Modelagem nebulosa c-regressão para combinação de previsões em séries temporais

Fernanda Janoele[1], Leandro Maciel[2], Rosangela Ballini[2] e Fernando Gomide[1]

[1] Departamento Engenharia de Computação e Automação Industrial
Faculdade de Engenharia Elétrica e de Computação, Unicamp
Campinas, São Paulo, Brasil
{janoele,gomide}@dca.fee.unicamp.br
[2] Departamento de Teoria Econômica
Instituto de Economia, Unicamp
Campinas, São Paulo, Brasil
maciel@dca.fee.unicamp.br,ballini@unicamp.br

Resumo Atualmente existe um número considerável de algoritmos para previsão de séries temporais, mas nenhum deles mostra-se, individualmente, como um preditor dominante. A combinação de preditores apresenta-se como um promissor mecanismo de previsão em contraposição às previsões individuais. Este artigo avalia o caso da combinação de algoritmos de previsão baseados em inteligência computacional, em particular o algoritmo c-regressão nebuloso, com algoritmos estatísticos tradicionais. Considera-se metodologias de combinação distintas como média simples, média simples aparada, ponderação baseada na variância e no desempenho, e mínimos quadrados. Além disso, um modelo nebuloso baseado em regras para seleção dos pesos na combinação das previsões é avaliado. Experimentos computacionais consideraram dados da competição 2016 *Computational Intelligence in Forecasting.* Os resultados sugerem que a combinação do modelo nebuloso c-regressão com as técnicas estatísticas, considerando a ponderação dos preditores com base na variância, apresenta previsões mais acuradas em contraposição aos modelos individuais e aos demais métodos de combinação.

Keywords: Séries Temporais. Modelagem Nebulosa C-Regressão. Combinação de Métodos de Previsão.

1 Introdução

A previsão de séries temporais desempenha papel central no apoio de operações industriais, no planejamento estratégico e na tomada de decisões. A modelagem e previsão de séries temporais têm aplicações amplamente difundidas em diversas áreas, tais como economia, finanças, engenharia, biologia, hidrologia, agricultura e medicina [1].

Estudos têm sido conduzidos para avaliar o desempenho dos distintos algoritmos de previsão atualmente disponíveis [2]. Em uma grande variedade de cenários possíveis a respeito do comportamento futuro de determinado fenômeno, verifica-se a impossibilidade em se obter resultados consistentes que conduzam a escolha de um melhor método ou modelo de predição, isto é, nenhum algoritmo é capaz de gerar melhores previsões em todos os possíveis conjuntos de dados [3]. Na prática, um algoritmo pode apresentar melhor desempenho em uma classe específica de problemas ou dados.

A ideia em se combinar algoritmos de previsão surgiu como uma forma de superar tal limitação e tem sua origem na década de 1960 [4,5] em que se evidenciou uma redução significativa de erros de previsão por meio da combinação ponderada de distintos preditores. Alguns anos depois, [6] sugeriu uma abordagem Bayesiana para combinar previsões e desde então inúmeros estudos têm indicado o melhor desempenho de combinações de previsões em contraposição a previsões individuais ([7,8]).

Uma abordagem usual para combinação de preditores consiste na avaliação de potenciais e distintos métodos em uma amostra de dados de treinamento e a posterior seleção dos melhores algoritmos para realizar previsões em um conjunto de dados para teste. O crescente interesse de pesquisadores no problema de combinação de previsões resultou em um grande número de técnicas para tal. Grande parte delas consiste em uma combinação linear ponderada das previsões individuais. Métodos estatísticos tais como média simples, média aparada, mediana, etc, estão dentre os modelos de combinação mais simples e utilizados na prática. Por exemplo, os autores de [10] propõem um modelo de previsão de demanda que combina distintos métodos individuais, associando pesos aos preditores utilizando técnicas lineares e não-lineares baseadas em redes neurais nebulosas.

