



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS



Estudo da Estimativa de Precipitação Utilizando diferentes Métricas de Interpolação Espacial para a Combinação de dados de Satélite e Observações *in situ*

RELATÓRIO FINAL DE PROJETO DE INICIAÇÃO CIENTÍFICA (PIBIC/CNPq/INPE)

Igor Frassoni Guedes dos Santos (UNIP, Bolsista PIBIC/CNPq)
E-mail: igor.frassoni@cptec.inpe.br

João Gerd Zell de Mattos (DMD/CPTEC/INPE, Orientador)
E-mail: joao.gerd@cptec.inpe.br

Julho de 2013

Sumário

Resumo.....	3
Introdução	3
Desenvolvimento	4
Métodos de Coleta de Dados	4
Softwares e linguagens	5
Métodos de Interpolação	5
Resultados de Discussões	8
Conclusão	9
Referências Bibliográficas	9
Lista de Anexos	10

1.0. RESUMO

O projeto de pesquisa apresentado tem como objetivo realizar um estudo dos diferentes métodos de interpolação utilizados na combinação dos dados utilizados na estimativa de precipitação e, a partir deste estudo, desenvolver um produto para a geração de precipitação, baseado na combinação dos dados do satélite TRMM (*Tropical Rainfall Measuring Mission*) e de dados de estações meteorológicas.

Nos primeiros 3 meses da pesquisa o bolsista focou-se na compreensão dos métodos utilizados para a obtenção e armazenamento dos dados de precipitação, por meio da leitura de artigos científicos e desenvolvimento de rotinas em diferentes linguagens de programação (como fortran, GrADS e shell scripts), sendo avaliado através de apresentações e resumos.

Nos demais meses o bolsista iniciou a leitura de artigos na busca do entendimento das diferentes métricas utilizadas para combinar diferentes fontes de dados, dando ênfase na metodologia de Interpolação Ótima, descrita na literatura como uma das que apresenta melhores resultados para a interpolação de campos de precipitação.

Embora, o bolsista tenha se familiarizado com os dados de precipitação, compreendendo as formas de coleta, tratamento, armazenamento e disseminação dos dados de precipitação, bem como ter compreendido razoavelmente bem o tema os objetivos propostos, não foi possível concluir o projeto pois o bolsista apresentou algumas deficiências para o desenvolvimento da metodologia e aplicação das técnicas empregadas no projeto.

2.0. INTRODUÇÃO

O conhecimento da distribuição da precipitação é de grande importância para o planejamento de diversas atividades tanto econômicas como sociais, tais como agricultura, pecuária, geração de energia, entre outras. A baixa densidade de plataformas de coleta de dados meteorológicas, em algumas regiões, compromete significativamente os estudos relacionados aos sistemas meteorológicos. Sendo assim, variados produtos de precipitação obtidos por meio da combinação de estimativas de satélite e estações meteorológicas convencionais são gerados e disponibilizados a comunidade para o suprimento desta carência (veja Viltz et. al.). Entretanto, vários destes produtos apresentam problemas, que geralmente são associados aos métodos de interpolação utilizados. É comum verificar nas áreas de montanha um espalhamento horizontal incoerente nos campos de precipitação. Também, notam-se em algumas regiões, com uma rede esparsa de observações convencionais, áreas de intensa precipitação. Esse tipo de problema é causado principalmente pela metodologia de interpolação utilizada no processo. Em vista disso, o corrente projeto propôs um estudo sobre os métodos de interpolação utilizados na combinação dos dados de satélite e observações terrestres, com a finalidade de desenvolver um novo produto de precipitação onde os campos espaciais não apresentassem tais problemas.

O plano de trabalho, com seus respectivos períodos divide-se da seguinte forma:

1. (08-09/2012)-Estudar e entender as diferentes fontes de observação da precipitação;
2. (10/2012)-Fazer um levantamento dos dados de precipitação de satélite e observações in situ, organizá-los e se familiarizar com os dados;
3. (11-12/2012)-Estudar e entender os métodos de interpolação espacial de precipitação a serem utilizados e elaborar um relatório interno sobre as técnicas aprendidas;
4. (01-02-03-04/2013)-Elaborar rotinas computacionais para aplicação dos métodos nos

- conjuntos de dados obtidos;
5. (05-06/2013)-Avaliar estatisticamente os campos combinados;
 6. (07/2013)-preparar o relatório de atividade da Bolsa de Iniciação científica.

