

# Previsão de Múltiplas Séries Temporais de Energia Elétrica Industrial Regional por Programação Genética Multigênica



**Autores: D.S. Soares, A.S. Koshiyama, D.P. Paula e R. Valle**

Universidade Federal do Rio de Janeiro – COPPE

Laboratório de Sistemas Avançados de Gestão da Produção

**Cidade Universitária - Ilha do Fundão, Rio de Janeiro, RJ, Brasil**

## Introdução

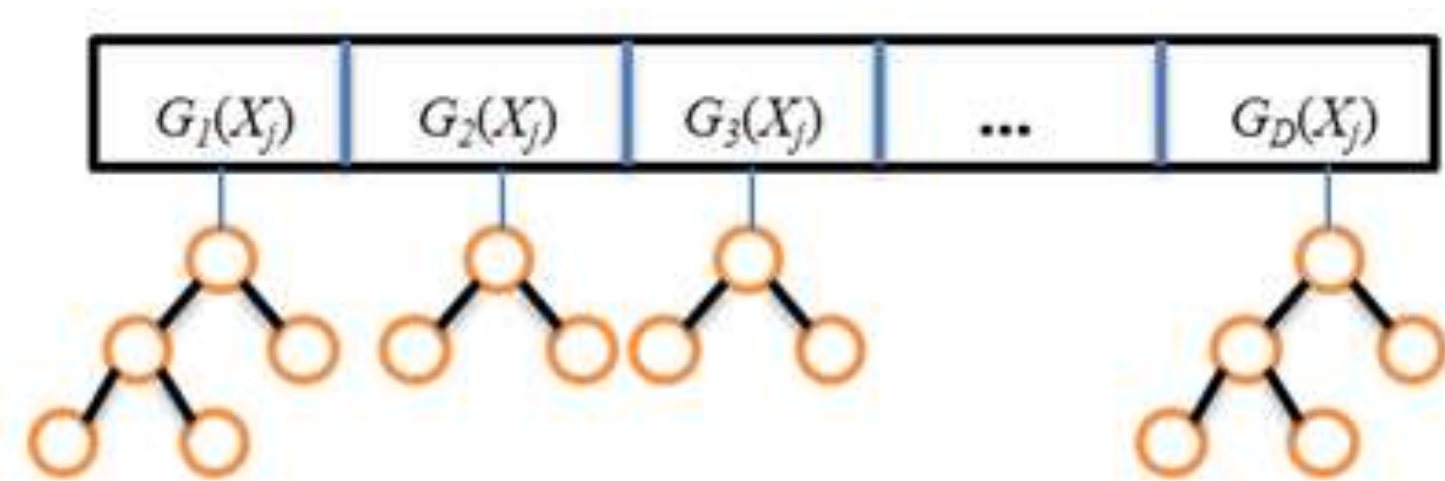
A gestão e planejamento do uso de recursos energéticos no Brasil tem sido foco de diversos trabalhos em diferentes áreas de pesquisa. Um dos maiores “players” do sistema energético brasileiro é o setor elétrico. Este integra uma grande gama de setores, desde a produção, até a distribuição da energia elétrica. Na área de previsão de séries temporais para o setor elétrico brasileiro, são encontrados fartamente trabalhos, principalmente usando modelos estatísticos (CHRISTO e SOUZA, 2006; MATTOS et al., 2008) e Redes Neurais Artificiais (FIGUEIREDO et al., 2007; DE ANDRADE, 2010). Um dos destaques são os modelos estocásticos Vetores Autorregressivos (VAR) (SCHMIDT e LIMA, 2004; IRFFI et al., 2009), que agregam informações provenientes de diferentes séries temporais, para efetuar a previsão destas todas conjuntamente. Uma alternativa é a partir do uso da Programação Genética (KOZA, 1992; LANGDON e POLI, 2002), um método advindo da inteligência computacional, que assume poucos pressupostos sobre o processo em análise. Esta ferramenta tem sido usada para previsão de séries temporais, obtendo resultados satisfatórios quando comparados a metodologias bem estabelecidas na literatura (LEE et al., 1997; JARA, 2011; SEO et al., 2013). É exíguo, em nível de Brasil, o uso desse método para auxiliar o setor elétrico, ainda mais em previsão simultânea de diversas séries históricas. Monitorar e detectar qualquer tipo de mudança na distribuição da base de dados;

