

GPF-Forecast: um Sistema Fuzzy-Genético para Previsão

A.S. Koshiyama, T. Escovedo, M.M.B.R. Vellasco e R. Tanscheit

Departamento de Engenharia Elétrica
Pontifícia Universidade Católica do Rio de Janeiro
Rio de Janeiro, RJ, Brazil
{adriano, tatiana, marley, ricardo}@ele.puc-rio.br

Resumo— Este trabalho apresenta um Sistema Fuzzy-Genético para Previsão, denominado GPF-Forecast (*Genetic Programming Fuzzy for Forecasting*). Este modelo difere da abordagem tradicional de utilizar a meta-heurística genética como uma forma de aprender regras do tipo “se-então”. Além disso, não necessita da otimização dos coeficientes do consequente da regra, comum em modelos do tipo Takagi-Sugeno-Kang geralmente utilizados em problemas de previsão. O GPF-Forecast explora o potencial da Programação Genética Multigênica na busca das melhores variáveis de entrada e a forma funcional que mais bem se adequa ao problema. Ao mesmo tempo, proporciona ao usuário um entendimento linguístico da solução auferida. A base teórica do modelo é apresentada e são realizados quatro testes, visando a uma avaliação preliminar do modelo frente a métodos de previsão bem estabelecidos na literatura. Análises estatísticas das previsões realizadas e um exemplo da compreensão linguística fornecida pelo modelo acompanham os resultados.

Palavras-Chave— Previsão; Sistemas Fuzzy-Genéticos; Programação Genética Multigênica.

I. INTRODUÇÃO

Diariamente são encontradas situações de incerteza sobre o valor de um ativo, gestão de um processo, entre outros, em um dado instante de tempo futuro (um mês à frente, por exemplo). Uma forma de antecipar ou mitigar essa incerteza de forma sistemática é fazer uso de modelos de previsão. Neste sentido, a abordagem tradicional evoca o uso de modelos estocásticos, como Holt-Winters, Auto-Regressivos Integrados de Médias Móveis (ARIMA), etc. [1,2]. Porém a metodologia de estimação destes tende a ser restritiva sobre as características da série temporal sob análise, requisitando em alguns casos estacionariedade e normalidade do processo [3]. Modelos inteligentes de previsão, como Redes Neurais e Programação Genética [4,5] são capazes de fornecer soluções eficazes quando o uso de modelos estatísticos é inviável ou quando estes fornecem resultados pouco acurados.

Busca-se, com estes modelos descritos, previsões acuradas, negligenciando-se a extração de conhecimento da série temporal sob análise. Este conhecimento é relevante em áreas como Econometria [3] e Climatologia [6], por exemplo, em que informações sobre como os momentos passados da série temporal afetam o presente são vitais para a compreensão do impacto de uma medida econômica, ou sobre a ocorrência de um evento natural. Sistemas Fuzzy

para Previsão (SFP) [7] surgem como uma abordagem que fornece acurácia e também interpretabilidade linguística dos resultados obtidos. No âmbito dos SFP, empregam-se comumente modelos Neuro-Fuzzy [8,9] e Fuzzy-Genéticos [10,11]. No que tange os modelos Fuzzy-Genéticos, a meta-heurística genética é usada para encontrar os parâmetros de um SFP (quantidades relativas às funções de pertinência, base de regras, etc.), predominando abordagens do tipo Takagi-Sugeno-Kang, que também demandam a otimização dos parâmetros dos consequentes das regras [12]. Assim, esta abordagem tende a se tornar computacionalmente custosa.

Este trabalho apresenta um novo Sistema Fuzzy-Genético para previsão, denominado GPF-Forecast (*Genetic Programming Fuzzy for Forecasting*), como uma alternativa em termos de modelagem de séries temporais para os SFP baseados em regras. Este modelo é estimado via Programação Genética Multigênica [13], unindo a capacidade de seleção e aproximação de funções da meta-heurística à acepção linguística dos Sistemas de Inferência Fuzzy. Assim, a interpretabilidade, neste modelo, difere daquela dos SFP convencionais, fornecendo, em consequência, outros tipos de informação ao usuário. A próxima seção apresenta os fundamentos da Programação Genética Multigênica. A terceira seção exhibe o modelo GPF-Forecast, seus aspectos teóricos e os passos para estimação e previsão. A quarta seção apresenta os estudos de caso, resultados e discussões. Considerações finais e sugestões de trabalhos futuro encerram o trabalho.

II. PROGRAMAÇÃO GENÉTICA MULTIGÊNICA

Programação Genética (PG) [5,14] 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 deles denotado por uma estrutura em árvores onde é codificada uma equação matemática que descreve um vetor de saída Y_i a partir de um conjunto de J variáveis de entradas X_{ij} ($j=1, \dots, J$). Com base nessa estrutura realiza-se a PG tradicional. Já a Programação Genética Multigênica (PGMG) [13,15] 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).

