

Séries Temporais Fuzzy do Tipo-2

Arthur Caio V. e Pinto* Petrônio C. L. Silva**
 Frederico G. Guimarães** Eduardo P. de Aguiar*

* Programa de Pós Graduação em Engenharia Elétrica, Universidade
 Federal de Juiz de Fora - UFJF, MG, Brasil (e-mail:
 arthur.vargas@engenharia.ufjf.br), (e-mail:
 eduardo.aguiar@engenharia.ufjf.br).

** Machine Intelligence and Data Science (MINDS) Laboratory,
 Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil,
 (e-mail: petronio.candido@ifnmg.edu.br)(e-mail:
 fredericoguimaraes@ufmg.br).

Abstract: In many time series practical forecasting problems, it is possible to find challenging behaviors and patterns within the data, making it difficult to analyze and forecast. This scenario has motivated, in recent years, the development of new forecasting methods, combining predictive performance and low processing costs. One of these methods is based on fuzzy time series. Thus, this work presents the first stage of implementation of a type-2 fuzzy time series model, called T2FTS, capable of taking advantage of the best treatment of uncertainties that type-2 fuzzy logic presents. This initial study discusses how the partition methodology affects the accuracy of the model, in addition to comparing the calculated metrics with other known forecasting techniques. The results indicate that the T2FTS forecasting model in this state of development is already highly competitive with the main existing forecasting methods, even surpassing some well-known regression methods in literature.

Resumo: Em muitos problemas práticos de previsão de séries temporais, é possível encontrar comportamentos e padrões desafiadores dentro dos dados, dificultando sua análise e previsão. Esse cenário motivou, nos últimos anos, o desenvolvimento de novos métodos de previsão, aliando desempenho preditivo e baixo custo de processamento. Um desses métodos é o baseado em séries temporais fuzzy. Dessa forma, este trabalho apresenta a primeira etapa de implementação de um modelo de série temporal fuzzy do tipo-2, chamado T2FTS, capaz de aproveitar o melhor tratamento de incertezas que a lógica fuzzy do tipo-2 apresenta. Este estudo inicial discute como a metodologia de partição afeta a precisão do modelo, além de comparar as métricas calculadas com outras técnicas de previsão conhecidas. Os resultados indicam que o modelo de previsão T2FTS neste estado de desenvolvimento já é altamente competitivo com os principais métodos de previsão existentes, chegando a superar alguns dos métodos de regressão conhecidos da literatura.

Keywords: Type-2 Fuzzy Systems; Prediction methods; Time Series Analysis; Type-2 Fuzzy Time Series; Sliding window.

Palavras-chaves: Sistemas Fuzzy Tipo-2; Métodos de previsão; Análise de Séries Temporais; Séries temporais fuzzy do tipo-2; Janela deslizante.

1. INTRODUÇÃO

A coleta de dados em intervalos regulares cria uma estrutura chamada de série temporal. A previsão de séries temporais é uma metodologia de *Machine Learning*, que consiste na análise de dados e padrões temporais para prever eventos futuros. Essa técnica fornece suposições sobre tendências futuras com base em dados históricos de séries temporais, que podem ser quase precisos, dependendo do modelo utilizado e das características da série temporal analisada. Bose and Mali (2019) descrevem o duplo objetivo da análise de séries temporais, que é “entender ou modelar o mecanismo estocástico que dá origem a uma série observada e prever os valores futuros de uma série com base em seu histórico de dados”.

A previsão de séries temporais permite analisar os principais padrões que a compõem, como tendência, sazonalidade, ciclicidade e irregularidade. A análise de séries temporais é usada para várias aplicações, como análise do mercado de ações, reconhecimento de padrões, previsão de terremotos, previsão econômica, entre outras. Tem sido realizada por meio de uma variedade de técnicas diferentes, como modelos auto-regressivos integrados de médias móveis (ARIMA), abordagens com inteligência artificial e redes neurais artificiais (RNA). É possível encontrar comportamentos e padrões desafiadores dentro dos dados de séries temporais, dificultando sua análise e previsão. Esse cenário motivou, nos últimos anos, o desenvolvimento de novos métodos de previsão, aliando desempenho e baixo custo computacional. Um desses métodos é o de séries

temporais fuzzy (do inglês, *Fuzzy Time Series* - FTS), que utilizam lógica fuzzy do tipo-1 para a construção do modelo, e tem sido utilizadas para previsão de séries econômicas (Tak (2020)), ambientais (Alyousifi et al. (2020)), entre outras.

