

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

SUZAN: Modelo Fuzzy para Classificação Espectral de Dados de Supernovas Ia

Luís Ricardo Arantes Filho ¹

Instituto Nacional de Pesquisas Espaciais

Lamartine Nogueira Frutuoso Guimarães ²

Instituto de Estudos Avançados

Reinaldo Roberto Rosa³

Instituto Nacional de Pesquisas Espaciais

Francisca Joamila Brito do Nascimento⁴

Instituto Nacional de Pesquisas Espaciais

Resumo. Supernovas são cataclismos cósmicos. Estas explosões representam o ciclo final de evolução das estrelas. A classificação do fenômeno é complexa e está condicionada à perícia de um astrônomo que identifica as peculiaridades do espectro óptico e das linhas espectrais. Desta forma, este trabalho apresenta SUZAN (Sistema fUZZy Avaliador de superNOvas) um método de classificação de supernovas utilizando o paradigma da Lógica Fuzzy, simulando o especialista humano no momento da análise espectral e da classificação espectral. SUZAN avalia um grande número de espectros obtendo seu melhor desempenho para supernovas Ia em sua fase de luminosidade máxima.

Palavras-chave. Supernovas, Fuzzy, classificação espectral, supernovas Ia.

1 Introdução

Supernovas (SN) são caracterizadas por explosões extremas no cosmos. Estas explosões derivam-se do ciclo final de evolução de uma estrela massiva ou de reações termonucleares em estrelas anãs brancas e podem gerar quantidades de energia e luminosidade equiparadas ao brilho total de uma galáxia. A importância da classificação deste fenômeno está na constatação das explosões de Supernovas do tipo Ia (SNIa) como sendo uma classe de explosões com luminosidade máxima pouco variável de objeto a objeto, o que as torna régua padronizáveis para medidas de distâncias cosmológicas. Este importante passo científico rendeu o prêmio Nobel de Física de 2011 aos astrofísicos Saul Perlmutter, Brian P. Schmidt e Adam G. Riess [9, 10] que constataram a expansão acelerada do universo através da observação de supernovas distantes.

¹luisricardoengcomp@gmail.com

²lamar.guima@gmail.com

³rrrosa.inpe@gmail.com

⁴francisca.nascimento@inpe.br

Esta importante descoberta foi derivada da análise das curvas de luz (magnitude em função do tempo) das SNIa, esta análise só é possível num período de 40 a 60 dias após a explosão de cada SN. Em contrapartida, é possível identificar as SNIa pelo espectro (fluxo de radiação em função do comprimento de onda), que é observado assim que determinado o alvo pelos telescópios e espectrógrafos [3,4,8]. A análise das linhas espectrais de emissão e absorção (picos e vales) permite identificar a composição química do objeto, as SNIa são ricas em Silício (Si) e Enxofre (S) e não apresentam traços de Hidrogênio (H) ou Hélio (He) [1, 2, 4]. Neste artigo descrevemos um sistema baseado em Lógica Fuzzy modelado pela inferência de Takagi-Sugeno [11] para classificação automática de um grande volume de dados espectrais de SNIa. Os dados utilizados nesta pesquisa pertencem ao acervo *The CfA Supernova Archive* [2,5] que contém espectros de diversos tipos de SN, principalmente SNIa. Apresentamos o sistema SUZAN (Sistema fUZzy Avaliador de superNovas) que executa a análise espectral de SN de forma semelhante ao astrônomo especialista, sendo capaz de avaliar e classificar as SNIa com maior precisão quando o espectro está no período próximo ao brilho máximo da explosão.

A justificativa para a adoção do paradigma da lógica fuzzy está no desenvolvimento de um sistema capaz de simular o conhecimento especialista sem a necessidade de um treinamento de dados prévio para identificação de padrões, utilizando portanto, toda a amostra de dados na validação do modelo. O desenvolvimento da SUZAN foi pensado como uma alternativa ao CIntIa (Classificador Inteligente de supernovas do Tipo Ia) proposto por [8] e posteriormente aprimorado por [3] (CINTIA2) que utiliza o conceito de redes neurais artificiais e por isso, necessita de uma grande quantidade de exemplos para relacionar corretamente os padrões de SN. Estes sistemas necessitam que os modelos sejam treinados com uma proporção de 80% dos dados originais, validando os resultados de classificação em 20% dos dados. O sistema SUZAN é estabelecido em um cenário em que a análise de SN demanda agilidade, e é por isso que a avaliação espectral é construída já que os espectros são um dos primeiros tipos de dados em que é possível realizar uma análise de componentes químicos em SN.