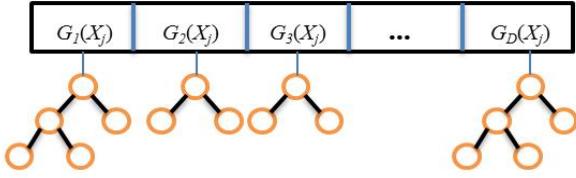


Fig. 1. Exemplo típico de um indivíduo multigênico.

Cada uma das D equações (G_d , $d=1, \dots, D$) é uma solução parcial para o problema, sendo que a saída final \hat{Y}_t pode ser resultado de uma combinação linear dos G_d . A partir disto é possível avaliar a qualidade global do indivíduo. É fácil verificar que, quando $D=1$, a PGMG reduz-se à PG clássica.

Em relação aos operadores genéticos, a operação de mutação na PGMG é similar à efetuada na PG clássica. No caso da operação de cruzamento, é necessário efetuar uma distinção no nível em que a operação é realizada: é possível aplicar o cruzamento no baixo e no alto nível. O baixo nível é o espaço onde é possível manipular as estruturas (terminais e funções) das equações presentes em um indivíduo. No caso, tanto a mutação quanto o cruzamento de baixo nível na PGMG são semelhantes ao que é efetuada na PG clássica.

Um exemplo de cruzamento de alto nível para um indivíduo com cinco equações é apresentado na Figura 2. O alto nível é o espaço em que se manipulam de forma macro as equações presentes no indivíduo. Logo, verifica-se que, a partir de dois pontos aleatórios, são permutadas equações (no caso três) de um indivíduo para o outro. Este ponto de corte aleatório pode ser simétrico - quando troca-se o mesmo número de equações entre indivíduos - ou assimétrico - caso contrário. A saída resultante é mais afetada pelo cruzamento de alto nível do que pela operação de cruzamento de baixo nível e pela mutação.

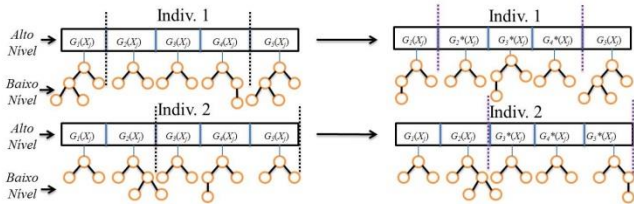


Fig. 2. Exemplo de cruzamento de alto-nível entre indivíduos.

A próxima seção aborda o modelo GPF-Forecast, que, a partir das bases apresentadas da PGMG, utiliza essa meta-heurística em consórcio com a Lógica Fuzzy para previsão, construindo-se, assim, em um Sistema Fuzzy-Genético.

III. MODELO GPF-FORECAST

Considere-se a série temporal Y_t como o vetor de saída (*target*) indexado no período de tempo t ($t=1, \dots, T$), e suas respectivas k defasagens ($k=1, \dots, t-1$) como o vetor de entrada Y_{t-k} . Em linhas gerais o modelo GPF-Forecast consiste em quatro etapas: fuzzificação, inferência, defuzzificação (previsão) e avaliação. O procedimento de inferência tem início na fuzzificação das entradas Y_{t-k} e saída Y_t . Em seguida, são sintetizadas as funções de cada

indivíduo usando-se a PGMG e efetuando-se a defuzzificação (previsão) por alguma heurística. Esta solução é então avaliada e o processo evolutivo é repetido o número de vezes que for necessário para que a solução atenda a critérios estabelecidos pelo usuário. O diagrama em blocos da Figura 3 ilustra o modelo GPF-Forecast.

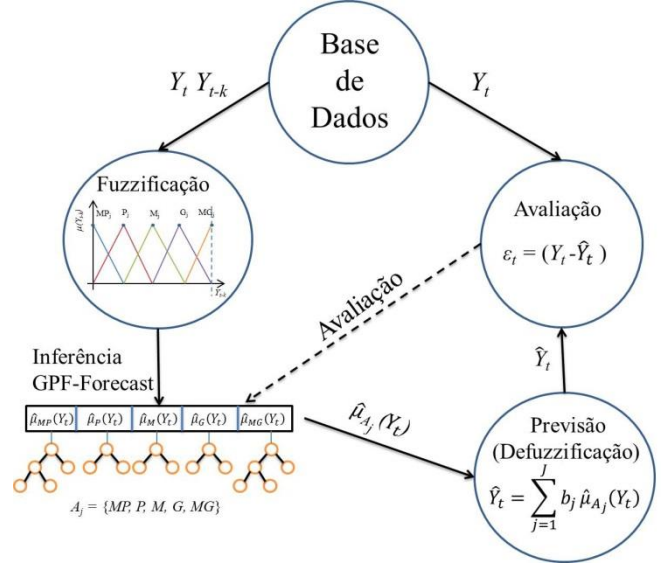


Fig. 3. Diagrama do modelo GPF-Forecast.

