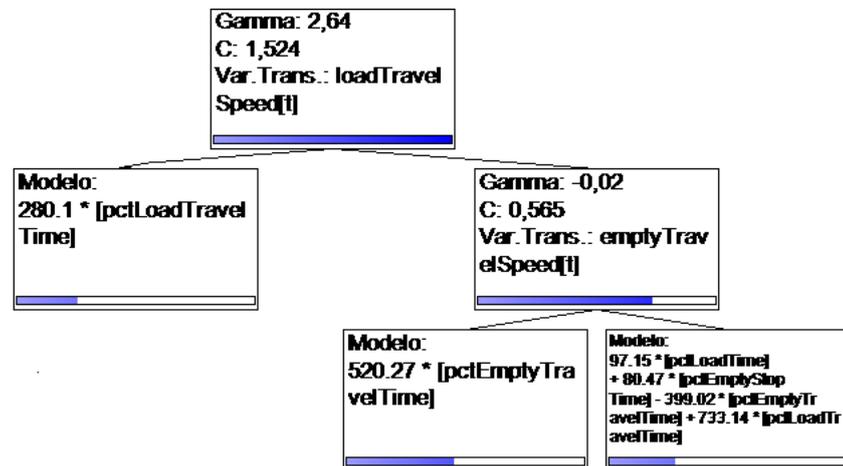


Figura 4.8: Árvore estimada pelo modelo STAR-Tree

As barras nos nós representam a pertinência média capturada por cada um dos nós estimados. É possível perceber que os nós de decisão estimados produziram divisões suaves entre os regimes. As variáveis de transição selecionadas são bastante correlacionadas com o regime de trabalho dos equipamentos. Por trás de uma velocidade baixa ou alta em tempo de carga ou andando descarregado, por exemplo, estão diferentes regimes de trabalho. Os equipamentos podem estar trabalhando em frentes com alta ou baixa inclinação de terreno, em períodos chuvosos ou não, dentre outras condições de trabalho que podem configurar diferentes regimes de operação, com diferentes padrões de consumo de combustível. A Figura 4.9 mostra a evolução da pertinência capturada pelos nós de decisão a cada dado da série:

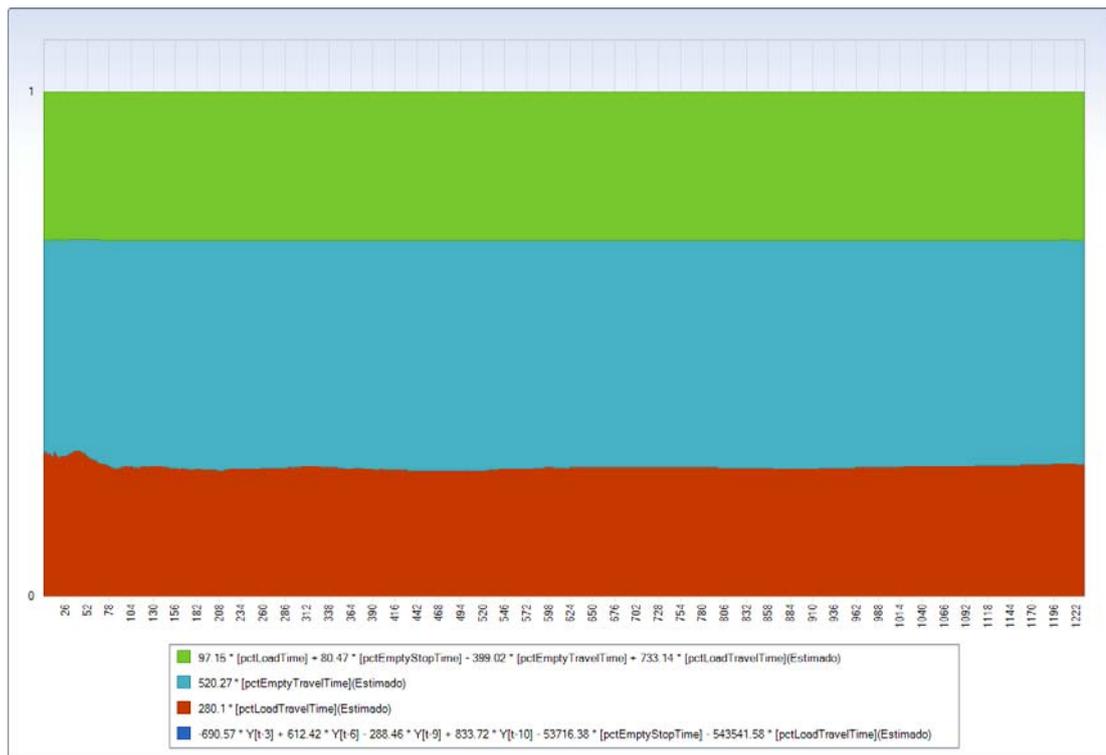


Figura 4.9: Pertinência acumulada capturada pelos nós de decisão a cada dado da série