## Objetivos

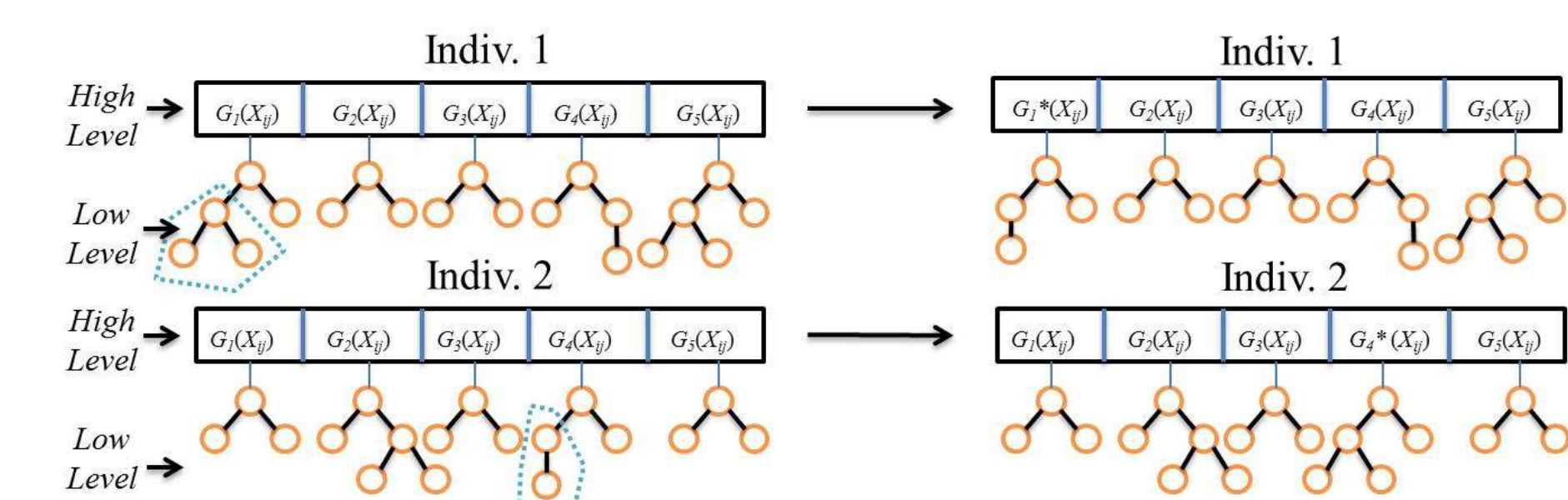
- Propor um modelo para a previsão de diversas séries temporais simultaneamente, visando auxiliar os tomadores de decisão no setor elétrico brasileiro;
- Utilizar uma variante da Programação Genética clássica, denominada Programação Genética Multigênica (HINCHLIFFE et al., 1996; SEARSON et al., 2007). Assim, deseja-se estabelecer um modelo pouco restritivo, de fácil uso, estimação e acurado.