2 Modelagem Fuzzy no sistema de Inferência de Takagi-Sugeno

O modelo Fuzzy permite a graduação na pertinência de elementos à determinada classe. A função que define o grau de pertencimento de um determinado elemento em um conjunto fuzzy é denominada função de pertencimento e é definida segundo o critério de um especialista ou empiricamente. Em relação à classificação quando são utilizadas várias entradas de dados (padrões), muitas funções de pertencimento devem ser utilizadas para melhor representar as regras que definem se um padrão pertence a determinado conjunto ou não [11]. O modelo de Inferência de Takagi-Sugeno consiste em um tipo de inferência capaz de descrever funções de forma aproximada e é muito utilizado em aplicações de controle. As etapas para o desenvolvimento do modelo são: a seleção de entradas; a modelagem das regras e a escolha das funções de pertencimento; a ativação das regras; e a inferência. Desta forma, este modelo permite que se tenha uma relação linear de entrada

e saída (input-output), As funções de associação de saída neste modelo, são lineares ou constantes [11]. As regras deste modelo são descritas de acordo com a definição (2.1).

Definição 2.1. *Se k_1 é B_1 e k_2 é $B_2 \dots k_n$ é B_n , Então $y = f_i(k_1, k_2, \dots, k_n)$.*

Em que B_1, \dots, B_n são os conjuntos fuzzy dos antecedentes enquanto que o consequente é uma função das variáveis de entrada k_1, k_2, \dots, k_n e n é o número de regras que são ativadas. As funções representam os polinômios gerados pela inferência do modelo vista em (1). A saída gerada por este sistema de inferência é polinomial [11].

$$y = \frac{\sum_{i=1}^n w_i f_i(k_1, k_2, \dots, k_n)}{\sum_{i=1}^n w_i}. \quad (1)$$

A ativação das regras é dada por um operador de implicação e é definido pela variável w_i . Este operador é definido para interpretar as informações das regras se-então interpretando os conectivos E, OU e CASO CONTRÁRIO. O operador de implicação é definido em (2.2).

Definição 2.2. $\phi[\mu_A(x), \mu_B(y)] \equiv \phi[\mu_A(x) \wedge \mu_B(y)]$.

No modelo de Takagi-Sugeno de ordem zero as regras são definidas conforme (2.3).

Definição 2.3. *Se k_1 é B_1 e k_2 é $B_2 \dots k_n$ é B_n , Então $y = S$.*

Tem-se um modelo de Takagi-Sugeno de ordem zero se o polinômio é uma constante. Neste ponto, S é definida como um consequente, ou seja, uma variável de saída que é dependente das variáveis de entrada k_n . Esta variável assume valores constantes e pode ser representada também por termos linguísticos e seus respectivos valores. A saída gerada é um polinômio de ordem zero. Quando a saída é uma constante as funções de associação são definidas como *singletons* [11], e os métodos de implicação e agregação das regras são fixos. A equação (2) descreve a inferência para o modelo de Takagi-Sugeno de ordem zero.

$$y = \frac{\sum_{i=1}^n w_i S_i}{\sum_{i=1}^n w_i}. \quad (2)$$

O método de inferência de Takagi-Sugeno de ordem zero aplicado neste trabalho foi escolhido pela forma como as saídas são geradas na avaliação fuzzy. Este modelo permite que o custo computacional para realizar as operações seja reduzido, pois os consequentes são polinômios de grau zero.