Apesar de modelos de séries temporais fuzzy apresentarem um avanço no campo de previsão, Mendel (2007) mostra que existem limitações na capacidade dos sistemas fuzzy do tipo-1 de modelar e minimizar o efeito de incertezas no processo de previsão. Essa limitação é melhor abordada pela lógica fuzzy do tipo-2, que devido à sua estrutura de construção possui desempenho melhor no tratamento de incertezas, tão presentes em séries temporais. Essa característica tem sido abordada em estudos recentes que aliam séries temporais à lógica fuzzy do tipo-2 como em Luo et al. (2019), Lee (2019) e Shukla and Muhuri (2019). Dessa forma, esta pesquisa propõe um modelo de Type-2 Fuzzy Time Series, chamado T2FTS, capaz de aproveitar o melhor tratamento de incertezas que a lógica fuzzy do tipo-2 apresenta para construir modelos de previsão. A versão final do T2FTS contará com mecanismos auxiliares de inferência como ponderamento de regras e probabilísticos, que aumentam a acurácia do resultado final. Este artigo apresenta a primeira etapa de implementação do T2FTS, utilizando lógica fuzzy do tipo-2 para a construção do modelo, porém ainda sem a adição de tais mecanismos de aprimoramento. As métricas calculadas pelo T2FTS neste estado de desenvolvimento já o tornam competitivo com os principais métodos de previsão existentes e demonstram seu potencial para melhoria quando mecanismos avançados forem implementados.

2. FORMULAÇÃO DO PROBLEMA

2.1 Séries Temporais Fuzzy

Os fundamentos da teoria sobre séries temporais fuzzy, que tem sido amplamente utilizada nos últimos anos, foram primeiramente definidos por Song and Chissom (1993). Os valores das séries temporais fuzzy são conjuntos fuzzy e existe uma relação entre as observações no momento atual e momentos anteriores. Existem diferentes métodos fuzzy propostos para resolver problemas de previsão de séries temporais, desde os mais simples até os mais avançados, que envolvem mecanismos de ponderamento, probabilísticos, entre outros, que melhoram o desempenho dos modelos.

As séries temporais fuzzy fornecem uma representação diferente de uma série temporal. Se uma série temporal convencional é composta por observações representadas por números reais, séries temporais fuzzy são compostas por conjuntos fuzzy. Esses conjuntos fuzzy estão todos dentro do universo do discurso, que é definido pelo intervalo dos valores observados na série temporal original.

Seja $Y(t)$ ($t = 0, 1, 2, \dots$) um subconjunto de R e o universo de discurso no qual os conjuntos fuzzy $\tilde{A}_i(t)$ ($i = 1, 2, \dots$) estão definidos e seja $F(t)$ uma coleção de \tilde{A}_i ($i = 1, 2, \dots$). Então, $F(t)$ é chamada de uma série temporal fuzzy definida em $Y(t)$ ($t = 0, 1, 2, \dots$).

As várias categorias de métodos de previsão geralmente variam principalmente de acordo com a ordem e a variação

temporal. A ordem indica quantos atrasos (amostras) são usados na modelagem das séries temporais. Sejam os dados da série temporal fuzzy $F(t)$, os modelos de primeira ordem usam dados de $F(t-1)$ para prever $F(t)$ e os modelos de alta ordem usam $F(t-1); F(t-2); \dots; F(t-n)$ para prever $F(t)$. A principal desvantagem do uso de modelos de alta ordem é que os custos computacionais são mais altos. Representando $F(t-1)$ e $F(t)$ por A_i e A_j , respectivamente, um modelo de primeira ordem pode ser expresso na forma da Equação (1):

$$A_i \rightarrow A_j \quad (1)$$

Quando à variação temporal, se em qualquer momento t , a relação entre $F(t-1)$ e $F(t)$ é independente de t , então $F(t)$ é chamada de série temporal fuzzy invariante no tempo. Caso contrário, é chamada de série temporal fuzzy com variação temporal (Bose and Mali (2019)).

A própria natureza das séries temporais faz com que sua análise e previsão seja afetada por inúmeras variáveis que adicionam incertezas ao modelo. Fatores como tendência, sazonalidade e movimentos cíclicos acabam por dificultar a construção de um modelo de previsão preciso e conseqüentemente, o processo de previsão. Nesse sentido, a estrutura de funcionamento de sistemas fuzzy do tipo-2 pode contribuir para a construção de um sistema mais robusto e tolerante a ruídos e incertezas nos dados.