Percebe-se também por meio da Figura 4.10 que as funções de pertinência estão bem representadas nos dados. Isto demonstra que os regimes estão bem representados nas séries das variáveis de transição.

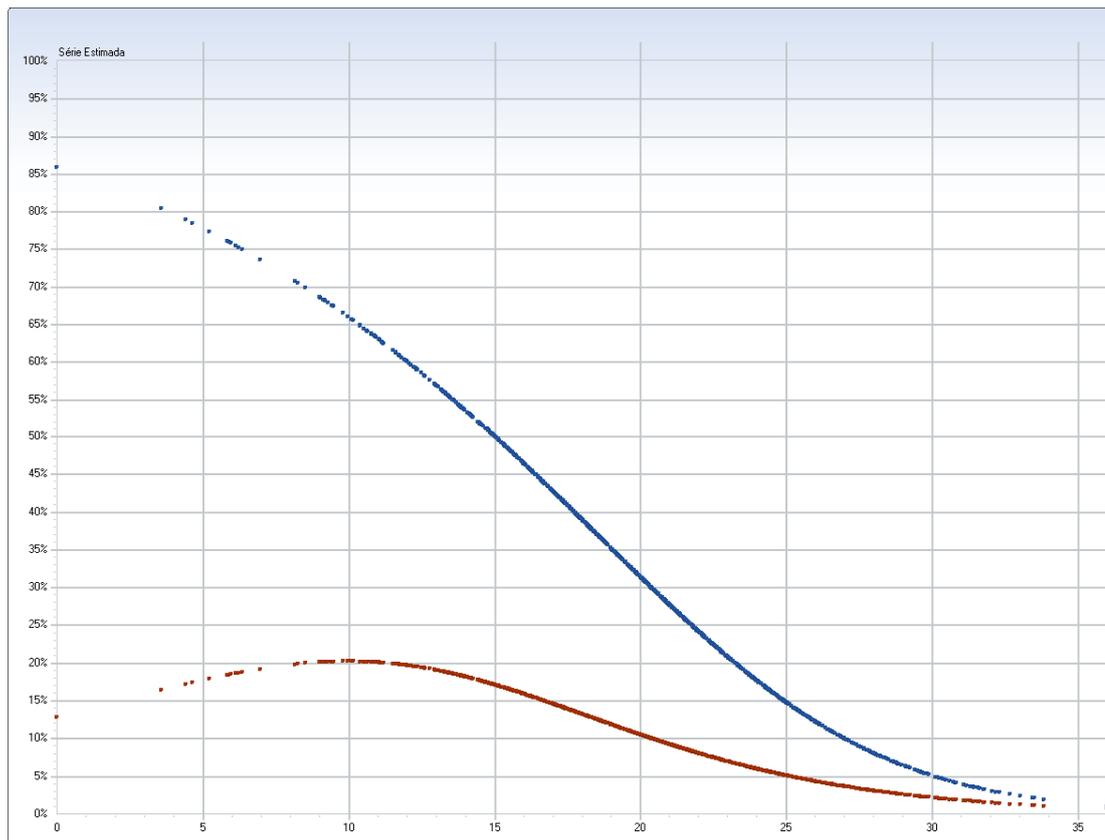


Figura 4.10: Funções de pertinência resultantes da estimação

Os parâmetros lineares estimados foram sumarizados na tabela 4.4, com seus respectivos intervalos de confiança:

Tabela 4.4: Parâmetros lineares resultantes da estimação do modelo STAR-Tree

Parâmetro	Estimativa	Intervalo de confiança (95%)
Percentual do tempo andando carregado (1.1)	280,101	(277,132 ; 283,069)
Percentual do tempo andando vazio (1.2.1)	520,268	(516,169 ; 524,367)
Percentual do tempo carregando (1.2.2)	97,152	(90,764 ; 103,54)
Percentual do tempo parado vazio (1.2.2)	80,469	(66,047 ; 94,891)
Percentual do tempo andando vazio (1.2.2)	-399,016	(-396,969 ; -401,062)
Percentual do tempo andando carregado (1.2.2)	733,141	(724,487 ; 741,795)

Uma quantidade razoável de parâmetros foi eliminada dos modelos terminais através dos sucessivos testes de hipótese. Três modelos que somados poderiam inicialmente ter cinquenta e quatro variáveis acabaram por ter um total de seis variáveis.