A tarefa de cada componente é explicitada passo a passo nas próximas subseções.

A. Fuzzificação

O processo de fuzzificação mapeia entradas precisas em graus de pertinência a conjuntos fuzzy [17]. Para a definição dos conjuntos fuzzy de Y_t (e, conseqüentemente, de Y_{t-k}), costuma-se recorrer ao auxílio de um especialista que compreenda a dinâmica de Y_t . Quando não se dispõe de um especialista, normalmente opta-se por dispor os conjuntos fuzzy de maneira uniforme e normalizada nos universos de discurso das entradas analisadas [17].

Seja A_{jk} o j -ésimo conjunto fuzzy para a k -ésima defasagem Y_{t-k} (quanto $k=0$, tem-se A_j). Considere-se, para exemplificação, um caso em que $j=1, 2, \dots, 5$. É, possível definir $A_{jk} = \{MP_k, P_k, M_k, G_k, MG_k\}$, tal que cada conjunto fuzzy manifeste linguisticamente, por exemplo, os conceitos muito pequeno (MP_k), pequeno (P_k), médio (M_k), grande (G_k) e muito grande (MG_k) para a k -ésima entrada. Por exemplo, A_{11} significa o conjunto fuzzy MP de Y_{t-1} , ou MP_1 . Na presença de variáveis categóricas, sugere-se que estas sejam inseridas na base de dados codificadas no formato 1 de N, para tornar o mais claro possível a aceção fornecida pelo modelo. A partir da fuzzificação das entradas inicia-se o processo de inferência GPF-Forecast.

B. Inferência GPF-Forecast

Considere-se que Y_t e Y_{t-k} tenham sido fuzzificados de forma uniforme e normalizada. Logo, tem-se as funções de pertinência $\mu_{MP_k}(Y_{t-k})$, $\mu_{P_k}(Y_{t-k})$, $\mu_{M_k}(Y_{t-k})$, $\mu_{G_k}(Y_{t-k})$ e $\mu_{MG_k}(Y_{t-k})$, para $k=0,1,\dots,t-1$. Deseja-se inferir quais e quantos conjuntos fuzzy de Y_{t-k} relacionam-se com os conjuntos fuzzy de Y_t . Em outras palavras, deseja-se saber como os conceitos linguísticos das variáveis de entradas estão ligados aos da variável de saída.

A partir de $\mu_{A_{jk}}(Y_{t-k})$ será possível extrair conclusões sobre a dinâmica de cada $\mu_{MP}(Y_t)$, $\mu_P(Y_t)$, $\mu_M(Y_t)$, $\mu_G(Y_t)$ e $\mu_{MG}(Y_t)$, e, indiretamente, sobre Y_t . Logo, busca-se um mapeamento tal que:

$$\mu_{MP}(Y_t) = f_1\left(\mu_{A_{jk}}(Y_{t-k})\right) + \varepsilon_{t1} = \hat{\mu}_{MP}(Y_t) + \varepsilon_{t1} \quad (1)$$

$$\mu_P(Y_t) = f_2\left(\mu_{A_{jk}}(Y_{t-k})\right) + \varepsilon_{t2} = \hat{\mu}_P(Y_t) + \varepsilon_{t2} \quad (2)$$

$$\mu_M(Y_t) = f_3\left(\mu_{A_{jk}}(Y_{t-k})\right) + \varepsilon_{t3} = \hat{\mu}_M(Y_t) + \varepsilon_{t3} \quad (3)$$

$$\mu_G(Y_t) = f_4\left(\mu_{A_{jk}}(Y_{t-k})\right) + \varepsilon_{t4} = \hat{\mu}_G(Y_t) + \varepsilon_{t4} \quad (4)$$

$$\mu_{MG}(Y_t) = f_5\left(\mu_{A_{jk}}(Y_{t-k})\right) + \varepsilon_{t5} = \hat{\mu}_{MG}(Y_t) + \varepsilon_{t5} \quad (5)$$

de tal forma que o erro ε_{tj} entre cada $\mu_{MP}(Y_t)$, $\mu_P(Y_t)$, $\mu_M(Y_t)$, $\mu_G(Y_t)$ e $\mu_{MG}(Y_t)$ observado e cada grau de pertinência (entre 0 e 1) estimado pelos conjuntos fuzzy de entrada ($\hat{\mu}_{MP}(Y_t)$, $\hat{\mu}_P(Y_t)$, $\hat{\mu}_M(Y_t)$, $\hat{\mu}_G(Y_t)$ e $\hat{\mu}_{MG}(Y_t)$) seja minimizado.

Para estabelecer a função de ligação $f_p(\cdot)$ é necessário definir a forma funcional e escolher quais conjuntos fuzzy das variáveis de entrada devem figurar em cada mapeamento dos conjuntos fuzzy da variável de saída. Esta decisão é complexa, pois envolve escolhas adversas, entre as quais:

1. Forma funcional: linear ou não linear; idêntico para todas as equações ou diferente em algumas.
2. Seleção dos conjuntos fuzzy: quais, quanto e em quais mapeamentos $f_p(\cdot)$ de saída eles devem ingressar.

