


ANÁLISE DE DESEMPENHO DA INFERÊNCIA FUZZY PARA AVALIAR APARTAMENTOS: ESTUDO DE CASO NA ZONA URBANA DE BELÉM

The performance analysis by fuzzy inference to appraisal apartments: cases into urban zone at Belém/PA

Érico Gaspar Lisboa

<http://orcid.org/0000-0002-2837-1490> 

Universidade da Amazônia (UNAMA / PPDUM),
Belém, Brasil.

erico.lisboa@unama.com

Wankes Solony de Carvalho Chaves Junior

<http://orcid.org/0000-0002-4685-9043> 

Divisão de Obras e Projetos do Tribunal Regional
do Trabalho da 8ª Região, Belém, Brasil.

wankes.chaves@trt8.jus.br

RESUMO

A modelagem fuzzy é utilizada em estudos científicos para valorar imóveis, cujo desempenho é associado a comparações por inferência estatística, envoltória de dados e redes neurais artificiais. Embora apresentem bons desempenhos, a modelagem fuzzy não é normalizada, pelo que ainda carece de estudos que possam definir modelos a serem adotados. Assim, o objetivo deste trabalho é analisar desempenhos de modelos fuzzy para valorar apartamentos. Considerou-se tipos de “fuzzificação” triangular (mftri), trapezoidal (mftrap) e gaussiana (mfgau) para área privativa, número de vagas de garagem, padrão construtivo, renda per capita por bairros, localização e valor unitário. Testou-se métodos clássicos de inferência fuzzy do tipo Mamdani (IFTM) e Larsen (IFTL), cujas regras foram construídas por consulta à especialistas. Métodos de defuzzificação centróide e bisetor foram avaliados. Para comparar aos valores observados, analisou-se a variabilidade fuzzy por intervalo interquartil (IQR). Definiu-se que, quanto maior IQR maior a média de erro absoluto percentual (MAPE) da estimação fuzzy. Por 81 dados de apartamentos situados na zona urbana de Belém/PA, obteve-se correlação linear positiva entre IQR e MAPE. A diferença entre inferência fuzzy e estatística foi aceitável entre intervalos de $\pm 10\%$. Deste modo, a valoração de 10 apartamentos na mesma localidade indicou que, modelos fuzzy com menores IQR e MAPE apresentaram mais diferenças aceitáveis. Pode-se concluir que, mesmo dependentes de regras, para utilizar modelos fuzzy deve-se priorizar “mfgau” e/ou “mftri”, IFTM e/ou IFTL e método centróide. Portanto, este estudo se apresentou como auxílio à utilização e eventual normalização de modelos fuzzy à valoração de bens imóveis.

Palavras-Chave: Avaliação; Apartamentos; Estatística; Fuzzy; Erros.

ABSTRACT

Fuzzy modeling has been used in scientific studies to value real estate, whose performance is associated with comparisons by statistical inference, data envelope and artificial neural networks. Although they present good results, due to the fact that it is not standardized, fuzzy modeling still lacks studies that can define models to be adopted. Thus, the objective of this article is to analyze the performance of fuzzy models to value apartments. Triangular (mftri), trapezoidal (mftrap) and gaussian (mfgau) fuzzification types were considered for: private area, number of parking spaces, constructive standard, per capita income by neighborhood, location and unit value. Classical fuzzy inference methods of the Mamdani (IFTM) and Larsen (IFTL) type were tested, whose rules were built by specialists. Centroid and bisector defuzzification methods were evaluated. To compare the observed values, the fuzzy variability by interquartile range (IQR) was analyzed. It was defined that the higher the IQR, the higher the mean absolute percentage error (MAPE) of the fuzzy estimation. Through 81 data from apartments located in the urban area of Belém/PA, a positive linear correlation was obtained between IQR and MAPE. The difference between fuzzy and statistical inference was acceptable between $\pm 10\%$ intervals. Thus, the valuation of 10 apartments in the same location indicated that fuzzy models with lower IQR and MAPE showed more acceptable differences. It can be concluded that, even depending on rules, to use fuzzy models one should prioritize “mfgau” and/or “mftri”, IFTM and/or IFTL and centroid method. Therefore, this study was presented as an aid to the use and eventual normalization of fuzzy models for the valuation of real estate.

Keywords: Appraisal; Apartments; Statistical; Fuzzy; Error.

Preenchimento dos Editores

INFORMAÇÕES SOBRE O ARTIGO

Submetido em 06/03/2023
Publicado em 25/04/2023

Comitê Científico Interinstitucional
Editor-Responsável: Carlos Augusto Zilli
Avaliado pelo Sistema Double Blind Review
(SEER/OJS – Versão 3)



1. INTRODUÇÃO

Os procedimentos de cálculo para estimar o valor de um bem imóvel são normatizados pela ABNT NBR 14.653 e suas partes integrantes. Com ênfase nas partes 1 e 2 desta norma, o método comparativo direto de mercado (MCDDM) deve ter sua utilização preferenciada.

O MCDDM é desenvolvido por tratamento de fatores homogeneizantes ou científico. Em geral, pelas limitações de alguns fatores, face as restrições quanto a validação, abrangências regionais na sua aplicação; condicionam que, na prática da Engenharia de Avaliações, é usual adotar o tratamento científico.

A metodologia científica é caracterizada pela aplicação da estatística de regressão linear (simples ou múltipla). Como refere a NBR 14.653/2 (ABNT, 2011), a prática da regressão linear recorre ao método dos mínimos quadrados ordinários (MQO) para calcular o(s) estimador(es) da(s) variável(is) independente(s), produzindo uma função que estima o valor da variável dependente (geralmente associado ao valor total ou unitário de um bem imóvel), associado a uma parcela de erros.

Para aplicar a estatística de regressão linear, se deve analisar o cumprimento de alguns requisitos, os ditos pressupostos ou axiomas de Gauss-Markov. Casos estes pressupostos não sejam violados, o teorema de Gauss-Markov considera que os estimadores calculados pelo método MQO sejam BLUE (*Best Linear Unbiased Estimators*). Assim, o modelo estatístico produzido pode estimar o valor médio, mínimo e máximo de um bem imóvel, por um intervalo de confiança admissível e com nível de significância satisfatória.

Entretanto, além de complexo, o rigor axiomático da inferência estatística é condição limitante para a utilização do modelo de regressão linear para extrapolações. Ainda assim, embora os estimadores sejam aferidos na tentativa de minimizar a parcela de erros, a própria imprecisão dos dados coletados no mercado imobiliário também é outro limitador da utilização do modelo de regressão linear.

Deste modo, recorrer a modelagens que possam superar tais limitações, tem levado muitos trabalhos a utilizar “sistemas inteligentes”. Assim, a inferência “fuzzy” (difusa ou nebulosa) supervisionados por regras (*expert system*), redes neurais artificiais e algoritmos genéticos tem sido frequentemente empregados para a valoração de bens imóveis.

A este propósito, a inferência “fuzzy” tem sido frequentemente utilizada para valorar imóveis, ainda que em estudos científicos, pelo que se adota, majoritariamente, as proposições de Mamdani (1974) e Takagi-Sugeno (1985), nomeados como inferência “fuzzy” do tipo Mamdani (IFTM) e do tipo Sugeno-Takagi (IFST).

Entretanto, quando se recorre ao sistema IFTM, é comum a utilização do método centro de gravidade para “defuzzificação” (detalhados no item 2). Por outro lado, quando se recorre ao método IFST é comum utilizar, de forma híbrida, algoritmo genético ou redes neurais artificiais, para calibração (aprendizagem) das regras fuzzy. Pela literatura nacional, um grupo de trabalhos são sinteticamente referenciados pela Tabela 1.

Tabela1: Revisão de literatura sobre aplicação de inferência “fuzzy” na avaliação de imóveis.

Literatura	Aplicação de inferência “fuzzy” para avaliações de bens imóveis
Gonzáles (2002); Gonzáles e Formoso (2006)	Desenvolveram um modelo fuzzy baseado em IFST, com regras ajustadas por algoritmo genético, sendo combinados funções de pertinência trapezoidal e triangular das variáveis independentes: área total e privativa, padrão de qualidade da construção, idade, data de venda e endereço. Constituíram-se dois modelos IFST, e 3 regras foram baseadas em equações lineares (advindo de modelos hedônicos) função das variáveis independentes. Os autores concluíram que a inferência estatística foi mais precisa, ao passo que a inferência fuzzy obteve melhor desempenho.
Duarte et al, (2007)	Desenvolveram um modelo fuzzy baseado em IFTM, cuja fuzzificação da área privativa, número de dormitórios, localização, padrão construtivo, idade e valor foi por função de pertinência trapezoidal. O modelo foi operacionalizado por 61 regras atribuídas de modo que, pelo método centróide (defuzzificação), obteve-se o valor de imóveis do tipo apartamento. Os autores não realizaram uma análise de desempenho da inferência fuzzy.
Mello e Santello (2012)	Desenvolveram um modelo fuzzy para auxiliar a tomada de decisão agregando-lhe o método de árvore de decisão para avaliação de apartamentos em zonas urbanas e glebas de terras. Com ênfase à valoração de apartamentos, o modelo fuzzy recorreu a fuzzificação das variáveis independentes por funções de pertinência triangular, e método IFTM e defuzzificação pelo método centróide. Os autores não realizaram análise de desempenho entre as modelagens utilizadas.
Santos Junior (2014)	Desenvolveram um modelo fuzzy baseado em IFST, agregando-lhe a redes neurais artificiais. A fuzzificação por função de pertinência triangular foi associado a área útil, lazer, bairro e preço de apartamento em bairros no Rio de Janeiro. Os autores não realizaram um estudo comparativo sobre o desempenho do modelo neuro-fuzzy.
Pelli e Bolinelli Junior (2015)	Desenvolveram um modelo fuzzy baseado em IFTM para valoração de apartamento na cidade de Belo Horizonte/MG. Para tanto, os autores realizaram a fuzzificação da área privativa, padrão e localização e valor, por função de pertinência trapezoidal, triangular e triangular/trapezoidal. Utilizaram o método centróide como defuzzificador. E, compararam a inferência fuzzy com a estatística, obtendo uma diferença de 1,09%. Entretanto, utilizaram diferentes variáveis independentes no modelo estatístico.

Lima (2017)	Desenvolveu um modelo fuzzy baseado em IFTM, a partir da fuzzificação por funções trapezoidais das variáveis área, padrão, idade, local, cidade, vagas de garagem e valor unitário de apartamento na cidade de Belém/PA. E, por 29 regras, o modelo recorreu a defuzzificação pelo método centróide. O autor analisou dois grupos de dados, de modo que a avaliação de apartamentos por modelo fuzzy apresentou desvios em torno de 15% quando compara a valores observados.
Malaman e Amarin (2017)	Desenvolveram um modelo fuzzy baseado em IFTM, pela fuzzificação por funções trapezoidais das variáveis distância até o centro da cidade, topografia, padrão construtivo, zona comercial, e coordenadas do bem imóvel e valor. E, por 12 regras, o modelo recorreu a defuzzificação pelo método centróide. Os autores efetuaram uma comparação com os valores advindos da inferência estatística através do cálculo da raiz do erro médio quadrático (REMQ). O REMQ, quando comparado o modelo fuzzy agregou 22,03% e por inferência estatística o valor de REMQ foi 7,50%. Os autores concluíram que, para o contexto da planta genérica de valores, a inferência estatística obteve melhor desempenho que a inferência fuzzy para avaliação de imóveis.

Duarte et al., (2007) referiram que uma das vantagens de se utilizar a inferência fuzzy é não recorrer a dados (amostras) de mercado para constituição de um modelo para a valoração de bens imóveis (como é a modelagem por inferência estatística). Estes mesmos autores ainda referiram como outra vantagem o fato de que, algumas variáveis estudadas na Engenharia de Avaliações possuem relações muito complexas, dúbias, vagas, imprecisas, de modo que a inferência fuzzy é uma concepção que, por dedução, se consegue melhor captar estas relações em detrimento de deduções indutivas (em face da complexidade e do rigor axiomático).

Em geral, por uma breve revisão da literatura nacional especializada foi possível constatar que, a priori, é quase unânime recomendar a modelagem por inferência fuzzy para a valoração de bens imóveis – exceto por Malaman e Amarin (2017) – pelo fato de serem mais simples, dedutivas, claras e que podem, de certo modo, relacionar o pensamento humano aos dados factuais mercadológicos de imóveis.

Entretanto, a modelagem fuzzy não está descrita e detalhada em norma técnica destinada a avaliação de bens imóveis do tipo apartamento, de modo que se pudesse definir, com clareza:

- Quais as funções de pertinências devem ser adotadas para a fuzzificação das variáveis?
- Qual o método de inferência fuzzy e o método de defuzzificação (caso seja adotado método clássico de inferência) devem ser efetivamente utilizados?
- Quais modelos fuzzy tem melhor desempenho para estimar valores de bens imóveis (em comparação aos valores observados)?
- Quais modelos fuzzy podem ser utilizados como método inferencial para valorar um bem imóvel?
- Qual o procedimento, mesmo que seja minimamente simples e inteligível, deve ser adotado para aferir desempenho e definir as configurações fuzzy a serem utilizadas?

Assim, o objetivo deste trabalho é analisar desempenhos de modelos fuzzy para valorar apartamentos, a fim de auxiliar a sua utilização e eventual normatização. Para tanto, desenvolve-se doze modelos fuzzy, por combinações de funções de pertinência, métodos clássicos de inferência e métodos de defuzzificação, adotando-se procedimento para aferir desempenhos. Para utilizar o modelo fuzzy como um método inferencial, calcula-se a diferença entre os resultados obtidos pelos modelos fuzzy e pela inferência estatística, considerada ser aceitável uma diferença entre $\pm 10\%$.

A metodologia está detalhada no item 3, sendo precedido de um sintético referencial teórico para o entendimento da modelagem fuzzy (fuzzificação – inferência – defuzzificação), e da consagrada modelagem por inferência estatística dissertado no item 2. Os resultados e discussões, bem como as conclusões e recomendações estão respectivamente apresentados pelos itens 4 e 5 deste artigo.

2. REFERENCIAL TEÓRICO

O referencial teórico que embasou esta pesquisa encontra-se sintetizado em dois tópicos: quanto aos modelos para inferência fuzzy; e, modelo de regressão linear, sendo apresentado sintética conceituação e dissertado sobre os principais pressupostos.

2.1. INFERÊNCIA “FUZZY” (“FUZZIFICAÇÃO – INFERÊNCIA – “DEFUZZIFICAÇÃO”)

A inferência “fuzzy” é baseada na teoria dos conjuntos difusos (ZADEH, 1965). A característica desta teoria se configura por converter qualquer valor numérico em conjunto “fuzzy”, relacionando este valor numérico (X) por uma função de pertinência - $\mu(x)$ (*membership function*). O conjunto “fuzzy” é formado por pares, X e $\mu(x)$, pertencente no intervalo (0,1), e que caracteriza o grau de pertinência de X (por atributo linguístico).

Diferentemente da lógica booleana (aristotélica) (Figura 1A), cujos valores de X não pertence em 0 (“falso” – $F(X) = 0$) ou 1 (“verdadeiro” – $F(X) = 1$); pela lógica “fuzzy” a função $\mu(X)$ considera que X está contido entre 0 e 1, de modo que o valor de X pode ser “falso” ($\mu(x) = 0$) ou “verdadeiro” ($\mu(x) = 1$) ou nem totalmente “falso” e nem totalmente “verdadeiro” ($0 < \mu(x) < 1$) (Figura 1B).

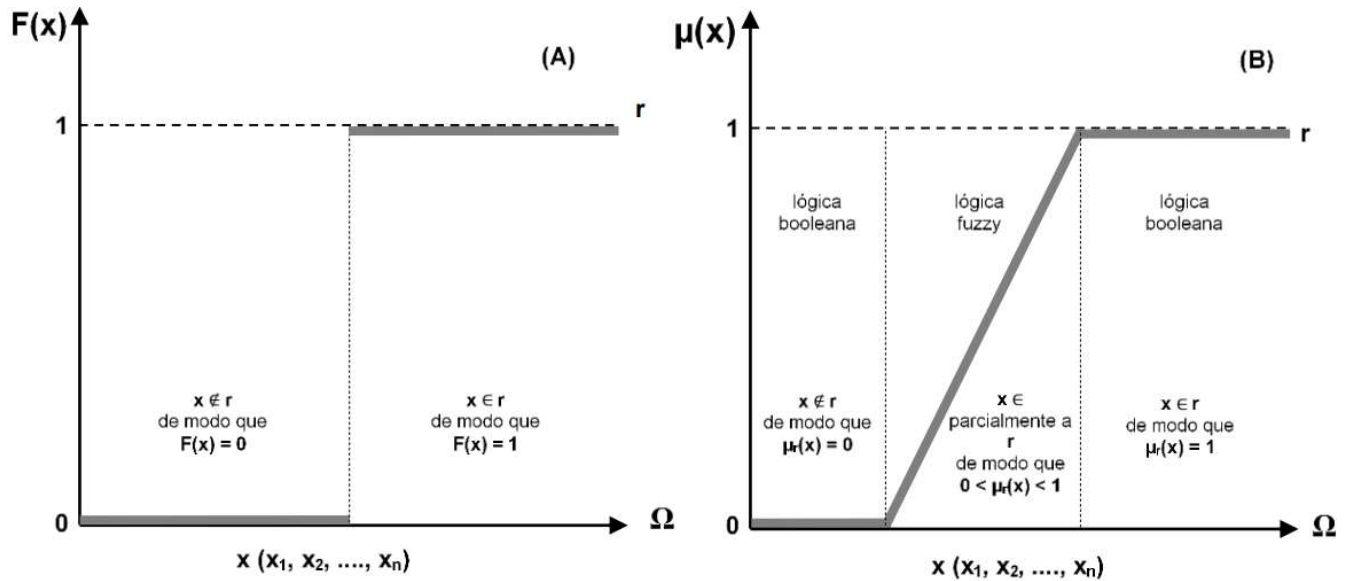


Figura 1: Representação cartesiana: (A) lógica booleana; (B) lógica fuzzy (Lisboa, 2019).

A função $\mu(x)$ é traduzida por um número real que pode ser chamado de número fuzzy. Estes números fuzzy são definidos como sendo um conjunto (r) de um conjunto de números reais, se existir pelo menos um x onde $\mu_r(x) = 1$, caracterizando uma suposição de normalização. Bauer et al., (1998) e Sandri e Correia (1999) referiram sobre três aspectos importantes para a representação dos números fuzzy: (i) suporte, (ii) corte de nível; e, (iii) núcleo. O suporte (*Supp*) de um conjunto fuzzy (r) é igual ao intervalo sobre o qual $\mu_r(x)$ não é nulo, sendo expresso por:

$$Supp(r) = \{x; \mu_r(x) > 0\} \tag{1}$$

A quantidade de elementos de um conjunto *fuzzy* (r), num universo do discurso (Ω), pode ser medida por uma forma discreta, que representa uma função característica de um conjunto clássico; e, contínua, que representa as características de um conjunto *fuzzy*, expressas, respectivamente, por: $|r| = \sum_{x \in \Omega} \mu_r(x)$ e $|r| = \int_{x \in \Omega} \mu_r(x) dx$, em que $|r|$ se refere a um conjunto (r) com elementos x (x_1, x_2, \dots, x_n) que pertencem a (r).