## Programação Genética

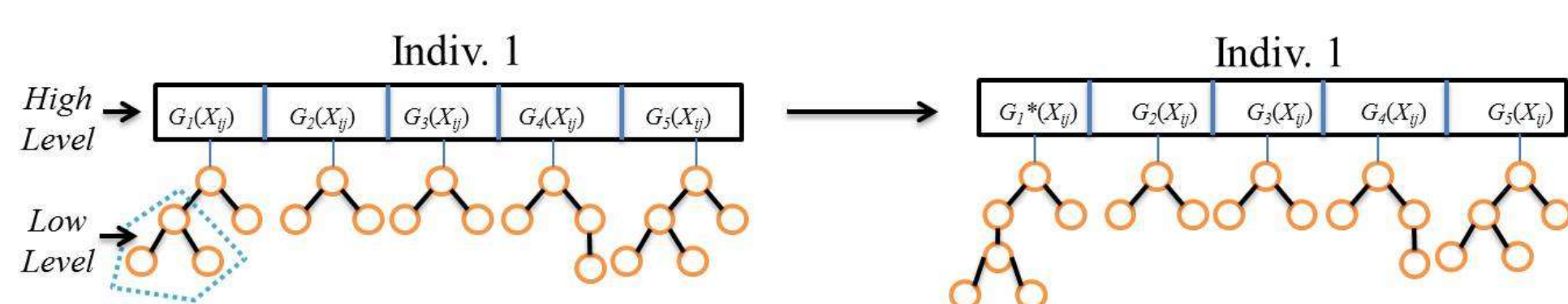
Programação Genética (PG) (KOZA, 1992; LANGDON e POLI, 2002) é um método pertencente à Computação Evolutiva, biologicamente inspirado e que sintetiza programas de computador para desenvolver alguma tarefa. Tipicamente, a PG usa uma população de indivíduos, cada um denotado por uma estrutura em árvores onde é codificada uma equação matemática que descreve um vetor de saída  $Y_i$ , usando um conjunto de  $J$  atributos de entradas  $X_{ij}$  ( $j=1,..., J$ ). Em comparação, a Programação Genética Multigênica (PGMG) (HINCHLIFFE et al., 1996; SEARSON et al., 2007) denota um indivíduo como um complexo de estruturas em árvores (equações), também chamados de genes, que da mesma forma recebem  $X_{ij}$ , buscando predizer  $Y_i$  (figura 1).



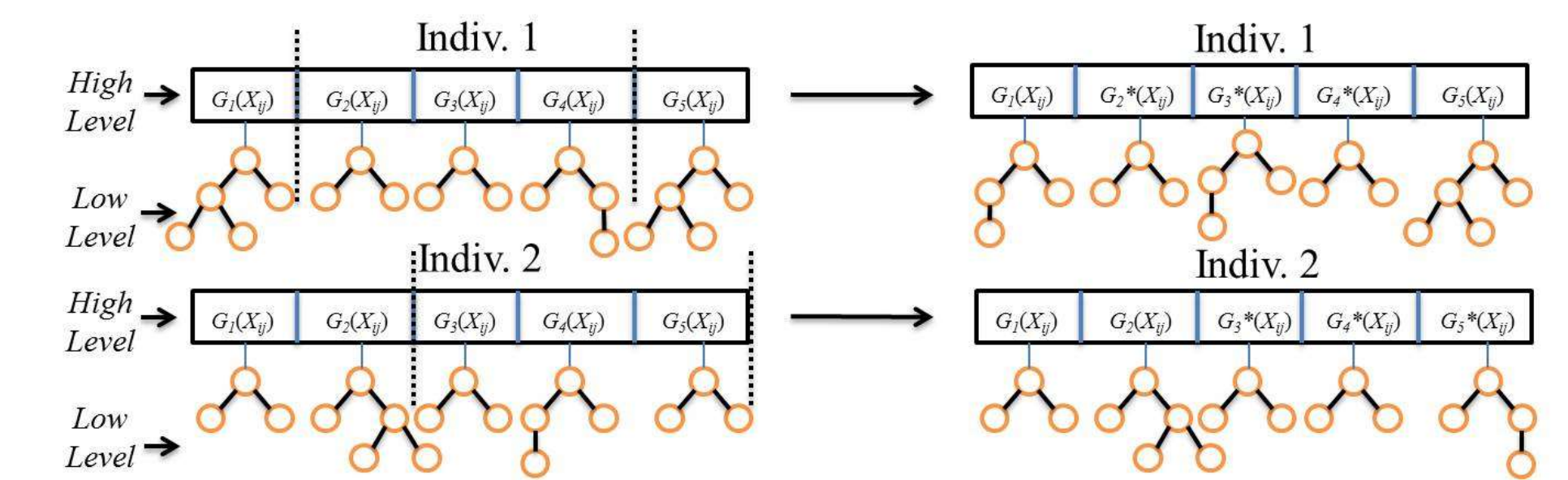
**Figura 1. Exemplo de Indivíduo Multigênico**