3. METODOLOGIA

3.1 Sistemas Fuzzy do Tipo-2

Um conjunto fuzzy é uma função com diferentes graus de pertinência ao conjunto. Seja U o universo de discurso, que é discreto e finito, o conjunto fuzzy \tilde{A} pode ser definido pela Equação (2):

$$\tilde{A} = \left\{ \frac{\mu_{\tilde{A}}(x_1)}{x_1} + \frac{\mu_{\tilde{A}}(x_2)}{x_2} + \dots \right\} = \sum_{n=i} \frac{\mu_{\tilde{A}}(x_i)}{x_i} \quad (2)$$

em que μ_A é a função de pertinência de \tilde{A} , $\mu_A: U \rightarrow [0,1]$, e $\mu_{\tilde{A}}(x_i)$ é o grau de pertinência do elemento x_i no conjunto fuzzy \tilde{A} , que representa o grau com que x_i pertence ao conjunto fuzzy \tilde{A} . Quando U é contínuo e infinito, então o conjunto fuzzy \tilde{A} pode ser definido pela Equação (3), como explica Bose and Mali (2019).

$$\tilde{A} = \left\{ \int \frac{\mu_{\tilde{A}}(x_i)}{x_i} \right\}, \forall x_i \in U \quad (3)$$

Enquanto em conjuntos fuzzy do tipo-1 o grau de pertinência é caracterizado por um único valor, chamado *crisp value*; em conjuntos fuzzy do tipo-2, o grau de pertinência é considerado um conjunto fuzzy, o que permite incorporar o tratamento de incertezas nos dados e na própria construção dos conjuntos fuzzy, tornando-o mais adequado para lidar com variações e ruídos nas entradas. Agora seja $\tilde{A}(U)$ o conjunto de conjuntos fuzzy em U . Um conjunto fuzzy do tipo-2 \tilde{A} em X é um conjunto fuzzy cujos graus de pertinência são eles próprios fuzzy. Isso implica que $\mu_{\tilde{A}}(x)$

é um conjunto fuzzy em U para todos os valores de x , ou seja, $\mu_{\tilde{A}}: X \rightarrow \tilde{A}(U)$ (Equação (4)):

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x) \mid \mu_{\tilde{A}}(x) \in \tilde{A}(U) \forall x \in X\} \quad (4)$$

Um exemplo de conjunto fuzzy do tipo-2 pode ser visto na Figura 1. É possível ver que um conjunto fuzzy do tipo-2 é definido por uma função de pertinência superior (UMF) e função de pertinência inferior (LMF). Ambas são equivalentes a uma função de pertinência do tipo-1 tradicional. A função de pertinência inferior é menor ou igual à função de pertinência superior para todos os valores de entrada possíveis. A região entre o UMF e o LMF é chamada de mancha de incerteza (FOU) (Bajestani and Zare (2011)).

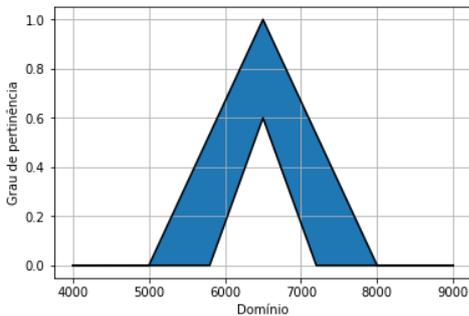


Figura 1. Funções de pertinência superior e inferior de um conjunto fuzzy do tipo-2.

Bajestani and Zare (2011) apresentam um exemplo que ajuda a entender melhor o conceito de conjuntos fuzzy do tipo-2: Considere o conjunto fuzzy mostrado na Figura 2

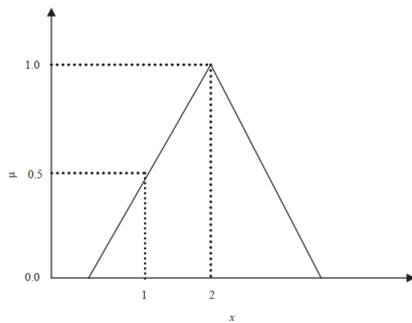


Figura 2. Pertinências em um conjunto fuzzy do tipo-1. Fonte: Bajestani and Zare (2011).