A análise dos resíduos produziu os seguintes resultados:

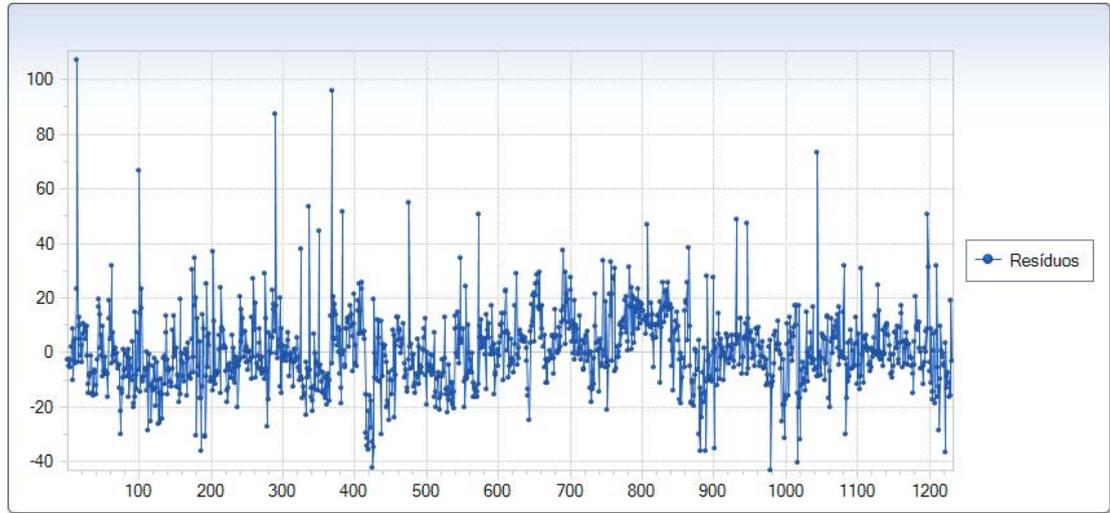


Figura 4.11: Série de resíduos da estimação do modelo STAR

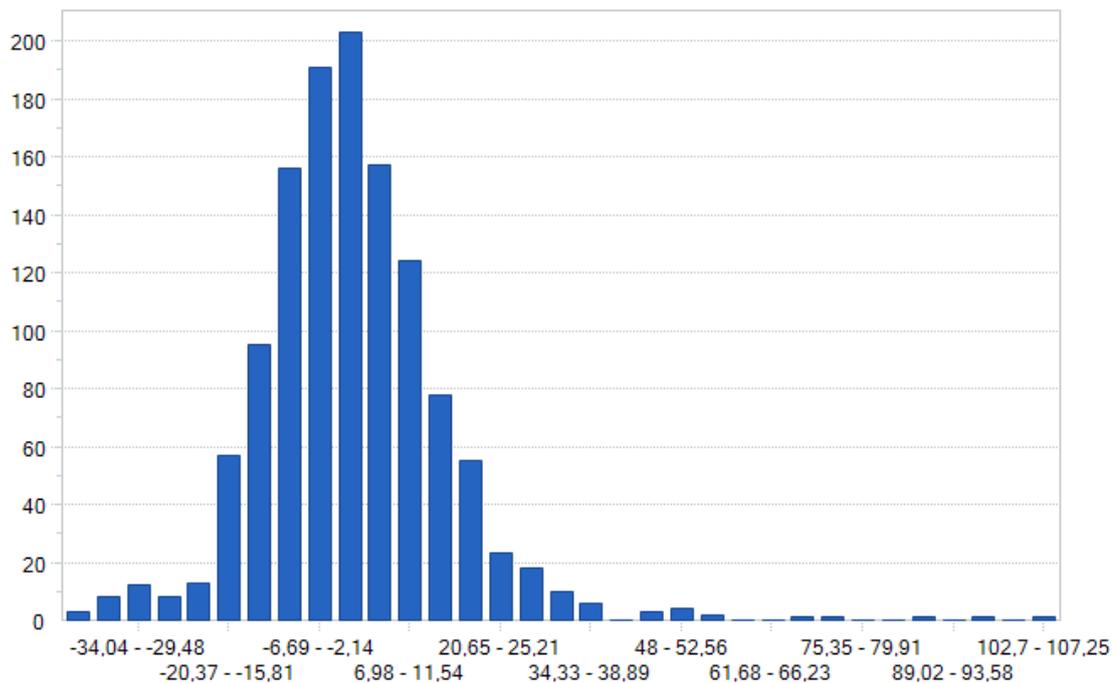


Figura 4.12: Histograma dos resíduos da estimação

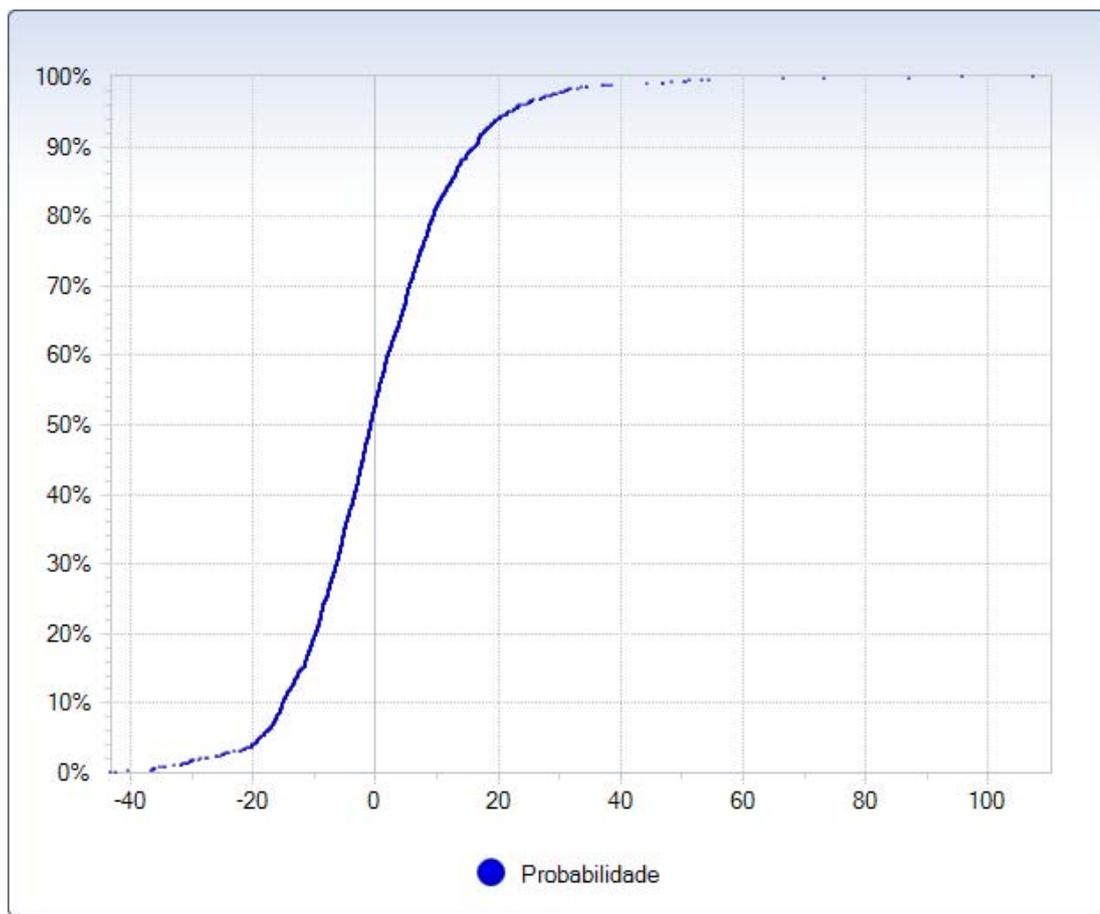


Figura 4.13: Probabilidade empírica acumulada dos resíduos

É possível notar um excesso de curtose no histograma dos resíduos. Foi aplicado o teste de Jarque-Bera para verificar a hipótese nula de normalidade dos resíduos. A hipótese foi rejeitada com nível de significância de 5%.