Além destes dois fatores, deve-se ter em conta que uma função com muitos parâmetros e concebida de forma não linear torna a aceitação do modelo mais obscura, apesar de possivelmente tornar a predição mais acurada. Logo, deve-se buscar ao máximo a preservação da inteligibilidade linguística do modelo, acompanhada de um mínimo de perdas em precisão. Uma forma de possibilitar o alcance de ambos os objetivos, visando a sintetizar automaticamente a forma funcional e simultaneamente escolher os conjuntos fuzzy das variáveis de entrada, é o uso da PGMG.

Uma vez que na PGMG um indivíduo representa um complexo de programas, é possível definir que cada indivíduo possui um número de programas necessário para produzir um número de saídas necessárias. No caso de cinco conjuntos fuzzy associados à variável de saída, é necessário manter cinco programas por indivíduo, com cruzamento de alto nível simétrico e com inicialização estrita de cinco programas por indivíduo.

Por exemplo, para a PGMG sintetizar a equação (5), considere-se o conjunto de terminais (funções de pertinência dos conjuntos associados às variáveis de entrada e constantes) e funções mostrados na Tabela 1 e ilustrados na figura 4.

TABELA 1. TERMINAIS E FUNÇÕES DE ENTRADA PARA O MODELO GPF-FORECAST.

Terminais	Funções
$\mu_{MP_k}(Y_{t-k})$, $\mu_{P_k}(Y_{t-k})$, $\mu_{M_k}(Y_{t-k})$, $\mu_{G_k}(Y_{t-k})$ e $\mu_{MG_k}(Y_{t-k})$ e constantes geradas aleatoriamente no intervalo entre [a,b].	Soma, subtração, raiz-quadrada, potência quadrática e produto.

No conjunto de terminais estão presentes os cinco conjuntos fuzzy de entrada de cada defasagem Y_{t-k} , além de constantes efêmeras geradas pela PGMG no intervalo [a,b], onde a e b são números reais.

Do conjunto de funções, as operações de soma e subtração denotam o sentido em que variações expressas pelos termos linguísticos das variáveis de entrada implicam em variações positivas – afirmação, ou negativas – negação, no grau de compatibilidade do termo linguístico da saída. As operações de raiz quadrada e de potência quadrática funcionam como os intensificadores e redutores de Zadeh [16], seja intensificando ou diluindo o sentido de uma expressão a partir de transformações aplicadas na região fuzzy definida por cada função de pertinência (de grande para pouco grande, de muito pequeno para muitíssimo pequeno, etc.). A função produto possui duas funcionalidades:

1. Produto tipo 1: Associar uma constante a um $\mu_{A_{jk}}(Y_{t-k})$;
2. Produto tipo 2: Combinar um $\mu_{A_{jk}}(Y_{t-k})$ com outro conjunto fuzzy de entrada.

Sem perda de generalidade, considere-se $\hat{\mu}_P(Y_t) = 0,8 \mu_{MG_1}(Y_{t-1})$ como exemplo da primeira característica do produto. Então, se $\mu_{MG_1}(Y_{t-1}) = 1$, segue-se que $\hat{\mu}_P(Y_t) = 0,8$. Logo, a constante usada em conjunto com a função produto tipo 1 denota o nível da transmissão que a ativação específica de um conjunto fuzzy de entrada acarreta em graus de pertinência ao de saída.

Para o segundo caso, seja a expressão abaixo:

$$\hat{\mu}_{MG}(Y_t) = \mu_{MP_1}(Y_{t-1})\mu_{P_1}(Y_{t-1}) + \mu_{MG_1}(Y_{t-1})\mu_{P_1}(Y_{t-1}) + \mu_{MG_1}(Y_{t-1})\mu_{M_2}(Y_{t-2}) \quad (6)$$

a primeira parte de (6), $\mu_{MP_1}(Y_{t-1})\mu_{P_1}(Y_{t-1})$, considerando-se partições uniformes e normalizadas (veja exemplo na figura 4), pode ser interpretada linguisticamente como “aproximadamente entre muito pequeno e pequeno”. Na segunda parte da expressão, $\mu_{MG_1}(Y_{t-1})\mu_{P_1}(Y_{t-1}) = 0$, pois os conjuntos não são adjacentes e, portanto o produto interno entre eles é nulo (ver figura 4).