A utilização de modelos de aprendizagem de máquina para aumentar a precisão de previsões em técnicas de combinação é uma abordagem que tem sido considerada sistematicamente [3]. O objetivo neste caso é o de identificar um conjunto de padrões que descrevam as séries temporais para, a partir deles, determinar os pesos dos preditores individuais. Os resultados evidenciam a superioridade de técnicas baseadas em ranqueamento em contraposição as demais avaliadas em [3].

A combinação de previsões pontuais e intervalares em modelos de suavização exponencial com base em pesos obtidos com base no critério de informação de Akaike foi examinado em [11] na previsão de séries temporais distintas. O autor sugere que as combinações simples e ponderada não se superam uma à outra constantemente e, em muitos casos, combinações simples apresentam resultados inferiores em relação a previsões individuais selecionadas de acordo com o critério de informação.

No contexto de preços de energia elétrica [12] avalia a utilização da combinação de previsão para realizar uma análise *backtesting* nos preços da eletricidade em três principais mercados europeus e norte-americanos. Os resultados demostram benefícios do uso da combinação de previsões dos métodos individuais na obtenção de previsões mais precisas.

Similarmente, [13] desenvolve um método de combinação baseado nas cadeias de Markov e uma combinação pela média ponderada variável no tempo (HM-

TWA) para prever o consumo mensal de energia elétrica na China. A avaliação mostrou que o método HM-TWA superou os modelos individuais e métodos de combinação tradicionais. A eficácia do método HM-TWA foi verificada por meio de comparações com modelos de combinações existentes.

Em inteligência computacional, [14] sugere uma combinação linear de previsão usando pesos determinados por uma rede neural. A rede neural reconhece padrões de ponderação dos modelos individuais, utilizando seus registros de previsão passados para produzir as previsões. Usando oito séries temporais, o autor mostra que a abordagem produz previsões significativamente mais precisas quando comparada aos modelos individuais e combinações lineares.

O objetivo deste artigo é sugerir uma abordagem de combinação de algoritmos de previsão derivados da inteligência computacional e dos métodos clássicos de previsão de séries temporais. Em particular, explora-se a modelagem nebulosa c-regressão (FCR). O modelo FCR é baseado em regras nebulosas do tipo Takagi-Sugeno (TS). A base de regras emerge da estrutura de agrupamento: cada grupo se traduz em uma regra nebulosa funcional.

As combinações de previsões neste trabalho usam o modelo nebuloso c-regressão, ARIMA, passeio aleatório, de alisamento exponencial e método theta como metodologias individuais. A combinação das previsões é realizada pela média simples, média simples aparada, ponderação baseada na variância e no desempenho, mínimos quadrados e um modelo nebuloso baseado em regras para seleção dos pesos na combinação das previsões. A ideia é comparar a acurácia das combinações das previsões pelos modelos de inteligência computacional e técnicas estatísticas de previsão de séries temporais tradicionais frente as previsões individuais dos modelos.

Os experimentos computacionais foram realizados com os dados fornecidos pela competição de previsão 2016 *Computational Intelligence in Forecasting* (CIF-2016). O conjunto de dados consiste de 72 séries temporais de frequências mensais e comprimentos distintos. O objetivo da competição é motivar o desenvolvimento de uma metodologia única e consistente para prever séries temporais de diferentes naturezas para diferentes horizontes de previsão.

Este artigo está organizado da seguinte forma. Na seção 2 desenvolve-se o modelo nebuloso c-regressão. As abordagens de previsão e combinação de previsão são discutidas na seção 3. Os experimentos computacionais são apresentados na seção 4. A seção 5 conclui o artigo resumindo os resultados obtidos e sugerindo temas para investigação futura.

2 Modelagem Nebulosa C-Regressão

Essa seção detalha o modelo nebuloso c-regressão (FCR). O FCR considera protótipos afins para a construção de um sistema de inferência funcional baseado em regras nebulosas.