A partir destas expressões é que se obtém uma representação aproximada de um conjunto *fuzzy* (r) em Ω por meio de um conjunto clássico em Ω , através de subconjuntos denominados de corte de nível (α) ou corte – α (r_α), sendo expressos por: $r_\alpha = \{r \in \Omega | \mu_r(x) > \alpha\}$. Algumas formas de funções de pertinência compatíveis com r_α originam o que se denomina de núcleo, $N(r)$, que agrupa elementos x do subconjunto *fuzzy* (r) com a altura $H(r)$, e que é constituído pelos elementos x com o maior grau de compatibilidade em relação ao subconjunto *fuzzy* (r). O núcleo, $N(r)$, e a altura, $H(r)$, são definidos por:

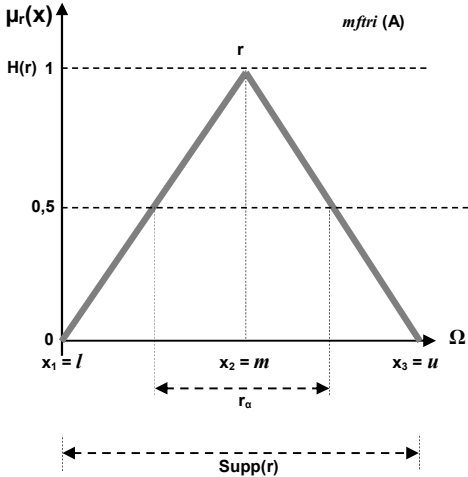
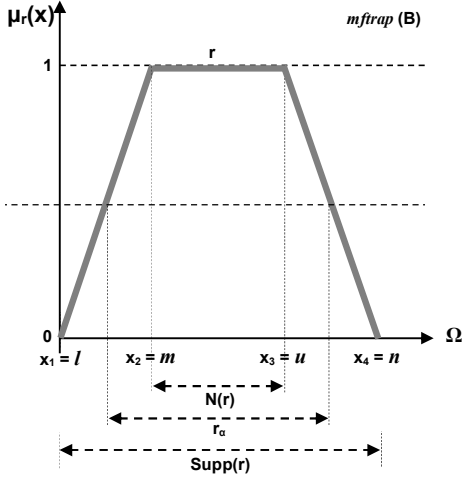
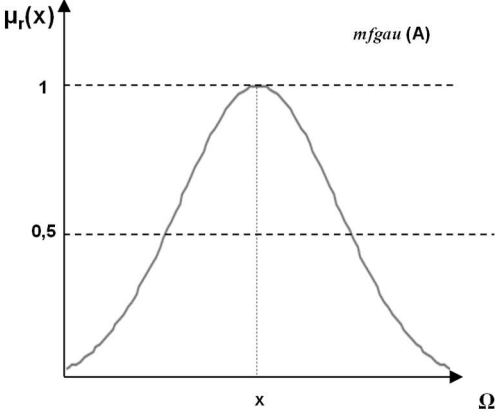
$$N(r) = \{x \in \Omega | \mu_r(x) = 1\} \tag{2}$$

$$H(r) = Supp_{(r \in \Omega)} \mu_r(x) \tag{3}$$

Estes aspectos constituem a representação de um conjunto *fuzzy*. Portanto, o conjunto suporte de um conjunto *fuzzy* (r) é o subconjunto dos elementos x em Ω , tal que $\mu_r(x) > 0$ (equação (1)). Um conjunto *fuzzy* cujo conjunto suporte é um único ponto de Ω para $\mu_r(x) = 1$ é chamado de um conjunto unitário.

Assim, as formas mais usuais que os conjuntos *fuzzy* pode assumir são triangulares, trapezoidais e gaussianas, associadas a uma função de pertinência ilustradas pela Tabela 2.

Tabela 2: Funções de pertinência definidas por números fuzzy e suas representações cartesianas.

Função de pertinência	Funções que definem um número fuzzy	Representação cartesiana
Triangular	$\mu_r(x, x_1, x_2, x_3) = \begin{cases} 0, & x \leq x_1 \\ \frac{x - x_1}{x_2 - x_1}, & x_1 \leq x \leq x_2 \\ \frac{x_3 - x}{x_3 - x_2}, & x_2 \leq x \leq x_3 \\ 0, & x_3 \leq x \end{cases}$	
Trapezoidal	$\mu_r(x, x_1, x_2, x_3, x_4) = \begin{cases} 0, & x \leq x_1 \\ \frac{x - x_1}{x_2 - x_1}, & x_1 \leq x \leq x_2 \\ 1, & x_2 \leq x \leq x_3 \\ \frac{x_4 - x}{x_4 - x_3}, & x_3 \leq x \leq x_4 \\ 0, & x_4 \leq x \end{cases}$	
Gaussiana	$\mu_r(x, \sigma, x_1) = e^{-\frac{(x-x_1)^2}{2\sigma^2}}$	

A função de pertinência triangular (*mftri*) considera (*r*) um conjunto *fuzzy* quando seu *Supp* (*r*) é o intervalo entre x_1 (*l*) e x_3 (*u*) [x_1, x_3]. O elemento x_1 refere-se ao suporte inferior, e x_3 o suporte superior. Neste caso, *N*(*r*) é igual ao elemento x_2 (*m*) posto que o mesmo não apresenta intervalo.

Portanto, um intervalo que compreende a configuração triangular é um conjunto *fuzzy* normalizado em números reais, tal que a função que descreve $\mu_r(x)$ entre *N*(*r*) suporte inferior e superior são monotônicas (crescente e decrescente).

Zadeh (1965) considerou que se estabelece um conjunto *fuzzy* ($r = (x_1, x_2, x_3)$), onde estes elementos são parâmetros escalares, sendo regido por uma forma triangular que é uma função de um vetor x_n . Deste modo, $x_1 \leq x_2 \leq x_3$ é um número *fuzzy* triangular definido por funções escritas na Tabela 2.

A função de pertinência trapezoidal (*mftrap*) considera (*r*) um conjunto *fuzzy*, quando seu *Supp* (*r*) é o intervalo entre x_1 e x_4 [x_1, x_4]. O elemento x_1 (*l*) refere-se ao suporte inferior e x_4 (*n*) ao superior. Neste caso,

$N(r)$ é definido pelo intervalo entre o núcleo inferior, $x_2 (m)$, e o superior, $x_3 (u)$. O intervalo que compreende a configuração “*mfttrap*” é um conjunto *fuzzy* normalizado em números reais, tal que a função que descreve $\mu_r(x)$ está inscrita entre *Supp* (r) e $N(r)$ inferior e superior.

Considerando um conjunto $(r) = (x_1, x_2, x_3, x_4)$, cujos seus elementos são parâmetros escalares regidos por forma trapezoidal, a função de um vetor x , em que $x_1 \leq x \leq x_2$ e $x_3 \leq x \leq x_4$ representam, respectivamente, uma pertinência crescente e decrescente, considerando $r_\alpha = 0,5$, de modo que um número *fuzzy* trapezoidal e suas funções de pertinências escritas na Tabela 2.

Tamaki et al., (1998) afirmaram que todas as funções de pertinência podem ser aproximadas pela forma trapezoidal. Entretanto, a função de pertinência na forma gaussiana (*mfgau*) é também muito utilizada para estimativas probabilísticas. A função gaussiana depende de dois parâmetros, σ e x_1 , em que uma simétrica similaridade da função densidade de probabilidade da estatística descritiva.

Assim, quando a representação cartesiana da forma “*mfgau*” assume o valor $\mu_r(x)$ igual 1, a referida função corresponde ao elemento x_1 . Os demais valores de x são limitados pela curtose. Deste modo, a simetria, σ , que é equivalente ao desvio-padrão, não torna nulo o resultado da equação que define $\mu_r(x)$.

Portanto, converter valores numéricos em conjuntos *fuzzy*, relacionando X e $\mu(X)$, na forma “*mfttrap*”, “*mfttr*” ou “*mfgau*”, se refere a “fuzzificação”. Deste modo, o conjunto *fuzzy* é formado por pares, X e $\mu(X)$, contida no intervalo $(0,1)$, e que caracteriza graus de pertinência de X (por atributo linguístico).

Entretanto, operações entre números *fuzzy*, independente das formas de pertinência, se caracterizam por propriedades algébricas entre conjuntos *fuzzy* que não seguem rigorosamente a teoria dos conjuntos clássica, como representado pela Figuras 1A e 1B.

Simões e Shaw (2007) referiram que, as principais operações entre conjuntos *fuzzy* são baseadas na teoria clássica dos conjuntos, de modo que as mais utilizadas são operações de união e interseção. Pela teoria de conjuntos *fuzzy*, a intersecção entre dois conjuntos *fuzzy* (r) e (s), $\subseteq \Omega$ (em que $x \in \Omega$) é denotada por $(r) \cap (s)$. A união entre dois conjuntos *fuzzy* (r), (s) $\subseteq \Omega$ (em que $x \in \Omega$) é denotada por $(r) \cup (s)$. Propriedades dos conjuntos clássicos que não se verificam para conjuntos *fuzzy* são: $(r) \cap (\bar{r}) \neq \emptyset$ (conjunto vazio) e $(r) \cup (\bar{r}) \neq \Omega$.

Zadeh (1965) sugeriu a propriedade da soma algébrica para união (*fuzzy algebraic sum*) e o produto algébrico para intersecção *fuzzy* (*fuzzy algebraic product*). Com a finalidade de generalizar as respectivas operações algébricas de união e intersecção *fuzzy*, definiu-se operadores norma triangular (*norma-t*) e co-norma triangular (*co-norma-t* ou *norma-s*). Portanto, sejam (r) e (s) dois conjuntos *fuzzy* em x (em que $x \in \Omega$), com funções $\mu_{r(x)}$ e $\mu_{s(x)}$, as operações de união e intersecção são definidas, respectivamente:

$$\mu_{r \cup s}(x) = \mu_r(x) \oplus \mu_s(x) \quad (4)$$

$$\mu_{r \cap s}(x) = \mu_r(x) \otimes \mu_s(x) \quad (5)$$

Em que “ \otimes ” é o operador “*norma-t*”; e, “ \oplus ” é o operador “*co-norma-t*” ou “*norma-s*”. Pedrycz (1989) referiu que as “*normas-t*” incluem o “*Min*” e “ \otimes ” e *normas-s* incluem o “*Máx*” e “ \oplus ”. Sandri e Correa (1999) descreveram outros operadores *fuzzy*, e.g. operação gama (*gamma operation*). Gomide et al., (1995) acrescentaram as operações de implicação utilizadas para a operacionalização de regras de inferência do tipo “SE” <premissa> e “ENTÃO” <conclusão>. Assim, operações de implicação classificam-se como: “*Min*”, produto “*Prod*”, “Lukasiewicz” e “*Variant of the System of Standard Sequences*” (VSS).

Referidas por Mamdani (1974), a operação de implicação de “*Min*” definiu-se como uma simplificação de operadores de “*Máx*” – “*Min*”. A operação de implicação de produto “*Prod*” foi aplicada como sucesso por Holmblad e Ostergaard (1981). A operação de implicação “Lukasiewicz” baseia-se na lógica de multivalores, incluindo uma soma algébrica entre conjuntos. Rescher (1969) definiu a implicação “VSS” como aquela que é efetuada por conectivo “OU” e implementada por uma operação de “*Máx*”.

Neste sentido, Gomide et al., (1995) esclareceram que é possível operar com conectivos “E” e “OU”, e com operadores de implicação “SE” e “ENTÃO”. Assim, pela “fuzzificação”, a operacionalização de regras pode ser sistematizada por uma forma de raciocínio ou inferências *fuzzy*.

Lee (1990a, b) conceituaram a inferência *fuzzy* como um processo utilizado para obter uma conclusão por um conjunto de regras que recorre aos operadores de implicação “SE” e “ENTÃO”. Estas regras são úteis para processar a lógica *fuzzy* e podem ser classificadas: “*modus ponens*” (MP) e “*modus tollens*” (MT).

Pelas regras MP: “SE” (r) implica num grau “ ψ ”, e (r) é verdadeiro, “ENTÃO” um grau “ ψ ” é verdadeiro. O argumento tem duas premissas: (i) “SE” e “ENTÃO” ou afirmação condicional, i.e. (r) implica num grau ψ ,

(ii) a premissa é que o conjunto (r), o antecedente do pedido condicional, é verdadeiro. Por estas premissas conclui-se que o grau ψ , conseqüente da afirmação condicional, também é verdadeiro.

Pelas regras MT: “SE” (r) implica num grau ψ , e (r) é falso, “ENTÃO” um grau ψ é falso. O argumento tem duas premissas: (i) “SE” e “ENTÃO” ou afirmação condicional, i.e. (r) implica num grau ψ , (ii) a premissa é que (r) é falso. Por estas premissas conclui-se que o grau ψ também é falso. Estas regras (R) auxiliam a inferência *fuzzy* e apresentam-se definidas por três tipos.

Tabela 3: Tipos, características e definição de regras.

R	Característica	Definição das regras
1	conseqüente <conclusão> constante	“SE” (r é ψ_1), “E”, ..., “E” (r é ψ_i), “ENTÃO” (k é ψ_j)
2	conseqüente <conclusão> por combinação linear	“SE” (r é ψ_1), “E”, ..., “E” (r é ψ_i), “ENTÃO” (k é $g(x_1, x_2, \dots, x_n) = b_0 + b_1.x_1 + b_2.x_2 + \dots + b_n.x_n - (b_0, b_1, b_2, b_n: \text{parâmetros})$).
3	conseqüente <conclusão>	“SE” (r é ψ_1), “E”, ..., “E” (r é ψ_i), “ENTÃO” (k é s)

A literatura sobre inferência *fuzzy* classifica a modelagem de regras como clássica e de interpolação. A modelagem clássica lhe é característica regras do tipo 1 e 3. A inferência de interpolação lhe é característica a regra do tipo 2. Seja uma regra: R: “SE” x_1^* é ψ_1 “E” x_2^* é ψ_2 “E”, ..., “E” x_i^* é ψ_i , “ENTÃO” (k) é ψ_j ; em que “k” é variável conseqüente (saída). Os termos x_1, x_2, \dots, x_i são as variáveis antecedentes (entrada), sendo $i = 1, 2, \dots, n$. O termo ψ_i refere-se as variáveis linguísticas que pode assumir i-termos para cada regra R, e ψ_j para variáveis linguísticas de saída. A inferência destes modelos é baseada em cinco etapas.

i. Seja x_i uma variável de entrada, definida no universo do discurso Ω , a realização de x_i é definida como o valor $x_i^* \in \Omega$, em que x_i^* é um valor de um número *fuzzy*.

ii. Sendo j-ésima premissa ($j = 1, 2, \dots, m$) da i-ésima regra ($i = 1, 2, \dots, n$) com x_i^* a compatibilidade de x_i^* com ψ_i da regra R é:

$$\alpha_{\psi_i,j} = \mu_{\psi_i,j}, \text{ onde } 1 \leq i \leq n, e 1 \leq j \leq m \tag{6}$$

iii. Com as premissas de uma dada regra avaliada, a compatibilidade global (α_j) da regra R, com x_i^* é dado pelo operador *norma-t*, igual:

$$\alpha_j = \otimes (\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{ij}) \text{ onde } 1 \leq j \leq m \tag{7}$$

iv. O valor de α_j é relacionado com k originando um conjunto conseqüente k' por operador de implicação:

$$\mu_{k'}(x) = I [\alpha_j, \mu_k(x)] \tag{8}$$

v. O operador (∇) faz agregação de contribuições das regras R acionadas para um conjunto *fuzzy* k':

$$\mu_{\psi'}(k') = \nabla [\mu_{\psi_1'}(k), \mu_{\psi_2'}(k), \dots, \mu_{\psi_i'}(k)] \tag{9}$$

Em que m é o número de termos ψ_j . O termo ∇ é associado a um operador *s-norma*. Embora estas etapas sejam atribuídas às inferências *fuzzy* por métodos clássicos, nos métodos de interpolação, apenas as etapas A), B) e C) são válidas. Neste trabalho serão apresentadas e utilizadas inferências *fuzzy* por método clássico proposto por Mamdani (1974) e Larsen (1981) (Figura 2 e 3).

Exemplificam-se regras R_1 e R_2 : R_1 é dada por $\alpha_1 = (\psi_{1,1} \otimes \psi_{1,2})$, originando (k_1); e, R_2 é dada por $\alpha_2 = (\psi_{2,1} \otimes \psi_{2,2})$ produzido (k_2). A inferência *fuzzy* do tipo Mamdani é inicialmente processado por operação de intersecção (*Min*) entre conjuntos (r_1, r_2) e (s_1, s_2). Efetua-se esta operação pela inserção das variáveis de entrada (r^* e s^*). A partir da compatibilidade de regras efetuada por α_1 e α_2 , obtém-se os conjuntos (k_1) e (k_2) por meio das respectivas funções de pertinência:

$$\mu_{\psi_1}(k) = \otimes [\mu_{\psi_{1,1}}(r), \mu_{\psi_{1,2}}(s)] \tag{10}$$

$$\mu_{\psi_2}(k) = \otimes [\mu_{\psi_{2,1}}(r), \mu_{\psi_{2,2}}(s)] \tag{11}$$

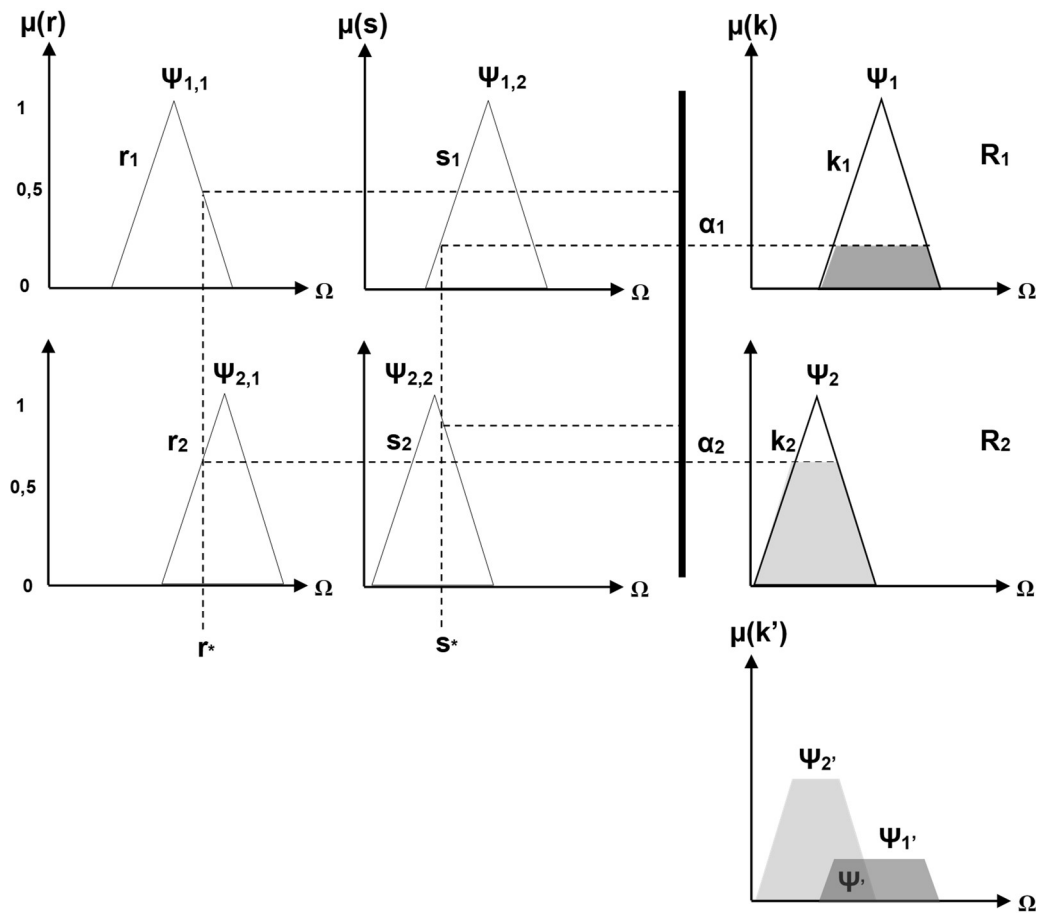


Figura 2: Inferência Fuzzy por método clássico proposto por Mamdani (1974) (Lisboa, 2019).