Para a terceira parte, $\mu_{MG_1}(Y_{t-1})\mu_{M_2}(Y_{t-2})$, apesar de possivelmente existir um Y_{t-1} e um Y_{t-2} tal que $\mu_{MG_1}(Y_{t-1})\mu_{M_2}(Y_{t-2}) > 0$, a interpretação linguística é mais complexa. Pode-se entender como “quanto maior for a correspondência (ou proximidade) de Y_{t-1} a MG_1 e Y_{t-2} a M_2 , maior será o grau de pertinência dessa entrada em $\hat{\mu}_{MG}(Y_t)$ ”. De forma mais simplificada: “se Y_{t-1} é MG_1 e Y_{t-2} é M_2 então Y_t é MG com grau de pertinência estimado em $\mu_{MG_1}(Y_{t-1})\mu_{M_2}(Y_{t-2})$ ”. Em outras palavras, este segundo caso do produto modela a interação entre conceitos de diferentes variáveis, revelando inter-relações e peculiaridades do comportamento destas, que, dependendo da situação, podem ser significativas para o analista.

C. Defuzzificação (Previsão)

O procedimento de defuzzificação associa a um conjunto fuzzy de saída um valor preciso, a partir de alguma heurística decisória (média dos máximos, centro de gravidade, etc.) [18].

No caso do modelo GPF-Forecast a defuzzificação é realizada pelo método da altura:

$$\hat{Y}_t = \frac{\sum_{j=1}^n w_j \hat{\mu}_{A_j}(Y_t)}{\sum_{j=1}^n \hat{\mu}_{A_j}(Y_t)} \quad (7)$$

onde \hat{Y}_t é a saída defuzzificada estimada, A_j é o conjunto que possui os termos linguísticos da variável Y_t , e w_j representa os respectivos centroides de cada conjunto fuzzy de saída. Considerando-se os conjuntos fuzzy da figura 4, tanto para as variáveis de entrada como para a de saída, a equação (7) torna-se:

$$\hat{Y}_t = \frac{\sum_{j=1}^5 b_j \hat{\mu}_{A_j}(Y_t)}{\sum_{j=1}^5 \hat{\mu}_{A_j}(Y_t)} \quad (8)$$

em que $\sum_{j=1}^5 b_j$ são $b_1 = b_{MP}$, $b_2 = b_P$, $b_3 = b_M$, $b_4 = b_G$ e $b_5 = b_{MG}$ são os parâmetros de centro de cada função de pertinência triangular (vértice superior) associadas aos conjuntos fuzzy de saída, e $\sum_{j=1}^5 \hat{\mu}_{A_j}(Y_t)$ são $\hat{\mu}_{A_1}(Y_t) = \hat{\mu}_{MP}(Y_t)$, $\hat{\mu}_{A_2}(Y_t) = \hat{\mu}_P(Y_t)$, $\hat{\mu}_{A_3}(Y_t) = \hat{\mu}_M(Y_t)$, $\hat{\mu}_{A_4}(Y_t) =$

$\hat{\mu}_G(Y_t)$ e $\hat{\mu}_{A_5}(Y_t) = \hat{\mu}_{MG}(Y_t)$ são as saídas de cada equação gerada pela PGMG.

D. Avaliação

A partir da defuzzificação é possível computar o Erro Quadrático Médio (EQM) ou qualquer outra métrica de erro entre o valor da saída estimada \hat{Y}_t e a observada Y_t , de forma a avaliar um determinado indivíduo na população da PGMG. Outra forma viável de avaliação é o Erro Fuzzy Absoluto Total (EFAT), dado por:

$$EFAT = \sum_{j=1}^n \frac{|\mu_{A_j}(Y_t) - \hat{\mu}_{A_j}(Y_t)|}{\mu_{A_j}(Y_t)} \quad (9)$$

esta medida é relevante pela dimensionalidade restrita de cada variável (intervalo [0,1]), além de não mascarar possíveis distorções geradas pelo processo de defuzzificação tradicional. Assim, é possível minimizar o EFAT, e, a seguir computar os erros resultantes de outras métricas (EQM, Erro Absoluto Médio Percentual (MAPE), etc.).

IV. RESULTADOS E DISCUSSÕES

A. Base de Dados e Procedimentos de Análise

De forma a avaliar o modelo GPF-Forecast foram efetuadas comparações com as bases de dados e modelos estimados em [19]. Fez-se uso de três bases de dados - venda de passagens aéreas mensais [2], captura anual de lincas no Canadá e cotação em R\$ diária da ação VALE5PN - de nove modelos de previsão e de dois de combinação de previsão. Efetuaram-se comparações somente com os modelos que obtiveram melhores resultados: modelo Ingênuo, Holt-Winters Aditivo (HWA) e Multiplicativo (HWM), um Sazonal Auto-Regressivo Integrado com Médias Móveis (SARIMA), Perceptron de Múltiplas Camadas (MLP) e a PG clássica. Além das três bases de dados, foi adicionada mais uma sobre demanda mensal de carga elétrica da Light SA, de dezembro de 2002 a abril de 2011. A previsão efetuada foi do tipo *multi-step*, com horizonte de previsão de 12 passos à frente para a série de Passagens Aéreas [19], 14 para a de captura de lincas [19], 5 para a de cotação da VALE5PN [19] e 15 para a série de demanda de carga elétrica, sendo que este mesmo número de dados foi omitido destas séries para fins de avaliação na fase de previsão *multi-step*.

