



INSTITUTO NACIONAL DE PESQUISA ESPACIAL – INPE

PROVA DISCURSIVA

TG32

PROCESSAMENTO DE DADOS METEOROLÓGICOS E GEORREFERENCIADOS; DESENVOLVIMENTO DE SOFTWARE NA ÁREA DE METEOROLOGIA



SUA PROVA

- Além deste caderno contendo 5 (cinco) questões discursivas **com as respectivas folhas de rascunho**, você receberá do fiscal de prova as folhas de textos definitivos;



TEMPO

- Você dispõe de **4 (quatro) horas** para a realização da prova;
- **2 (duas) horas** após o início da prova, é possível retirar-se da sala, sem levar o caderno de questões;
- A partir dos **30 (trinta) minutos** anteriores ao término da prova é possível retirar-se da sala **levando o caderno de questões**.



NÃO SERÁ PERMITIDO

- Qualquer tipo de comunicação entre os candidatos durante a aplicação da prova;
- Anotar informações relativas às respostas em qualquer outro meio que não seja no caderno de questões e nas folhas de textos definitivos;
- Levantar da cadeira sem autorização do fiscal de sala;
- Usar o sanitário ao término da prova, após deixar a sala.

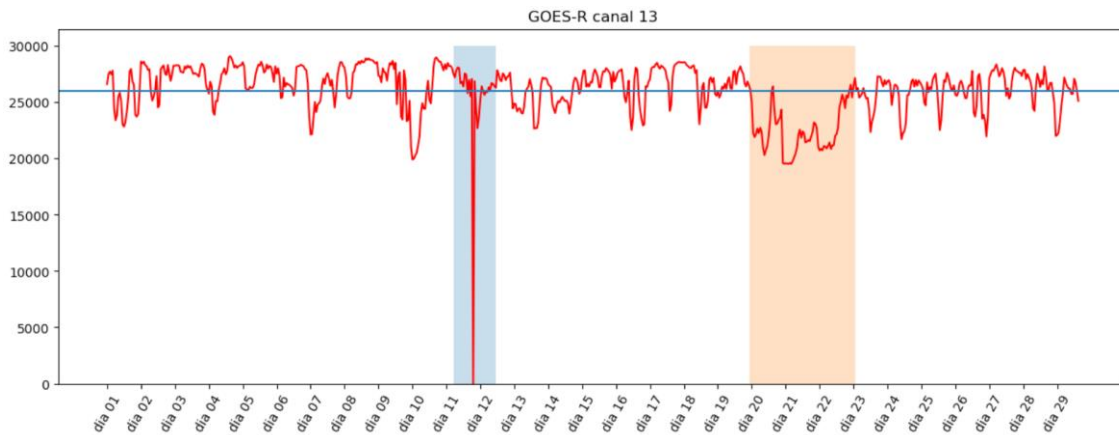


INFORMAÇÕES GERAIS

- Verifique se seu caderno de questões está completo, sem repetição de questões ou falhas. Caso contrário, **notifique imediatamente o fiscal da sala**, para que sejam tomadas as devidas providências;
- Confira seus dados pessoais, especialmente nome, número de inscrição e documento de identidade e leia atentamente as instruções para preencher as folhas de textos definitivos;
- Para o preenchimento das folhas de textos definitivos, use somente caneta esferográfica, fabricada em material transparente, com tinta preta ou azul;
- Assine seu nome apenas no(s) espaço(s) reservado(s) no cartão de respostas;
- Caso você tenha recebido caderno de cargo **diferente** do impresso em suas folhas de textos definitivos, o fiscal deve ser **obrigatoriamente** informado para o devido registro na ata da sala;
- O preenchimento das folhas de textos definitivos é de sua responsabilidade e **não será permitida a troca de folha de texto definitivo em caso de erro cometido pelo candidato**;
- Para fins de avaliação, serão levadas em consideração apenas os textos das folhas de textos definitivos;
- A FGV coletará as impressões digitais dos candidatos na lista de presença;
- Os candidatos serão submetidos ao sistema de detecção de metais quando do ingresso e da saída de sanitários durante a realização das provas.
- **Boa prova!**

Questão 1

De acordo com o guia simplificado de referência dos satélites GOES-R (*Geostationary Operational Environmental Satellites-R*), a banda número 13 (infravermelho) possui 2 km de resolução espacial e está centrada no comprimento de onda 10.3 μm . Por meio de um *script*, foram obtidas imagens a cada 1 hora, entre 01 e 29 de fevereiro de 2024, correspondentes à banda 13. O gráfico a seguir mostra uma série temporal da média dos valores de uma sub-região das imagens, para cada hora.