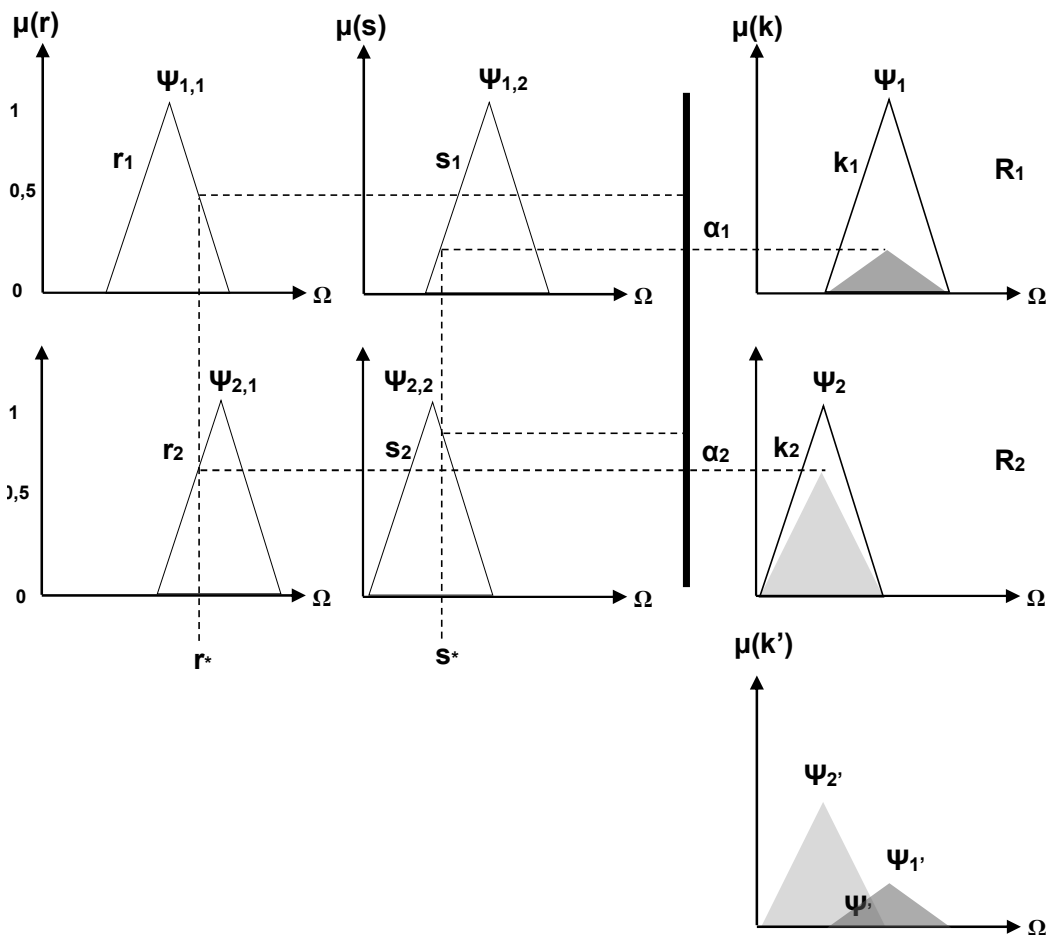


Figura 3: Inferência Fuzzy por método clássico proposto por Larsen (1981) (Lisboa, 2019).

Para obter-se um número *fuzzy*, a associação entre (k_1) e (k_2) é processada pela operação *Máx* (união) produzindo um valor (k') final, em que (k_1') = $\alpha_1 \otimes (k_1)$ e (k_2') = $\alpha_2 \otimes (k_2)$ é expresso por:

$$\mu_{\Psi}(k') = \mu_{\Psi_1'}(k) \oplus \mu_{\Psi_2'}(k) = \{[\alpha_1 \otimes \mu_{\Psi_1}(k)] \oplus [\alpha_2 \otimes \mu_{\Psi_2}(k)]\} \quad (12)$$

A inferência fuzzy de Mamdani é do tipo “*Min-Máx*”.

Exemplificam-se R_1 e R_2 aferidas respectivamente por $\alpha_1 = (\psi_{1,1}, \psi_{1,2})$ e $\alpha_2 = (\psi_{2,1}, \psi_{2,2})$. Pela inserção dos valores (r^*) e (s^*) a inferência Larsen é inicialmente processada pela operação de *Prod* entre conjuntos (r_1, r_2) e (s_1, s_2). Pela compatibilidade de regras efetuada por α_1 e α_2 , obtém-se os conjuntos (k_1) e (k_2), em função das respectivas funções de pertinência:

$$\mu_{\Psi_1}(k) = [\mu_{\Psi_{1,1}}(r) \times \mu_{\Psi_{1,2}}(s)] \quad (13)$$

$$\mu_{\Psi_2}(k) = [\mu_{\Psi_{2,1}}(r) \times \mu_{\Psi_{2,2}}(s)] \quad (14)$$

Para obter-se um número *fuzzy*, a associação entre (k_1) e (k_2) é processada pela operação *Máx* (união) produzindo o valor (k') final, em que (k_1') = $\alpha_1(k_1)$ e (k_2') = $\alpha_2(k_2)$, é expresso por:

$$\mu_{\Psi}(k') = \mu_{\Psi_1'}(k) \oplus \mu_{\Psi_2'}(k) = \{[\alpha_1 \mu_{\Psi_1}(k)] \oplus [\alpha_2 \mu_{\Psi_2}(k)]\} \quad (15)$$

A inferência fuzzy de Larsen é do tipo “*Prod-Máx*”. Importante considerar que, Tsukamoto (1979) e Sugeno-Takagi (1985) propuseram inferências fuzzy por métodos de interpolação que recorrem as regras do tipo 2 (Tabela 3). As regras tipo 1 e 3 são processadas pelos modelos clássicos de Mamdani e Larsen, cujos resultados do valor (k_1 e k_2) são diferentes e ainda são números fuzzy. Deste modo, requer-se um processo para transformá-los num número determinístico (*crisp value*), nomeado de “defuzzificação”.

A “defuzzificação” é uma operação inversa a lógica fuzzy, de modo que os métodos de centro de gravidade ou centróide são bastantes utilizados. Para aplicar o método centróide três etapas são cumpridas:

- i. Determinar o ponto centróide ψ_1' e ψ_2' para saída de k_1 e k_2 , respectivamente;
- ii. Calcular a área que corresponde ambos os conjuntos para cada saída ativada, e.g. para uma função de pertinência trapezoidal tem-se: k_1 e k_2 as bases menores de ψ_1' e ψ_2' , respectivamente, e a base maior ψ_1' é k_1 e de ψ_2' é k_2 de tal modo que:

$$\Psi_1' = \frac{\mu_{\Psi_1'}(k) \cdot (k_1' + k_1'')}{2} \quad (16)$$

$$\Psi_2' = \frac{\mu_{\Psi_2'}(k) \cdot (k_2' + k_2'')}{2} \quad (17)$$

- iii. Calcular a média ponderada dos pontos centróides pelas respectivas áreas de ψ_2' e ψ_1' , em que:

$$k' = \frac{k_1 \cdot \Psi_1' + k_2 \cdot \Psi_2'}{\Psi_1' + \Psi_2'} \quad (18)$$

Assim, o valor de saída obtido pelo método centróide é ilustrado pela Figura 4A. Como ilustrado pela 4B, considerando a exemplificação dos resultados advindo do modelo de Mamdani ou de Larsen, pode-se calcular o valor de k' (associado ao valor estimado \hat{Y}_i) que se refere ao ponto de encontro das bissetrizes dos ângulos internos de um conjunto fuzzy:

$$k' = \frac{k_1 \cdot \mu_{\Psi_1'}(k) + k_2 \cdot \mu_{\Psi_2'}(k)}{\mu_{\Psi_1'}(k) + \mu_{\Psi_2'}(k)} \quad (19)$$

Lima (2016) considerou que a bissetriz é a linha vertical que dividirá uma região em duas sub-regiões de mesma área, as quais nem sempre são linhas coincidentes com a linha de centroide (Figura 4B).

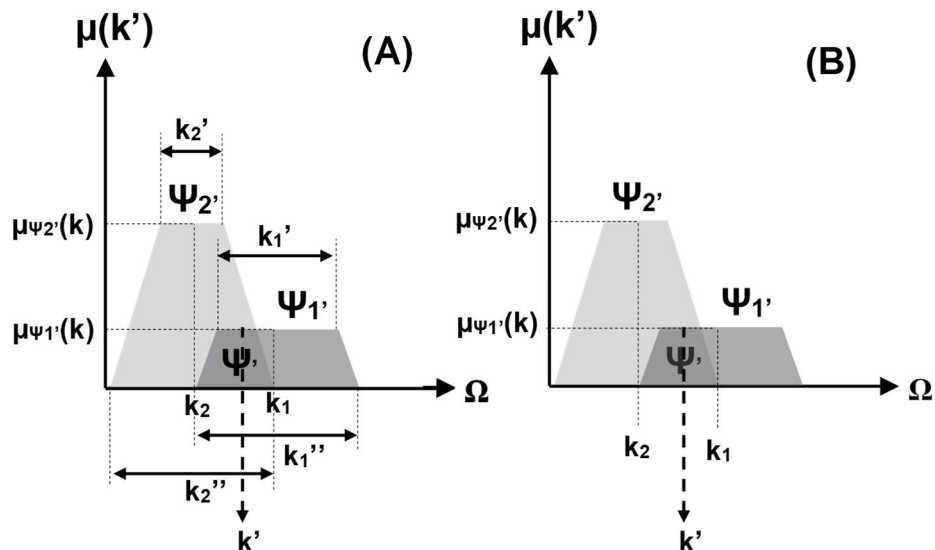


Figura 4: Métodos de “defuzzificação”: (A) Centróide; (B) Bissetor (Lisboa, 2019)

Importa considerar que uma das características adotadas pelos métodos clássicos de inferência *fuzzy* refere-se à necessidade de utilizar métodos de “defuzzificação”. Outros métodos de “defuzzificação” são referidos por Mizumoto (1982); Lee (1990b); Ross (2004), e.g., método da média dos máximos, média ponderada dos máximos, do critério máximo ou mínimo. A utilização destes métodos de defuzzificação são dispensados nos modelos *fuzzy* de interpolação.

Notadamente que a inferência *fuzzy* pode ser tratada como um sistema. Por esta configuração torna-se possível a construção de um sistema de inferência *fuzzy* (*fuzzy inference system* – FIS). Lee (1990b); Driankov et al., (1993); Sandri e Correia (1999) referiram que um FIS é composto dos elementos: “fuzzificação”, base do conhecimento, inferência e “defuzzificação” (Figura 5).

Grima e Verhoef (1999) referiram que o FIS é construído por três etapas. A primeira etapa se caracteriza pela seleção de variáveis de entrada (r^* e s^*) e saída (k'), baseadas nos conhecimentos científicos (*state of art*) e capacidade de julgamento. A literatura apresenta dois tipos de modelos monovariáveis e multivariáveis.

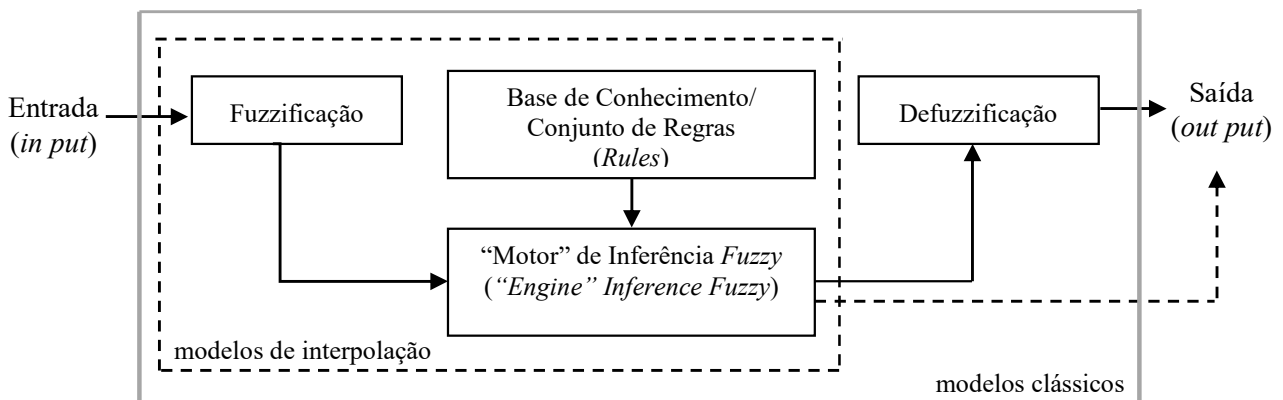


Figura 5: Ilustração de um “Fuzzy Inference System – FIS” (Lisboa, 2019).

Os modelos monovariáveis produzem uma relação causa-efeito por apenas uma variável de entrada e de saída. Os multivariáveis produzem três tipos de relação causa-efeito: múltiplas entradas e uma saída (MISO - *multiple inputs, single outputs*); múltiplas saídas (SIMO - *single inputs, multiple outputs*); e, múltiplas entradas e saídas (MIMO - *multiple inputs, multiple outputs*) (LISBOA, 2019).

A segunda etapa caracteriza-se pela relação entre entrada e saída através de regras. Deste modo, utiliza-se uma sentença do tipo “SE” – “ENTÃO”, com o eventual recurso aos operadores “E” e “OU” no antecedente da sentença, processadas pelos modelos de inferência *fuzzy*.

A terceira etapa se caracteriza pelo processo de “defuzzificação”. Como anteriormente referido, este processo transforma a variável linguística de saída em valores determinísticos. Este procedimento é apenas adotado para modelos de inferência *fuzzy* clássicos.

2.2. INFERÊNCIA ESTATÍSTICA

A inferência estatística aplicada na avaliação de bens imóveis, normatizada pela NBR 14.653/2 (ABNT, 2011), é implementada por quatro procedimentos básicos: seleção de variáveis independentes, predição, estimação de regressores (parâmetros) das variáveis independentes e a inferência propriamente dita.

A seleção de variáveis independentes (X_i) ($i = 1, 2, \dots, n$) consiste em verificar quais podem influenciar na variação da variável dependente (Y_i). A análise de regressão pode auxiliar neste processo para eliminar X_i que não tenha importante contribuição para estimação de Y_i . Adita-se a este processo de seleção a micronumerosidade da amostra, referido no subitem 9.2.1 da NBR 14.653/2 (ABNT, 2011).

Para este processo de seleção, algumas técnicas são utilizadas entre X_i e Y_i , tais como: determinação do coeficiente de determinação (R^2) e ajustado (R^2_a); o cálculo do quadrado médio dos erros; estimativa do C_p de Mallows (baseado no erro quadrático médio dos valores ajustados); informação de Akaike e Bayesiano; etc. Convém referir que, adicionalmente, a seleção automática pode ser realizada pelo processo de “Forward”, “Backward” e “Stepwise”.

A predição consiste na obtenção de um modelo matemático, cuja variação da variável Y_i seja explicada por uma ou n-variáveis X_i , onde $Y_i = f(X_1, X_2, \dots, X_n)$. Pela extração de dados amostrais de uma população é possível estabelecer parâmetros de X_i , tornando possível explicar o comportamento de Y_i por:

$$\hat{Y}_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon_i \tag{20}$$

Em que $\beta_0, \beta_1, \beta_2, \beta_n$ se refere aos regressores (ou parâmetros) das variáveis X_1, X_2, \dots, X_n . O termo ε_i se refere ao erro na estimativa de Y_i , representado por \hat{Y}_i . Assim é esperado que a diferença (ΔY) entre Y_i e \hat{Y}_i seja igual a 0. Por razões algébricas eleva-se ao quadrado: $\Delta Y^2 = (Y_i - \hat{Y}_i)^2$, onde ΔY^2 é substituído pelo termo “S”, convencionalmente adotada para referir sobre o desvio quadrático total, cuja incerteza seja expressa por:

$$\Delta Y^2 = \sigma = \sqrt{\frac{S}{N}} = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{N}} \tag{21}$$

Em que “N” seja o número de amostras. A otimização do ajuste consiste em minimizar o termo “ σ ” (erro padrão), pelo que, correspondentemente, minimiza “S”. Assim, pelos dados amostrais, pode-se, depois de devidos ajustes, levar à obtenção de uma relação satisfatória entre a variável dependente (Y_i) e independentes (X_1, X_2, \dots, X_n), por uma função \hat{Y}_i . Para tanto, se deve estimar os parâmetros $\beta_0, \beta_1, \beta_2, \beta_n$ pelo método dos mínimos quadrados ordinários (MQO). Considerando que a função depende de “n” parâmetros, β_n , ao se recorrer a minimização do erro por meio da equação (20), pode-se escrever a seguinte expressão:

$$S(\beta_0; \beta_1; \beta_2; \dots; \beta_n) = \sum_{i=1}^n [Y_i - \beta_0 - \beta_1 X_1 - \beta_2 X_2 - \dots - \beta_n X_n]^2 \tag{22}$$

Para que $S(\beta_0; \beta_1; \beta_2, \dots, \beta_n)$ seja mínimo se considera a condição: $\frac{\partial S(\beta_0; \beta_1; \beta_2; \dots; \beta_n)}{\partial \beta_0} = 0, \frac{\partial S(\beta_0; \beta_1; \beta_2; \dots; \beta_n)}{\partial \beta_1} = 0, \frac{\partial S(\beta_0; \beta_1; \beta_2; \dots; \beta_n)}{\partial \beta_2} = 0, \dots, \frac{\partial S(\beta_0; \beta_1; \beta_2; \dots; \beta_n)}{\partial \beta_n} = 0$ (ALMEIDA, 2015). Recorrendo a derivação parcial desta condição, se constrói um sistema de equações lineares, cujas incógnitas sejam $\beta_0; \beta_1; \beta_2, \dots, \beta_n$, de n+1 equações com n+1 incógnitas que, quando solucionado, fornece os parâmetros da função de ajuste \hat{Y}_i escrito pela matriz:

$$\begin{bmatrix} n & \sum X_1 & \sum X_2 & \dots & \sum X_n \\ \sum X_1 & \sum X_1 X_1 & \sum X_2 X_1 & \dots & \sum X_n X_1 \\ \sum X_2 & \sum X_1 X_2 & \sum X_2 X_2 & \dots & \sum X_n X_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum X_n & \sum X_1 X_n & \sum X_2 X_n & \dots & \sum X_n X_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum Y_i X_1 \\ \sum Y_i X_2 \\ \vdots \\ \sum Y_i X_n \end{bmatrix} \tag{23}$$

A resolução desta matriz pode ser feita pelo regramento de Cramer, calculando seu determinante (Δ), e os determinantes de cada um dos regressores ($\Delta\beta_0, \Delta\beta_1, \Delta\beta_2, \Delta\beta_n$), de modo que a relação entre eles se obtém os respectivos parâmetros: $\beta_0 = \Delta\beta_0 / \Delta$; $\beta_1 = \Delta\beta_1 / \Delta$; $\beta_2 = \Delta\beta_2 / \Delta$; $\beta_n = \Delta\beta_n / \Delta$.