2.1 Modelagem da SUZAN

A modelagem fuzzy do sistema SUZAN ocorre com a correta identificação das linhas espectrais do espectro de SN conforme a normalização de dados feita em [3]. Esta normalização compreende uma filtragem dos dados espectrais e o ajuste do *redshift* para avaliar as linhas espectrais do objeto em repouso. Este ajuste é feito para que as linhas espectrais de elementos de supernovas estejam ajustadas o mais próximo das linhas espectrais de elementos que são obtidas no laboratório. A partir da correta localização das linhas

espectrais de SN são abstraídos dois parâmetros de características: as larguras (W) e intensidades de pico (p_k), como visto na Fig. 1. Na Fig.1 os valores de F_λ e F_c correspondem respectivamente aos valores de fluxo do espectro da linha de absorção e do envoltório para delimitação da área da linha espectral. As variáveis λ_1 e λ_2 correspondem aos limites da linha espectral e recebem os valores do comprimento de onda. Ainda na Fig. 1, h é a altura baseada nos fluxos do espectro, ou seja, $h = (F_c - F_\lambda)$. Desta forma é possível calcular os valores de W e p_k , o calculo é feito de forma relativa para as linhas de emissão.

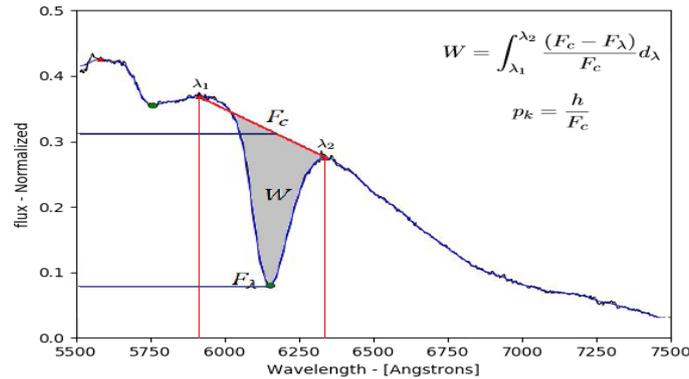


Figura 1: Cálculo das características W e p_k relacionadas às linhas espectrais.

O algoritmo fuzzy (regras fuzzy) define se uma linha espectral tem características estreitas, medianas ou largas. Quanto maior for o valor de W e mais intenso for p_k então a linha se enquadra como sendo "larga", gradativamente quando os valores são pequenos as características assumem o valor de mediana, estreita, declínio e nulas (valores baixos de p_k e W). As linhas avaliadas pela SUZAN como largas, medianas e estreitas são utilizadas na análise de elementos químicos. Assim são definidas duas outras variáveis fuzzy (*Distância* e *Elementos fuzzy*) para avaliar e definir os elementos Si, S, He e H para cada linha espectral selecionada anteriormente. A Fig.2 (a,b e c) ilustra como são modelados os parâmetros W , p_k e de *Distância* para funções de pertencimento. A modelagem Fuzzy (construção de funções de pertencimento, termos linguísticos e regras fuzzy) seguiu um critério para definir o quanto uma linha espectral é "larga" e com "grande intensidade" e o quanto uma linha é "estreita" e com "baixa intensidade". Para definir o universo de discurso de cada função foi definida a largura e a intensidade média de cada linha nas regiões do Si, S, H e He. A Fig. 2 (d) mostra todos os espectros da base de dados de SNIa e o padrão médio (linha azul) em diferentes períodos (fases espectrais definidas em relação ao brilho máximo da explosão de SN, por exemplo, a definição Fase: +4 refere-se a 4 dias após o pico da explosão de SN, Fase: -4 refere-se a 4 dias antes do pico da explosão).

A variável *Distância* define a proximidade (comprimento de onda) da linha selecionada para o valor de comprimento de onda dos elementos que caracterizam as SN. Esta variável foi criada, pois mesmo com o ajuste de *redshift* os valores de elementos dos espectros são imprecisos em relação aos valores de elementos de laboratório. A variável *Elementos fuzzy* foi definida como uma região de comprimento de onda que representa um dos elementos que aparecem nos espectros de SN (Si, S, He e H), desta forma quanto mais próxima a