2.1 Estrutura do modelo nebuloso

Um modelo nebuloso Takagi-Sugeno (TS) consiste em uma coleção de regras da forma:

$$\mathcal{R}_i: \text{ Se } \mathbf{x} \text{ é } \mathcal{A}_i \text{ então } y_i = f(\theta_i, \mathbf{x}), \tag{1}$$

em que $\mathcal{R}_i$ denota a i-ésima regra nebulosa, com $i = 1, 2, \ldots, c$, c é o número de regras, $\mathbf{x} = [x_1, \ldots, x_q]^T \in \Re^q$ o vetor das variáveis de entrada, $\mathcal{A}_i$ um conjunto nebuloso associado a entrada na i-ésima regra com a função de pertinência $\mathcal{A}_i(\mathbf{x}) : \Re^q \to [0, 1], \theta_i = [\theta_{i1}, \theta_{i2}, \ldots, \theta_{iq}]^T$ é o vetor de parâmetros do consequente da regra i. $y_i \in \Re$ a saída do i-ésimo subsistema, e $f_i(\cdot)$ uma função afim das variáveis de entrada associada à i-ésima regra.

A saída do modelo TS é calculada utilizando a média ponderada da contribuição individual de cada regra:

$$y = \sum_{i=1}^{c} \left(\frac{\mathcal{A}_i(\mathbf{x}) y_i}{\sum_{j=1}^{c} \mathcal{A}_j(\mathbf{x})} \right). \tag{2}$$

A expressão (2) pode ser reescrita utilizando graus de ativação normalizados:

$$y = \sum_{i=1}^{c} \lambda_i y_i = \sum_{i=1}^{c} \lambda_i \mathbf{x}_e^T \theta_i \tag{3}$$

onde

$$\lambda_i = \frac{\mathcal{A}_i(\mathbf{x})}{\sum_{j=1}^{c} \mathcal{A}_j(\mathbf{x})}, \tag{4}$$

corresponde ao grau de ativação normalizado da regra i, θ_i é o vetor de parâmetros do consequente da i-ésima regra, e $\mathbf{x}_e = [1, \mathbf{x}^T]^T$ o vetor de entrada expandido.

A identificação de um modelo TS consiste em dois passos. O primeiro consiste em determinar a estrutura do modelo, ou seja, os c's conjuntos nebulosos que compõem os antecedentes das regras. O segundo passo consiste em estimar os parâmetros dos consequentes. Para cada modelo afim local, os parâmetros podem ser estimados com técnicas baseadas em quadrados mínimos ordinários. A determinação da estrutura do modelo consiste em encontrar as regiões nebulosas, ou equivalentemente, a estimação dos centros dos grupos de acordo com a escolha de sua representação. A cada grupo com função de pertinência $\mathcal{A}_i(x)$, corresponde uma regra.

2.2 Agrupamento com protótipos afins

Para desenvolver modelos nebulosos é necessário particionar o espaço de entrada/saída. Quando a modelagem é feita a partir de dados, o particionamento pode ser feito usando métodos de agrupamento. Neste trabalho utilizou-se funções afins como protótipos [15]. No modelo nebuloso c-regressão, a distância entre um determinado $\mathbf{x}_t$ e o i-ésimo protótipo é o erro de estimação é dada por:

$$d_{it}^2 = \left(y_t - \mathbf{x}_{et}^T \theta_i \right)^2, \imath = 1, \ldots, c, \tag{5}$$

onde θ_i corresponde aos parâmetros do consequente da i-ésima regra nebulosa e $\mathbf{x}_{et}$ é o vetor de entrada $\mathbf{x}_t$ expandido. A saída de um modelo nebuloso Takagi-Sugeno é :