Pelo método MQO importa obter estimadores lineares não viesado do tipo BLUE (*Best Linear Unbiased Estimator*), i.e., os quais tenham a menor variância entre os estimadores lineares não-viesados, de tal modo que as hipóteses de Gauss-Markov não sejam violadas, as quais sejam: o modelo de regressão é linear nos parâmetros; a amostragem aleatória; erros são homocedásticos e tem distribuição normal, e as variáveis independentes não devem ser colineares entre si.

Iniciado a inferência estatística, inclui-se neste processo a análise da variância (ANOVA) que afere o desempenho do modelo obtido e a significância dos parâmetros das variáveis independentes e do referido modelo. As medidas de desempenho estão relacionadas a mensuração dos erros (resíduos) de estimação da função linear pela soma do quadrado dos erros (SQE) e devido a regressão (SQR) iguais a:

$$SQE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (24)$$

$$SQR = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 \quad (25)$$

Em que \bar{Y} é o valor médio da i-amostras. A soma dos quadrados totais (STQ) é igual à soma de SQE e SQR. No geral, pela relação entre SQR e STQ se calcula o coeficiente de determinação (R^2), indicando o quão preciso é o modelo \hat{Y}_i para explicar a variável Y_i , e o R^2 ajustado, detalhados em Almeida (2015).

A significância dos parâmetros das variáveis independentes é aferida pelo teste de Student (teste-T). Esta significância (α) está associada a uma taxa de erro tipo I (rejeita a hipótese nula – $H_0: \beta_n = 0$, cuja variável X não é relevante para explicar Y) e erro do tipo II (não rejeita H_0 , e aceita a hipótese alternativa $H_1: \beta_n \neq 0$, cuja variável X é relevante para explicar Y).

O subitem 9.2.1 a NBR 14653/2 (ABNT, 2011) recomenda avaliar a significância através de três graus de fundamentação, os quais estão associadas a um erro de 10% (grau III), 20% (grau II) e de 30% (grau I). Para tanto, efetua-se uma comparação entre o indicativo tabelado da função densidade de probabilidade “T” de Student (T_{tabelado}) e um indicativo função do regressor da variável X_i estimado pelo método MQO por:

$$T_{\text{calculado}} = \frac{\hat{\beta}_i}{\sigma_{\hat{\beta}_i}} \quad (26)$$

Em que $\hat{\beta}_i$ é um regressor estimado pelo método MQO para a variável X_i . E o termo $\sigma_{\hat{\beta}_i}$ se refere ao erro padrão da regressão aferido por:

$$\sigma_{\hat{\beta}_i} = \sqrt{\frac{SQE}{n - k - 1}} \quad (27)$$

Em que k se refere ao número de variáveis independentes consideradas na modelagem estatística inferencial. Portanto, se o $T_{\text{calculado}}$ for superior ao T_{tabelado} aceita-se a hipótese H_1 , rejeitando a hipótese H_0 .

Pelo teste de Fisher-Snedecor (teste-F) se avalia a significância do modelo de regressão linear obtido. Neste teste, a taxa de erro tipo I (rejeita a hipótese $H_0: \beta_n = 0$, o modelo não contribui para explicar Y) e erro tipo II (não rejeita H_0 , e aceita a hipótese $H_1: \beta_n \neq 0$, o modelo é relevante para explicar Y) são relacionados a 1%, 2% e 5%.

Estes erros estão respectivamente associados a graus de fundamentação: erro de 1% (grau III), 2% (grau II) e de 5% (grau I), referido pelo subitem 9.2.1 da NBR 14653/2 (ABNT, 2011). Para aferi-los, compara-se um indicativo tabelado da função densidade de probabilidade qui-quadrado (F_{tabelado}) e outro indicativo, $F_{\text{calculado}}$, que é função da média do quadrado dos erros devido a regressão e dos erros.

$$F_{\text{calculado}} = \frac{MQR}{MQE} \quad (28)$$

O termo MQR se refere ao quadrado médio dos erros devido a regressão que é uma relação entre SQR e o número de variáveis independentes (k). A média do quadrado dos erros (MQE) é obtido pela relação entre SQE e o termo $(n - k - 1)$. Se o $F_{\text{calculado}}$ for superior ao F_{tabelado} aceita-se a hipótese H_1 , rejeitando a hipótese H_0 , em que o modelo de regressão tem fortes evidências de explicar o valor de Y .

No que se refere a avaliação dos pressupostos de Gauss-Markov, quanto as duas primeiras hipóteses, em que se deve considerar os parâmetros lineares e as amostras aleatórias, Dantas (2005) referiu que, nos casos de dados imobiliários, as variáveis independentes estão relacionadas com as características fixas de cada elemento tomado como referência, de tal modo que esta hipótese é atendida.

Assim sendo, Dantas (2005) também considerou que as amostras sejam aleatórias, se pressupõe que os modelos para estimar valores de imóveis são lineares nos parâmetros das variáveis independentes, não existindo relações lineares entre as variáveis independentes.

Considerando que os erros são variáveis aleatórias com valor esperado nulo e variância de cada termo residual seja constante e igual ao desvio-padrão (σ^2), em que: $\text{Var}(\hat{Y}_i | X_i) = \text{Var}(\epsilon_i) = \sigma^2$, os erros são denominados homocedásticos. A violação desta hipótese H_0 ($H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$), implica em considerar que pelo menos a variância de um termo residual σ_i^2 seja diferente ($i = 1, 2, \dots, n$), rejeitando H_0 , pelo que $\text{Var}(\hat{Y}_i | X_i) = \text{Var}(\epsilon_i) = \sigma_i^2$, em que os erros sejam heterocedásticos.

Para aferir se esta condição não seja violada é usual recorrer a análise gráfica que relaciona erros (ΔY) e valores estimados pelo modelo (\hat{Y}_i), se visualizando a variação uniforme dos erros, sem tendências. Convém adotar a versão padronizada dos erros, relacionado ao Z-score da distribuição normal, igual a relação da diferença entre $(\Delta Y - \bar{\Delta Y})$ e $\sigma_{\Delta Y}$, onde $\bar{\Delta Y}$ e $\sigma_{\Delta Y}$ seja a média e o desvio-padrão dos erros. Além da análise gráfica, pode-se recorrer aos testes de Park ou White, referido pelo Anexo A, subitem A.2.1.3 da NBR 14.653/2 (ABNT, 2011).

Pelos erros padronizados também é possível avaliar se estão enquadrados à distribuição normal pela comparação da frequência relativa nos intervalos de $[-1; +1]$, $[-1,64; +1,64]$ e $[-1,96; +1,96]$, respectivamente associados as probabilidades de 68%, 90% e 95%. Esta hipótese pode ser graficamente validada igualmente pela condição anterior, i.e, sem tendências pela relação entre erros padrões e \hat{Y}_i . Alguns testes podem ser utilizados para validar esta hipótese, e.g., teste Qui-quadrado, de Komolgorovi-Smirnov, Jarque-Bera, Cramer-Von Mises, de Anderson-Darling.

A hipótese de não colinearidade entre as variáveis explicativas considera que estas sejam linearmente independentes. Para diagnosticar a aceitação da hipótese de não-colinearidade, inicialmente, recorre-se a análise da matriz de correlações. Esta matriz espelha as dependências lineares de primeira ordem entre as variáveis independentes através dos valores dos coeficientes de correlação ou de Person "R".

No âmbito da Engenharia de Avaliações, o Anexo A 2.1.5.2. da NBR 14653/2 (ABNT, 2011) refere que, ao avaliar a matriz de correlação, é necessário ter atenção especial para "R" > 0,80. Montgomery; Peck; Vining (2006) alertaram que, quando mais de dois regressores estão envolvidos na dependência linear, a matriz de correlação não é eficiente.

A referida normativa alertou sobre a possibilidade de ocorrer multicolinearidade, mesmo quando a matriz de correlação apresentar "R" << 0,80, de modo a recomendar a correlação de cada variável com subconjuntos de outras variáveis independentes por regressões auxiliares (GUJARATI; PORTER, 2011). Além da matriz de correlação e das regressões auxiliares, Gazola (2002) sugeriu a utilização de outros métodos e avaliações para a detecção de multicolinearidade, como pela avaliação do coeficiente R^2 do modelo de regressão e cálculo do Fator de Inflação da Variância (FIV).

Importa ressaltar que amostras pequenas e faltas de variabilidade das variáveis independentes podem causar problemas que são no mínimo tão graves quantos aqueles decorrentes da multicolinearidade, nomeadamente os problemas da micronumerosidade da amostra.

Adicionalmente ao processo de inferência, deve-se analisar os dados discrepantes ("outliers") e influenciante. Os dados discrepantes podem ser analisados pelo gráfico que relaciona erros padronizados e \hat{Y}_i , cujos erros padrões superiores ao intervalo $[-2; +2]$ sugerem que sejam discrepantes.

Os "outliers" podem ser avaliados pela distância de Mahalamobis. Importa ressaltar que, se $n-1$ elementos estiveram aleatoriamente distribuídos em intervalo de erro padrão $[-1; +1]$, o enésimo elemento, que pode corresponder a $[-1,5; +1,5]$, mesmo abaixo de $[-2; +2]$, pode ser "outlier". Os dados influenciantes podem ser avaliados pela distância de Cook.

3. METODOLOGIA

O desenvolvimento da metodologia foi baseado em quatro etapas: (i) pesquisa de dados imobiliário de oferta referente a bens imóveis apartamentos; (ii) modelagem por inferência fuzzy; (iii) modelagem por inferência estatística; e, (iv) análise comparativo e de desempenho entre inferências fuzzy e estatística para valoração de um bem imóvel urbano do tipo apartamento.

3.1. A PESQUISA DE DADOS IMOBILIÁRIOS

A pesquisa de dados imobiliário foi realizada considerando três aspectos. O primeiro definido pela preparação de dados de imóveis do tipo apartamento, localizada na zona urbana da cidade de Belém/PA, utilizado para financiamento imobiliário de uma instituição bancária do Pará, realizada entre 2019 e 2020.

O segundo aspecto definiu as variáveis independentes mais utilizadas pelas empresas credenciadas para realizar as avaliações imobiliárias do referido financiamento por uma instituição financeira local. Deste modo, a amostra foi composta por 81 dados relacionadas as variáveis independentes: (a) área privativa (AP), (b) número de vagas de garagens (VG), (c) padrão construtivo (PC), (d) renda familiar (RP); e, (e) localização (LOC); e a própria variável dependente de valor dos apartamentos (f) (Figura 6).

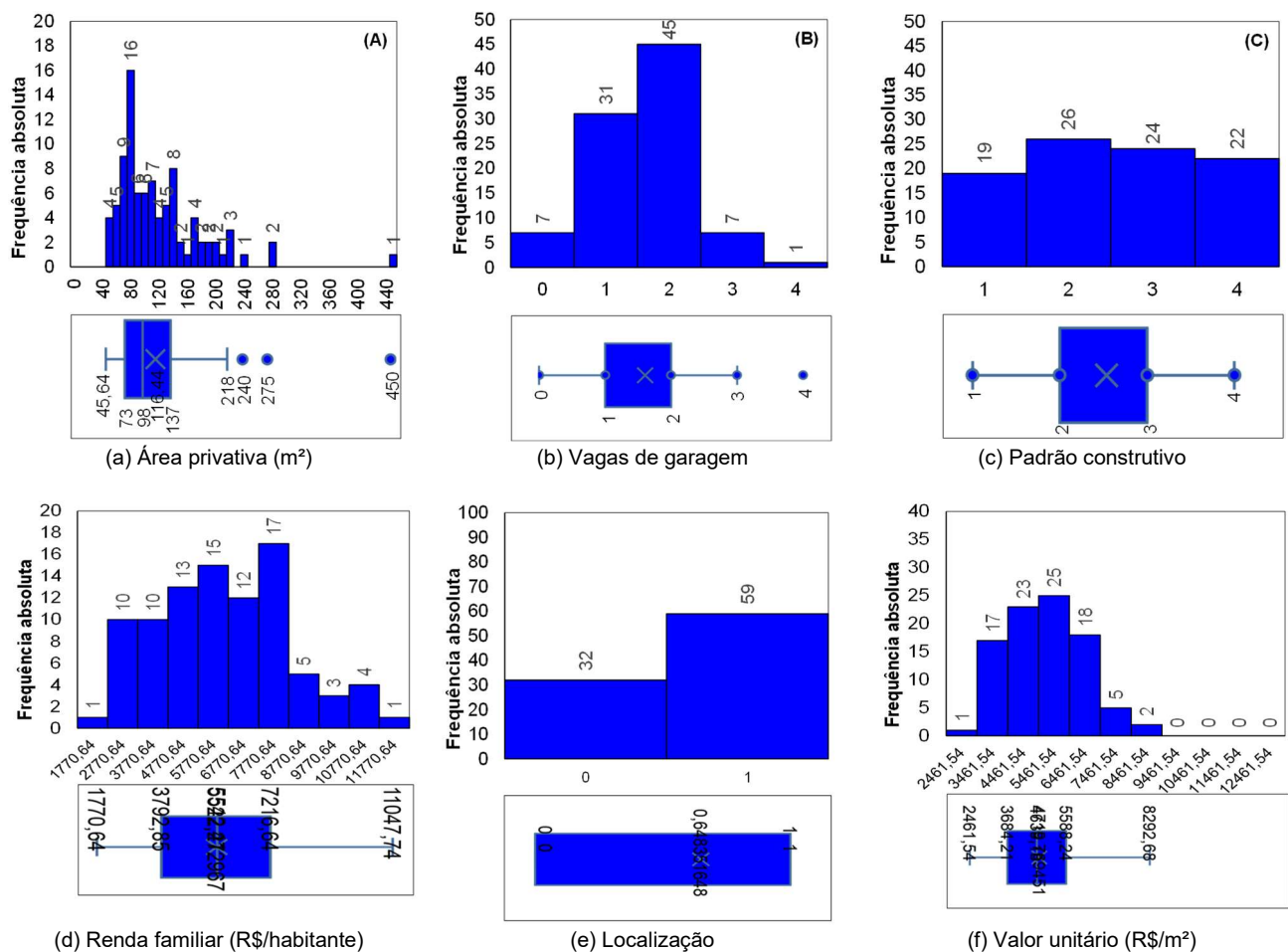


Figura 6: Estatística descritiva dos dados amostrais.











Definidos e organizados os dados referentes as variáveis independentes e dependente, recorreu-se a modelagem por duas metodologias para avaliar apartamentos. A primeira foi desenvolvida com base na teoria dos conjuntos fuzzy (ZADEH, 1965), de modo que foram utilizados métodos clássicos de inferência fuzzy. A segunda modelagem foi aplicada a partir da inferência estatística de regressão linear multivariada, cujos regressores foram estimados pelo método MQO, de tal modo que a avaliação dos pressupostos deste procedimento é fundamentada pela NBR 14.653/2 (ABNT, 2011).

Em seguida, pelos valores observados no mercado, fez-se comparações com valores estimados pelas modelagens fuzzy para aferir o desempenho médio pelo erro absoluto percentual (MAPE). Assim, é

pretendo avaliar quais configurações de modelos fuzzy, no âmbito do processo de fuzzificação, inferência e defuzzificação, apresentam melhor desempenho para estimar valor de imóveis do tipo apartamento.

Para estabelecer praticidade deste estudo, o terceiro aspecto da pesquisa de dados imobiliários considerou a avaliação de 10 apartamentos por inferência estatística (Tabela 4). A Figura 7 ilustra a espacialização dos dados amostrais e avaliados na zona urbana da região metropolitana Belém (RMB).

Tabela 4: Imóveis que foram disponibilizados à leilão (imóveis de garantias bancárias).

ID	1	2	3	4	5	6	7	8	9	10
APTO.										
AP	80	40,89	57,97	40,47	43,59	53,37	51,05	61,85	45,0	89,73
VG	0	0	0	0	0	1	1	1	0	0
PC	3	1	1	1	1	2	2	2	2	1
RP	3.710,31	1.770,64	1.770,64	1.770,64	1.770,64	4.779,52	4.779,52	4.007,96	4.007,96	3.627,94
LOC	1	0	0	0	0	1	1	1	1	0

AP: em m²; RP em R\$/habitante. As unidades 2, 3, 4 e 5, pertencem ao mesmo prédio, igualmente a 6 e 7, e 8 e 9.

Portanto, efetua-se o cálculo da diferença entre os valores produzidos pela inferência estatística em comparação aos modelos fuzzy. Deste modo, pode-se observar o quão precisos foram os modelos produzidos pela inferência fuzzy que, de forma dedutiva, se propuseram em avaliar apartamentos que foram aferidos pelo rigor axiomático da inferência estatística, que é normatizado pela NBR 14653/2 (ABNT, 2011).

3.2. PROPOSTA DE MODELAGEM POR INFERÊNCIA FUZZY

O processo de modelagem por inferência fuzzy foi desenvolvido pela implementação de três fases. A primeira está associada a “fuzzificação” das variáveis independentes e dependente. Nesta fase, recorre-se as funções de preferências, de modo que se definiu três tipos de configurações:

(A) Variáveis independentes e dependente quantitativas contínuas assumiram uma função de pertinência trapezoidal (*mfttrap*), e as variáveis independentes quantitativas discretas (incluído aqueles que foram associadas aos códigos alocados) assumiram uma função de pertinência triangular (*mfttri*);

(B) Variáveis independentes e dependente quantitativas contínuas, variáveis independentes quantitativas discretas (incluído os que foram associadas aos códigos alocados) assumiram uma função de pertinência gaussiana (*mfgau*);

(C) Variáveis independentes e dependente quantitativas contínuas assumiram uma função “*mfgau*”, e as variáveis independentes quantitativas discretas (incluído aqueles que foram associadas aos códigos alocados) assumiram uma função “*mfttri*”.