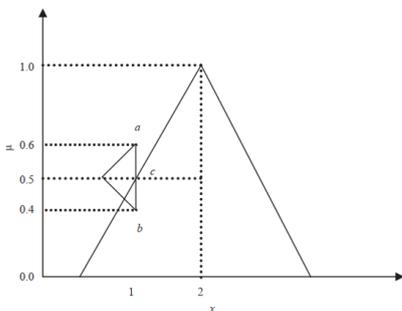


Figura 3. Pertinências em um conjunto fuzzy do tipo-2. Fonte: Bajestani and Zare (2011).

e os respectivos graus de pertinência para um valor de $x = 1$ e $x = 2$. Tem-se um valor *crisp* de pertinência igual a 0,5 para $x = 1$ e um valor *crisp* de pertinência igual a 1,0 para $x = 2$. Agora, no caso do conjunto fuzzy do tipo-2 mostrado na Figura 3, para o mesmo valor de x , pode haver vários graus de pertinência devido às funções de pertinência superior e inferior. Por exemplo, há um conjunto fuzzy triangular (0,4; 0,5; 0,6) como um grau de pertinência para $x = 1$ na Figura 3.

Sistemas de lógica fuzzy (SLF) utilizam conjuntos fuzzy do tipo-1 ou do tipo-2 para produzir valores *crisp* de saída a partir de valores *crisp* de entrada. Um SLF do tipo-2 é similar a um SLF do tipo-1. Segundo Mendel (2007) a principal diferença estrutural é que o bloco defuzzificador de um SLF do tipo-1 é substituído por um bloco processador de saída em um SLF do tipo-2, composto por um mecanismo redutor de tipo seguido de um defuzzificador, como mostra a Figura 4. O bloco processador de saída é responsável por realizar as operações que irão converter o conjunto fuzzy resultante em um valor *crisp* que possa ser utilizado.

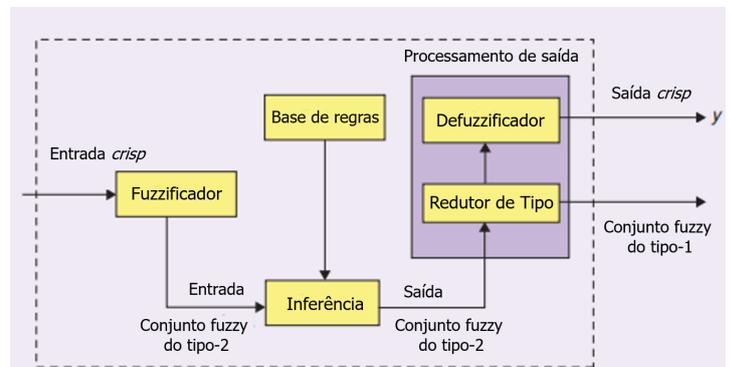


Figura 4. Estrutura de um SLF do tipo-2. Fonte: Adaptado de Mendel (2007).

3.2 Séries Temporais Fuzzy do Tipo-2

Esta seção apresenta as etapas para a implementação do modelo de previsão T2FTS, implementado na linguagem Python. O procedimento de previsão é dividido em etapas, ilustradas na Figura 5, que podem ser descritas como:

- (1) **Diferenciação da série temporal:** Os dados de séries temporais geralmente exigem alguma preparação antes de serem modelados, geralmente usando uma ou mais transformações de dados, a fim de se tornar a série estacionária. Operações de diferenciação podem ser usadas para remover a tendência e a estrutura sazonal das séries temporais, a fim de aprimorar o processo de previsão. Após a diferenciação, o modelo é treinado e testado e, em seguida, os valores previstos devem ser ajustados de volta à escala da série temporal original usando alguma transformação de diferenciação inversa. Isso é necessário para que as métricas de desempenho calculadas estejam na mesma escala da variável de saída.
- (2) **Definição do Universo de Discurso:** Suponha que uma série temporal seja representada por $Y(t)$ e o universo de discurso seja $U = [\min(Y(t)) - f_1, \max(Y(t)) + f_2]$, em que $\min(Y(t))$ e $\max(Y(t))$ são



Figura 5. Fluxograma do modelo proposto T2FTS.

os valores mínimo e máximo do conjunto de dados históricos da série temporal. Aqui, f_1 e f_2 são fatores escolhidos para expandir o universo do discurso. Para esta pesquisa, os fatores utilizados foram $f_1 = 0.1 \cdot \min(Y(t))$ e $f_2 = 0.1 \cdot \max(Y(t))$.