$$y_i = \mathbf{x}_e^T \theta_i, \ i = 1, \ldots, c. \tag{6}$$

A distância (5) é utilizada da mesma forma como no algoritmo de agrupamento FCM (*Fuzzy C-Means*) [16]. A minimização das funções objetivo dos modelos c-regressão resulta na determinação dos graus de pertinência de $\mathbf{x}_t$ ao grupo i μ_{it} de acordo com:

$$\mu_{it} = \left(\sum_{j=1}^{c} \left(\frac{d_{it}}{d_{jt}} \right)^{\frac{2}{\eta - 1}} \right)^{-1}, \tag{7}$$

onde η é um número real, geralmente $\eta = 2$. Os graus de pertinência são organizados em uma matriz $U = \{\mu_{it}\}$, $i = 1, \ldots, c$ graus e $t = 1, \ldots, N$, onde N é o número de dados.

2.3 Estimativa dos parâmetros dos consequentes

Usando (6), reescrevemos a expressão (3) como:

$$y = \sum_{i=1}^{c} \lambda_i \mathbf{x}_e^T \theta_i, \tag{8}$$

onde θ_i é o vetor de parâmetros.

Se o conjunto de dados consiste em N observações, então cada vetor de entrada de regressão expandida é $\lambda_{it} \mathbf{x}_{et}^T$. Os dados de entrada expandidos e ponderados podem ser acomodados em uma matriz do conjunto de dados, de dimensão $(q + 1) \times N$, Ψ_i^T onde:

$$\Psi_i = [\lambda_{i1} \mathbf{x}_{e1}, \ldots, \lambda_{iN} \mathbf{x}_{eN}] \tag{9}$$

O vetor com os dados de saída, de dimensão $N \times 1$, pode ser escrito como:

$$\mathbf{y}_i = [y_{i1}, \ldots, y_{iN}]^T \tag{10}$$

Consequentemente:

$$\mathbf{y}_i = \Psi_i^T \theta_i \tag{11}$$

Portanto, os parâmetros do modelo nebuloso afim são obtidos de acordo com o algoritmo de quadrados mínimos ponderados da seguinte forma [17]:

$$\theta_i = (\Psi_i \Psi_i^T)^{-1} \Psi_i \mathbf{y}_i, i = 1, \ldots, c. \tag{12}$$

2.4 Algoritmo FCR

Esta seção resume os passos da modelagem nebulosa c-regressão. Neste artigo o algoritmo foi implementado no Matlab ®.

Modelagem Nebulosa C-Regressão

1. Escolha c e o valor de terminação ε.
2. Inicialize θ_i, $i = 1, \ldots, c$, e U.
3. Para $t = 1, \ldots, N$ **faça**
4. Compute $d_{it}^2 = \left(y_t - \mathbf{x}_{et}^T \theta_i\right)^2, \; i = 1, \ldots, c$
5. Compute $\mu_{it} = \left(\sum_{j=1}^{c} \left(\frac{d_{it}}{d_{jt}}\right)^{\frac{2}{\eta-1}}\right)^{-1}$
6. Atualize $\theta_i = \left(\Psi_i \Psi_i^T\right)^{-1} \Psi_i \mathbf{y}_{it}$
7. fim para
8. se $||\Delta U|| \leq \varepsilon$ **então** retorne ao passo 3.
9. fim se

3 Métodos de previsão e de combinação de previsão

Esta seção detalha os modelos ARIMA, passeio aleatório, de alisamento exponencial e método theta e as metodologias de combinação média simples, média simples aparada, ponderação baseada na variância e no desempenho, e mínimos quadrados.

3.1 Métodos de Previsão

Abordagens mais simples para a previsão de séries temporais são muitas vezes surpreendentemente robustas como, por exemplo, a suavização exponencial ([3], [18]). Este trabalho utiliza, além do algoritmo nebuloso c-regressão, detalhado na seção 2, algoritmos derivados de técnicas estatísticas de séries temporais tradicionais como ARIMA, passeio aleatório, alisamento exponencial e método theta.