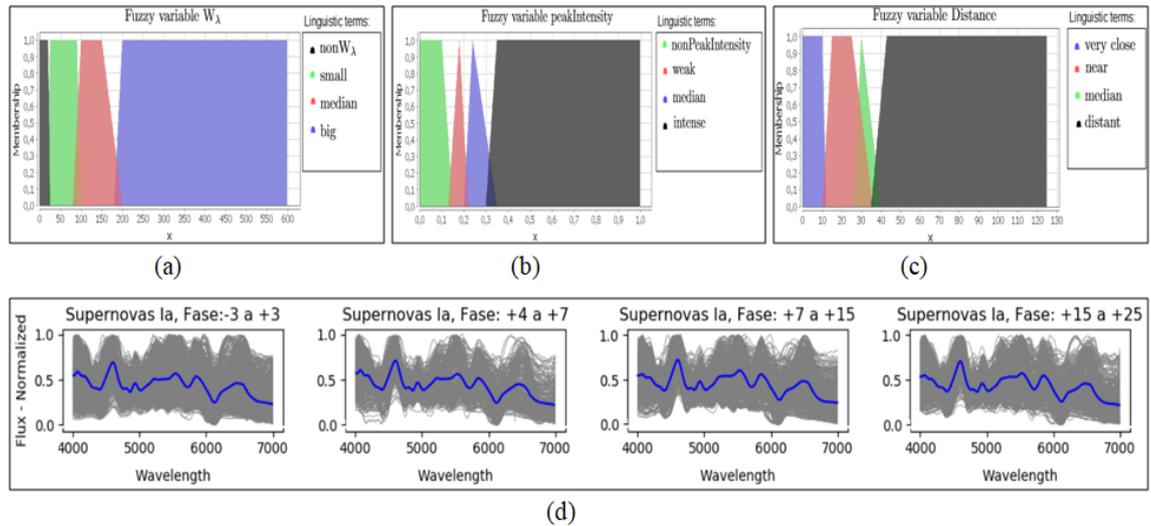


Figura 2: Modelagem das Funções de Pertencimento. (a) Corresponde a modelagem da variável fuzzy W , (b) corresponde a modelagem da variável fuzzy p_k e (c) corresponde à variável fuzzy de *Distância*. Em (d) é ilustrado o padrão médio dos espectros de SNIa em diferentes fases de observação, que foi utilizado como critério para modelagem das variáveis fuzzy.

linha avaliada está desta região maior é o grau de pertencimento a este elemento. A Fig. 3 indica como estes elementos definem a classificação dos tipos clássicos de SN. Quando SUZAN identifica a presença Si e o S no espectro e a ausência do H e do He, **então** o espectro é classificado como **SNIa**.

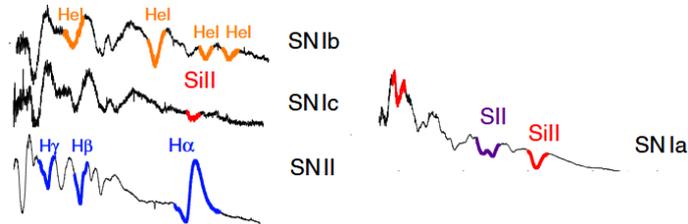


Figura 3: Esquema de classificação de elementos em SN [7].

3 Resultados

Foram avaliados os dados de SNIa do acervo The CfA Supernova Archive [5]. A Tab. 1 descreve os dados brutos que foram submetidos a análise da SUZAN, principais referências e características temporais dos espectros, bem como, o desempenho da classificação para cada base.

A análise espectral de SNIa pode ser feita com precisão no intervalo de 5 dias em relação ao brilho máximo da explosão, ou seja, quando a curva de luz da SN atinge seu

Tabela 1: Dados espectrais de SN e desempenho da SUZAN na classificação das SNIa.

Principal referência	Tipo de SN	Qtde. Espectros	Fase Espectral (dias)	Acertos(%)
[1]	SNIa	47	-2.5 a +2.5	100
[2]	SNIa	2603	-15 a +611	71,92
[6]	SNIa	432	-13 a +243	76,15
Total	-	3082	-	-

ponto máximo são contados 2.5 dias antes do máximo e 2.5 dias depois do máximo (-2.5 dias a +2.5 dias). Neste período as características do Si e do S são mais evidentes, como apontam [1,4] e é neste período em que os astrônomos especialistas executam a classificação de maneira visual. Desta forma, para verificar a eficiência da SUZAN em avaliar as SNIa apenas neste período foram correlacionados espectros que mantinham suas características principais de elementos dentro do período de tempo de -2.5 a +2.5 dias conforme a Tab. 2.