- (3) **Particionamento do universo de discurso:** O universo do discurso é dividido em k intervalos regulares sobrepostos. Cada intervalo é associado a um conjunto fuzzy. Considerando n intervalos (I_1, I_2, \dots, I_n), cada intervalo é associado a um termo linguístico \tilde{A}_i atribuído a uma variável linguística \tilde{A} . A Figura 6 apresenta o particionamento do universo do discurso usando diferentes números de partições. O número correto de intervalos (conjuntos fuzzy) para os modelos de séries temporais fuzzy tem sido objeto de estudo e não é um parâmetro fácil de ser definido pois afeta diretamente a acurácia do modelo. Assim, é recomendado o uso de técnicas de otimização para encontrar o número ótimo de intervalos do modelo de previsão.
- (4) **Fuzzificação dos dados:** Considerando como entradas os valores da série temporal, é feito um mapeamento de cada um desses dados para os conjuntos fuzzy correspondentes. Assim, a série original $Y(t)$ é transformada na série fuzzificada $F(t)$.
- (5) **Extração dos Grupos de Relação entre Conjuntos Fuzzy (FLRG):** Como uma série temporal é uma sequência de amostras, a relação entre uma amostra e a próxima pode ser analisada e identificada por uma regra fuzzy. Por exemplo: se uma amostra aleatória em um momento $y(t)$ é fuzzificada como \tilde{A}_i e $y(t+1)$ é fuzzificada como \tilde{A}_j , então uma regra fuzzy pode ser criada: $\tilde{A}_i \rightarrow \tilde{A}_j$, em que \tilde{A}_i é chamado de antecedente e \tilde{A}_j é chamado de consequente. Esta

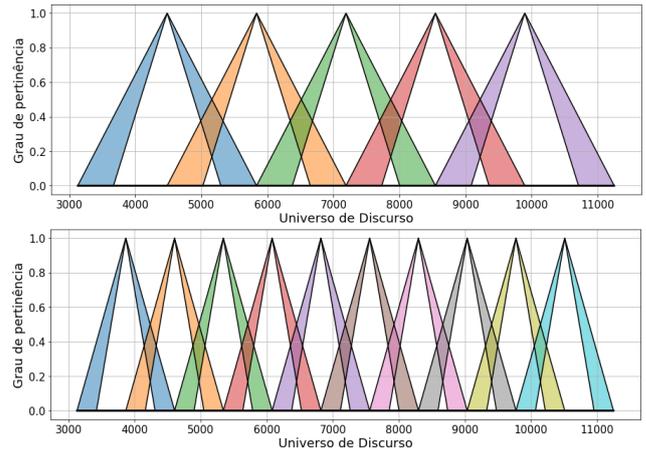


Figura 6. Diferentes formas de particionamento de conjuntos fuzzy do tipo-2.

regra fuzzy pode ser lida como SE $y(t)$ for \tilde{A}_i ENTÃO $y(t+1)$ é \tilde{A}_j . Por exemplo, se o valor da série temporal para o dia 09/06/2008 for fuzzificado como \tilde{A}_{11} e o valor da série em 10/06/2008 for fuzzificado como \tilde{A}_{14} , então é possível estabelecer a seguinte regra: SE $y(t)$ é \tilde{A}_{11} ENTÃO $y(t+1)$ é \tilde{A}_{14} , ou seja, $\tilde{A}_{11} \rightarrow \tilde{A}_{14}$.

- O processo de extração de regras é então feito para cada amostra da série temporal fuzzificada. O próximo passo é agrupar todas as regras com o mesmo antecedente e criar um Grupo de Relacionamento Lógico Fuzzy (FLRG) para cada conjunto fuzzy. Os FLRGs gerados são a base do modelo treinado, portanto o procedimento de previsão usa o modelo gerado para fazer previsões sobre amostras futuras.
- (6) **Inferência:** O procedimento de inferência consiste em encontrar as regras que são ativadas por cada uma das amostras fuzzificadas e seus respectivos FLRGs. O processamento dos antecedentes é executado utilizando o operador mínimo, que é aplicado para selecionar os valores mínimos de UMFs e LMFs de todos os antecedentes para cada regra. Esse processo produz um conjunto fuzzy de saída para cada regra ativada. Ao final, cada amostra ativou vários conjuntos fuzzy e o processamento do consequente é realizado usando o operador máximo. A agregação é então executada, uma etapa em que o objetivo é encontrar um conjunto fuzzy do tipo-2 agregado aplicando o método de agregação aos UMFs e LMFs dos conjuntos fuzzy de saída (consequentes) de todas as regras ativadas.
- (7) **Defuzzificação:** O processamento da saída do sistema é feito pelo mecanismo redutor de tipo, e depois pelo defuzzificador. O conjunto fuzzy agregado do tipo-2 é então reduzido a um conjunto fuzzy do tipo-1 utilizando o método do Algoritmo iterativo aprimorado com condição de parada (EIASC). A defuzzificação é a última parte do processo de previsão, onde os valores obtidos pelo algoritmo de redução de tipo são defuzzificados e valores *crisp* são gerados para cada amostra do conjunto de dados. A defuzzificação foi realizada usando o método Centróide, que retorna o centro de gravidade do conjunto fuzzy do tipo-1 ao longo do eixo x .