O estabelecimento dos números fuzzy, quanto ao suporte ($Supp(r)$), corte de nível (α) e nível ($N(r)$), foram empiricamente definidos, de modo a produzir intervalos que, a priori, traduziram a atribuição linguística de AP, VG, PC, RP e LOC, e do valor unitário observado de um apartamento, considerando:

- Área privativa (AP): Expresso em metro quadrado (m²), o atributo está associado a uma área privativa pequena “P” de 0 – 50 m². A área privativa média “M” se refere a um intervalo entre 100 – 150 m². Entre as áreas privativas maiores que 50 m² e inferiores a 100 m² é um intervalo nebuloso, não é uma área privativa “P” e nem “M”. Para intervalos superiores a 200 m² a área privativa é alta “A”. Entretanto, AP maiores a 150 m² e inferiores a 200 m² é um intervalo nebuloso, não é enquadrada como “M” e nem como “A”;

- Vagas de garagem (VG): Expresso em quantidade (número de vagas), o atributo está associado a poucas vagas de garagem “P” igual a 1. O número VG regular “R” é 3. Entre o número VG maior que 1 e inferior a 3 é um intervalo nebuloso. Para um número VG igual ou superior a 5 é “A”. Entretanto, para número VG menor que 5 e maior que 3 é um intervalo nebuloso, não é um número de vagas “A” e nem “R”;

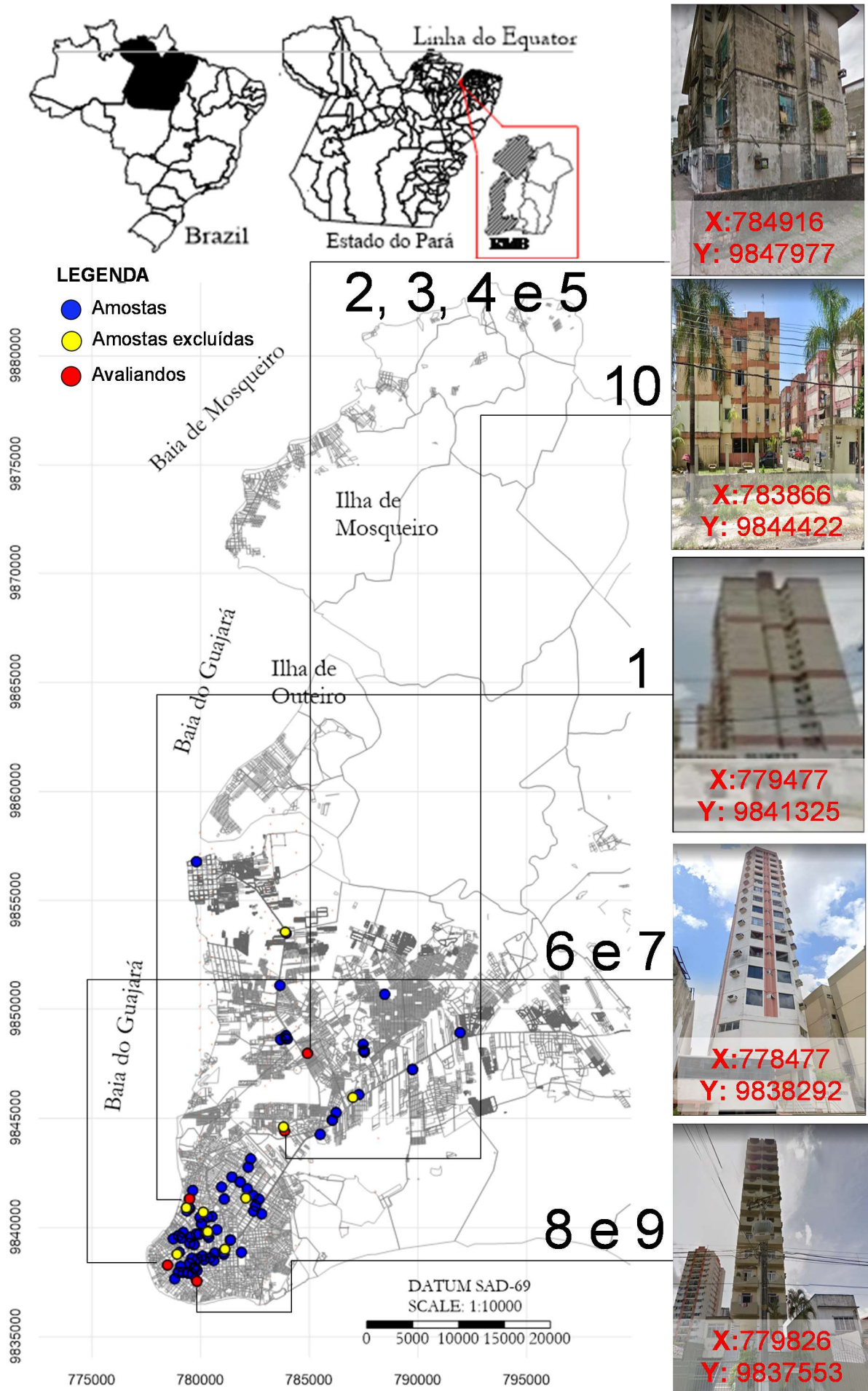


Figura 7: Distribuição espacial dos dados amostrais e dos avaliandos (autoria própria).

- Padrão construtivo (PC): Expresso em código alocado, PC baixo “B” é igual a igual a 0. O PC médio “M” é igual a 2. Entre o PC maior que 0 e inferior a 2 é um intervalo nebuloso. O PC é alto “A” para um valor 4. Entretanto, para um valor referente a PC maior que 2 e menor que 4 é intervalo nebuloso, não é um padrão “M” e nem “A”;

- Renda *per capita* por bairro (RP): Expresso em reais por habitante (R\$/hab), este atributo é associado a uma renda baixa “B” de 0 – R\$ 2.000,0/ hab. A renda média “M” se refere a valores entre 3.000,0 – R\$ 4.500/hab. Rendas maiores que R\$ 2.000,0/hab e inferiores a R\$ 3.000,0/hab é um intervalo nebuloso, não é “B” e nem “M”. Valores iguais ou superiores a R\$ 6.000,0/hab é alta (A). Entretanto, para rendas maiores que R\$ 4.500/hab e inferiores a R\$ 6.000,0/hab é um intervalo nebuloso, não é “M” e nem “A”;

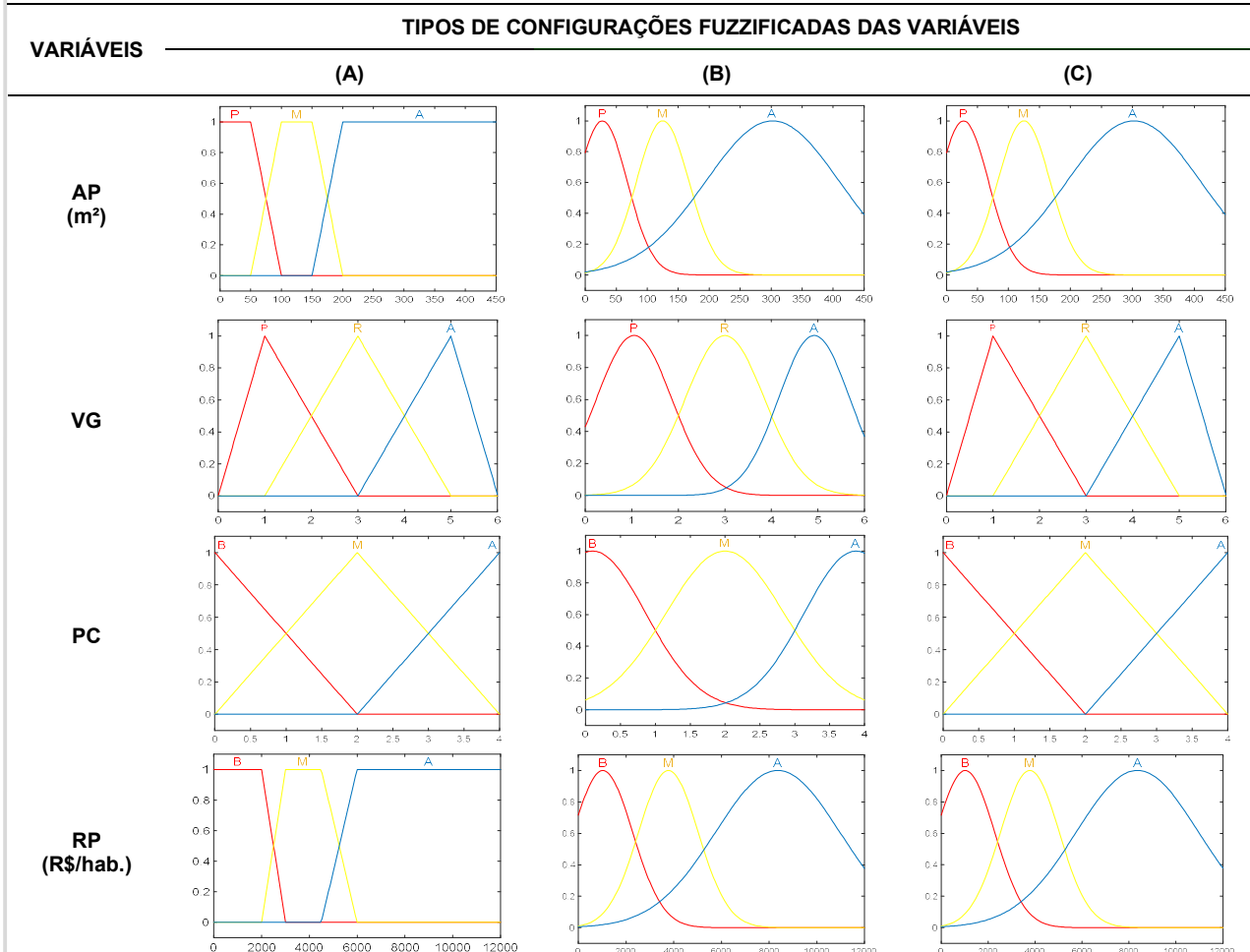
- Localização (LOC): Expresso em código alocado, a localização péssima “P” é 0. A localização regular “R” é igual a 0,5. A localização maior que 0 e inferior a 0,5 é um intervalo nebuloso. A localização é boa “B” para um valor igual 1. Entretanto, para um valor referente a localização maior que 0,5 e menor que 1 é um intervalo nebuloso, não é um padrão “R” e nem “B”;

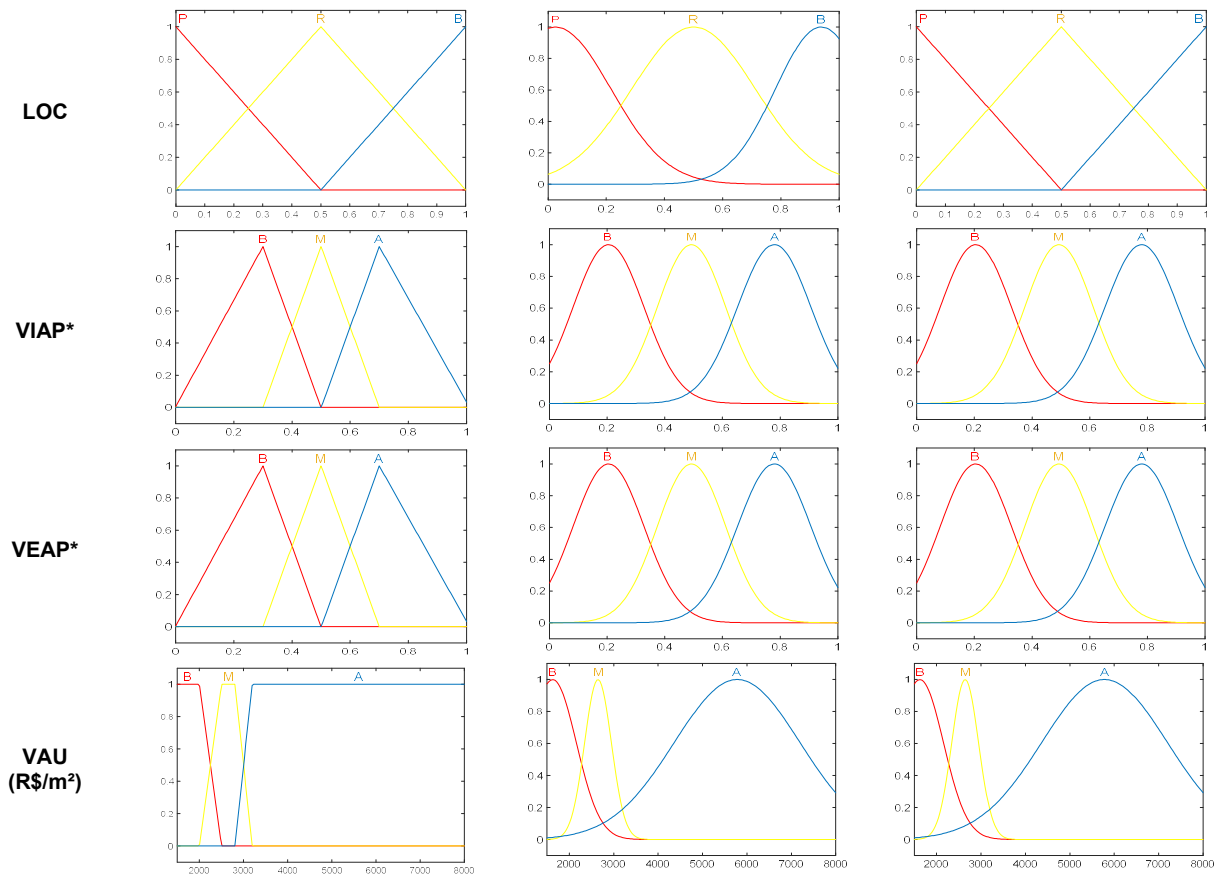
- Valor unitário da amostra do bem imóvel do tipo apartamento (VUA): Expresso em reais por área privativa (R\$/m²), este atributo é baixo “B” quando variar entre R\$ 1.500,0/m² – R\$ 2.000,0/m². O VAU é médio “M” quando variar entre 2.500,0 – R\$ 2.800/m². Entre VUA maiores que R\$ 2.000,0/m² e inferiores a R\$ 2.500,0/ m² é um intervalo nebuloso, não é “B” e nem “M”. Para valores iguais ou superiores a R\$ 3.200,0/ m² é alta “A”. Entretanto, valores maiores que R\$ 2.800/ m² e inferiores a R\$ 3.200,0/ m² é um intervalo nebuloso, não é uma “M” e nem “A”.

Importa referir que, para configuração (B), a “fuzzificação” pela função “*mfgau*”, considerou-se que os referidos intervalos definidos para AP, VG, PC, RP, LOC, VIAP, VEAP e VAU, conforme os termos da Tabela 2, cujos parâmetros σ e x_i ajustaram-se às referidas variáveis quantitativas e qualitativas. As configurações das fuzzificações de todas as variáveis estão identificadas e ilustradas na Tabela 5.

A segunda fase se reverte integralmente em aplicar os métodos de inferência fuzzy de Mamdani (IFTM) e Larsen (IFTL). Os métodos IFTM e IFTL foram processados por regras do tipo “SE”-“ENTÃO”, que se constituem de atributos linguísticos associadas as variáveis.

Tabela 5: “Fuzzificação” das variáveis independentes e dependente.





*Embora integrantes do processo de obtenção do VAU, os termos VIAP e VEAP não são consideradas como variáveis propriamente dita. TIPO (A): VIAP e VEAP a função é “*mfr*”; TIPO (B) e TIPO (C): VIAP e VEAP a função é “*mfgau*”.

Para diminuir o esforço computacional do processamento de regras, as variáveis independentes foram divididas em duas: (i) intrínsecas ao bem imóvel do tipo apartamento (AP, VG e PC); (ii) extrínseca ao apartamento e a edificação (LOC e RP). Assim, condicionou-se que a valoração de apartamentos (VAP) é associado ao valor intrínseco (VIAP) e extrínseco (VEAP), cujas formas de fuzzificações estão na Tabela 5.

- Valor intrínseco (VIAP): Expresso valores entre 0 e 1, o valor intrínseco de um bem imóvel do tipo apartamento é qualificado em grau “B”, “M” e “A”. Neste trabalho em específico, considerou-se que o VIAP é função da AP, VG e PC, cujo processamento é efetuado de maneira dedutiva por 27 regras;

- Valor extrínseco (VEAP): Expresso valores entre 0 e 1, o valor extrínseco de um apartamento é qualificado em grau “B”, “M” e “A”. Neste trabalho considerou-se que o VEAP é função da RP e LOC, cujo processamento é efetuado de maneira dedutiva através de 9 regras.

Deste modo, o VAU (VAP) é função do VIAP e VEAP, cujo processamento é feito de forma dedutiva por 9 regras (Figura 8). Todos as regras foram obtidas pelas propostas de Afshar et al., (2007) e Lisboa (2019). As regras foram estabelecidas pela consulta a 5 especialistas, que preferiram guardar o anonimato, por aplicação de um simples questionário, a fim de conjugar operações “SE” - “ENTÃO” e caracterizar VIAP, VEAP e VAP (Tabela 6, 7 e 8).

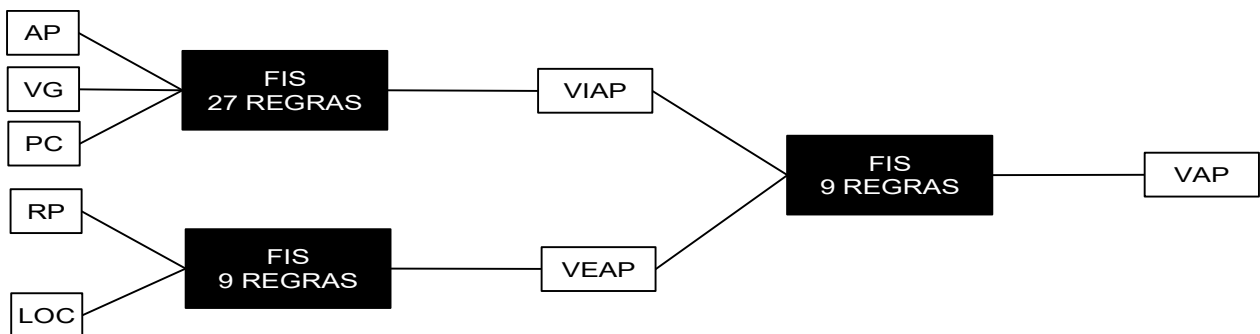


Figura 8: Modelo de Inferência Fuzzy para a valoração de um bem imóvel do tipo apartamento (autoria própria).

Tabela 6: Regramentos para valoração intrínseca do bem imóvel apartamento (VIAP).