No gráfico são apontadas 2 regiões (região 1 entre os dias 11 e 12; região 2 entre os dias 20 e 23), cujos valores apresentam diferenças perceptíveis em relação aos outros dias. O código a seguir, escrito na linguagem Python, foi utilizado para gerar o gráfico.

```

1. import os
2. import numpy as np
3. from netCDF4 import Dataset
4. import matplotlib.pyplot as plt
5. base_folder = '../goes/'
6. files = [f for f in os.listdir(base_folder)]
7. time_series, time_line = [], []
8. row, column, delta = 1500, 1500, 10
9. for file in sorted(files):
10.    if file.endswith('.nc'):
11.        nc = Dataset(base_folder + file)
12.        raster = np.array(nc.variables['Band13'][:])
13.        time_series.append(np.mean(raster[column:column+delta, row:row+delta]))
14.        time_line.append(f'dia {file[-9:-7]}')
15.        raster = None
16.        nc.close()
17. plt.figure(figsize=(15,5))
18. plt.title('GOES-R canal 13')
19. plt.plot(time_series, 'r-')
20. plt.axhline(np.mean(time_series))
21. ticks = np.arange(start=0, stop=len(time_line), step=24)
22. plt.xticks(ticks, np.array(time_line)[ticks], rotation=60)

```

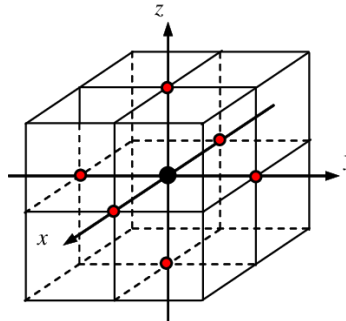
Com base no gráfico e no código:

- Apresente uma justificativa para os valores encontrados na região 1.
- Observe os dias 20 a 23 de fevereiro, destacados na região 2, e apresente um conjunto de passos que permita identificar essa região e situações similares em outras séries temporais.
- Determine a área, em km, da sub-região das imagens que foi utilizada para gerar a série temporal mostrada no gráfico.
- Indique, no código, onde estão as informações necessárias para determinar a área do item anterior.

Questão 2

Muitos Modelos Numéricos de Previsão de Tempo e Clima (MNPTC) são baseados na solução de conjuntos de equações diferenciais discretizadas por elementos finitos ou volumes finitos, onde se tem uma grade computacional tridimensional cobrindo a região de interesse de simulação, a qual pode representar todo o globo terrestre para os chamados modelos globais, ou uma determinada região geográfica para os chamados modelos regionais.

Considere que em determinados trechos do código-fonte se tem um padrão do tipo *stencil* (arranjo geométrico de um grupo nodal relacionado a um ponto de interesse) para cálculo de propriedades físicas de interesse, tal como ilustrado na figura a seguir:



As quantidades usuais de pontos de grade nas três dimensões dependem da área de cobertura geográfica e da resolução espacial adotada, porém, é possível afirmar que valores típicos para o tamanho da grade variam de várias centenas de pontos de grade nas dimensões horizontais ao longo dos eixos X e Y, e de várias dezenas de pontos na dimensão vertical referente ao eixo Z.

Supondo um contexto onde se tem um MNPTC tal qual descrito acima e uma infraestrutura computacional, formada por um *cluster* de alto desempenho com servidores (nós computacionais) interconectados por uma rede de dados, onde cada servidor/nó computacional contém dezenas de núcleos computacionais, encapsulados em processadores do tipo *multicore*, responda ao que se pede a seguir.

- A) Descreva como poderia ser modelada uma estratégia para paralelização do código-fonte em Fortran do citado modelo para ser utilizado com eficiência computacional no equipamento descrito. Explique como os padrões de programação paralela OMP (OpenMP) e MPI podem ser utilizados em função da divisão de domínio entre as instâncias (unidades de computação) paralelas, justificando a escolha.
- B) Comumente em MNPTCs se tem necessidade de trocas de dados entre as unidades de execução paralela, cite as diferenças relacionadas às trocas de dados entre os diferentes padrões de programação paralela OMP e MPI que ocorrem entre as diversas instâncias paralelas de execução (processos e/ou *threads*), considerando a divisão de domínio adotada para a paralelização do item acima.
- C) Considerando que se deseja maximizar o desempenho computacional, informe se existe a possibilidade de se melhorar o desempenho computacional utilizando chaves de compilação, ao se compilar o código-fonte Fortran.

Importante citar que a resposta esperada será na forma de um texto com as informações requisitadas, a qual pode ou não conter trechos de códigos ou pseudocódigos.

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30

Questão 3

Você recebeu uma série histórica de arquivos com histórico de precipitação e temperatura para o Brasil. A série está completa e sem falhas. Os arquivos estão em formato NetCDF com padrão de nome `precip_YYYY-MM-DD.nc`, onde YYYY-MM-DD é uma data com ano (YYYY), mês (MM) e dia (DD) respectivamente. A série se inicia em 01/01/2000 termina em 31/12/2023, inclusive.