4. RESULTADOS

O Índice de Capitalização da Bolsa de Valores de Taiwan (TAIEX), uma série econômica geralmente usada na literatura de previsão de séries temporais, foi utilizada nesta pesquisa para avaliar o desempenho do modelo proposto. A base de dados é composta por 5200 amostras, referentes ao período de 1995 a 2015, e contém os dados dos índices de ações, como valor de fechamento, valor de abertura e também o índice médio diário por dia útil, entre outros. Nesta pesquisa, foi utilizado o índice médio diário. A Figura 7 mostra o conjunto de dados, com a linha vermelha dividindo o conjunto de treinamento do conjunto de testes.

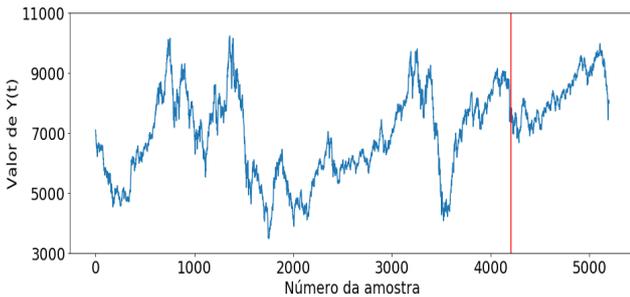


Figura 7. Série temporal correspondente ao índice médio diário por dia útil da base de dados TAIEX.

Foi aplicada uma metodologia de validação cruzada chamada de janela deslizante, usando um conjunto de trabalho de 1000 instâncias e um incremento deslizante de 200 amostras. Os dados foram divididos em dois conjuntos; o primeiro, o conjunto de treinamento, continha 80% dos dados, e o segundo, o conjunto de teste, os 20% restantes. Os resultados foram então comparados aos modelos Naïve (repetição do último valor), ARIMA, traditional FTS - FTS (Song and Chissom (1993)), Conventional FTS - CFTS (Chen et al. (1996)), Weighted FTS - WFTS (Yu (2005)), Improved Weighted FTS - IWFTS (Ismail and Efendi (2011)), Exponentially Weighted FTS - EWFTS (Sadaei (2013)), High Order FTS - HOFTS (Severiano et al. (2017)), Hwang (Hwang et al. (1998)), Trend Weighted FTS - TWFTS (Cheng et al. (2006)) e Probabilistic Weighted Fuzzy Time Series - PWFTS (Silva et al. (2019)).

A métrica de comparação utilizada foi o *Root Mean Square Error* (RMSE), que é a métrica de precisão padrão usada para avaliar o desempenho dos métodos de previsão de séries temporais. O RMSE é descrito pela Equação (5), em que y é o valor original, \hat{y}_i é o valor previsto e n o número de amostras.

$$RMSE = \sqrt{\left(\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}\right)} \quad (5)$$

Foram gerados modelos T2FTS de primeira, segunda e terceira ordem, variando também o número de partições do universo de discurso por meio de um algoritmo de *Grid search* (busca em grade), um método que testa todas as combinações possíveis e encontra a combinação ótima para os parâmetros. Os modelos foram testados tanto com os dados originais quanto com o dados diferenciados.

A Tabela 1 mostra os resultados do modelo T2FTS, apresentando os valores RMSE para as diferentes configurações testadas para os dados sem diferenciação. Já a Tabela 2 mostra os valores RMSE para o modelo quando os dados foram diferenciados. Os valores de RMSE presentes nas tabelas correspondem ao valor médio do RMSE obtido durante todas as instâncias da janela deslizante. As métricas deixam claro que a diferenciação dos dados melhora significativamente o procedimento de previsão, pois os componentes de tendência e sazonalidade são atenuados. O melhor modelo na Tabela 1 (dados sem diferenciação) obteve um valor de RMSE de 142,19, enquanto o melhor modelo na Tabela 2 (dados diferenciados) obteve um valor de RMSE de 87,04. Também é possível observar os efeitos da escolha correta do número de partições e da ordem do sistema ao criar um modelo de previsão.