**Figura 2. Exemplo de Cruzamento de Baixo Nível**



**Figura 3. Exemplo de Mutação**



**Figura 4. Exemplo de Cruzamento de Alto Nível**

## Programação Genética Multigênica Aplicada a Previsão de Múltiplas Séries Temporais Simultaneamente

Considere  $Y_t = [Y_{t1}, ..., Y_{tJ}]$  o vetor de  $j$  séries temporais,  $j=1,...,J$ , em um determinado instante de tempo  $t$  qualquer,  $t=1,...,T$ . Além disso, defina por  $Y_{t-k}$  as  $k$ -defasagens da série temporal original,  $k=1,...,K$ . Busca-se então um conjunto de funções  $f_j(\cdot)$ , tal que dado valores de  $Y_{t-k}$  seja possível inferir sobre valores de  $Y_t$ . Formalmente:

$$Y_{t1} = f_1(Y_{t-k}) + \varepsilon_{t1} = \bar{Y}_{t1} + \varepsilon_{t1} \quad (2)$$

$$Y_{t2} = f_2(Y_{t-k}) + \varepsilon_{t2} = \bar{Y}_{t2} + \varepsilon_{t2} \quad (3)$$

...

$$Y_{tJ} = f_J(Y_{t-k}) + \varepsilon_{tJ} = \bar{Y}_{tJ} + \varepsilon_{tJ} \quad (4)$$

onde,  $\varepsilon_{tj}$   $j=1,...,J$ , é o erro entre o valor previsto e o real para a  $j$ -ésima série temporal. Logo, deseja-se minimizar cada  $\varepsilon_{tj}$ , visando tornar o valor previsto o mais próximo do real. Portanto, torna-se necessário minimizar a seguinte função de avaliação (objetivo):

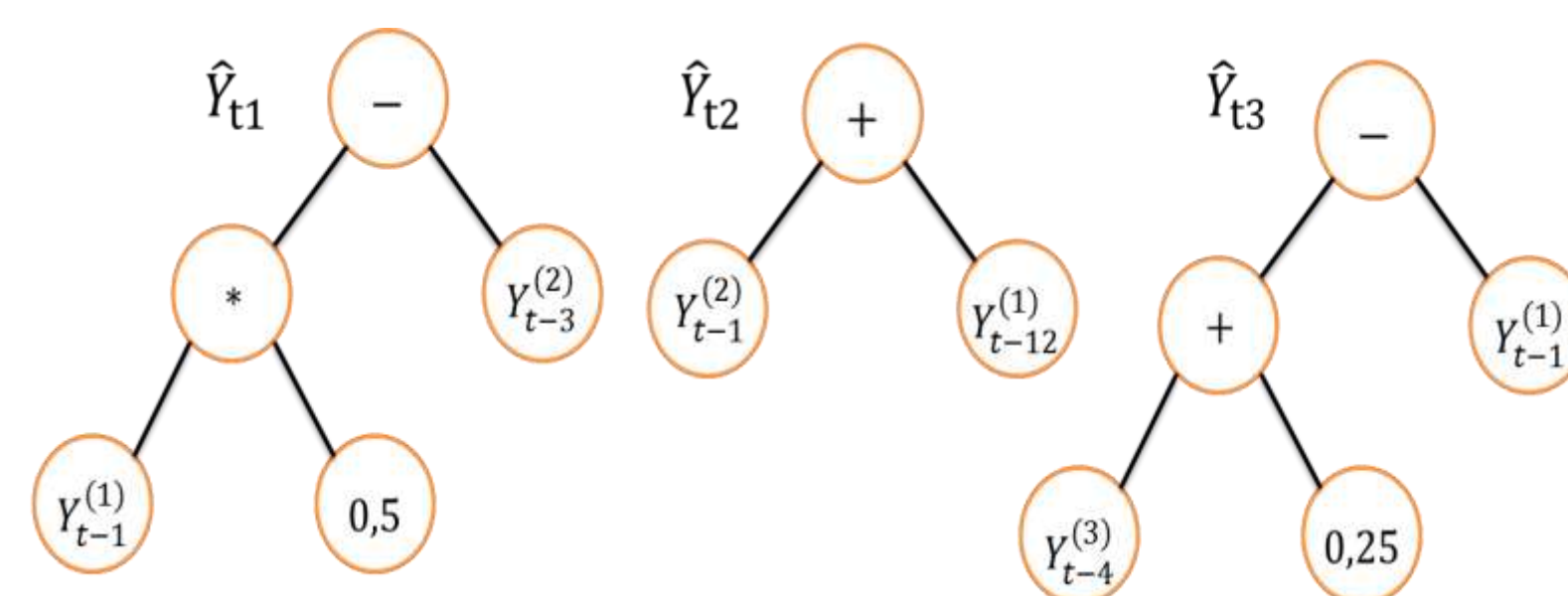
$$EAg = g(\varepsilon_{t1}, \varepsilon_{t2}, ..., \varepsilon_{tJ}) \quad (5)$$

