

PONTIFÍCIA UNIVERSIDADE CATÓLICA  
DO RIO DE JANEIRO



**Alexandre José dos Santos**

**Modelos Vetoriais Auto-Regressivos com Transição  
Suave Estruturados por Árvores – STVAR-Tree**

**Dissertação de Mestrado**

Dissertação apresentada como requisito parcial para  
obtenção do título de Mestre pelo Programa de Pós-  
Graduação em Engenharia Elétrica da PUC-Rio.

Orientador: Prof. Álvaro de Lima Veiga Filho

Rio de Janeiro  
Setembro de 2009



**Alexandre José dos Santos**

**Modelos Vetoriais Auto-Regressivos com Transição  
Suave Estruturados por Árvores – STVAR-Tree**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

**Prof. Álvaro de Lima Veiga Filho**

Orientador  
Departamento de Engenharia Elétrica  
PUC-Rio

**Prof. Reinaldo Castro Souza**

Departamento de Engenharia Elétrica  
PUC-Rio

**Prof. Joel Maurício Corrêa da Rosa**

Departamento de Estatística  
UFF

**Prof. José Eugenio Leal**

Coordenador Setorial do Centro Técnico Científico  
PUC-Rio

Rio de Janeiro, 11 de setembro de 2009

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

### **Alexandre José dos Santos**

Graduou-se em Ciências Estatísticas na Escola Nacional de Ciências Estatísticas – ENCE (Rio de Janeiro, Brasil). Durante o mestrado, trabalhou com modelos estatísticos lineares e não-lineares, tanto univariados quanto multivariados e, ainda, com modelagem de inteligência artificial, todos com aplicações no mercado brasileiro de energia elétrica.

#### Ficha Catalográfica

Santos, Alexandre José dos

Modelos Vetoriais Auto-Regressivos com Transição Suave Estruturados por Árvores – STVAR-Tree / Alexandre José dos Santos ; orientador: Álvaro de Lima Veiga Filho. – 2009.

121 f. : il. ; 30 cm

Dissertação (Mestrado em Engenharia Elétrica)– Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009.

Inclui bibliografia

1. Engenharia elétrica – Teses. 2. Modelos não-lineares multivariados. 3. Árvore de regressão. 4. STVAR-Tree. 5. Teste LM. 6. Vasão de rios. 7. Preço spot. I. Veiga Filho, Álvaro de Lima. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

## Agradecimentos

A Deus, em primeiro lugar, pelo dom da vida.

Ao professor Álvaro Veiga, pela orientação, incentivo e amizade.

À PUC-Rio, pela estrutura e auxílio, imprescindíveis para a realização deste trabalho.

À CAPES, pelo suporte financeiro.

Aos professores do Departamento de Engenharia Elétrica, pela contribuição a minha formação, em especial aos professores Álvaro Veiga, Cristiano Fernandes e Reinaldo Souza.

Aos colegas do Departamento de Engenharia Elétrica, que estiveram disponíveis para ajudar em todos os momentos.

Aos meus pais, pelo amor incondicional e por não terem medido esforços para me propiciar a melhor formação possível.

A minha irmã e ao meu sobrinho, pelo amor, companhia, incentivo e confiança.

Aos meus amigos, que estiveram presentes ao longo destes anos, pelo apoio, compreensão e perdão por não ter outro assunto, que não o mestrado e, no final, a dissertação.

## Resumo

Santos, Alexandre; Veiga, Álvaro. **Modelos Vetoriais Auto-Regressivos com Transição Suave Estruturados por Árvores – STVAR-Tree**. Rio de Janeiro, 2009. 121p. Dissertação de Mestrado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Esta dissertação tem como objetivo principal introduzir uma formulação de modelo não-linear multivariado, a qual combina o modelo STVAR (Smooth Transition Vector Autoregressive) com a metodologia CART (Classification and Regression Tree) a fim de utilizá-lo para geração de cenários e de previsões. O modelo resultante é um Modelo Vetorial Auto-Regressivo com Transição Suave Estruturado por Árvores, denominado STVAR-Tree e tem como base o conceito de múltiplos regimes, definidos por árvore binária. A especificação do modelo é feita através do teste LM. Desta forma, o crescimento da árvore é condicionado à existência de não-linearidade nas séries, que aponta a divisão do nó e a variável de transição correspondente. Em cada divisão, são estimados os parâmetros lineares, por Mínimos Quadrados Multivariados, e os parâmetros não-lineares, por Mínimos Quadrados Não-Lineares. Como forma de avaliação do modelo STVAR-Tree, foram realizados diversos experimentos de Monte Carlo com o objetivo de constatar a funcionalidade tanto do teste LM quanto da estimação do modelo. Bons resultados foram obtidos para amostras médias e grandes. Além dos experimentos, o modelo STVAR-Tree foi aplicado às séries brasileiras de Vazão de Rios e Preço Spot de energia elétrica. No primeiro estudo, o modelo foi comparado estatisticamente com o Periodic Autoregressive (PAR) e apresentou um desempenho muito superior ao concorrente. No segundo caso, a comparação foi com a modelagem Neuro-Fuzzy e ganhou em uma das quatro séries. Somando os resultados dos experimentos e das duas aplicações conclui-se que o modelo STVAR-Tree pode ser utilizado na solução de problemas reais, apresentando bom desempenho.