Tabela 1. Valores de RMSE obtidos pelo modelo T2FTS sem diferenciação dos dados.

Partições	1ª Ordem	2ª Ordem	3ª Ordem
1	1158,12	1158,12	1158,12
2	1158,12	1158,12	1158,12
3	815,64	812,65	811,17
4	550,31	507,54	506,31
5	454,27	452,84	429,00
6	421,57	393,66	375,95
7	350,65	349,05	351,08
8	311,35	291,41	291,69
9	296,25	286,52	274,94
10	288,99	287,67	273,34
11	257,76	264,01	271,84
12	234,80	242,72	246,56
13	225,78	221,56	220,31
14	207,98	212,10	212,75
15	183,74	186,08	201,90
16	181,97	187,34	191,23
17	174,95	182,55	167,06
18	171,18	154,69	160,01
19	150,77	152,92	154,71
20	142,19	145,66	146,09

Tabela 2. Valores de RMSE obtidos pelo modelo T2FTS com diferenciação dos dados.

Partições	1ª Ordem	2ª Ordem	3ª Ordem
1	113,12	151,46	182,40
2	113,12	151,46	182,40
3	113,12	151,50	182,53
4	113,01	151,16	182,51
5	111,11	149,52	180,84
6	108,47	141,30	171,46
7	107,57	138,08	169,14
8	100,50	139,46	166,28
9	95,26	137,00	167,25
10	99,10	133,47	164,28
11	99,66	134,89	166,52
12	100,70	136,91	163,81
13	94,92	134,73	162,55
14	93,70	132,57	159,96
15	92,58	133,21	162,89
16	90,63	128,04	159,12
17	89,71	129,63	156,62
18	87,89	124,78	156,20
19	87,04	128,66	157,90
20	90,54	126,00	154,77

Para um maior número de partições (como 40 ou 80), observa-se uma melhora no desempenho dos modelos mas

também aumenta o seu gasto computacional, dessa forma o número de partições do sistema deve ser escolhido pensando-se também na parcimônia do modelo. Fica evidente que o modelo que obteve melhor desempenho foi o modelo com dados diferenciados de primeira ordem com 19 partições (destacado em negrito na Tabela 2).

Os resultados obtidos pelo modelo T2FTS são comparados a outros modelos de previsão da literatura, na Tabela 3. Todos os modelos foram treinados e testados usando a metodologia de janela deslizante e classificados de acordo com seu valor RMSE.

Tabela 3. Valores de RMSE calculados
Fonte: Adaptado de: Silva et al. (2019).

Modelo	Ordem	Partições	RMSE	Colocação
ARIMA	1,1,0	-	87,32	4
CFTS (Chen et al. (1996))	1	31	95,08	10
EWFTS (Sadaei (2013))	1	7	88,70	7
FTS (Song and Chissom (1993))	1	13	90,73	8
HOFTS (Severiano et al. (2017))	3	5	91,06	9
Hwang (Hwang et al. (1998))	3	5	105,54	11
IWFTS (Ismail and Efendi (2011))	1	5	87,65	6
Naive	1	-	127,24	12
PWFTS (Silva et al. (2019))	3	3	83,72	1
TWFTS (Cheng et al. (2006))	1	7	85,35	2
T2FTS	1	19	87,04	3
WFTS (Yu (2005))	1	5	87,43	5

Pode-se notar que o modelo T2FTS proposto é altamente competitivo com métodos de regressão conhecidos e outros métodos mais avançados de fuzzy do tipo-1. Considerando o valor RMSE do modelo T2FTS (RMSE = 87,04), ele foi classificado em terceiro lugar na lista da Tabela 3.

É importante mencionar que o modelo fuzzy do tipo-2 proposto neste trabalho não é um modelo ponderado ou probabilístico (como os outros a que foi comparado), características que serão implementadas em um futuro próximo e que provavelmente irão melhorar a acurácia do modelo.