3.0. DESENVOLVIMENTO

Inicialmente, foi efetuada a elaboração de um resumo sobre assuntos relacionados a precipitação, mais especificamente sobre as formas de coleta de dados de precipitação. Este processo teve por finalidade a compreensão dos diferentes instrumentos utilizados para a coleta de precipitação, bem como os motivos pelos quais a rede de observações é espaça e também os tipos de erros que podem estar presentes.

- 3.1. Métodos de coleta de dados

3.1.1. Pluviômetros/pluviógrafos

São basicamente instrumentos destinados a medida de precipitação ocorrida em determinado local. Diferem-se entre si apenas no fato de armazenamento dos dados coletados, onde o pluviógrafo possui a capacidade de gravar os dados coletados, já o pluviômetro não possui tal artifício. Estes instrumentos são amplamente utilizados ao redor do mundo, por serem baratos e de fácil manuseio, porém, esse sistema de coleta de dados torna-se muito dependente da quantidade de pluviômetros e de seu espalhamento em torno da região a ser feita a estimativa.

Além de depender da quantidade e de seu espalhamento na região, estes tipos de equipamentos possuem uma taxa de erro associada a eles, visto que são equipamentos onde geralmente utiliza-se de contato humano para a coleta desses dados (exceto quando os dados são enviados automaticamente para o banco de dados, PCD's). Há também os erros associados ao equipamento, onde este pode estar descalibrado, ou sofrer interferência de algum fator externo do ambiente, como catástrofes naturais, onde este tipo de equipamento pode sofrer danos.

3.1.2. Satélites

A precipitação por satélite pode ser estimada a partir de radares, sensores infravermelhos, micro-ondas, entre outros. Este é um método de estimativa indireta, pois o satélite não detecta exatamente as gotas de chuva, o que ele faz é utilizar-se de algoritmos que associam a espessura e/ou a temperatura de brilho da nuvem com a taxa de chuva de nuvens precipitantes. Um dos dados utilizados neste projeto, foi proveniente do produto de estimativa de precipitação produzida pelo TRMM (Tropical Rainfall Measuring Mission), o produto 3B42RT. Este produto utiliza-se da combinação de sensores de micro-ondas e infravermelho para estimar a precipitação.

O erro associado a este tipo de equipamento pode ser gerado de diferentes formas, por ser um equipamento do satélite, este pode sofrer danos no trajeto da Terra ao espaço, além de muitos outros tipos de interferência que possam vir a ocorrer no espaço, tanto físicas como magnéticas, por exemplo.

- 3.2. Softwares e Linguagens

Após o breve estudo sobre a obtenção dos dados de precipitação, deu-se início ao trabalho direto com os dados, por meio de alguns softwares e linguagens, como o GrADS (*Grid Analysis and Display System*), Matlab, Fortran e Shell Script.

3.2.1. GrADS

GrADS (Grid Analysis and Display System) é uma ferramenta de manipulação e visualização de dados relacionados a ciência terrestre. Esta ferramenta foi utilizada para a introdução do bolsista aos dados de precipitação.

3.2.2. Matlab

Matlab é um software de alta performance voltado para o cálculo numérico. Foi utilizado na elaboração de rotinas computacionais a fim de gerar o produto proposto.

3.2.3. Fortran

Fortran (Formula Translator) é uma linguagem de programação voltada especialmente para cálculo numérico, e foi utilizada para criar rotinas que seriam usadas no projeto.