O Passeio Aleatório (RW) pressupõe que, em cada instante de tempo, a série dá um passo aleatório a partir de seu último valor, como se segue:

$$\hat{y}_{t+1} = y_t + \epsilon_t, \tag{13}$$

onde $\hat{y}_t$ e y_t são o valor previsto e valor atual da série temporal y em t, respectivamente, e $\epsilon_t \sim \mathcal{N}(0,1)$.

Na Suavização Exponencial Simples (ES) um parâmetro $0 \leq \alpha \leq 1$ controla o ajuste e pode ser encontrado por meio da minimização do erro. O método ES utiliza a seguinte expressão:

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t. \tag{14}$$

O Método Theta (MT) [19] decompõe uma série temporal em componentes de curto e longo prazo com base num coeficiente Θ, modificando a curvatura da série. A expressão do MT é:

$$\hat{y}_{t+1}(\Theta) = \hat{a} + \hat{b}t + \Theta y_t, \tag{15}$$

onde $\hat{a}$ e $\hat{b}$ são valores escolhidos como indicado em [20].

O ARIMA é desenvolvido em [21]. Um modelo ARIMA(p, d, q) pode ser representado como segue:

$$\begin{aligned} \Delta^d y_t = \omega_0 + \omega_1 \Delta^d y_{t-1} + \ldots + \omega_p \Delta^d y_{t-p} + \\ \phi_0 + \phi_1 \epsilon_{t-1} + \ldots + \phi_q \epsilon_{t-p} + \epsilon_t, \end{aligned} \tag{16}$$

onde ω_i e ϕ_j, $i = 1, \ldots, p$, $j = 1, \ldots, q$ são os parâmetros, $\epsilon \sim \mathcal{N}(0, 1)$ e d é o grau de diferenciação.

3.2 Combinação de previsão

Esta seção apresenta as técnicas de combinação de previsões dos métodos individuais: média simples, média simples aparada, ponderação baseada na variância e no desempenho, mínimos quadrados e um modelo nebuloso baseado em regras para seleção dos pesos na combinação das previsões. O objetivo das combinações é reduzir os erros de previsão, explorando as potencialidades de cada método e reduzindo suas limitações.

A combinação pela média aritmética simples (S-AVG) atribui pesos iguais para todos os métodos, tornando o método livre de erros de determinação dos pesos:

$$w_j = \frac{1}{M}, \ \forall \ j, \tag{17}$$

onde M é o número de algoritmos de previsão $w_j \in \mathbb{R}$ é o peso do j-ésimo algoritmo individual.

A média aparada (S-AVGT) não considera 20% das piores previsões, o que está entre 10-30% recomendado em [3] e [22].

Considerando $\mathbf{z} = [z_1, z_2, \ldots, z_N]^T$ o conjunto de dados, N as previsões obtidas do k-ésimo algoritmo $\hat{\mathbf{z}}^k = [\hat{z}_1^k, \hat{z}_2^k, \ldots, \hat{z}_N^k]^T$, com $k = 1, \ldots, M$, onde M é o número de métodos de previsão, o procedimento de previsão pela combinação linear é:

$$\hat{z}_t = w_1 \hat{z}_t^1 + w_2 \hat{z}_t^2 + \ldots + w_M \hat{z}_t^M, \tag{18}$$

onde w_k, $k = 1, 2, \ldots, M$ são os pesos, $w_k \geq 0$, $\forall \ k$ e $\sum_{k=1}^{M} w_k = 1$.

A ponderação baseada na variância (VB-AVG) atribui os pesos aos algoritmos de previsão utilizando o desempenho das previsões passadas. Seja $e_t^k = \left(z_t - \hat{z}_t^k\right)$

o erro de previsão do k-ésimo algoritmo em t, $k = 1, \ldots, M$. Os pesos da VB-AVG são encontrados por meio dos erros quadráticos de previsão normalizados:

$$w_k = \left(\frac{\sum_{t=1}^{N} (e_t^k)^2}{\sum_{t=1}^{N} \left((e_t^1)^2 + (e_t^2)^2 + \ldots + (e_t^M)^2 \right)} \right)^{-1}. \tag{19}$$

Em erros de previsão VB-AVG são calculados usando conjuntos de treinamento e validação dos dados. Este esquema atribui menor peso aos métodos que apresentam maiores erros de previsão da amostra, e vice-versa [14].