## Palavras-chave

Modelos Não-Lineares Multivariados; Árvore de Regressão; STVAR-Tree; Teste LM; Vazão de Rios; Preço Spot.

## Abstract

Santos, Alexandre; Veiga, Álvaro (advisor). **Tree-Structure Smooth Transition Vector Autoregressive Models – STVAR-Tree**. Rio de Janeiro, 2009. 121p. MSc. Dissertation. Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

The main goal of the dissertation is to introduce a nonlinear multivariate model, which combines the model STVAR (Smooth Transition Vector Autoregressive) with the CART (Classification and Regression Tree) method and use it for generating scenarios and forecasting. The resulting model is a Tree-Structured Vector Autoregressive model with Smooth Transition, called STVAR-Tree, which is based on the concept of multiple regimes, defined by binary tree. The model specification is based on Lagrange Multiplier tests. Thus, the growth of the tree is conditioned on the existence of nonlinearity in the time series, which indicates the node to be split and the corresponding transition variable. In each division, linear parameters are estimated by Multivariate Least Squares, and nonlinear parameters by Non-Linear Least Squares. As a way of checking the STVAR-Tree model, several Monte Carlo experiments were performed in order to see the functionality of both the LM test and the model estimation. Best results were obtained with medium and large samples. Besides, the STVAR-Tree model was applied to Brazilian time series of Rivers Flow and electricity spot price. In the first study, the model was statistically compared to the Periodic Autoregressive (PAR) model and had a much higher performance than the competitor. In the second case, the model comparison was with Neural-Fuzzy Modeling and the STVAR-Tree model won in one of the four series. Adding both the experiments and the two applications results we conclude that the STVAR-Tree model may be applied to solve real problems, having good results.

## Keywords

Multivariate Non Linear Models; Regression Tree; STVAR-Tree; LM Test; Rivers Flow; Spot Price.

## Sumário

1	Introdução	12
1.1.	Vazão de Rios	14
1.2.	Preço Spot de energia elétrica	16
1.3.	Organização da dissertação	17
2	Modelos Lineares Multivariados	18
2.1.	Introdução	18
2.2.	Modelo VAR	19
2.3.	Modelo VEC	25
3	Modelos Não-Lineares	29
3.1.	Modelos Não-Lineares Univariados	29
3.2.	Modelos Não-Lineares Multivariados	35
3.3.	Metodologia CART	47
4	Metodologia	52
4.1.	Introdução	52
4.2.	Modelo STVAR-Tree	52
4.3.	Modelo PAR	62
4.4.	Sistema Neuro-Fuzzy	63
4.5.	ANFIS: Adaptive Neuro-Fuzzy Inference System (Sistema Adaptativo de Inferência Neuro-Fuzzy)	63
5	Experimentos e Aplicações	66
5.1.	Experimentos de Monte Carlo	66
5.2.	Aplicações a dados reais	78
6	Conclusão	114
	Referências Bibliográficas	116

## Lista de figuras

Gráfico 2.1: Séries temporais bivariadas co-integradas	26
Gráfico 3.1: Função logística com parâmetros fixos	32
Figura 3.1: Exemplo de árvore com ausência de alguns nós	49
Figura 3.2: Exemplo de um modelo gerado por uma árvore de regressão	50
Figura 4.1: Arquitetura ANFIS equivalente	64
Figura 4.2: Arquitetura ANFIS para o modelo Sugeno	65
Figura 5.1: Arquitetura STVAR-Tree simulada	68
Figura 5.2: Mediana de $\Phi_i, i = 1,2$ do modelo STVAR-Tree simulado, com $\alpha = 0,1\%$	75
Figura 5.3: Mediana de $\Phi_i, i = 1,2$ do modelo STVAR-Tree simulado, com $\alpha = 1\%$	76
Figura 5.4: Mediana de $\Phi_i, i = 1,2$ do modelo STVAR-Tree simulado, com $\alpha = 5\%$	77
Figura 5.5: Mediana de $\Phi_i, i = 1,2$ do modelo STVAR-Tree simulado, com $\alpha = 10\%$	77
Figura 5.6: Mediana de $\Phi_i, i = 1,2$ do modelo STVAR-Tree simulado, com $\alpha = 15\%$	78
Figura 5.7: Árvore estimada do modelo 7	94
Figura 5.8: FAC e FACP do logaritmo de ENA Sudeste	95
Figura 5.9: FAC e FACP do logaritmo de ENA Sul	95
Figura 5.10: FAC e FACP do logaritmo de ENA Nordeste	95
Figura 5.11: FAC e FACP do logaritmo de ENA Norte	95
Figura 5.12: Árvore estimada do modelo 12	110



