



## ANÁLISE DE CORRELAÇÃO ESTATÍSTICA ENTRE DADOS DA REDE EMBRACE E IMAGENS SOLARES

**ORIENTADOR:** Ms. Tiago da Silva Almeida

**NOME DO ALUNO:** Caio Henrique Machado

**CAMPUS:** Campus Universitário de Palmas – CUP

**CURSO:** Ciência da Computação

**LOCAL DE EXECUÇÃO:** Campus Universitário de Palmas – CUP, Bloco III, Sala 107

**GRANDE ÁREA DO CONHECIMENTO:** Ciências Exatas e da Terra

### FONTES DE FINANCIAMENTO:

**NOME DO GRUPO DE PESQUISA:** GCC – Grupo de Computação Científica

**PALAVRAS-CHAVE:** INPE, EMBRACE, NASA, SOHO, Tempestade Solar, Clima Espacial, Redes Neurais Artificiais, Aprendizado de Máquina.

### EQUIPE EXECUTORA:

Nome	Categoria
Tiago da Silva Almeida	Coordenador
Caio Henrique Machado	Aluno

### GRANDES ÁREAS DO CONHECIMENTO DA UFT:

**CIÊNCIAS EXATAS E DA TERRA**

## 1 – INTRODUÇÃO

A maioria das tempestades solares produzem apenas pequenas perturbações na Terra como pequenas variações momentâneas da alimentação das redes de energia elétrica, interferência na comunicação entre aparelhos eletrônicos, necessidade de novo traçado de rotas para aeronaves, perda de alguns satélites e belíssimas auroras boreal e austral. Mas uma tempestade solar pode ter a capacidade de causar grandes desastres na Terra caso se torne maior. A tempestade solar de setembro de 1859, que ficou conhecida como *Carrington Flare* (Labareda de Carrington) foi um exemplo (HOFF, 2003). Outro exemplo aconteceu em julho de 2000, quando o governo japonês lançou um satélite telescópio, parte de programa de astronomia, que sofreu danos irreversíveis após ser pego por uma tempestade solar (NORMILE, 2000).

Existem hoje dois sistemas de captura de informações relacionadas à influência solar na Terra e ao comportamento do sol que são de interesse deste trabalho: a Rede Embrace de Magnetômetros (Embrace MagNet) (DENARDINI, 2018), uma rede de sensores posicionados estrategicamente por toda a América Latina; e os satélites posicionados orbitando a Terra, capturando dados através de sensores gerando imagens, no projeto de estudos solar da NASA, SOHO (Solar and Heliospheric Observatory) (RAYMOND et al., 2001).

A previsibilidade dos dados é fundamental para o sistema de alertas de tempestade solares, já existente no Brasil, por meio da Embrace MagNet. Além de ser uma fonte de dados muito importante na crescente pesquisa e entendimento dos fenômenos eletromagnéticos do nosso planeta e do sistema Solar.

Pensando nisso, o objetivo deste trabalho é demonstrar e mensurar o grau de confiabilidade da correlação entre os dados dos dois sistemas de captura de informações solares, utilizando um modelo de RNA (Redes Neurais Artificiais). Se houver essa correlação, consideramos ser possível fazer modelos preditivos permitindo nos preparar para eventuais perturbações. Tal preparação seria fundamental na prevenção de danos uma vez que as tempestades solares viajam na velocidade da luz, dando-nos uma janela de tempo muito pequena de ação (em torno de oito minutos para chegar a Terra).

## 2 – OBJETIVOS

O objetivo deste trabalho é obter dados dos sistemas EMBRACE e SOHO de captura de informações solares e traçar um comparativo fazendo uso de redes neurais artificiais para averiguar a concordância entre as informações e estabelecer e mensurar um grau de confiabilidade dessa correlação. Os passos se darão da seguinte maneira:

- Estudar os principais índices de medição eletromagnética da Terra (índice Pk);
- Estudar os principais modelos de RNA e aprendizado profundo aplicado ao problema e implementar o modelo mais eficaz;
- Criar um modelo baseado em RNA tendo como dado de entrada as imagens solares;
- Coletar e classificar os dados;
- Analisar o grau de erro do modelo;
- Quantificar de forma estatística por meio de teste de hipótese a correlação entre as variáveis;

## 3 – METODOLOGIA

Um dos principais objetivos da ciência espacial é determinar a origem da atividade eruptiva do sol, a qual eventualmente vai ditar a variabilidade magnetosférica e ionosférica da Terra. O Estudo e Monitoramento Brasileiro de Clima Espacial (EMBRACE) desenvolveu a Rede de Magnetômetros Embrace (Embrace MagNet) para cobrir a maior parte do setor longitudinal leste da América do Sul, onde a UFT tem a satisfação de sediar um dos sensores, tendo como um dos objetivos fornecer informações para serem usadas como estimativa dos níveis de perturbações causadas pelas tempestades geomagnéticas. Para isto, o programa desenvolveu o Índice K (Ksa) da América do Sul que é usado para caracterizar a magnitude das tempestades geomagnéticas ajudando a decidir se os alertas geomagnéticos devem ser repassados para seus usuários que são afetados por estas perturbações.

Os principais usuários afetados pelas tempestades geomagnéticas são redes de energia, operações de aeronaves, usuários de sinais de rádio que refletem ou atravessam a ionosfera, e observadores da aurora.