A ponderação baseada no desempenho (O-AVG) [6] é semelhante à ponderação baseada na variância. No entanto determina os pesos dos métodos a partir do número de vezes que um método apresentou melhor desempenho no passado.

Na combinação por mínimos quadrados (OLS-AVG) as previsões são combinadas da seguinte forma:

$$z_t = w_1 \hat{z}_t^1 + w_2 \hat{z}_t^2 + \ldots + w_M \hat{z}_t^M + \epsilon_t, \tag{20}$$

onde $\epsilon_t \sim \mathcal{N}(0, 1)$. A combinação de previsões $\hat{z}^t$ em t é produzida resolvendo (20) usando o algoritmo de mínimos quadrados ordinários para calcular os pesos $\hat{w}_k$, $k = 1, \ldots, M$:

$$\hat{z}_t = \hat{w}_1 \hat{z}_t^1 + \hat{w}_2 \hat{z}_t^2 + \ldots + \hat{w}_M \hat{z}_t^M, \tag{21}$$

A combinação através do mecanismo *Fuzzy Rule-Based Ensemble*(FRBE) é a adotada pelo programa *lfl* do ambiente estatístico R. FRBE é uma técnica que produz as previsões combinando quatro algoritmos: RW, ES, MT e ARIMA. A previsão é calculada como uma média ponderada das previsões individuais. Os pesos dos métodos individuais são determinados em função das características das séries temporais, como frequência, comprimento, assimetria, curtose, tendência, sazonalidade, coeficiente de variação e estacionaridade. Estas características são usadas para a determinar as bases de regras nebulosas cujo processamento é feito no contexto da *Perception-based Logical Deduction* [23].

4 Experimentos Computacionais

O critério de avaliação de desempenho e de comparação adotado é o *Symmetric Mean Absolute Percentage Error*. O conjunto de dados consiste das 72 séries temporais de frequências mensais fornecidas pela CIF-2016. Cada série temporal tem comprimento e horizonte de previsão específicos. Destaca-se, ainda, a avaliação de robustez dos métodos individuais e combinações.

O *Symmetric Mean Absolute Percentage Error* (SMAPE) é calculado usando a expressão:

$$\text{SMAPE} = \frac{1}{T} \sum_{t=1}^{T} \frac{|\hat{y}_t - y_t|}{(|y_t| + |\hat{y}_t|)/2}, \tag{22}$$

onde $\hat{y}_t$ é o t-ésimo valor previsto, y_t o t-ésimo valor real, e T o horizonte de previsão.

As entradas do algoritmo FCR usam valores atuais e passados da série. O número de valores passados para cada séries temporal é selecionado no intervalo $[1, 15]$ de acordo com os critério de informação Bayesiano (BIC) [24]. O θ_i foi inicializado como um vetor de zeros e U aleatoriamente. O FCR processa todas as séries temporais usando a primeira diferença ($\Delta^1 y_t = y_t - y_{t-1}$) a fim de remover a não-estacionaridade. O número de agrupamentos c do FCR foi selecionado no intervalo $[1, 10]$ através de simulações para selecionar o de melhor desempenho. Definimos $\varepsilon = 10^{-3}$.

Os experimentos computacionais relatados neste trabalho removeram as T últimas amostras das 72 séries temporais, onde T é o horizonte de previsão indicado para cada série temporal para testes. O restante dos dados da amostra são usados para estimar os parâmetros dos modelos. Os algoritmos de previsão foram executados para prever os valores da série até T e em seguida calcula-se o SMAPE.