Cada arquivo NetCDF contém a precipitação acumulada por hora e temperatura horária média para todos os pontos da grade. O conteúdo de cada arquivo é o seguinte:

Dimensões:

* (lat : 400, lon : 400, time : 24)

Coordenadas:

* lat : latitude (float)

* lon : longitude (float)

* time : data e hora do início da medição (datetime)

Variáveis:

* prec : a precipitação acumulada em 1 hora, em milímetros (float)

* temp : a temperatura média durante a hora, em Kelvin (float)

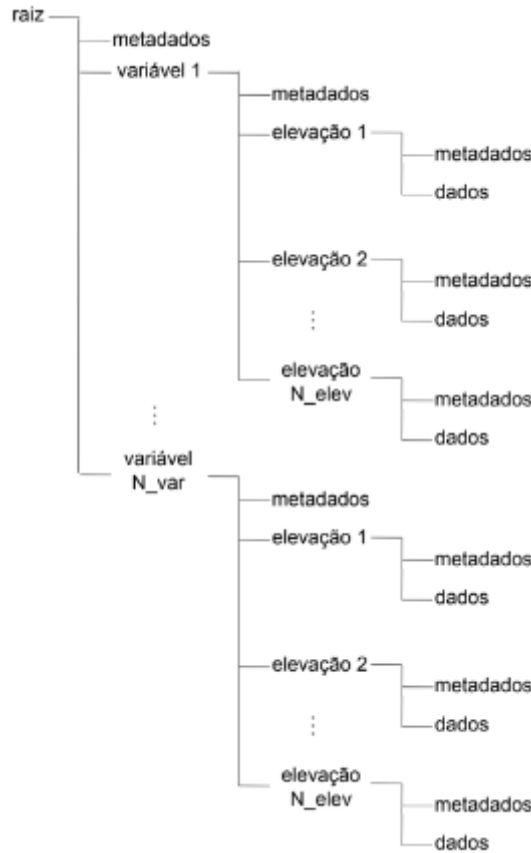
Crie um *script* em Python, utilizando apenas as bibliotecas `xarray` e `matplotlib`, que realize as seguintes operações:

- A) Calcular a precipitação mensal acumulada para cada ponto da grade de latitude-longitude para os meses de Dezembro.**
- B) Calcular o Z-Score de Dezembro de 2023 em relação ao conjunto dos demais meses de Dezembro.**
- C) Gerar um gráfico para o Z-Score calculado e salve no arquivo PNG “z-score.png”.**

Questão 4

Considere um radar meteorológico de superfície efetuando uma varredura volumétrica com número de variáveis observadas (ex.: Z, ZDR, V etc.) N_{var} , número de elevações (*sweeps*) N_{elev} , e 360 azimutes em cada elevação.

Num arquivo com formato HDF5, o dado de uma varredura volumétrica do radar pode ser estruturado da seguinte maneira:



Com base na descrição do dado apresentada, e assumindo que todas as matrizes de dados (ex.: /raiz/variável_3/elevação_2/dados) foram gravadas no arquivo HDF5 com seus elementos sendo do tipo UINT16 (unsigned integer, com 2 bytes), responda ao que se pede a seguir.

- A) Descreva as informações mais importantes contidas nos metadados localizados em /raiz/metadados.
- B) Descreva as informações mais importantes contidas nos metadados localizados em /raiz/variável x/, onde x é o índice da variável (indo de 1 até N_{var}).
- C) Descreva as informações mais importantes contidas nos metadados localizados em /raiz/variável x/elevação y/metadados, onde x é o índice da variável (indo de 1 até N_{var}) e y é o índice da elevação (indo de 1 até N_{elev}).
- D) Descreva quais das informações contidas nos metadados descritos nos itens anteriores são necessárias para a correta interpolação e conversão dos dados de uma grade polar para uma grade cartesiana geolocalizada, de modo a gerar um CAPPI (Constant Altitude Plan Position Indicator) em uma determinada altitude.

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

Questão 5

Considere um banco de dados com extensão espacial e a existência de duas tabelas: uma contendo a malha municipal brasileira referente a 2018, e outra contendo as notificações dos casos de dengue compiladas e atualizadas diariamente pelas secretarias municipais de saúde. A inserção das notificações ocorre a partir da confirmação dos casos na respectiva cidade. Você é, então, o tecnologista designado para implementar um conjunto de consultas SQL para analisar os dados coletados e visualizar os resultados dessas análises.

Neste contexto:

- A) Descreva uma solução para estabelecer o relacionamento entre as tabelas de notificações e de malha municipal, exibindo todos os municípios da malha em um dia pré-definido, mesmo quando não há notificações para algumas cidades.**
- B) Transcreva a sintaxe SQL para essa solução, empregando os nomes de tabelas e atributos indicados no item (A).**

Realização