REGRAS	“SE”	AP	“E”	VG	“E”	PC	“ENTÃO”	VIAP
1	“SE”	P	“E”	P	“E”	B	“ENTÃO”	B
2	“SE”	P	“E”	P	“E”	M	“ENTÃO”	B
3	“SE”	P	“E”	P	“E”	A	“ENTÃO”	M
4	“SE”	P	“E”	R	“E”	B	“ENTÃO”	B
5	“SE”	P	“E”	R	“E”	M	“ENTÃO”	M
6	“SE”	P	“E”	R	“E”	A	“ENTÃO”	M
7	“SE”	P	“E”	A	“E”	B	“ENTÃO”	B
8	“SE”	P	“E”	A	“E”	M	“ENTÃO”	A
9	“SE”	P	“E”	A	“E”	A	“ENTÃO”	A
10	“SE”	M	“E”	P	“E”	B	“ENTÃO”	B
11	“SE”	M	“E”	P	“E”	M	“ENTÃO”	B
12	“SE”	M	“E”	P	“E”	A	“ENTÃO”	B
13	“SE”	M	“E”	R	“E”	B	“ENTÃO”	B
14	“SE”	M	“E”	R	“E”	M	“ENTÃO”	M
15	“SE”	M	“E”	R	“E”	A	“ENTÃO”	A
16	“SE”	M	“E”	A	“E”	B	“ENTÃO”	M
17	“SE”	M	“E”	A	“E”	M	“ENTÃO”	A
18	“SE”	M	“E”	A	“E”	A	“ENTÃO”	A
19	“SE”	A	“E”	P	“E”	B	“ENTÃO”	B
20	“SE”	A	“E”	P	“E”	M	“ENTÃO”	M
21	“SE”	A	“E”	P	“E”	A	“ENTÃO”	M
22	“SE”	A	“E”	R	“E”	B	“ENTÃO”	M
23	“SE”	A	“E”	R	“E”	M	“ENTÃO”	A
24	“SE”	A	“E”	R	“E”	A	“ENTÃO”	A
25	“SE”	A	“E”	A	“E”	B	“ENTÃO”	M
26	“SE”	A	“E”	A	“E”	M	“ENTÃO”	A
27	“SE”	A	“E”	A	“E”	A	“ENTÃO”	A

Tabela 7 - Regramentos para valoração extrínseca do bem imóvel apartamento (VEAP).

REGRAS	“SE”	RP	“E”	LOC	“ENTÃO”	VEAP
1	“SE”	B	“E”	P	“ENTÃO”	B
2	“SE”	B	“E”	R	“ENTÃO”	B
3	“SE”	B	“E”	B	“ENTÃO”	B
4	“SE”	M	“E”	P	“ENTÃO”	M
5	“SE”	M	“E”	R	“ENTÃO”	M
6	“SE”	M	“E”	B	“ENTÃO”	A
7	“SE”	A	“E”	P	“ENTÃO”	M
8	“SE”	A	“E”	R	“ENTÃO”	A
9	“SE”	A	“E”	B	“ENTÃO”	A

Tabela 8: Regramentos para valoração do bem imóvel apartamento (VAP).

REGRAS	“SE”	VIAP	“E”	VEAP	“ENTÃO”	VAP
1	“SE”	B	“E”	B	“ENTÃO”	B
2	“SE”	B	“E”	M	“ENTÃO”	M
3	“SE”	B	“E”	A	“ENTÃO”	M
4	“SE”	M	“E”	B	“ENTÃO”	B
5	“SE”	M	“E”	M	“ENTÃO”	M
6	“SE”	M	“E”	A	“ENTÃO”	A
7	“SE”	A	“E”	B	“ENTÃO”	B
8	“SE”	A	“E”	M	“ENTÃO”	A
9	“SE”	A	“E”	A	“ENTÃO”	A

Definido as regras, considerando os métodos IFTM e IFTL, associados as “fuzzificação” TIPO (A), (B) e (C), define-se as seguintes tipologias (Figura 9):

(A) – (1): Variáveis “fuzzificadas” por “*mfttrap*” e “*mftri*”, cuja inferência é IFTM;

(B) – (4): Variáveis “fuzzificadas” por “*mfgau*”, cuja inferência é IFTM;

(C) – (7): Variáveis “fuzzificadas” por “*mfgau*” e “*mftri*” e cuja inferência é IFTM;

(A) – (2): Variáveis “fuzzificadas” por “*mfttrap*” e “*mftri*”, cuja inferência é IFTL;

(B) – (5): Variáveis “fuzzificadas” por “*mfgau*”, cuja inferência é IFTL;

(C) – (8): Variáveis “fuzzificadas” por “*mfgau*” e “*mftri*” e cuja inferência é IFTL

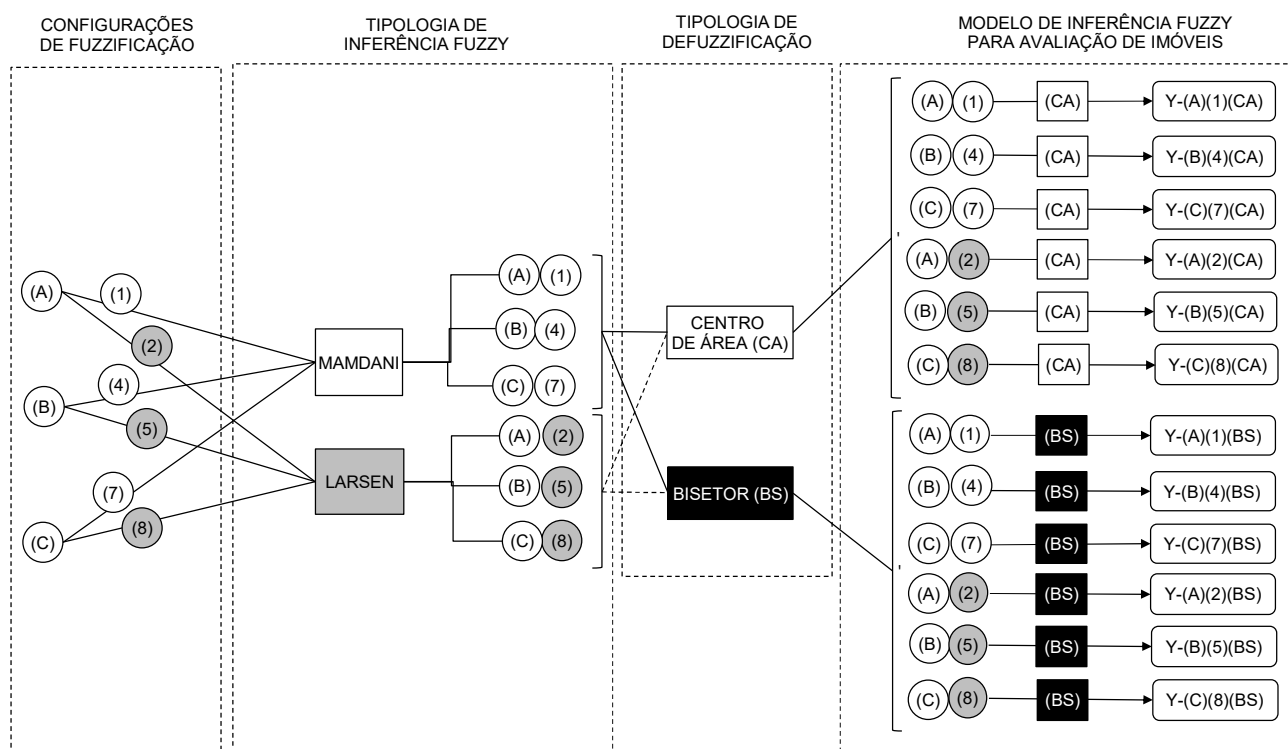


Figura 9: Modelos de inferência fuzzy considerados (autoria própria).

A terceira fase é implementada para os modelos que recorrem aos métodos IFTM e IFTL, posto que se faz necessário processar os métodos de “defuzzificação”. Portanto, definiu-se que este processo será operacionalizado pelos métodos centroide (CA) e bisetor (BS), sendo produzido doze modelos de inferência fuzzy para a avaliação de um bem imóvel do tipo apartamento.

Pela adoção das configurações de “fuzzificação”, e pela aplicação dos métodos de inferência IFTM e IFTL, e “defuzzificação”, foram constituídos doze modelos de inferência fuzzy (Figura 9), os quais sejam:

- F1 - (A) – (1) – (CA): Variáveis “fuzzificadas” por “*mfttrap*” e “*mftri*”, IFTM, “defuzzificação” CA;
- F2 - (B) – (4) – (CA): Variáveis “fuzzificadas” por “*mfgau*”, IFTM, “defuzzificação” CA;
- F3 - (C) – (7) – (CA): Variáveis “fuzzificadas” por “*mfgau*” e “*mftri*”, IFTM, “defuzzificação” CA;
- F4 - (A) – (2) – (CA): Variáveis “fuzzificadas” por “*mfttrap*” e “*mftri*”, IFTL, “defuzzificação” CA;
- F5 - (B) – (5) – (CA): Variáveis “fuzzificadas” por “*mfgau*”, IFTL, “defuzzificação” CA;
- F6 - (C) – (8) – (CA): Variáveis “fuzzificadas” por “*mfgau*” e “*mftri*”, IFTL, “defuzzificação” CA;
- F7 - (A) – (1) – (BS): Variáveis “fuzzificadas” por “*mfttrap*” e “*mftri*”, IFTM, “defuzzificação” BS;
- F8 - (B) – (4) – (BS): Variáveis “fuzzificadas” por “*mfgau*”, IFTM, “defuzzificação” BS;
- F9 - (C) – (7) – (BS): Variáveis “fuzzificadas” por “*mfgau*” e “*mftri*”, IFTM, “defuzzificação” BS;
- F10 - (A) – (2) – (BS): Variáveis “fuzzificadas” por “*mfttrap*” e “*mftri*”, IFTL, “defuzzificação” BS;
- F11 - (B) – (5) – (BS): Variáveis “fuzzificadas” por “*mfgau*”, IFTL, “defuzzificação” BS;
- F12 - (C) – (8) – (BS): Variáveis “fuzzificadas” por “*mfgau*” e “*mftri*”, IFTL, “defuzzificação” BS.

Todo o processo metodológico foi completamente operacionalizado pelo suplemento “*Fuzzy Logic Toolbox*” do aplicativo computacional MatLab® (versão R2020a, MathWorks, Inc., Natick, nMA, USA).

3.3. MODELAGEM POR INFERÊNCIA ESTATÍSTICA

O processo de modelagem por inferência estatística seguiu as etapas descritas pelo subitem 2.2, nomeadamente quanto a seleção das variáveis independentes, predição, estimação dos parâmetros e inferência propriamente dita. Assim, recorreu-se ao software SisDEA®, de modo a verificar: não violação

das hipóteses de Gauss-Markov; o desempenho do modelo obtido pela análise do coeficiente R^2 ; e, a significância dos parâmetros das variáveis independentes (teste-T) e do referido modelo (teste-F).

Considerando o atendimento das hipóteses de Gauss-Markov, a avaliação da homocedasticidade será realizado pela análise gráfica entre os erros padronizados e o valor estimado pelo modelo estatístico. Importa ressaltar que o modelo estatístico obtido é aquele advindo do processo de otimização efetuado pelo algoritmo do software SisDEA®, a partir da transformação automática das variáveis.

A normalidade dos erros é aferida pelos enquadramentos aos percentuais nos intervalos $[-1; +1]$, $[-1,64; + 1,64]$ e $[-1,96; + 1,96]$, respectivamente associados as probabilidades de 68%, 90% e 95%, conforme é referida pela NBR 14.653/2 (ABNT, 2011). Para tanto, recorre-se a regra dos 3σ para o enquadramento dos erros à distribuição normal padrão: 1σ [63% - 73%], $1,64\sigma$ [85% - 95%] e $1,96\sigma$ [95% - 100%]. Caso não haja enquadramento, recorre-se ao teste qui-quadrado para validar esta hipótese.

Para avaliação do cumprimento da hipótese de não-colinearidade, inicialmente, recorre-se a análise da matriz de correlações. Ainda assim, avaliou-se a presença de erros padrões atípicos e influenciadores pela análise gráfica resultante do cálculo da distância de Cook. E, finalmente, todo o processo estatístico inferencial será enquadrado nas especificações normativas, quanto ao grau de fundamento e precisão da NBR 14653/2 (ABNT, 2011).

3.4. DESEMPENHO E DIFERENÇA ENTRE AVALIAÇÃO POR INFERÊNCIA FUZZY E ESTATÍSTICA

A aferição de desempenho dos modelos de inferência fuzzy ($\hat{Y}_{fuzzy(i)}$) desenvolvido para avaliar bens imóveis do tipo apartamento foram realizados pelo cálculo do erro absoluto médio percentual (MAPE) por:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left(\left| \frac{Y_i - \hat{Y}_{fuzzy(i)}}{Y_i} \right| \right) \times 100 \quad (29)$$

Em que Y_i se refere a valor unitário de uma i -amostra de um bem imóvel do tipo apartamento (em que $i = 1, 2, \dots, n$). E, $\hat{Y}_{fuzzy(i)}$ se refere ao valor unitário estimado pelos modelos de inferência fuzzy. Se $\frac{\hat{Y}_{fuzzy(i)}}{Y_i} > 1$ o MAPE indicará desempenho do modelo fuzzy satisfatório. E, se $\frac{\hat{Y}_{fuzzy(i)}}{Y_i} < 1$ o MAPE indicará desempenho fuzzy não satisfatório, sendo estabelecido dois patamares: (i) não aceitabilidade em que $5\% \leq MAPE \leq 10\%$; e, (ii) aceitabilidade em que $MAPE < 5\%$. Deste modo, é possível avaliar quais configurações de modelos fuzzy, no âmbito do processo de fuzzificação, inferência e defuzzificação, apresentaram melhor desempenho para estimar valor de apartamento.

Precede-se ao cálculo do MAPE uma avaliação inicial das distribuições dos modelos fuzzy por gráficos *box-plot* e cálculo do intervalo interquartil – IQR (relação entre 3º e 1º quartil), a fim de diagnosticar a variabilidade das estimativas, sem a influência de “outliers”, em relação aos valores observados.

Deste modo, pode-se prognosticar que, quanto maior for a variabilidade das estimativas dos modelos fuzzy, maior será a agregação de erros em comparação ao valor observado, de modo a comprometer o seu desempenho. Por outro lado, propõem-se avaliar a diferença entre a avaliação de imóveis por modelos de inferência fuzzy e estatística através da seguinte expressão:

$$D(\%) = \left(\frac{\hat{Y}_{fuzzy(i)} - \hat{Y}_{estatístico(i)}}{\hat{Y}_{estatístico(i)}} \right) \times 100 \quad (30)$$

Esta medida afere o aumento (+) ou diminuição (-) do valor de um apartamento por inferência fuzzy em comparação a inferência estatística. Empiricamente, por consulta aos especialistas, adotou-se $D < \pm 10\%$ como minimamente aceitável. Este procedimento foi utilizado para aferir o quanto a inferência fuzzy (simples) mais se afasta ou se aproxima da inferência estatística (complexa).

4. RESULTADOS E DISCUSSÕES

4.1. DESEMPENHO DAS ESTIMATIVAS PELOS MODELOS DE INFERÊNCIA FUZZY

Os modelos de inferência fuzzy proposto neste trabalho, que recorreram aos métodos IFTM e IFTL, estimaram, a partir da configuração da fuzzificação das variáveis independentes AP, VG, PC, RP e LOC, e dos métodos de defuzzificação adotados, os valores de bens imóveis do tipo apartamento.

Pela estimação destes valores, quando comparados aos valores de apartamentos observados no mercado de imóveis, foi possível aferir os desempenhos dos modelos fuzzy. A distribuição por gráficos *box-plot*, apresentou algumas importantes características da estimação por inferência fuzzy (Figura 10).

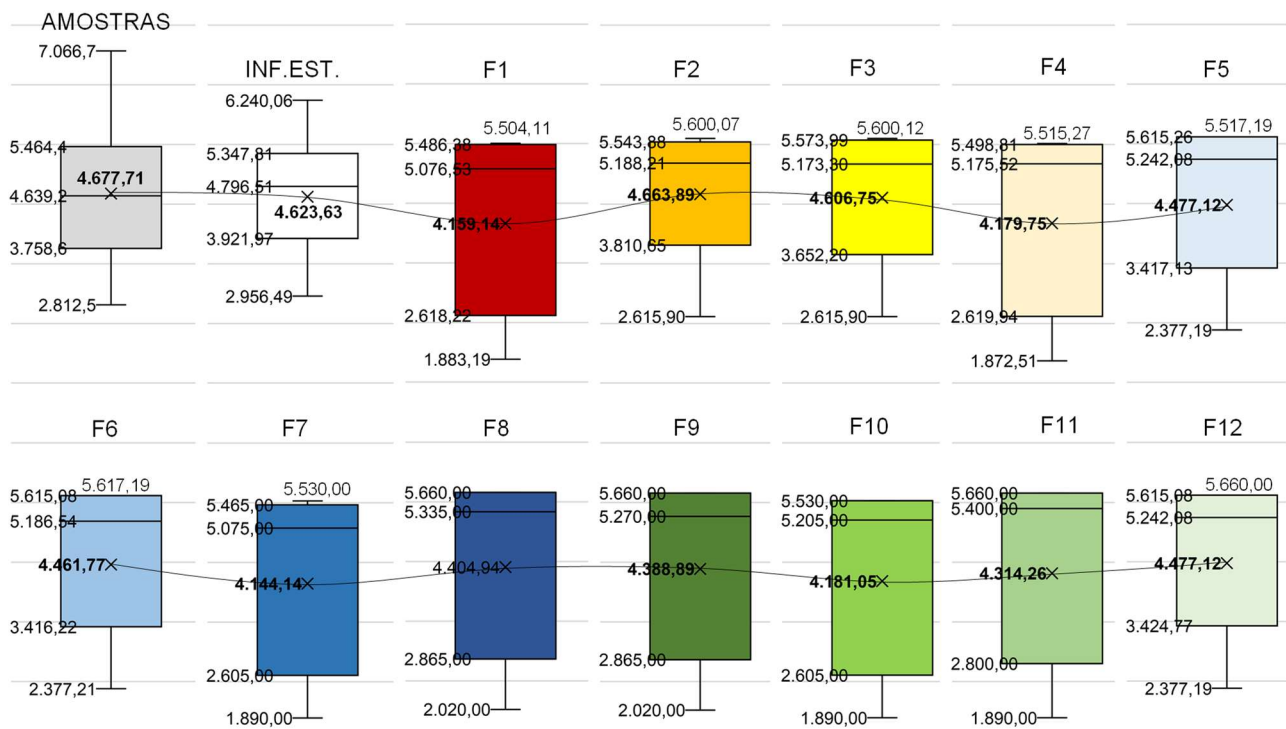


Figura 10: Gráficos box-plot para avaliar a estimação dos modelos fuzzy (autoria própria).

Observou-se que os modelos fuzzy F1, F4, F7, F8, F9, F10, F11 e F12 apresentaram variabilidades significativas e similares entre si, quando comparados a variabilidade da amostra (IQR = R\$ 1.705,88/ m²). O maior IQR entre os referidos modelos fuzzy foi de R\$ 2.925,00/m² (F10) e o menor de R\$ 2.795,00/m² (F8 e F9). Por outro lado, para a estimação do valor de apartamentos na zona urbana da cidade de Belém/PA, os modelos fuzzy F5 e F6 apresentaram alguma variabilidade, respectivamente associado ao cálculo do IQR de R\$ 2.198,13/ m² e R\$ 2.224,86/m², ao passo que a variabilidade dos modelos F2 e F3 foram as que mais se aproximaram da variabilidade dos valores observados no mercado, como verifica-se na Tabela 9.