em que  $EAg$  é a métrica de Erro Agregado e  $g(\cdot)$  é uma função de agregação. Esta função de agregação pode tomar diferentes formas (CALVO et al., 2002; YAGER, 2008), e deve ser definida pelo usuário. Como cada  $\varepsilon_{tj}$  pode tomar escalas distintas, como por exemplo, quando  $Y_{tj}$  refere-se à temperatura (°C) e a demanda de carga elétrica (GWh). Então, antes de calcular a métrica  $EAg$  é necessário tornar cada  $\varepsilon_{tj}$  adimensional. Se todo  $Y_{tj}$  sempre tomar valores maiores do que zero é possível definir:

$$\varepsilon_{tj} NORM = \frac{|Y_{tj} - \bar{Y}_{tj}|}{Y_{tj}} \quad (6)$$

com  $\varepsilon_{tj} NORM$  torna-se plausível a agregação pela função  $g(\cdot)$  entre os erros de cada  $j$ -ésima série temporal. Quando  $Y_{tj}$  pode assumir valores negativos é possível usar a métrica MASE (HYNDMAN e KOEHLER, 2006).

Exemplo de solução:



**Figura 5. Exemplo de Solução encontrada pela PGMG**

A partir de figura 5, tem-se as seguintes equações

$$\bar{Y}_{t1} = 0,5 * Y_{t-1} - Y_{t-3} \quad (7)$$

$$\bar{Y}_{t2} = 0,5 * Y_{t-1} - Y_{t-2} \quad (8)$$

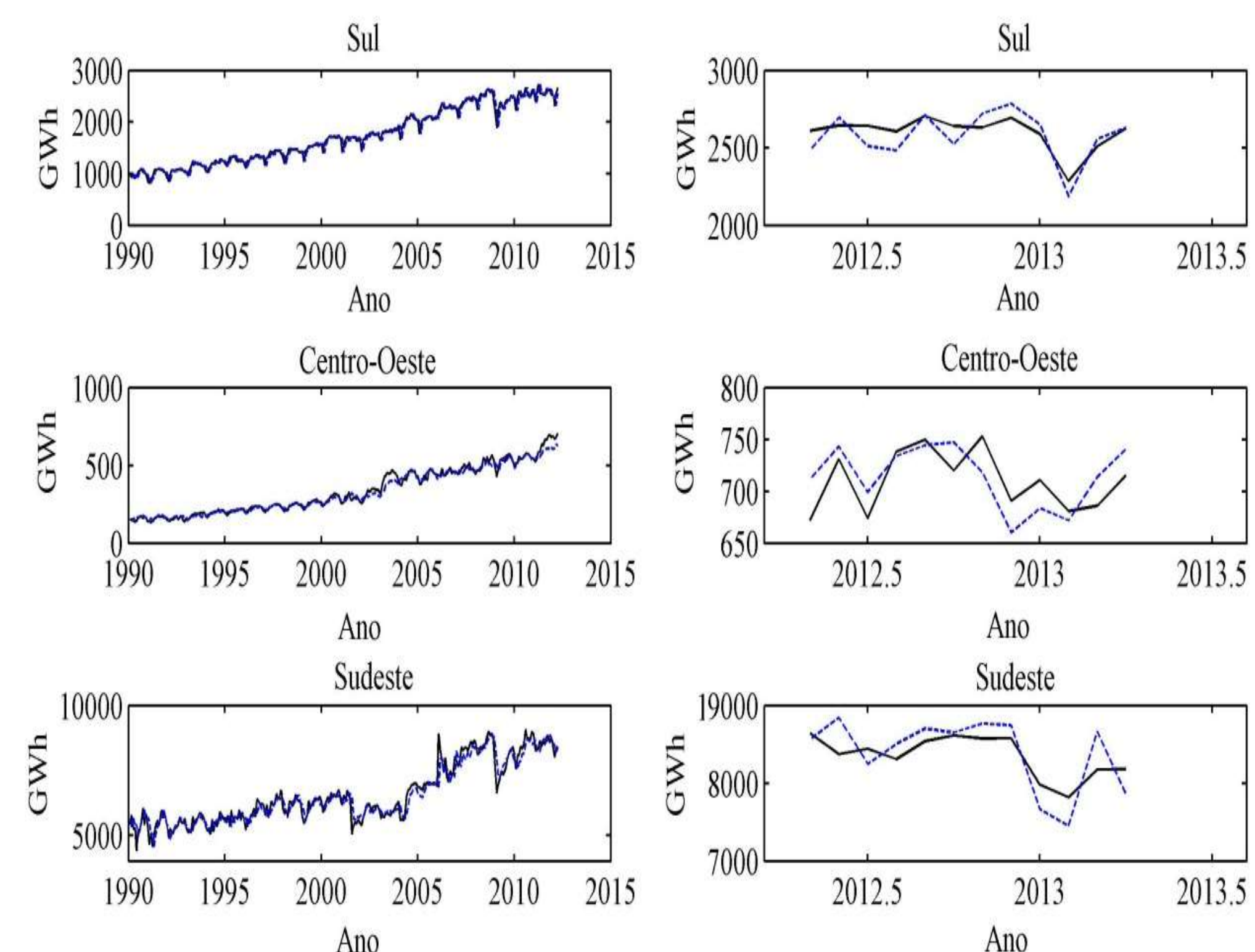
$$\bar{Y}_{t3} = 0,5 * Y_{t-4} - Y_{t-1} \quad (9)$$