A tabela a seguir resume as estatísticas de avaliação do erro na estimação do modelo STAR-Tree:

Tabela 4.5: Estatísticas de avaliação do modelo STAR-Tree

Estatística	Valor
MAE	10,064
MAPE	9,153
s.e.	14,040

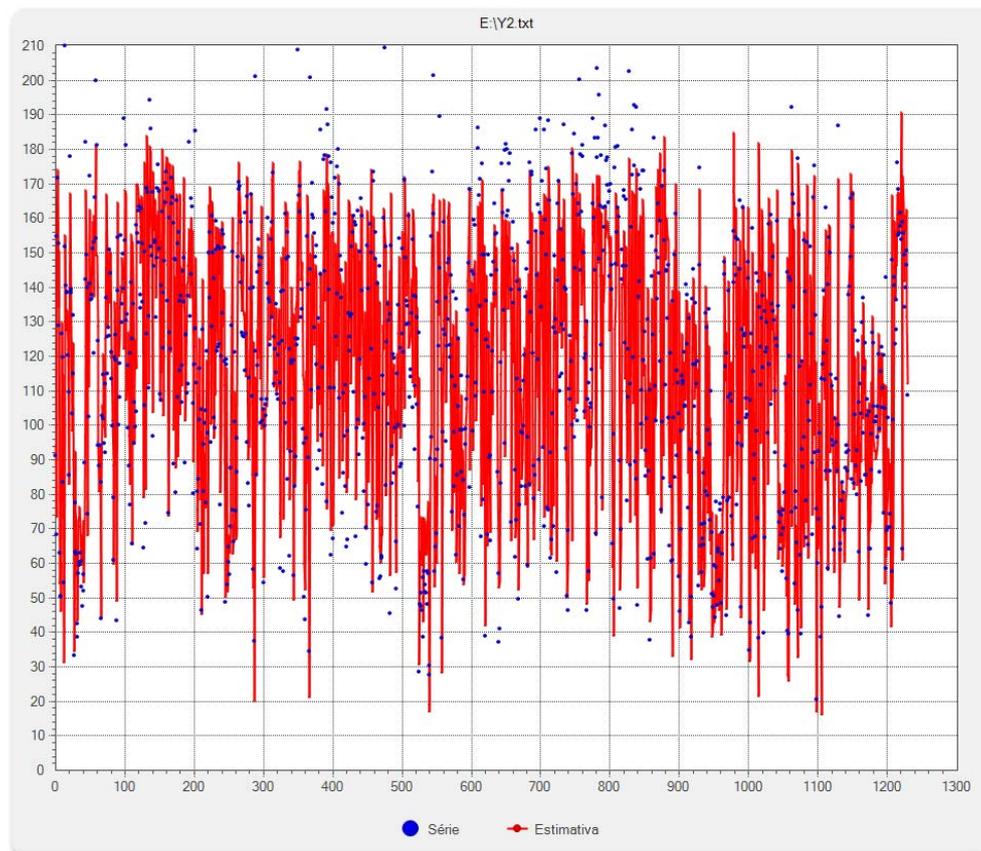


Figura 4.14: Estimação do modelo STAR-Tree

4.5

Conclusões

Ao utilizar as mesmas variáveis que o modelo ARX, o modelo STAR-Tree conseguiu atingir resultados superiores ao daquele modelo. Nenhum dos modelos selecionou variáveis auto-regressivas.

O modelo STAR-Tree detectou dois diferentes regimes através das variáveis de velocidade com o equipamento descarregado e velocidade com o equipamento carregado. Nenhuma variável auto-regressiva foi escolhida como variável de transição. A variável de carga não foi selecionada como variável de transição em nenhum nó de decisão da árvore, porém foi utilizada nos modelos dos nós terminais. Isto pode ter relação com o fato de que, independentemente das condições de trabalho, os equipamentos sempre trabalham com bastante carga. As velocidades foram detectadas como fatores mais importantes neste sentido. O modelo STAR-Tree gerou estatísticas de avaliação superiores às do modelo

ARX. No entanto, o tempo levado pelo modelo STAR-Tree para estimar uma série deste tamanho com uma grande quantidade de variáveis – dez auto-regressivos e oito variáveis exógenas – foi muito grande. O modelo ARX foi estimado em questão de segundos, enquanto que o modelo STAR-Tree foi estimado em torno de três horas. Deve-se isso às inversões de matrizes de tamanho razoavelmente grande, às gerações de grid de resultados e à grande quantidade de avaliação de candidatas a variáveis de transição e eliminação de variáveis seguidas de reestimação do modelo. Porém, claramente há espaço para revisões na implementação de forma a reduzir o tempo total gasto na estimação do modelo STAR-Tree.