## Lista de tabelas

Tabela 5.1: Probabilidades em um teste de hipóteses	66
Tabela 5.2: Experimento Monte Carlo – Simulação de um VAR	67
Tabela 5.3: Experimento Monte Carlo – Simulação de um STVAR-Tree	69
Tabela 5.4: Experimento Monte Carlo – Estimação dos parâmetros lineares	71
Tabela 5.5: Experimento Monte Carlo – Estimação da diagonal principal da matriz de covariâncias dos erros, $\Sigma_\varepsilon$	71
Tabela 5.6: Experimento Monte Carlo – Estimação dos parâmetros não-lineares $\gamma$ e $c$	73
Tabela 5.7: Experimento Monte Carlo – Estimação da diagonal da matriz de covariâncias dos erros, $\Sigma_\varepsilon$	75
Tabela 5.8: Divisão dos dados	79
Tabela 5.9: Estatísticas Descritivas das séries de ENA	79
Tabela 5.10: Estatísticas Descritivas do logaritmo natural das séries de ENA	80
Tabela 5.11: Matriz de Correlação do logaritmo natural das séries de ENA	81
Tabela 5.12: Teste Estatístico ADF para o logaritmo das séries de ENA	82
Tabela 5.13: Número de nós terminais e de parâmetros dos modelos estimados	83
Tabela 5.14: Critérios de Informação	84
Tabela 5.15: Estatísticas descritivas dos resíduos – <i>in-sample</i>	85
Tabela 5.16: Estatísticas de erro dos modelos – <i>in-sample</i>	86
Tabela 5.17: Testes de normalidade dos resíduos – <i>in-sample</i>	87
Tabela 5.18: Estatísticas descritivas dos resíduos (RC) – <i>out-of-sample</i>	88
Tabela 5.19: Estatísticas de erro dos modelos (RC) –	

<i>out-of-sample</i>	89
Tabela 5.20: Testes de normalidade dos resíduos (RC) – <i>out-of-sample</i>	89
Tabela 5.21: Estatísticas descritivas dos resíduos – <i>out-of-sample</i>	90
Tabela 5.22: Estatísticas de erro dos modelos (MM) – <i>out-of-sample</i>	90
Tabela 5.23: Testes de normalidade dos resíduos (MM) – <i>out-of-sample</i>	91
Tabela 5.24: Estatísticas descritivas dos resíduos (ARC) – <i>out-of-sample</i>	92
Tabela 5.25: Estatísticas de erro dos modelos (ARC) – <i>out-of-sample</i>	92
Tabela 5.26: Testes de normalidade dos resíduos (ARC) – <i>out-of-sample</i>	93
Tabela 5.27: Modelos PAR(p) para cada um dos sub-mercados	96
Tabela 5.28: Estimativa dos parâmetros dos modelos PAR(p)	96
Tabela 5.29: Estatísticas Descritivas das séries de resíduos	96
Tabela 5.30: Comparação dos modelos STVAR-Tree e PAR(p)	97
Tabela 5.31: Divisão dos dados	97
Tabela 5.32: Estatísticas Descritivas das séries de Preço spot	98
Tabela 5.33: Estatísticas Descritivas do logaritmo natural das séries de Preço spot	99
Tabela 5.34: Matriz de Correlação do logaritmo natural das séries de Preço spot	99
Tabela 5.35: Teste Estatístico ADF	100
Tabela 5.36: Teste Co-integração de Johansen	101
Tabela 5.37: Número de nós terminais e de parâmetros dos modelos estimados	102
Tabela 5.38: Critérios de Informação	103
Tabela 5.39: Estatísticas descritivas dos resíduos – <i>in-sample</i>	103
Tabela 5.40: Estatísticas de erro dos modelos – <i>in-sample</i>	104
Tabela 5.41: Teste de normalidade dos resíduos – <i>in-sample</i>	104

Tabela 5.42: Estatísticas descritivas dos resíduos (RC) – <i>out-of-sample</i>	105
Tabela 5.43: Estatísticas de erro dos modelos (RC) – <i>out-of-sample</i>	106
Tabela 5.44: Testes de normalidade dos resíduos (RC) – <i>out-of-sample</i>	106
Tabela 5.45: Estatísticas descritivas dos resíduos (MM) – <i>out-of-sample</i>	107
Tabela 5.46: Estatísticas de erro dos modelos (MM) – <i>out-of-sample</i>	108
Tabela 5.47: Testes de normalidade dos resíduos (MM) – <i>out-of-sample</i>	108
Tabela 5.48: Estatísticas descritivas dos resíduos (ARC) – <i>out-of-sample</i>	109
Tabela 5.49: Estatísticas de erro dos modelos (ARC) – <i>out-of-sample</i>	109
Tabela 5.50: Testes de normalidade dos resíduos (ARC) – <i>out-of-sample</i>	110
Tabela 5.51: Variáveis disponíveis de acordo com a ordem da defasagem	111
Tabela 5.52: Modelos selecionados pela modelagem Neuro-Fuzzy	112
Tabela 5.53: Estatísticas de erro dos modelos – <i>in-sample</i> e <i>out-of-sample</i>	112
Tabela 5.54: Comparação dos modelos STVAR-Tree e Neuro-Fuzzy	113