Por exemplo a partir de (7) tem-se que o modelo para prever a primeira série temporal  $\bar{Y}_{t1}$  relaciona o seu passado mais recente  $Y_{t-1}$  e a terceira defasagem da série temporal  $Y_{t-3}$ .

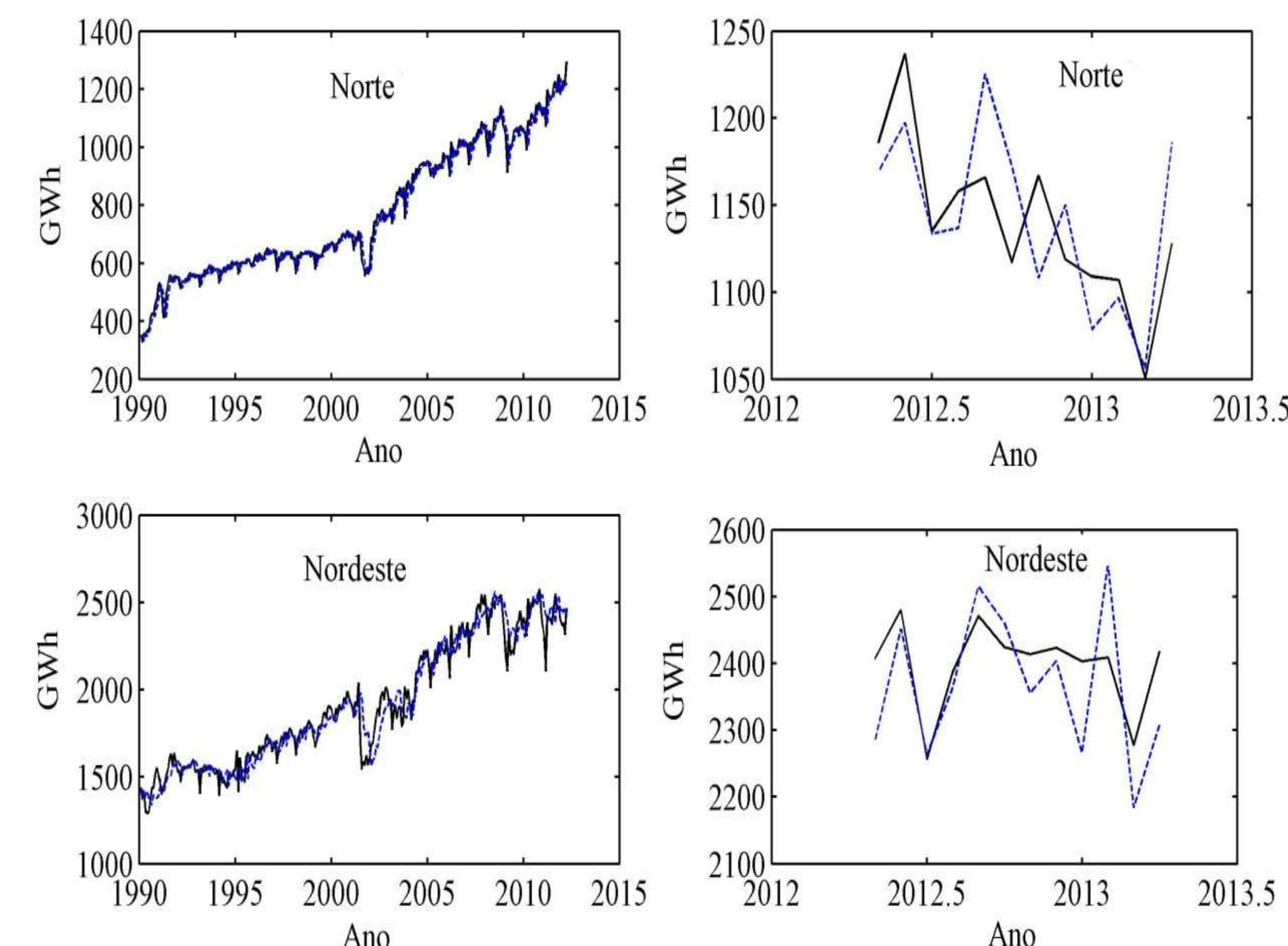
## Estudo de Caso

**Tabela 1. MAPE dos modelos PGMG e VAR.**

Série Temporal	PGMG		VAR	
	In-Sample	Out-Sample	In-Sample	Out-Sample
Sudeste	3,96%	3,06%	2,77%	4,01%
Sul	5,14%	3,20%	4,24%	3,67%
Norte	3,50%	2,85%	3,24%	3,02%
Centro-Oeste	3,33%	2,82%	2,98%	2,91%
Nordeste	3,78%	3,02%	3,08%	3,05%



**Figura 6. Valores observados in-sample e out-sample, ajustados e previstos.**



**Figura 7. Valores observados in-sample e out-sample, ajustados e previstos.**

## Conclusões

Este trabalho apresentou um modelo para previsão simultânea de diferentes séries temporais, baseado na Programação Genética Multigênica, visando assim possibilitar identificar e utilizar as possíveis inter-relações que podem favorecer a acurácia da previsão como um todo. Além disso, este modelo é automaticamente sintetizado, bastando ao usuário definir à priori informações que este tenha como relevante, que então, a Programação Genética Multigênica encarrega-se de estimar e efetuar a previsão. Ainda, o custo computacional é relativamente baixo, além de o modelo exigir pressupostos bem menos restritivos que os modelos estocásticos.

## Referências Bibliográficas

CHRISTO, Eliane da Silva; SOUZA, Reinaldo Castro. Uma abordagem estatística para a previsão de potência reativa em sistemas elétricos. Pesquisa Operacional, v. 26, n. 2, p. 361-381, 2006.

Figueiredo, K., Barbosa, C. H., Vellasco, M. M. B. R., Pacheco, M. A. C., & Contreras, R. J. (2007). Modelo de previsão de vazão com informação de precipitação utilizando redes neurais. Revista Brasileira de Recursos Hídricos, Julho/Setembro, 69-82.

KOZA, J.R. Genetic programming: on the programming of computers by means of natural selection. Michigan: MIT Press, 1992  
HINCHLIFFE, M.P.; WILLIS, M.J.; HIDEN, H.; THAM, M.T.; McKAY, B.; BARTON, G.W. Modeling chemical process systems using a multi-gene genetic programming algorithm. In: Genetic Programming: Proceedings of the First Annual Conference (late breaking papers), 56-65. Michigan: The MIT Press, 1996