Para fins de comparação foram realizadas 15 execuções independentes do GPF-Forecast, com o critério de parada em 50000 avaliações [19]. Essas 15 execuções visam à redução dos efeitos gerados pelos geradores de números aleatórios nos resultados obtidos. A partir dos resultados dessas 15 execuções foram computadas médias para os valores previstos na fase de previsão *multi-step*. A seguir foi computado o MAPE para cada modelo e os Erros Quadráticos do modelo GPF-Forecast e do melhor (ou segundo melhor) modelo para cada base de dados, visando à aplicação do teste de Diebold-Mariano [20] para avaliar se o

modelo GPF-Forecast gerou previsões mais acuradas ou não, no horizonte de previsão estudado (previamente foi aplicado o teste de normalidade de Jarque-Bera [3]). Considerou-se um nível de significância de 5% e, para a defuzzificação o método da altura (7).

As funções de pertinência foram dispostas de forma uniforme, conforme mostrado na Figura 4. Os parâmetros, terminais e operações matemáticas usadas no GPF-Forecast são apresentados na Tabela 2. Em sua maioria, os valores são semelhantes aos de [19], com o objetivo de se aproximar ao máximo das condições enfrentadas pela PG usada nesse estudo. Caso ocorram empates na busca pelo melhor indivíduo, a pressão lexicográfica [23] é um critério de desempate, onde o melhor indivíduo é o que possuir menor número de nós. As parametrizações efetuadas para os demais modelos de previsão são exibidas em [19]. O modelo GPF-Forecast foi implementado em MATLAB R2010a [21], e as análises estatísticas foram executadas no pacote R [22].

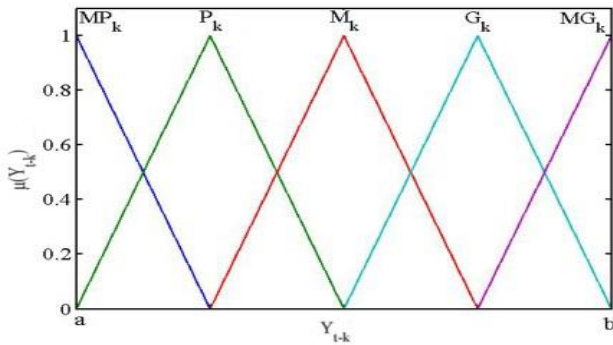


Fig. 4. Conjuntos fuzzy para a k-ésima defasagem.

TABELA 2. PARÂMETROS GENÉRICOS DO MODELO GPF-FORECAST.

Parâmetros	Valores
Tamanho da População	100
Máxima Profundidade da Árvore	7
Número de Gerações	500
Tamanho do Torneio	2
Taxa de Cruzamento de Alto Nível	65%
Taxa de Cruzamento de Baixo Nível	85%
Taxa de Mutação	10%
Taxa de Clonagem	5%
Taxa de Elitismo	1%
Pressão Lexicográfica [23]	Sim
Terminais	$\mu_{MP_k}(Y_{t-k}), \mu_{P_k}(Y_{t-k}), \mu_{M_k}(Y_{t-k}), \mu_{G_k}(Y_{t-k})$ e $\mu_{MG_k}(Y_{t-k})$ e constantes geradas aleatoriamente no intervalo entre [0,1].
Operações Matemáticas	Soma, subtração e produto.

B. Resultados Obtidos

A Tabela 3 apresenta os principais resultados e análises estatísticas efetuadas para a comparação entre os modelos. Verifica-se que o modelo GPF-Forecast foi aquele que obteve menor erro de previsão na série de demanda de carga

elétrica quando comparado aos demais (p-valor < 0,05). Nas demais séries, os modelos concorrentes obtiveram melhores resultados (p-valor < 0,01), exceto na série de captura de lincos onde o GPF-Forecast obteve o melhor resultado em termos de MAPE, porém não substancialmente menor do que o modelo SARIMA (p-valor = 0,14). A figura 5 mostra os valores ajustados e previstos dos modelos GPF-Forecast, PG e MLP para a série de demanda de carga (os três melhores modelos).

TABELA 3. PRINCIPAIS RESULTADOS DOS MODELOS NA FASE DE PREVISÃO.