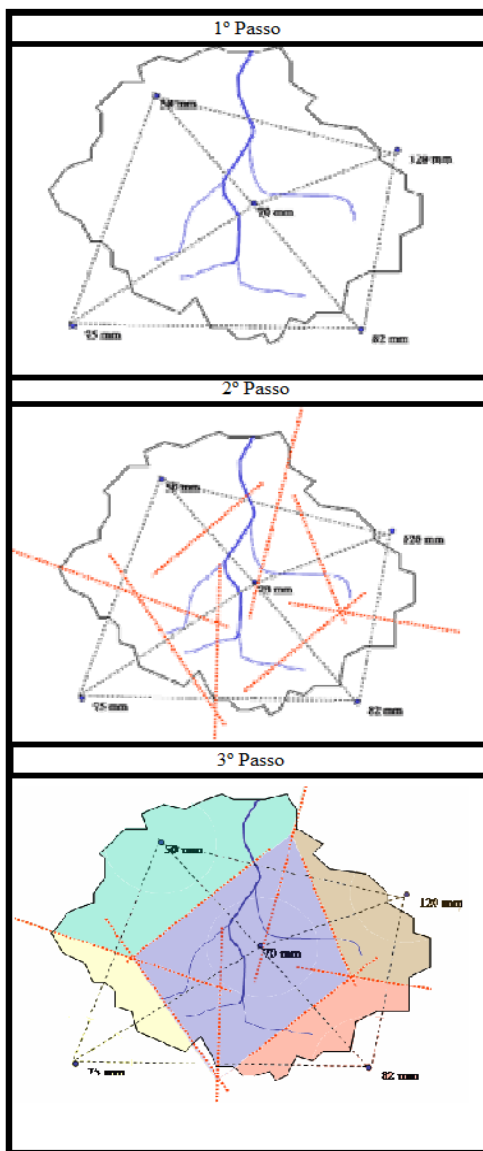
3.2.4. Shell Script

Shell script é uma linguagem de script utilizada no sistema operacional GNU/Linux, com diferentes interpretadores de comando. No presente projeto, foi utilizada para a implementação de rotinas computacionais a fim de atualizar o banco de dados do produto de precipitação proveniente do TRMM. Um exemplo pode ser visto no *anexo 1*.

- 3.3. Métodos de Interpolação

Após o entendimento do dados, sua forma de coleta, e um breve estudo dos softwares disponíveis para a visualização e manipulação dos dados de precipitação, iniciou-se os estudos a respeito das técnicas de interpolação existentes.

3.3.1. Método dos Polígonos de Thiessen Este método assume que em qualquer ponto da bacia a precipitação é igual a medida do posto mais próximo. Nesse método é definida a área de influência de cada posto pluviométrico dentro da bacia hidrográfica. Esta técnica permite determinar, através da construção de polígonos, qual a área mais próxima de um ponto numa área, ou seja, unem-se todos os pares de pontos com segmentos de reta, e após, traça-se a perpendicular a cada segmento de reta. Forma-se assim um polígono que delimita uma área em que todos os pontos possuem o mesmo valor. Um exemplo de como este método funciona pode ser visto na *figura 1*. Este é um método básico, exato, e que não leva em conta os erros associados a interpolação, porém,



de simples montagem e entendimento.

Figura 1: Exemplificação do método dos Polígonos de Thiessen

- Passo 1: Traçam-se linhas que unem os postos pluviométricos mais próximos entre si;
- Passo 2: Traçam-se linhas médias perpendiculares às linhas que unem os postos pluviométricos;
- Passo 3: Define-se a região de influência de cada posto pluviométrico a partir da área formada pela intersecção das mediatrizes das linhas que unem os postos pluviométricos;

3.3.2. Ponderação do Inverso das Distâncias

Este método utiliza-se do pressuposto que estações próximas exercem maior influência no ponto que se deseja interpolar. Assim, uma estação próxima possui maior peso no cálculo do que uma estação mais distante.

Calcula-se da seguinte forma:

$$\hat{Z}(s_0) = \sum_{i=1}^N \lambda_i Z(s_i) \quad ; \quad \lambda_i = \frac{d_{i0}^{-p}}{\sum_{i=1}^N d_{i0}^{-p}} \quad ; \quad \text{sendo} \quad \sum_{i=1}^N \lambda_i = 1$$

Onde:

$\hat{Z}(s_0)$ é o valor a ser interpolado para o local s_0 ;

N é o número de pontos observados a serem usados no valor a ser interpolado;

λ_i São os pesos colocados para cada ponto observado a ser utilizado;

$Z(s_i)$ é o valor observado no local s_i ;

d_{i0} é a distância entre o local a ser interpolado, s_0 , e cada um dos locais observados, s_i .

O valor “p” é determinado minimizando o erro médio quadrático, que é a estatística calculada por um procedimento de validação cruzada, neste procedimento, por sua vez, tem-se que cada ponto observado é removido e comparado com o predito para aquele local.