A Tabela 1 mostra as médias SMAPE dos métodos individuais e combinados para a previsão das 72 séries temporais. A melhor média obtida e o melhor desvio padrão estão em negrito. O desvio padrão do SMAPE é uma medida de robustez do método. Entre os métodos individuais, a suavização exponencial (ES) apresentou menor média SMAPE, seguida do método theta (MT). Métodos ARIMA e FCR apresentaram pior desempenho como mostra a Tabela 1.

Dentre as metodologias de combinação, a ponderação baseada na variância (VB-AVG) aparece como a melhor, seguida dos mínimos quadrados (OLS-AVG), como apresentado na Tabela 1. Exceto para a média simples e a combinação baseada em regras nebulosas, todas as combinações apresentaram desempenho melhor do que os algoritmos individuais. Os métodos de combinação também mostram menores valores de desvio padrão do SMAPE. Os resultados em termos de médias e desvios padrão SMAPE, indica a VB-AVG como o melhor método, seguido pela OLS-AVG e o O-AVG, respectivamente. Em geral, os resultados foram semelhantes.

Os resultados sugerem que a combinação do modelo nebuloso c-regressão com técnicas estatísticas tradicionais de previsão, considerando a ponderação dos preditores com base na variância, fornecem previsões mais acuradas em contraposição aos modelos individuais e aos demais métodos de combinação, considerando os dados da competição de previsão 2016. Os resultados fornecidos pela competição de previsão 2016 *Computational Intelligence in Forecasting* (CIF-2016) [30] confirmam a acurácia da combinação do modelo nebuloso c-regressão com técnicas estatísticas tradicionais, considerando a ponderação dos preditores com base na variância.

5 Conclusão

Este trabalho examinou o algoritmo nebuloso c-regressão e algoritmos estatísticos tradicionais de previsão. O trabalho combinou previsões dos modelos nebuloso c-

Tabela 1. SMAPE médio e desvio padrão para dos métodos individuais e combinações.

Método	SMAPE (Média)	SMAPE (Desvio padrão)
RW	0.12300	0.09776
ES	0.11309	0.09316
MT	0.11732	0.09531
ARIMA	0.15116	0.22491
FCR	0.13202	0.17960
S-AVG	0.11532	0.10563
S-AVGT	0.11267	0.10872
VB-AVG	**0.10004**	**0.08221**
O-AVG	0.11092	0.10553
OLS-AVG	0.11086	0.09514
FRBE	0.12061	0.12195

regressão, ARIMA, passeio aleatório, de alisamento exponencial e método theta com base em distintas metodologias de combinação, tais como, média simples, média simples aparada, ponderação baseada na variância e no desempenho, mínimos quadrados e um modelo nebuloso baseado em regras para seleção dos pesos na combinação das previsões. Os experimentos computacionais, para os dados da competição de previsão 2016 *Computational Intelligence in Forecasting*, indicam que a combinação do modelo nebuloso c-regressão com técnicas estatísticas tradicionais, considerando a ponderação dos preditores com base na variância, apresenta previsões mais acuradas em contraposição aos modelos individuais e aos demais métodos de combinação avaliados.

Referências

1. Maciel, L., Lemos, A., Ballini, R., Gomide, F.: Adaptive fuzzy c–regression modeling for time series forecasting. In: 16th World Congress of the International Fuzzy Systems Association and 9th Conference of the European Society for Fuzzy Logic and Technology (IFSA-EUSFLAT 2015), ser. Advances in Intelligent Systems Research. Gijón, Atlantic Press (2015)
2. Makridakis, S., Hibon, M.: The M3-Competition: results, conclusions and implications. International Journal of Forecasting. 16, 451–476 (2000)
3. Lemke, C., Gabrys, B.: Meta–learning for time series forecasting and forecast combination. Neurocomputing. 73, 2006–2016 (2010)
4. Bates, J., Granger, C.: The combination of forecasts. Operational Research Quarterly. 20, 451–468 (1969)
5. Crane, D., Crotty, J.: A two-stage forecasting model: exponential smoothing and multiple regression. Management Science. 6, B501–B507 (1967)
6. Bunn, D.: A Bayesian approach to the linear combination of forecasts. Operational Research Quarterly. 2, 325–329 (1975)
7. Stock, J., Watson, M.: Combination forecasts of output growth in a seven-country data set Journal of Forecasting. 23, 405–430 (2004)