Modelos	Passagens	Lincos	VALE5PN	Carga
Ingênuo	14,25%	32,76%	2,23%	13,03%
HWA	7,41%	26,36%	2,07%	6,31%
HWM	3,36%	27,16%	2,05%	7,83%
(S)ARIMA	3,65%	9,47%	2,19%	7,97%
MLP	5,90%	10,11%	1,64%	6,18%
PG	4,63%	18,32%	2,83%	6,15%
GPF-Forecast	5,09%	9,22%	2,02%	5,78%
Comparação	p-valor	Conclusão		
GPF-Forecast--HWM	<0,01	Diferença Significativa		
GPF-Forecast--(S)ARIMA	0,14	Inconclusivo		
GPF-Forecast--MLP	<0,01	Diferença Significativa		
GPF-Forecast--PG	<0,01	Diferença Significativa		

As equações (10-14) apresentam uma solução elaborada pelo modelo GPF-Forecast para a série de previsão da demanda de carga elétrica.

$$\hat{\mu}_{MP}(Y_t) = \mu_G(Y_{t-6}) \mu_P(Y_{t-12}) \quad (10)$$

$$\hat{\mu}_P(Y_t) = 0,88\mu_P(Y_{t-1})\mu_P(Y_{t-12}) - 0,25 \quad (11)$$

$$\hat{\mu}_M(Y_t) = \mu_P(Y_{t-5})^2 \quad (12)$$

$$\hat{\mu}_G(Y_t) = \mu_P(Y_{t-5}) \quad (13)$$

$$\hat{\mu}_{MG}(Y_t) = \mu_G(Y_{t-12})\mu_{MG}(Y_{t-12}) \quad (14)$$

Constata-se, da equação (10), que se a demanda de carga elétrica foi grande no sexto mês anterior e pequena no décimo segundo mês anterior (idêntico ao mês atual, mas do ano anterior) a demanda atual tenderá a ser muito pequena. De forma similar, se a demanda for pequena no mês antecedente e também pequena no mês idêntico ao atual, mas do ano anterior, então a demanda atual é pequena (11). A demanda atual será média (12) se a demanda há cinco meses for extremamente pequena, e, por sua vez, esta será grande atualmente (13) se for pequena no quinto mês anterior. Por fim, se a demanda for aproximadamente entre grande e muito grande no décimo segundo mês anterior, então a demanda por energia elétrica será muito grande (14).

Cabe ressaltar a aparente captação, por parte do modelo, das variações sazonais presentes na série. Isto pode ser observado nas equações (10) e (14) mais precisamente. Como a demanda de carga alterna uma temporada de menor fluxo no meio do ano (março e agosto) e maior intensidade nos demais meses (demandas de final de ano, verão, etc., ver Figura 6), há dois movimentos aproximadamente bem

definidos no ano. Pela equação (14) é possível observar que, se a demanda de energia elétrica for aproximadamente muito grande (o que costuma ocorrer entre os meses de setembro a fevereiro) no mês respectivo ao atual, mas do ano anterior, então a demanda de energia nesse instante será muito grande. Na equação (10) tem-se a ideia inversa, mais apropriada para os meses de demanda pequena.

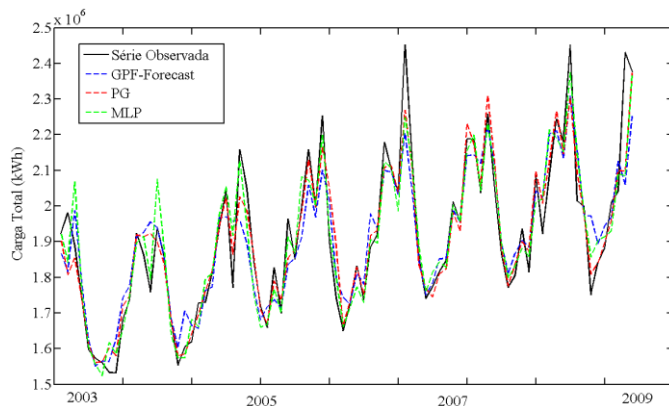


Fig. 5. Valores observados e ajustados na fase de treinamento dos modelos GPF-Forecast, PG e MLP.

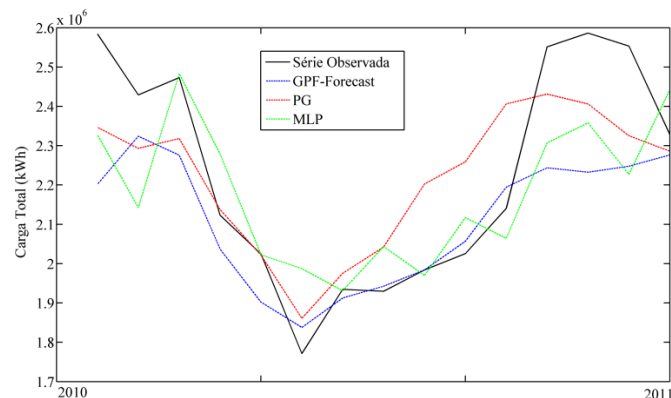


Fig. 6. Valores observados e previstos (*multistep*) pelos modelos GPF-Forecast, PG e MLP.