Este método de interpolação é exato, onde poucas decisões são tomadas em relação aos parâmetros do modelo. Porém, nele não é realizada uma avaliação da predição dos erros, que pode gerar o efeito “olho de boi” em torno da localização do dado, esse efeito, é caracterizado como um espalhamento incoerente da precipitação no campo ao redor do dado, um exemplo pode ser visto na *figura 2*.

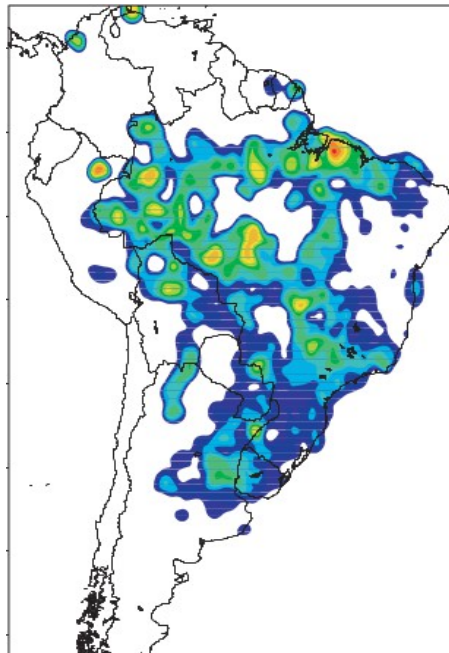


Figura 2: Exemplo de espalhamento incoerente da precipitação, próximo ao Norte do Estado do Pará

3.3.3. Interpolação Polinomial

A interpolação polinomial define que uma equação global seja adequada ao estudo da área de interesse, usando uma função polinomial algébrica ou trigonométrica. Este método de interpolação possui como desvantagem, em alguns casos, dependendo da ordem da função polinomial, o excessivo uso computacional necessário para sua aplicação, e possui como vantagem a remoção dos efeitos de tendências globais (i.e. Efeitos de vento, poluição, etc) de longa duração, criando superfícies que descrevem processos físicos.

3.3.4. Interpolação Multi-quadrática

Este método de interpolação define que a influência de cada amostra é representada por cones quadráticos como uma função das coordenadas desses pontos. A estimativa de um certo ponto é então obtida pela soma das contribuições de todos os cones quadráticos. As equações que a representam são:

$$h_o = \sum_{j=1}^n w_j h_j \quad ; \quad w_j = \sum_{i=1}^n \delta_{ij} d_{oi}$$

Onde:

n representa o número de estações

d_{oi} é a distância entre os pontos (x_o, y_o) e (x_i, y_i)

δ_{ij} representa um elemento da matriz inversa da distancia entre as estações ($n \times n$), com elementos d_{ji} , $j=1, \dots, n$ e $i=1, \dots, n$.

3.3.5. Interpolação Ótima

Este método de interpolação determina que os pesos aplicados as estações são determinados a partir da minimização da variância do erro da interpolação. Ou seja, faz com que diminuam as possibilidades do valor interpolado ser distante do esperado.

• 3.4. Resultados e Discussões

Variados métodos foram utilizados neste estudo, e o que se pode notar é que eles se assemelham em sua forma de determinar o local a ser interpolado, possuindo variações apenas na determinação dos pesos a serem aplicados as estações.

A partir destes métodos, pode-se perceber que eles funcionam de determinadas formas, e que podem ou não ser úteis, o que vai depender será a aplicação destinada a eles. Por exemplo, caso haja uma área com dados a serem interpolados, e ainda, esse processo seja somente para exemplificação, pode-se utilizar um método que menos dispendioso, de fácil execução, como o método dos Polígonos de Thiessen, por exemplo. Em suma, o que define a usabilidade de um método é sua aplicação.

Entre esses métodos, devido as características funcionais, o que se destaca para a finalidade deste projeto, é o método de interpolação ótima, visto que este minimiza a variância do

erro da interpolação, ou seja, diminui as possibilidades do valor interpolado ser distante do esperado. É claro que não deve-se descartar os outros métodos, visto que possuem suas particularidades, de modo que se tornam úteis dependendo da sua utilização.

O projeto não foi finalizado, faltando concluir as rotinas para a implementação do produto com base na interpolação ótima. Esta etapa não pôde ser finalizada devido a dificuldades sofridas pelo bolsista em compreender a técnica a ser utilizada para a elaboração da rotina computacional.