Tabela 9: Análise da variabilidade e desempenho dos modelos fuzzy em função do IQR e MAPE (%).

-	MÍNIMO	MÁXIMO	MÉDIA	1º Q	2º Q	3º Q	IQR	MAPE (%)
AMOSTRA	2.812,50	7.066,67	4.677,71	3.758,56	4.639,18	5.464,44	1.705,88■	-
INF. EST.	2.956,49	6.240,06	4.623,63	3.921,97	4.796,51	5.347,81	1.425,84■	-
F1	1.883,19	5.504,11	4.159,14	2.618,22	5.076,53	5.486,38	2.868,15■	8,85%▲
F2	2.615,90	5.600,07	4.663,89	3.810,65	5.188,21	5.543,88	1.733,24■	3,41%▼
F3	2.615,90	5.600,12	4.606,75	3.652,20	5.173,30	5.573,99	1.921,79■	1,95%▼
F4	1.872,51	5.515,27	4.179,75	2.619,94	5.175,52	5.498,81	2.878,87■	8,47%▲
F5	2.377,19	5.617,19	4.477,12	3.417,13	5.242,08	5.615,26	2.198,13■	1,68%●
F6	2.377,21	5.617,19	4.461,77	3.391,27	5.186,54	5.616,13	2.224,86■	2,00%●
F7	1.890,00	5.530,00	4.144,14	2.605,00	5.075,00	5.465,00	2.860,00■	9,59%▲
F8	2.020,00	5.660,00	4.404,94	2.865,00	5.335,00	5.660,00	2.795,00■	3,16%●
F9	2.020,00	5.660,00	4.388,89	2.865,00	5.270,00	5.660,00	2.795,00■	3,64%●
F10	1.890,00	5.530,00	4.181,05	2.605,00	5.205,00	5.530,00	2.925,00■	8,96%▲
F11	1.890,00	5.660,00	4.314,26	2.767,50	5.400,00	5.660,00	2.892,50■	5,90%▲
F12	1.890,00	5.660,00	4.314,26	2.767,50	5.400,00	5.660,00	2.892,50■	5,90%▲

(■) Variabilidade dos valores observados; (■) Variabilidade menor que a dos valores observados; (■) Variabilidade igual ou próxima a dos valores observados; (■) Variabilidade maior que a dos valores observados; (▼) O desempenho de modelos fuzzy é satisfatório, pois a relação entre o valor estimado “fuzzy” e o valor observado é menor que 1; (▲) O desempenho de modelos fuzzy é considerado não aceitável, MAPE é maior que 5% e menor ou igual a 10%; (●) O desempenho de modelos fuzzy é aceitável, pois o MAPE é menor que 5%.

Notadamente, o modelo de inferência fuzzy, cuja fuzzificação das variáveis independentes tenham sido caracterizadas por funções de pertinências triangular (*mftri*) e trapezoidal (*mftrap*), caracterizados como

TIPO (A), independentemente do método de inferência fuzzy (Mamdani ou Larsen) e defuzzificação (centróide ou bisetor), agregaram erros, cujo valor do MAPE foi entre 8,47% a 9,59%, nomeadamente os modelos F1, F4, F7 e F10.

Neste contexto, Surgelas et al., (2020) também recorreram a inferência do tipo Mamdani, adotando como variáveis a área privativa, a conservação do apartamento e do prédio e o valor unitário, cujas funções foram *mfgau*, *mftri* e *mfttrap*, respectivamente. O método centróide para defuzzificação foi utilizado, de modo que os autores aplicaram o referido modelo aplicado na cidade de Jelgava (Letônia), cujo MAPE estimado foi em torno de 9%, pelo que tal desempenho foi considerado como um bom resultado.

Entretanto, neste trabalho, modelos fuzzy que agregaram erros $> 5\%$ e $\leq 10\%$ não foram considerados como bons resultados, posto que, não se pode negligenciar a possibilidade de calibração das regras, ajustes dos “*ranges*” da fuzzificação, adoção de diferentes funções de pertinência, de métodos de inferência e defuzzificação para melhorar o desempenho (diminuir erros) dos métodos clássicos de inferência fuzzy.

Importa considerar que neste trabalho foi possível constatar que modelos fuzzy que adotaram apenas as configurações de fuzzificação das variáveis TIPO (A) foram aquelas que agregaram maiores erros para efetuar a valoração de bens imóveis do tipo apartamento.

Por outro lado, as configurações de fuzzificação das variáveis TIPO (B) e (C) foram aquelas em que se obtiveram desempenhos aceitáveis para efetuar a valoração de bens imóveis do tipo apartamento. As configurações TIPO (B) e (C) que recorreram ao método clássico de inferência fuzzy do tipo Mamdani (IFTM) e defuzzificação pelo método centróide agregaram menores erros, nomeadamente os modelos fuzzy F2 e F3, cujos valores do MAPE foram considerados como satisfatórios.

Embora a literatura frequentemente utilize o método IFTM, como inferência fuzzy, e o método centróide para a defuzzificação, outras combinações mostraram excelentes desempenhos. Nestes casos, os modelos fuzzy F5 e F6 apresentaram erro absoluto médio percentual de 1,68% e 2,0%, os quais utilizaram o método clássico de inferência fuzzy do tipo Larsen (IFTL) e bisetor (como defuzzificador), respectivamente.

Os modelos fuzzy F8 e F9 que utilizaram método IFTM e bisetor como método de defuzzificação, cujas configurações de fuzzificação tenha sido do TIPO (B) e (C), apresentaram desempenhos aceitáveis, sendo-lhes aferidos o valor de MAPE em 3,16% e 3,64%, respectivamente.

Ainda considerando as configurações TIPO (B) e (C), embora os percentuais de MAPE até 9% tenha sido considerado como bom resultado por Surgelas et al., (2020), os modelos F11 e F12 que utilizaram IFTL e bisetor apresentaram MAPE de 5,90% para estimativa do valor de apartamentos, em comparação aos valores observados no mercado. Oportuno ressaltar que, neste trabalho apenas se considerou desempenho aceitáveis erros menores ou iguais a 5%.

Deste modo, pode-se concluir que, quanto maior for a variabilidade das estimativas dos modelos fuzzy, em termos do cálculo do IQR, maior será a agregação de erros dos métodos clássicos de inferência fuzzy em comparação o valor observado, comprometendo o seu desempenho em termos de MAPE (Figura 11).

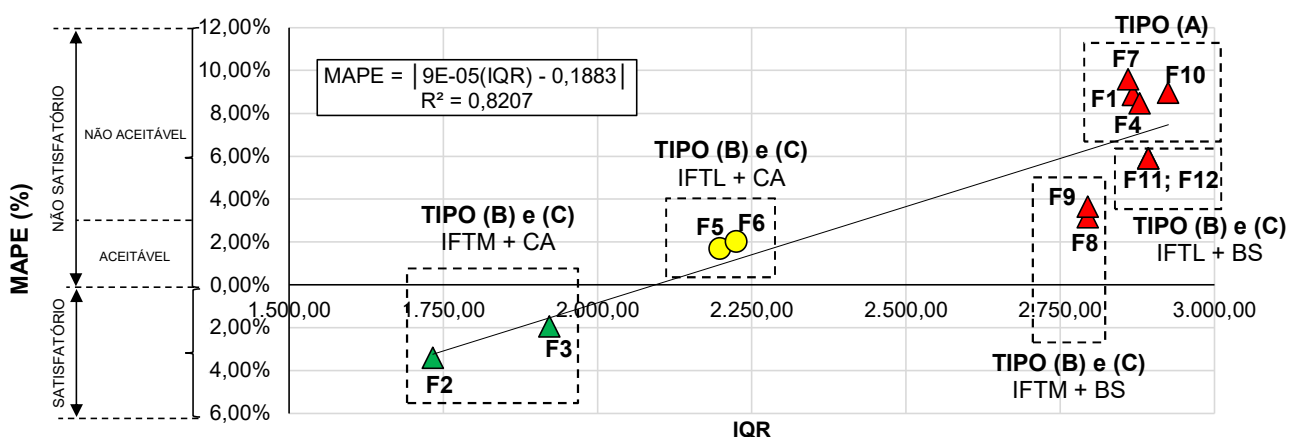


Figura 11: Correlação entre variabilidade (IQR) e os erros (MAPE) (autoria própria).

A correlação entre IQR e MAPE evidenciou uma tendência linear positiva. Assim, a priori, pode-se concluir que os modelos fuzzy que adotarem configurações de fuzzificação do TIPO (B) ou (C), com método IFTM e centróide (CA), apresentaram menor variabilidade e erros em que o MAPE foi estimado por uma relação de $\frac{\hat{Y}_{fuzzy(i)}}{Y_i} > 1$, de modo a apresentar desempenhos satisfatórios.

Por outro lado, considerando esta mesma tendência, fuzzificação do TIPO (B) ou (C), com método IFTL e centróide ou IFTM e bisetor, apresentaram variabilidade e erros aceitáveis. Enquanto modelos que adotarem fuzzificação do TIPO (B) ou (C), com método IFTL e bisetor (BS), ou aqueles que utilizarem apenas configurações de fuzzificação TIPO (A) apresentam com variabilidade significativa e erros não aceitáveis para a estimação de valores de bens imóveis do tipo apartamento.

4.2. DIFERENÇA NA AVALIAÇÃO DE IMÓVEIS POR INFERÊNCIA FUZZY E ESTATÍSTICA

O processo de inferência fuzzy e estatística para a avaliação de bens imóveis do tipo apartamento foram aplicados para valoração de dez unidades localizadas na zona urbana da cidade de Belém/PA, os quais foram identificados e as variáveis independentes consideradas pela Tabela 4.

Neste sentido, ao aplicar a metodologia da estatística inferencial proposta e descrita pelo subitem 2.2 (cuja avaliação dos pressupostos está detalhada no Anexo), obtiveram-se os valores mínimos, médios e máximos (Tabela 10).

Tabela 10: Valoração mínima, média e máxima dos apartamentos pelo modelo de inferência estatística.

ID	MODELAGEM POR INFERÊNCIA ESTATÍSTICA - $\hat{Y}_{estatística(i)}$ (R\$/m ²)									
	1	2	3	4	5	6	7	8	9	10
\hat{Y}_{min}	4.117,63	2.854,67	2.698,82	2.859,43	2.825,39	4.752,21	4.804,86	4.582,69	4.257,82	2.639,45
$\hat{Y}_{méd}$	4.296,05	2.946,51	2.780,43	2.951,73	2.914,62	4.899,22	4.961,97	4.701,92	4.428,46	2.718,59
$\hat{Y}_{máx}$	4.490,63	3.044,46	2.867,13	3.050,18	3.009,68	5.055,61	5.129,69	4.827,52	4.613,35	2.802,63

No geral, os laudos de avaliação pelo método MCDDM, que recorreram a inferência estatística para avaliar os referidos apartamentos, foram enquadrados no grau II de fundamentação e grau III de precisão conforme a NBR 14.653/2 (ABNT, 2011). Aplicando o processo metodológico descrito pelo subitem 2.1, obtiveram-se valores de avaliação dos apartamentos pelos 12 modelos de inferência fuzzy (Tabela 11).

Tabela 11: Valorações dos avaliandos produzidas pelos modelos de inferência fuzzy.

ID	MODELAGEM POR INFERÊNCIA FUZZY - $\hat{Y}_{fuzzy(i)}$ (R\$/m ²)									
	1	2	3	4	5	6	7	8	9	10
F1	5.054,53	1.887,21	1.922,59	1.887,21	1.887,21	4.257,93	4.029,87	4.729,79	4.144,96	2.612,18
F2	5.188,21	2.221,39	2.292,36	2.221,39	2.221,39	4.786,34	4.726,83	5.000,38	4.704,71	3.602,21
F3	5.452,92	2.649,42	2.649,42	2.649,42	2.649,42	4.550,76	4.480,89	4.820,10	5.531,53	3.255,44
F4	5.010,00	1.890,00	1.890,00	1.890,00	1.890,00	2.865,00	2.865,00	3.905,00	3.060,00	2.605,00
F5	5.188,09	2.009,40	2.010,50	2.009,39	2.009,47	4.060,22	3.961,11	4.503,42	3.746,79	3.197,89
F6	5.186,54	2.008,96	2.010,06	2.008,95	2.009,02	4.050,69	3.952,00	4.493,30	3.746,72	3.189,62
F7	5.175,52	1.872,51	1.872,51	1.872,51	1.872,51	4.171,43	4.016,54	4.681,42	3.945,20	2.619,94
F8	5.270,00	1.955,00	2.020,00	1.955,00	1.955,00	3.190,00	3.190,00	4.165,00	3.905,00	2.930,00
F9	5.595,00	2.020,00	2.020,00	2.020,00	2.020,00	2.995,00	2.995,00	3.905,00	5.595,00	2.800,00
F10	5.205,00	1.890,00	1.890,00	1.890,00	1.890,00	2.930,00	2.930,00	3.710,00	2.930,00	2.605,00
F11	5.400,00	1.890,00	1.890,00	1.890,00	1.890,00	2.800,00	2.800,00	2.930,00	2.800,00	2.735,00
F12	5.400,00	1.890,00	1.890,00	1.890,00	1.890,00	2.800,00	2.800,00	2.930,00	2.800,00	2.735,00

A comparação entre valores gerados pela modelo da inferência estatística e os modelos da inferência fuzzy foram expressos pelo cálculo do valor D(%) (equação 30) e organizados considerando três grupos. O primeiro em função das configurações de fuzzificação das variáveis – TIPO (A), (B) e (C). O segundo fez comparações com modelos IFTM e IFTL, associados a inferência fuzzy propriamente dita. O último grupo considerou as comparações apenas relacionados ao método de defuzzificação centróide e bisetor.

O cálculo de D(%) para configurações de fuzzificação TIPO (A) (F1, F4, F7 e F10) apontaram 7 estimações fuzzy com valores inferiores a $\pm 10\%$, 29 subestimções ($D > -10\%$) e 4 superestimções ($D > +10\%$) dos valores dos apartamentos (Tabela 12).

Tabela 12: Diferença entre valoração fuzzy e estatística – D(%).

AVALIANDO – BENS IMÓVEIS DO TIPO APARTAMENTO – CÁLCULO DE D(%)										
	1	2	3	4	5	6	7	8	9	10
F1	22,75(+)	33,89(-)	28,76(-)	34,00(-)	-3,21(-)	10,40(-)	16,13(-)	3,21(-)	-2,65(-)	-1,03(-)
F2	20,77(+)	24,61(-)	17,55(-)	24,74(-)	23,78(-)	-2,30(-)	-4,74(-)	6,35(-)	6,24(-)	32,50(+)
F3	26,93(+)	-10,0(-)	-4,71(-)	10,24(-)	-9,10(-)	-7,11(-)	-9,70(-)	2,51(-)	24,91(+)	19,75(+)
F4	20,47(+)	36,45(-)	32,65(-)	36,56(-)	35,75(-)	14,86(-)	19,05(-)	-0,44(-)	10,91(-)	-3,63(-)
F5	20,76(+)	31,80(-)	27,69(-)	31,92(-)	31,06(-)	17,13(-)	20,17(-)	-4,22(-)	15,39(-)	17,63(+)
F6	20,73(+)	31,82(-)	27,71(-)	31,94(-)	31,07(-)	17,32(-)	20,35(-)	-4,44(-)	15,39(-)	17,33(+)
F7	16,62(+)	35,86(-)	32,02(-)	35,97(-)	35,15(-)	41,52(-)	42,26(-)	16,95(-)	30,90(-)	-4,18(-)
F8	22,67(+)	33,65(-)	27,35(-)	33,77(-)	32,92(-)	34,89(-)	35,71(-)	11,42(-)	11,82(-)	7,78(-)
F9	30,24(+)	31,44(-)	27,35(-)	31,57(-)	30,69(-)	38,87(-)	39,64(-)	16,95(-)	26,34(+)	2,99(-)
F10	21,16(+)	35,86(-)	32,02(-)	35,97(-)	35,15(-)	40,19(-)	40,95(-)	21,10(-)	33,84(-)	-4,18(-)
F11	25,70(+)	35,86(-)	32,02(-)	35,97(-)	35,15(-)	42,85(-)	43,57(-)	37,69(-)	36,77(-)	0,60(-)
F12	25,70(+)	35,86(-)	32,02(-)	35,97(-)	35,15(-)	42,85(-)	43,57(-)	37,69(-)	36,77(-)	0,60(-)
Fuzzificação	Configuração TIPO (A): (+) 4 superestimções; (-) 29 subestimções; (-) 7 diferenças aceitáveis. Configuração TIPO (B): (+) 6 superestimções; (-) 27 subestimções; (-) 7 diferenças aceitáveis. Configuração TIPO (C): (+) 8 superestimções; (-) 23 subestimções; (-) 9 diferenças aceitáveis.									
Inferência	IFTM: (+) 10 superestimções; (-) 34 subestimções; (-) 16 diferenças aceitáveis. IFTL: (+) 08 superestimções; (-) 45 subestimções; (-) 07 diferenças aceitáveis.									
Defuzzificação	Centróide: (+) 11 superestimções; (-) 32 subestimções; (-) 17 diferenças aceitáveis. Bissetor: (+) 07 superestimções; (-) 47 subestimções; (-) 06 diferenças aceitáveis.									

Por outro lado, o valor D(%) considerando configurações TIPO (B) para os modelos fuzzy (F2, F5, F8 e F11), apontaram 7 estimções fuzzy com diferenças aceitáveis à estimção por inferência estatística, enquanto 27 subestimaram e 6 superestimaram.

Neste mesmo contexto, considerando uma configuração fuzzy do TIPO (C) (F3, F6, F9 e F12), 9 foram as diferenças aceitáveis, ao passo que 23 foram as valorações subestimadas e 8 superestimadas em comparação a modelagem estatística para a valoração de apartamentos na cidade de Belém/PA.

O valor D(%) indicaram que, ao utilizar o método clássico de inferência fuzzy do tipo Mamdani (IFTM) (F1, F2, F3, F7, F8 e F9), se passou a ter mais diferenças aceitáveis ($D < \pm 10\%$). Por outro lado, embora esta diferença diminuiu em relação àqueles modelos utilizaram o método clássico de inferência fuzzy do tipo Larsen (IFTL) (F4, F5, F6, F10, F11 e F12), superestimções passaram de 10 (com IFTM) para 8. Quando se recorreu ao método de defuzzificação centróide (F1, F2, F3, F4, F5 e F6), 17 valores por inferência fuzzy passaram a sugerir uma diferença aceitável à valoração dos apartamentos por inferência estatística, em detrimento de 6 valores cujo valor de D(%) foram menores a $\pm 10\%$ quando se utilizou o método bissetor (F7, F8, F9, F10, F11 e F12) (Tabela 12).