V. CONCLUSÕES

Este trabalho apresentou o modelo GPF-Forecast. Os resultados obtidos nos testes exibiram a capacidade do modelo de aliar acurácia com uma interpretabilidade linguística. As análises estatísticas demonstraram o potencial do modelo, quando comparado a métodos de previsão bem estabelecidos na literatura. Trabalhos futuros devem contemplar um maior conjunto de benchmarks (competição M3, ou NN3), visando a uma melhor análise do modelo. Além disso, pretende-se investigar a sua possível utilização em aplicações reais, como a de demanda de carga elétrica em companhias de distribuição, incorporando variáveis auxiliares (como a temperatura), com o objetivo de obter ganhos em acurácia e maior aceção do modelo.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] P.A. Morettin and C.M. Tolói, *Análise de Séries Temporais*. São Paulo: Editora Edgard Blücher, 2004.
- [2] G.E.P. Box, G.M. Jenkins and G.C. Reinsel, *Time series analysis: forecasting and control*. New Jersey: Wiley, 2011.
- [3] W. Enders, *Applied econometric time series*. New Jersey: John Wiley & Sons, 2008.
- [4] S. Haykin, *Neural networks: a comprehensive foundation*. New Jersey: Prentice Hall, 2007.
- [5] J.R. Koza, *Genetic programming: on the programming of computers by means of natural selection*. Massachusetts: The MIT Press, 1992.
- [6] A.F. Zuur, E.N. Ieno and G.M. Smith, *Analysing ecological data*. New York: Springer, 2007.
- [7] A.K. Palit and D. Popovic, *Computational intelligence in time series forecasting: theory and engineering applications*. New York: Springer-Verlag, 2005.
- [8] N. Kasabov, and Q. Song, DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. *IEEE Transactions on Fuzzy Systems*, vol.10, n.2, pp.144-154, 2002.
- [9] M.M.B.R. Vellasco, M.A.C. Pacheco, L.S. Ribeiro-Neto and F.J. de Souza, Electric load forecasting: evaluating the novel hierarchical neuro-fuzzy BSP model. *International journal of electrical power & energy systems*, vol.26, n.2, pp.131-142, 2004.
- [10] O. Cordon, F. Gomide, F. Herrera, F. Hoffmann, and L. Magdalena, Ten years of genetic fuzzy systems: current framework and new trends. *Fuzzy sets and systems*, vol.141, n.1, pp.5-31, 2004.
- [11] A. Tsakonas, Local and global optimization for Takagi-Sugeno fuzzy system by memetic genetic programming. *Expert Systems with Applications*, vol. 40, n.8, pp.3282-3298, 2013.
- [12] J.S.R. Jang, C.T. Sun and E. Mizutani, *Neuro-Fuzzy and Soft Computing*. New Jersey: Prentice-Hall, 1997.
- [13] M.P. Hinchliffe, M.J. Willis, H. Hiden, M.T. Tham, B. McKay and G.W. Barton, Modeling chemical process systems using a multi-gene genetic programming algorithm. In: *Genetic Programming: Proceedings of the First Annual Conference*. Massachusetts: MIT Press, 1996, pp. 56-65.
- [14] W.B. Langdon and R. Poli, *Foundations of Genetic Programming*. Heidelberg: Springer-Verlag, 2002.
- [15] D.P. Searson, M.J. Willis and G.A. Montague, Co-evolution of non-linear PLS model components. *Journal of Chemometrics*, vol. 2, pp. 592-603, 2007.
- [16] W. Pedrycz and F. Gomide, *Fuzzy systems engineering: toward human-centric computing*. New Jersey: Wiley, 2007.
- [17] F. Berlanga, A.J. Riveira, M. Del Jesus and F. Herrera, GP-COACH: Genetic Programming-based learning of Compact and Accurate fuzzy rule-based classification systems for High-dimensional problems. *Information Sciences*, vol.180, pp.1183-1200, 2010.
- [18] S. Roychowdhury and W. Pedrycz, A survey of defuzzification strategies. *International Journal of intelligent systems*, vol. 16, n. 6 pp.679-695, 2001.
- [19] A.S. Koshiyama, T. Escovedo, D.M. Dias, M.M.B.R. Vellasco and M.A.C. Pacheco, Combining Forecasts: A Genetic Programming Approach. *International Journal of Natural Computing Research*, vol.3, n.3, pp. 41-58, 2012.
- [20] F.X. Diebold and R.S. Mariano. Comparing predictive accuracy. *Journal of Business & economic statistics*, vol. 20, n.1, 2002.
- [21] MATLAB. MATLAB 7.10.0 (R2010a). Massachusetts: The MathWorks Inc., 2010.
- [22] R Core Team, *R: A language and environment for statistical computing*. Viena: R Foundation for Statistical Computing, 2012.
- [23] S. Luke and L. Panait, Lexicographic parsimony pressure. In: *Proceedings of the Genetic and Evolutionary Computation Conference*, W. B. Langdon et al. (Ed.). New York: Morgan Kaufmann Publishers, 2002, pp. 829-836.