4.0. CONCLUSÃO

O projeto apresentado foi um grande desafio proposto ao bolsista, visto que este não possuía conhecimento algum na área tratada, meteorologia e assimilação de dados, além dos softwares, este pôde aprender a utilizar diversos softwares como Fortran, Grads, além do sistema operacional GNU/Linux.

Por mais que o projeto de iniciação científica não tenha sido finalizado, este foi de grande valia para o bolsista, visto que este adquiriu conhecimentos sobre determinados assuntos que são de grande valia para a sua vida acadêmica. Além disso, o projeto ajudou a ampliar a visão de trabalho, a qual engloba diversos fatores antes desconhecidos em relação ao ambiente de trabalho de um instituto de pesquisa. Como o nome sugere, iniciação científica, este projeto foi útil para criar a noção de pesquisa no bolsista, para que, futuramente, este possa a vir desenvolver novos e melhores projetos.

5.0. REFERÊNCIAS BIBLIOGRÁFICAS

5.1. Tabios, G. Q. And Salas, J. D. (1985), A COMPARATIVE ANALYSIS OF TECHNIQUES FOR SPATIAL INTERPOLATION OF PRECIPITATION. JAWRA Journal of the American Water Resources Association, 21: 365-380. doi: 10.1111/j.1752-1688.1985.tb00147.x

5.2. Vila, Daniel A., Luis Gustavo G. De Gonçalves, David L. Toll, José Roberto Rozante, 2009: Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimates over Continental South America. J. Hydrometeor, 10, 533-543. doi: <http://dx.doi.org/10.1175/2008JHM1048.1>

5.3. Rozante, J. R. ; Moreira, D. S. ; de Gonçalves, L.G.G. ; Vila, Daniel. Combining TRMM and Surface Observation Precipitation: Technique and Validation over South American. Weather and Forecasting, 2010.

6.0. LISTA DE ANEXOS

7.1. Anexo 1: Rotina em Shell Script para atualização de dados de precipitação

Anexo 1: Rotina em Shell Script para atualização de dados de precipitação

```
#!/bin/bash

#
# Deverão ser passados 4 argumentos:
# 1 - Versão do TRMM [3B42RT/3B42_V6]
# 2 - Data Inicial [yyyymmddhh]
# 3 - Data Final [yyyymmddhh]
# 4 - Incremento de tempo [+ -ndnhm]

TRMM=${1}
DataInicial=${2} #"1998010100"
DataFinal=${3} #"2010123121"
Incremento=${4} #" +3h"

DIR="/letkf/backup_tariba/saldas/jgerd/FORCING/PRECIPITACAO/TRMM/${TRMM}"
ftp="ftp://disc2.nascom.nasa.gov/data/TRMM/Gridded/${TRMM}"

while [ ${DataInicial} -ne ${DataFinal} ]; do

Ano=${DataInicial:0:4}
amd=${DataInicial:0:8}
Mes=${DataInicial:4:2}
Dia=${DataInicial:6:2}
hor=${DataInicial:8:2}

Ano2=${DataInicial:2:2}
amd2=${DataInicial:2:8}

case ${TRMM} in
B42RT)
if [ ${DataInicial:0:4} -lt 2008 ]; then
file="3B42RT.${Ano}${Mes}${Dia}${hor}.bin"
else
file="3B42RT.'${Ano}'. '${Mes}'. '${Dia}'. '${hor}'z.bin"
fi
;;
3B42_V6)

```

```

file="3B42.${amd2}.${hor}z.6.precipitation.bin"
;;
esac

if [ -e ${DIR}/${Ano}/${amd}/${file} ];then
echo "${file} OK"
else

if [ ! -d "${TRMM}/${Ano}/${amd}"]; then mkdir -p ${TRMM}/${Ano}/${amd}; fi

wget -c -t0 -O ${DIR}/${Ano}/${amd}/${file} ${ftp}/${Ano}${Mes}/${file}

if [ ! -e ${DIR}/${Ano}/${amd}/${file} ]; then
echo "${file}" >> ${DIR}/${Ano}_notfound.txt
else
echo "${file} GET OK !!"
fi

fi

DataInicial=`/home/jgerd/bin/inctime ${DataInicial} ${Incremento} %y4%m2%d2%h2`
done

```