5. CONCLUSÃO

Este trabalho apresentou a primeira etapa de implementação de um modelo de série temporal fuzzy do tipo-2, o T2FTS, e avaliou seu desempenho comparando-o com outros métodos da literatura. Também foi discutido o importante papel da diferenciação dos dados originais, bem como da correta escolha do número de partições e da ordem do sistema de previsão, a fim de se aumentar a acurácia do resultado. O modelo proposto alcançou bons resultados que o colocam à frente de outros métodos já conhecidos e mais estudados na literatura, mesmo utilizando apenas lógica fuzzy do tipo-2, sem a adição de mecanismos ponderados ou probabilísticos (que será feita em trabalhos futuros). Mesmo estando atrás de técnicas de previsão mais avançadas, isso demonstra o potencial da lógica fuzzy do tipo-2 para superá-las quando forem aplicadas melhorias ao modelo, como as citadas acima. Para futuros trabalhos estão previstas a adição de mecanismos de ponderamento das regras fuzzy e algoritmos probabilísticos, com o objetivo de se aprimorar o modelo e alcançar melhores resultados, competitivos com os métodos mais avançados existentes.

AGRADECIMENTOS

Os autores agradecem o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (processo 433389/2018-4), Fundação de Amparo à Pesquisa do Estado de Minas Gerais - FAPEMIG (APQ-02922-18) e Universidade Federal de Juiz de Fora - UFJF pelo suporte financeiro.

REFERÊNCIAS

- Alyousifi, Y., Othman, M., Sokkalingam, R., Faye, I., and Silva, P.C. (2020). Predicting daily air pollution index based on fuzzy time series markov chain model. *Symmetry*, 12(2), 293.
- Bajestani, N.S. and Zare, A. (2011). Forecasting taieix using improved type 2 fuzzy time series. *Expert Systems with Applications*, 38(5), 5816–5821.
- Bose, M. and Mali, K. (2019). Designing fuzzy time series forecasting models: A survey. *International Journal of Approximate Reasoning*, 111, 78–99.
- Chen, S.M. et al. (1996). Forecasting enrollments based on fuzzy time series. *Fuzzy sets and systems*, 81(3), 311–319.
- Cheng, C.H., Chen, T.L., and Chiang, C.H. (2006). Trend-weighted fuzzy time-series model for taieix forecasting. In *International Conference on Neural Information Processing*, 469–477. Springer.
- Hwang, J.R., Chen, S.M., and Lee, C.H. (1998). Handling forecasting problems using fuzzy time series. *Fuzzy sets and systems*, 100(1-3), 217–228.
- Ismail, Z. and Efendi, R. (2011). Enrollment forecasting based on modified weight fuzzy time series. *Journal of Artificial Intelligence*, 4(1), 110–118.
- Lee, R.S. (2019). Chaotic interval type-2 fuzzy neuro-oscillatory network (cit2-fnon) for worldwide 129 financial products prediction. *International Journal of Fuzzy Systems*, 21(7), 2223–2244.
- Luo, C., Tan, C., Wang, X., and Zheng, Y. (2019). An evolving recurrent interval type-2 intuitionistic fuzzy neural network for online learning and time series prediction. *Applied Soft Computing*, 78, 150–163.
- Mendel, J.M. (2007). Type-2 fuzzy sets and systems: an overview. *IEEE Computational Intelligence Magazine*, 2(1), 20–29.
- Sadaei, H.J. (2013). *Improved models in Fuzzy Time Series for forecasting*. Ph.D. thesis, Universiti Teknologi Malaysia.
- Severiano, C.A., Silva, P.C., Sadaei, H.J., and Guimarães, F.G. (2017). Very short-term solar forecasting using fuzzy time series. In *2017 IEEE international conference on fuzzy systems (FUZZ-IEEE)*, 1–6. IEEE.
- Shukla, A.K. and Muhuri, P.K. (2019). Big-data clustering with interval type-2 fuzzy uncertainty modeling in gene expression datasets. *Engineering Applications of Artificial Intelligence*, 77, 268–282.
- Silva, P.C., Sadaei, H.J., Ballini, R., and Guimarães, F.G. (2019). Probabilistic forecasting with fuzzy time series. *IEEE Transactions on Fuzzy Systems*.
- Song, Q. and Chissom, B.S. (1993). Fuzzy time series and its models. *Fuzzy sets and systems*, 54(3), 269–277.

- Tak, N. (2020). Type-1 possibilistic fuzzy forecasting functions. *Journal of Computational and Applied Mathematics*, 370, 112653.
- Yu, H.K. (2005). Weighted fuzzy time series models for taiex forecasting. *Physica A: Statistical Mechanics and its Applications*, 349(3-4), 609–624.