Assim, para efeito de inferência, a modelagem fuzzy que recorreu às configurações que menor agregaram erros, quando comparados aos valores observados durante a análise da estimativa referidas no subitem 4.1, nomeadamente modelos F2 e F3 (ver Tabela 9), foram aquelas que obtiveram mais diferenças aceitáveis quando comparadas a inferência estatística (ver Tabela 12).

O efeito da super ou subestimção dos modelos fuzzy pode ser explicada pelo fato de que, as regras que os constituem devem ser construídas por consulta aos especialistas (ou por “*feelings*” científicos) que, por dedução, apresenta-se como o principal componente do modelo, cuja consequência é valor final de um bem imóvel. Este procedimento pode afastar os valores “dedutivos” dos valores “indutivos”.

A este propósito, Duarte et al., (2007) referiram que, ainda que a base de regras seja construída a partir das informações de mercado, a inferência fuzzy não recorre a dados (amostras) de mercado para constituição do modelo, de modo a aferir parâmetros das amostras por estimativas indutivas (como se propõe a inferência estatística), posto que não se trata de uma modelagem de tratamento de dados.

Por outro lado, por concepção, a inferência estatística tem como propósito a minimização de erros (pela aplicação do método MQO para estimativa dos regressores), de modo que, quanto mais o valor estimado pelo modelo se aproximar do valor observado, menores serão os erros agregados. Assim, estimativas de valor efetuadas por inferência estatística são influenciadas pelos dos parâmetros das variáveis independentes.

Em geral, Duarte et al., (2007) também referiram que, pelo fato de algumas variáveis quantitativas estudadas na Engenharia de Avaliações possuírem relações muito complexas entre si, bem como as variáveis qualitativas serem dúbias, vagas, imprecisas; a inferência fuzzy é uma concepção que, por dedução, consegue melhor captar estas relações em detrimento de deduções indutivas (diante da complexidade e do rigor axiomático).

Neste sentido, para realizar a inferência de valores de apartamentos, comparar modelos dedutivos (simples) com modelos indutivos (complexo, diante do rigor axiomático) foi uma consideração referencial. Entretanto, Duarte et al., (2007) ao apresentarem a aplicabilidade da modelagem fuzzy para avaliar imóveis, enfatizaram que a simplicidade é, ou deveria ser, um dos principais objetivos da Engenharia de Avaliações.

Em geral, a modelagem por inferência fuzzy é recomendável para aferir o valor de bens imóveis do tipo apartamentos, por serem mais simples, dedutivos, claros e que, para determinada finalidade, conseguem captar, essencialmente, o pensamento humano que podem também estar relacionados aos dados factuais mercadológicos propriamente ditos.

Finalmente, embora a modelagem fuzzy não esteja descrita e detalhada em norma técnica destinada a avaliação de bens imóveis, este estudo demonstrou que sua aplicação e desempenho, mesmo dependentes das regras, devem priorizar a adoção de funções de pertinência gaussianas para a fuzzificação de variáveis, do método clássico de inferência fuzzy Mamdani ou Larsen e do método centróide para defuzzificação. Assim, estas considerações podem auxiliar a utilização e enquadramento normativo de modelos fuzzy.

5. CONCLUSÕES

Este artigo se propõe analisar desempenhos de modelos fuzzy para valorar apartamentos. Para tanto, se desenvolveu doze modelos fuzzy com diferentes combinações que envolveram as formas da função de pertinência para cada uma das variáveis adotadas (fuzzificação), métodos clássicos de inferência fuzzy e métodos de defuzzificação.

Para comparar aos valores observados no mercado de apartamento situados na zona urbana da cidade de Belém/PA, variabilidade fuzzy teve relação linear com o desempenho da estimação pelos modelos fuzzy. E, quando comparado com resultados da inferência estatística, a diferença foi mais aceitável para modelos fuzzy que apresentaram menores variabilidades e desempenho satisfatório. Assim, com base nestes resultados, as principais conclusões deste estudo foram:

- As funções de pertinências a serem adotadas para o processo de fuzzificação das variáveis devem ser do tipo gaussiana. Entretanto, estas funções devem ser prioritariamente utilizadas para fuzzificação de variáveis quantitativas contínuas. Embora sejam recomendadas à função gaussiana para a fuzzificação das variáveis quantitativas discretas, a combinação com função de pertinência triangular é mais recomendada. A função de pertinência trapezoidal, quando utilizada de forma isolada não é recomendado.

- O método clássico de inferência fuzzy mais recomendado é do tipo Mamdani. Entretanto, a utilização do método de inferência de Larsen apresentou excelente desempenho.

- O método de defuzzificação a ser considerado é o centróide. A utilização do método bissetor não é recomendada.

- A fuzzificação de variáveis por funções de pertinência gaussianas, sendo utilizado método de inferência Mamdani e método centróide para defuzzificação, se apresentou como a estruturação de modelagem fuzzy mais conveniente a ser utilizada para a valoração de bens imóveis do tipo apartamento.

- A fuzzificação de variáveis qualitativas contínuas por funções de pertinência gaussianas, e variáveis quantitativas discretas por funções triangulares, sendo utilizado método de inferência Larsen e método centróide para defuzzificação, se apresentou como a estruturação de modelagem fuzzy alternativa a ser utilizada para a valoração de bens imóveis do tipo apartamento.

Pelas referidas conclusões, pode-se estabelecer graus de Fundamento quanto a adoção das funções de pertinência (“*mfgau*” – grau III; “*mftri*” – grau II; e, “*mfttrap*” – grau I), métodos clássicos de inferência fuzzy Mamdani (grau III), Larsen (grau II); e, defuzzificação pelo método centróide (grau III) e bissetor (grau I).

Importante ressaltar que estas conclusões foram baseadas nos procedimentos adotados que permitiram, a prior, aferir sobre a variabilidade dos modelos fuzzy pelo cálculo do IQR, da agregação de erros pelo cálculo do MAPE e pelo número de diferença aceitável entre a inferência fuzzy e estatística. Ainda assim, convém considerar que estas conclusões satisfazem a configuração de regras construídas pela consulta aos especialistas, bem como pela estruturação da modelagem fuzzy proposta para valoração de bens imóveis do tipo apartamento desenvolvidas neste artigo.

Assim, quaisquer que sejam as mudanças na configuração adotado pelos modelos fuzzy de melhor desempenho, podem comprometer e influenciar a valoração de um determinado apartamento. Finalmente, pode-se concluir que este estudo pode auxiliar na utilização e enquadramento normativo de modelos fuzzy.

A este propósito, pode-se considerar efetuar o enquadramento em graus de precisão de modelos fuzzy quando, comparados aos valores observados no mercado, relativamente associados ao cálculo do MAPE $< 5\%$ (grau III), $5\% < \text{MAPE} \leq 10\%$ (grau II); e, $\text{MAPE} > 10\%$ (grau I).

Neste sentido, é recomendável a adoção para trabalhos futuros:

- A adoção de outras variáveis independentes que possam ser integrantes a valoração intrínseca de apartamentos, tais como: andar, disponibilidade de salão de festa, depreciação física, etc, as quais esteja relacionada a atributos físicos das unidades.
- A adoção de outras variáveis independentes que possam ser integrantes a valoração extrínseca de apartamentos, tais como: proximidade a supermercados, shopping, amenidades e desamenidades, etc, as quais esteja relacionada a atributos circundantes (ou circunvizinhas) às unidades.

Para trabalhos futuros, pode-se recorrer ao desenvolvimento de metodologias que sejam considerados estudos comparativos entre os métodos de interpolação de inferência fuzzy (do tipo Tsukamoto e Takagi-Sugeno), com métodos neuro-fuzzy (aditados do método de interpolação de inferência fuzzy e redes neurais artificiais), com métodos clássicos de inferência fuzzy (do tipo Mamdani e Larsen), com inferência estatística, com envoltória de dados e com redes neurais artificiais.

REFERÊNCIAS

ABNT (Associação Brasileira de Normas Técnicas). **NBR 14.653-2: Avaliação de bens – Parte 2:** imóveis urbanos. Rio de Janeiro: ABNT, 2011.

AFSHAR, A.; MARINO, M. A.; EBTEHAJ, M.; MOOSAVI, J.. **Rule-based fuzzy system for assessing groundwater vulnerability.** Journal of Environmental Engineering, ASCE, v. 133, n. 5, p. 532-540. doi: 10.1061/(ASCE)0733-9372(2007)133:5(532), 2007.

ALMEIDA, R.N. **O Método dos Mínimos Quadrados: Estudo e Aplicações para o Ensino Médio.** Dissertação (Mestrado em Matemática) –Centro de Ciências e Tecnologia da Universidade Estadual Norte Fluminense Darcy Ribeiro, UENF, Campos dos Goytacazes -RJ. 2015.

BAUER, P., NOAUK, S; WINKLER, R. **A brief course in fuzzy logic and fuzzy control.** Courses offered by Department of knowledge-Based Mathematical Systems, Johannes Kepler University Linz, 1998. Disponível em: <http://www.esru.strath.ac.uk/Reference/concepts/fuzzy/fuzzy.htm>. Acesso em: 23 de set. de 2022.

DANTAS, R. A. **Engenharia de Avaliações: uma introdução à metodologia científica.** São Paulo: Pini, 2005.

DRIANKOV, D; HELLENDORRN. H; REINFRANK, M. **An Introduction to Fuzzy Control.** Springer-Verlag, 1993. Disponível em: <http://www.springer.com/gp/book/9783540606918>. Acesso em: 13 de ago. 2022.

DUARTE, A.A.A.M; ROTTMANN, E.; DUARTE, A.R.C. DE L. M.; COSTA JR., C.T., KAWASHIMA, M. E GALVÃO NETO, O. **Um novo enfoque na engenharia de avaliações: o método do entendimento racional.** In: COBREAP, XIV. 2007, Salvador/BA. Disponível em <<http://www.mrci.com.br>>. Acesso em: 20 de de set. 2022, 21 p.

GAZOLA, S. **Construção de Um Modelo de Regressão Para Avaliação de Imóveis.** Florianópolis, 2002.Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2002.

GOMIDE, F. A. C, GUDWIN R. R; TANSCHHEIT. R. **Conceitos fundamentais da teoria de conjuntos fuzzy, lógica fuzzy e aplicações.** Proceedings of 6th International Fuzzy Systems Association World Congress - IFSA95, Tutorials, pp. 1-38, 1995.

GONZÁLEZ, M. A. S. **Aplicação de Técnicas de Descobrimeto de Conhecimento em Bases de Dados e de Inteligência Artificial em Avaliação de Imóveis.** Tese Doutorado em Engenharia Civil –Programa de Pós-Graduação em Engenharia Civil, UFRGS, Porto Alegre -RS. 2002.

GRIMA, M.A., VERHOEF, P.N.W.. **Forecasting rock trencher performance using fuzzy logic.** International Journal of Rock Mechanics and Mining Sciences. Vol. 36, 413–432. doi: 10.1016/S0148-9062(99)00025-X, 1999.

GUJARATI, D. N.; PORTER, D. C. **Econometria básica.** 5. ed. Porto Alegre: AMGH, 2011. 924 p

HOLMBLAD, I.P; OSTERGAARD, J.J. **Fuzzy Logic Control: Operator Experience applied in Automatic Process Control.** FLS Review, F.L.Smidth & Co., Copenhagen, Denmark, n. 5: 11-16, 1981;

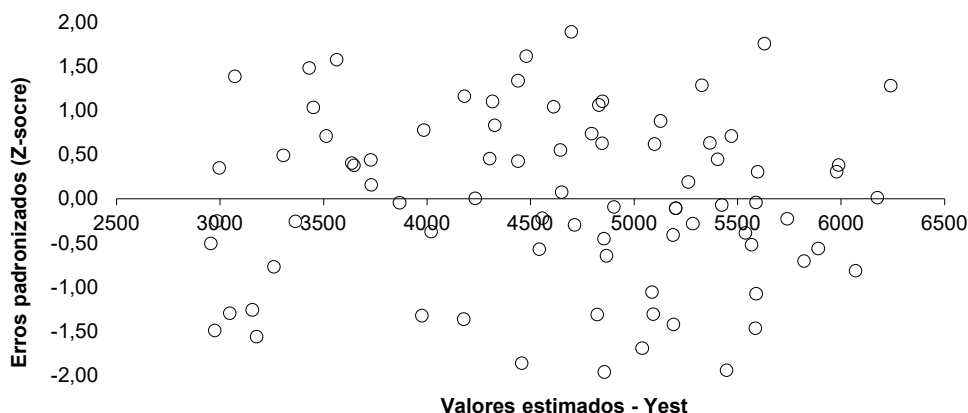
LARSEN, P. M. **Industrial applications offuzzy logic control. Fuzzy Reasoning and its Aplications.** London: Academic Press Inc., 335-342. doi: 10.1016/S0020-7373(80)80050-2, 1981.

LEE, C. C. **Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part I.** IEEE Transactions on Systems, Man, and Cybernetics, Vol. 20, No. 2, March/April, pp. 58404 – 418, 1990a. 5960. doi:10.1109/21.52551, 1990a.

- LEE, C. C. **Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part II.** IEEE Transactions on Systems, Man, and Cybernetics, Vol. 20, No. 2, March/April 1990, pp. 634-19 – 435. doi:10.1109/21.52551, 1990b.
- LISBOA, E.G. **Contribuições ao processo de análise de riscos de contaminação das águas subterrâneas: proposta fuzzy-probabilística.** (Tese de Doutorado em Engenharia Civil) – Faculdade de Engenharia da Universidade do Porto, 2019. Disponível em: <https://hdl.handle.net/10216/120610>. Acessada em: 18 set. 2022.
- MAMDANI, E.H.. **Application of Fuzzy Algorithms for Control of Simple Dynamic Plant.** Proceedings of the IEEE (Control and Science), V. 121: 298-316. doi: 10.1049/piee.1974.0328, 1974.
- MELLO, R; SANTELLO, R. **O Uso de Lógica Fuzzy para Avaliação de Imóveis e Exemplos de Aplicação.** Revista Produção Online, Florianópolis, 2012.
- MIZUMOTO, M.. **Fuzzy Inference using max-min composition in the Compositional Rule of Inference.** In: Approximate Reasoning in Decision Analysis, M.M. Gupta, E. Sanchez (Ed.), North-Holland: 67-76, 1982.
- MONTGOMERY, D.C., PECK, E.A., VINING, G.G. **Introduction to Linear Regression Analysis.** 4th Edition, John Wiley & Sons, Inc., Hoboken, 2006.
- PEDRYCZ, W. **Fuzzy Control and Fuzzy Systems.** - John Wiley and Sons Inc, NY, 1989. Disponível em: <https://dl.acm.org/citation.cfm?id=529624>. Acesso em: 21 de set. de 2022.
- RESCHER, N. **Many-valued logic,** McGraw-Hill, New York. 359 p, 1, 1969.
- ROSS, T. J. **Fuzzy Logic with Engineering applications.** 2nd Edition. England: Wiley, 2004.
- SANDRI, S.; CORREA, C. **Lógica nebulosa.** In: V Escola de Redes Neurais: Conselho Nacional de Redes Neurais, São José dos Campos. Anais. São José dos Campos: ITA. p.73-90, 1999. Disponível em: <http://www.ele.ita.br/cnrm/minicursos-5ern/log-neb.pdf>. Acesso em: 18 de jul. de 2014.
- SIMÕES, M. G; SHAW, I. S., **Controle e Modelagem Fuzzy.** São Paulo, Blucher, 2007
- SUGENO, M; TAKAGI, T. **Fuzzy Identification of Systems and its Applications to Modelling and Control.** IEEE Transactions on Systems, Man & Cybernetics, Vol. 15: 116-132. doi: 10.1109/TSMC.1985.6313399, 1985.
- SURGELAS V.; ARHIPOVA I; PUKITE V. **Civil Engineering Inspection for Real Estate Evaluation with the Use of Artificial Learning Algorithms and Fuzzy Logic.** Rural Sustainability Research, Vol.43, (Issue 338), pp. 75-82, 2020. <https://doi.org/10.2478/plua-2020-0010>.
- TAMAKI, F; KANAGAWA, A; OHTA, H. **Identification of membership functions based on fuzzy observation data.** J. Fuzzy Sets and Systems, v.93 (1998) 311-318. doi: 10.1016/S0165-0114(96)00204-7, 1998.
- TSUKAMOTO, Y. **An approach to fuzzy reasoning method, in Advances in Fuzzy Set Theory and Applications** (M. M. Gupta et al., Eds.), North-Holland, 137-149 p, 1979.
- ZADEH, L. **Fuzzy Sets: Information and Control,** vol. 8, pp 338-353. doi:10.1016/S0019-9958(65)90241-X, 1965.

ANEXO - I

- O software SisDEA®, por processo de otimização, selecionou 81 modelos estatísticos, a partir da transformação (1/x; Ln(x)) das variáveis.
- A avaliação da homocedasticidade: análise gráfica entre os erros padronizados e o valor estimado.



- A normalidade dos erros foi associada a probabilidades de 61%, 92% e 98%. A regra dos 3σ permitiu validar parcialmente esta hipótese.
- Pela análise da matriz de correlações não houve indicações colinearidade x1 (AP), x2 (VG), x3 (PC), x4 (RP) e x5 (LOC):

Transf.	Alias	x1	x2	x3	x4	x5
ln(x)	x1	0	0,59	-0,52	0,35	-0,17
x	x2	0,59	0	-0,37	0,26	-0,01
1/x	x3	-0,52	-0,37	0	-0,50	0,10
ln(x)	x4	0,35	0,26	-0,50	0	-0,04
x	x5	-0,17	-0,01	0,10	-0,04	0

- Não se detectou presença de erros padrões atípicos e influenciantes.
- O coeficiente R²: 0,9214 (regressão); 0,8819 (estimativa)/ coeficiente R: 0,849 (na regressão); 0,7778 (estimativa); Desvio padrão: 2,16379E-05 (regressão)/ 523,4744086 (estimativa).
- Análise de variância, teste-T, teste-F:

Variável	Média	Mínimo	Máximo	Coefficiente	teste-T	Sig(%)	transf
AP	4,64	3,82	6,11	5,81E-05	7,92	0,01	ln(x)
VG	1,64	0,00	4,00	-2,74E-05	-6,61	0,01	x
PC	0,48	0,25	1,00	1,70E-04	14,46	0,01	1/x
RP	8,55	7,48	9,31	-2,40E-05	-3,80	0,03	ln(x)
LOC	0,64	0,00	1,00	-1,46E-05	-2,86	0,55	x
VAU	0,00	0,00	0,00	1,33E-04	2,12	3,75	1/y

Fonte de Variação	Soma dos Quadrados	Graus de Liberdade	Quadrado Médio	teste-F
Explicada	1,97667E-07	5	3,95E-08	84,4375189
Não explicada	3,51148E-08	75	4,68E-10	
Total	2,32782E-07	80		

- Modelo por inferência estatística:

$$\hat{Y}_{estatística} = 1/(0.0001332724773 + 5.807987963E-005\ln(AP) - 2.738576761E-005(VG) + 0.000169840125/PC - 2.396764788E-005 \ln(RP) - 1.463439836E-005(LOC)$$
- Correlação entre valor observado (Y_i) e valor estimado (Ŷ_i) foi de R² de 0,7822.