

# Banco de Dados para Identificação de Nuvens Magnéticas por Aprendizado de Máquina

Milene O. Botelho<sup>1</sup> Marcos W. da S. Oliveira<sup>2</sup>

GPMCC/IFSP, São José dos Campos, SP

Rosemeire A. R. Oliveira<sup>3</sup>

GPMCC, São José dos Campos, SP

As Ejeções de Massa Coronal (CMEs) são grandes quantidades de plasma e campo magnético expelidas pelo Sol. No meio interplanetário, são chamadas de Ejeções de Massa Coronal Interplanetárias (ICMEs). As Nuvens Magnéticas (MCs) são uma classe de ICMEs com: i) baixo parâmetro beta do plasma, onde a pressão magnética domina; ii) campo magnético médio mais intenso que o do vento solar; iii) baixa temperatura dos íons; e iv) rotação angular significativa do vetor do campo magnético [2]. A análise dessas estruturas é essencial para entender seus impactos, como tempestades geomagnéticas, interferências em satélites e tecnologias terrestres.

Este estudo propõe a construção de um banco de dados para identificar e classificar MCs via aprendizado de máquina, utilizando hodogramas dos componentes do campo magnético interplanetário (IMF)  $B_x$ ,  $B_y$  e  $B_z$ , gerados a partir de dados reais medidos *in situ* por satélites e de dados sintéticos calculados pelo modelo livre de força (*force-free model*) [3]. Dessa forma, foram obtidos planos de máxima, média e mínima variância por meio do Método da Mínima Variância (MVA) aplicado aos dados reais, bem como utilizando rotações do modelo livre de força.

No modelo livre de força, a geometria Circular-Cilíndrica é adotada para o tubo de fluxo das MCs [3]. Assim, a orientação da nuvem magnética é definida pelo seu eixo, representado pelos ângulos  $\phi$  (longitude) e  $\theta$  (latitude) em coordenadas GSE. Esses ângulos foram utilizados para rotular as imagens dos dados sintéticos de acordo com os intervalos de classificação dos tipos de MCs, estabelecidos a partir de revisão bibliográfica [1] [2] [3]. Definiram-se intervalos para os tipos de MCs bipolares: SWN, SEN, NWS, NES e para os unipolares: WNE, WSE, ENW e ESW. Além disso, foi introduzido ruído gaussiano nas séries temporais obtidas, variando entre 0% e 50%. O ruído foi aplicado a fim de simular as condições dos dados obtidos pelos satélites, seguindo abordagem similar à adotada em [5]. Essa abordagem resultou em 2160 imagens de hodogramas sintéticos para o banco.

Os gráficos que compõem o banco derivam da plotagem dos componentes do campo magnético nos planos de máxima, média e mínima variância. Ou seja, cada subgráfico representa a relação entre  $B_z \times B_y$ ,  $B_y \times B_x$  e  $B_z \times B_x$ , sendo organizados em imagens lado a lado, sem a exibição de eixos. Desse modo, 325 imagens foram compostas de hodogramas de eventos reais identificados por [4] e rotulados com tipo de nuvem e ângulos  $\phi$  e  $\theta$  usando a metodologia em [3]. Com isso, o banco totaliza 2485 imagens. Esse número pode ser maior ou menor conforme os parâmetros do ruído, mas sem gerar grande variabilidade. De outro modo, pode-se alterar a resolução das imagens, e isso foi investigado considerando  $32 \times 32$  e  $64 \times 64$ . A Figura 1 apresenta imagens do banco de dados, destacando que possuem resolução e dimensões superiores às utilizadas no modelo.

---

<sup>1</sup>mileneoliveirabotelho@gmail.com

<sup>2</sup>oliveiramw@ifsp.edu.br

<sup>3</sup>meirematematica@gmail.com

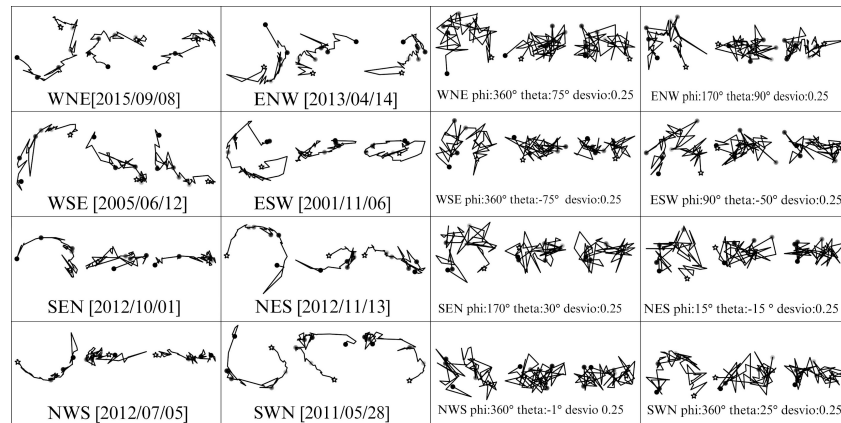


Figura 1: Comparação entre gráficos de eventos reais e sintéticos (à esquerda e à direita, respectivamente). Fonte: Autor.

As imagens de dados sintéticos e reais foram combinadas nos experimentos para a separação entre treino, teste e validação. Os resultados indicaram que a rede neural convolucional (CNN) com camadas convolucionais e de *pooling*, superou os modelos sequenciais simples (128 neurônios) e expandidos (256 neurônios com *dropout*), apresentando maior precisão e desempenho na classificação das MCs. A classificação binária, que distingue unipolares e bipolares, apresentou desempenho superior à classificação em oito categorias. O aumento da resolução das imagens de  $32 \times 32$  para  $64 \times 64$  pixels resultou em um maior tempo de processamento sem melhorias significativas na acurácia. Esses resultados indicam que a melhoria na classificação das MCs não depende exclusivamente da resolução das imagens. Estudos futuros podem investigar abordagens mais robustas, capazes de identificar com maior precisão os tipos de imagens analisadas. De outro modo, pode-se combinar diferentes modelos físicos para gerar hodogramas sintéticos e melhorar a identificação das MCs.

## Referências

- [1] V. Bothmer e R. Schwenn. “The structure and origin of magnetic clouds in the solar wind”. Em: **Annales Geophysicae** 16.1 (1997). DOI: 10.1007/PL00021390.
- [2] L. F. Burlaga. “Magnetic clouds and free force fields with constant alpha”. Em: **Journal of Geophysical Research** 93.A7 (1988). DOI: 10.1029/ja093ia07p0721.
- [3] R. A. Rosa Oliveira, M. W. da Silva Oliveira, A. Ojeda-González, V. Gil Pillat, E. Echer e T. Nieves-Chinchilla. “Resolving the Ambiguity of a Magnetic Cloud’s Orientation Caused by Minimum Variance Analysis Comparing it with a Force-Free Model”. Em: **Solar Physics** 296 (2021), p. 182. DOI: 10.1007/s11207-021-01921-2.
- [4] I. Richardson e H. Cane. **Near-Earth Interplanetary Coronal Mass Ejections Since January 1996**. Versão V2. Accessed 2024. 2024. DOI: 10.7910/DVN/C2MHTH.
- [5] L. F. G. dos Santos, T. Nieves-Chinchilla, M. Kirk, A. Narock e M. Nuñez. “Identifying Flux Rope Signatures Using a Deep Neural Network”. Em: **Solar Physics** 295 (2020). DOI: 10.1007/s11207-020-01697-x.