8. Timmermann, A.: Forecast combinations. In: Handbook of Economic Forecasting, Elliot, G., Granger, C., Timmermann, A. (eds.) New York: Elsevier 2006. vol. 18. pp. 135–196. Elsevier, north-holland (2006)
9. Andrawis, R., Atiya, A., Saavedra, A.: Combination of long term and short term forecasts, with application to tourism demand forecasting. International Journal of Forecasting. 27, 870–886 (2011)
10. Wang, F., Chang, K.: Adaptive neuro-fuzzy inference system for combined forecasts in a panel manufacturer. Expert Systems with Applications. 37, 8119–8126 (2010)
11. Kolassa, S.: Combining exponential smoothing forecasts using Akaike weights. International Journal of Forecasting. 27, 238–251 (2011)
12. Nowotarski, J., Raviv, E., Truck, S., Weron, R.: An empirical comparison of alternative schemes for combining electricity spot price forecasts. Energy Economics. 46, 395–412 (2014)
13. Zhao, W., Wang, J., Lu, H.: Combining forecasts of electricity consumption in China with time-varying weights updated by a highorder Markov chain model. Omega. 45, 80–91 (2014)
14. Adhikari, R.: A neural network based linear ensemble framework for time series forecasting. Neurocomputing. 157, 231-242 (2015)
15. Breiman, L.: Hinging hyperplanes for regression, classificaion and function approximation. IEEE Transactions on Information Theory. 39, 311–325. (1993)
16. Bezdek, J.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York (1981)
17. Ljung, L.: System Identification: Theory for the User. Englewood Cliffs, NJ: Prentice-Hall (1988)
18. Gardner, E.: Exponential smoothing: the state of the artpart ii. International Journal of Forecasting. 22 , 637-666 (2006)
19. Assimakopoulos V., Nikolopoulos, K.: The theta model: A decomposition approach to forecasting. International Journal of Forecasting. 16, 521-530 (2000)
20. Hyndman, R., Billah, B.: Unmasking the theta method. International Journal of Forecasting. 19, 287-290 (2003)
21. Box, G., Jenkins, G. : Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco (1976)
22. Jose, V., Winkler, R.: Simple robust averages of forecasts: some empirical results. International Journal of Forecasting. 24, 163-169 (2008)
23. Štěpnička, M., Burda, M.: Computational intelligence in forecasting –The results of the time series forecasting competition. Proceedings of the 6th World Congress of the International Fuzzy Systems Association and 9th Conference of the European Society for Fuzzy Logic and Technology (IFSA-EUSFLAT 2015), pp. 1–8. Gijón, Spain (2015)
24. Schwarz, G.: Estimating the dimension of model. The Annals of Statistics. 6, 461-464 (1978)
25. Angelov, P., Filev, D.: An approach to online identification of Takagi-Sugeno fuzzy models. IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics. 4, 484–498 (2012)
26. Forecasting: principles and practice, `http://robjhyndman.com/`
27. A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, `http://www.R-project.org/`
28. Linguistic Fuzzy Logic (R package on CRAN 2015), `http://cran.r-project.org/web/packages/lfl/`
29. Forecasting Functions for Time Series and Linear Models (R package on CRAN 2015), `https://github.com/robjhyndman/forecast`

30. Computational Intelligence in Forecasting (CIF), `http://irafm.osu.cz/cif/main.php`