Tabela 2: Avaliação da SUZAN para SNIa próximas ao brilho máximo.

Principal referência	Tipo de SN	Qtde. Espectros	Qtde de SN	Acertos(%)
[1]	SNIa	47	26	100
[2]	SNIa	634	462	96,21
[6]	SNIa	91	32	97,80
Total	-	772	520	95,85

4 Considerações sobre a Abordagem

Neste trabalho destaca-se a análise de SNIa no período em que as SNIa atingem seu brilho máximo, identificando aproximadamente 96% (conforme a Tab. 2) dos espectros de SNIa nesta fase. SUZAN possui uma imprecisão na classificação de SNIa quando as fases espectrais estão muito distantes da fase de brilho máximo da SN (conforme a Tab. 1), nestes longos períodos as características morfológicas dos espectros variam, pois é visualizado o espectro da nuvem de gás da explosão da estrela que se dissipa no decorrer do tempo. Estas variações nas características dos espectros indicam a presença de outros elementos, como o Ferro (Fe), tornando a classificação de SNIa ainda mais complexa. As funções de pertencimento e regras fuzzy da SUZAN foram modeladas para SNIa próximas ao período de -2.5 a +2.5 dias em relação à luminosidade máxima, pois é neste período em que o astrônomo observa as características dos elementos(Si, S, H, He) e executa a classificação. Apresentamos um modelo fuzzy capaz de modelar o conhecimento de um astrônomo e executar de maneira automática a classificação espectral de SNIa. A correta classificação e identificação de SNIa por sistemas automáticos propicia contribuição referente a estudos cosmológicos que avaliam a expansão acelerada do universo, haja visto que sistemas classificadores de SN como a SUZAN, proporcionam uma análise precisa sobre

um grande volume de dados em um curto período de tempo algo difícil para especialistas humanos.

Agradecimentos

Agradecemos ao INPE e ao IEAV pelo apoio à pesquisa, bem como, ao CNPq e à CAPES.

Referências

- [1] S. Blondin, K. S. Mandel, and R. P. Kirshner. Do spectra improve distance measurements of type ia supernovae?, *Astronomy & Astrophysics*, 526(A81), 2011. DOI: <https://doi.org/10.1051/0004-6361/201015792>.
- [2] S. Blondin, et al. The spectroscopic diversity of type ia supernovae, *The Astronomical Journal*, 143(5): 126, 2012. URL: <http://stacks.iop.org/1538-3881/143/i=5/a=1>.
- [3] F. J. B. do Nascimento, L. R. Arantes Filho, and L. N. F. Guimarães. Intelligent Classification of Supernovae Using Artificial Neural Networks. *Inteligencia Artificial*, 22(63), 39-60, 2019. DOI: <https://doi.org/10.4114/intartif.vol22iss63pp39-60>.
- [4] A. V. Filippenko. Optical spectra of supernovae. *Annual Review of Astronomy and Astrophysics*, v.35, n. 1, p. 309-355, 1997.
- [5] CfA. The CfA Supernova Archive. Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, 2018. URL: <https://www.cfa.harvard.edu/supernova/SNarchive.html>.
- [6] T. Matheson, et al. Optical spectroscopy of Type Ia supernovae. *The Astronomical Journal*, v. 135, n.4, p. 1598, 2008.
- [7] M. Modjaz, et al. Optical spectra of 73 stripped envelope core-collapse supernovae. *The Astronomical Journal*, v. 147, n. 5, p. 99, 2014.
- [8] M. Módolo. Classificação Automática de Supernovas Usando Redes Neurais Artificiais. Tese (Doutorado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2016. URL: <http://mtc-m21b.sid.inpe.br/col/sid.inpe.br/mtc-m21b/2016/04.20.20.44/doc/publicacao.pdf>.
- [9] S. Perlmutter, et al. Measurements of omega and lambda from 42 high-redshift supernovae. *The Astrophysical Journal*, v. 517, n. 2, p. 565, 1999.
- [10] A.G. Riess, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, v. 116, n. 3, p. 1009, 1998.
- [11] T. Takagi, M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE transactions on systems, man, and cybernetics*, n. 1, p. 116-132, 1985.