A principal ferramenta usada para este projeto será uma rede neural artificial. Redes neurais artificiais são técnicas de aprendizado de máquina que simulam os mecanismos de aprendizado de organismos biológicos. O sistema nervoso humano contém células chamadas de neurônios que são conectados uns aos outros pelos axiomas e dentritos, e as regiões onde os axiomas e dentritos se conectam são chamadas de sinapses.

Na busca por cada vez melhores resultados, pesquisadores e estudiosos das redes neurais e aprendizado de máquina alcançaram uma técnica denominada Aprendizado Profundo (Deep Learning), que consistia não só em processar os dados de forma análoga a um sistema neural biológico, mas também simulava uma capacidade cognitiva ao construir um processo mental mais elaborado fazendo uso de um número maior de camadas de processamento, chamadas de camadas ocultas (*Hidden Layers*). Esta Técnica tomava mais tempo, mas apresentava melhores resultados. Neste trabalho vamos estudar o ganho de usarmos um modelo de Aprendizado Profundo tentando mensurar o ganho na precisão contra o tempo gasto, decidindo assim fazer uso ou não da metodologia para o resultado final.

#### 4 – CRONOGRAMA DE EXECUÇÃO

As atividades do trabalho são divididas conforme a Tabela 1, enumeradas de A à J, e sua execução é de acordo com o cronograma da Tabela 2.

**Tabela 1 – Lista de atividades previstas.**

Atividades	Descrição
<b>A</b>	Estudo da teoria e prática de redes neurais artificiais.
<b>B</b>	Desenvolvimento e treinamento de uma rede neural artificial.
<b>C</b>	Estudo da teoria e prática de processamento de imagens e discretização das imagens dos satélites.
<b>D</b>	Alimentação da rede neural artificial com os dados das imagens e verificação de resultados.
<b>E</b>	Aplicação e desenvolvimento de algoritmo de aprendizado de máquina.
<b>F</b>	Coleta de dados da rede EMBRACE.
<b>G</b>	Estudo e criação do modelo de inferência linear entre as variáveis de estudo.
<b>H</b>	Análise estatística dos resultados.
<b>I</b>	Escrita de relatório parcial em forma de trabalhos parciais para publicação em congressos e/ou periódicos e como relatório de prestação de contas com os sistemas de gestão da UFT.
<b>J</b>	Escrita de relatório final em forma de trabalhos parciais para publicação em congressos e/ou periódicos e como relatório de prestação de contas com os sistemas de gestão da UFT.

**Tabela 2 – Cronograma de Atividades.**

Ano	2020/2021											
Mês	Jul	Ago	Set	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun
Atv.												
A	✓	✓										
B		✓	✓	✓	✓	✓	✓					
C			✓	✓	✓	✓	✓					
D					✓	✓	✓	✓	✓			
E						✓	✓	✓	✓	✓		
F							✓	✓	✓	✓	✓	
G									✓	✓	✓	
H										✓	✓	
I						✓	✓					
J											✓	✓

## 5 – RESULTADOS ESPERADOS

A rede neural fará o trabalho de comparar os dados amostrais utilizando conceitos de estatística. Faremos um estudo da melhor aplicação estatística entre Teste de Hipótese e Inferência Linear. Para aplicarmos Teste de Hipótese definimos a hipótese de que existe uma correlação entre as informações geradas pelos sistemas Embrace MagNet e SOHO, chamamos de Hipótese Nula ( $H_0$ ) e associamos uma projeção de valor teste  $x$ . Definimos também uma hipótese de que esse valor esteja errado, chamamos de Hipótese Alternativa ( $H_1$ ) e atribuímos valor diferente de  $H_0$  ( $H_1 \neq x$ ). Usamos evidências fornecidas pelas amostras através dos cálculos para rejeitar ou não as hipóteses. Tal procedimento envolve mensurar quão diferente o resultado obtido é ao assumirmos o  $H_0$  como verdadeiro. Para Inferência Linear fazemos um estudo da correlação dos dados sob dois pontos de vista: Quantificamos a força dessa relação entre os dados que vamos observar, chamamos de correlação; Explicitamos a forma dessa relação, chamamos de regressão. Obtemos representação gráfica das variáveis quantitativas que chamamos de diagrama de dispersão e calculamos o coeficiente da correlação linear que nos ajudará a obter o resultado final.

Ao final do projeto, esperamos demonstrar e mensurar a confiabilidade da correlação entre as informações de ambos os sistemas EMBRACE e imagens por satélites, além de podermos prever em séries temporais os dados da rede EMBRACE a partir das imagens de satélite conhecendo o grau de confiabilidade da correlação.

## 6 – REFERÊNCIAS BIBLIOGRÁFICAS

DENARDINI, S. S. C. L. C. A. R. J. M. A. V. B. P. R. F. M. A. G. M. A. C. M. J. A. B. A. L. P. N. J. S. J. L. H. L. R. A. P. F. B. N. P. A. B. N. G. A. S. P. T. O. B. C. M. The embrace magnetometer network for south america:network description and its qualification. *Radio Science*, v. 53, p. 288 – 302, 2018.

HOFF, L. T. . E. Z. . D. S. . J. Carrington flare of 1859 as a prototypical worst-case solar energetic particle event. *IEEE Transactions on Nuclear Science*, v. 50, p. 2307 – 2309, 2003.

NORMILE, D. Solar storm knocks out japanese satellite. *Science*, v. 289, p. 709, 2000.

RAYMOND, J. C. et al. Soho observations of a coronal mass ejection. *The Astrophysical Journal*, v. 553, p. 922–934, 2001.