

MANIPULAÇÃO E VISUALIZAÇÃO DE DADOS GERADOS PELO MODELO ATMOSFÉRICOS BRAZILIAN GLOBAL ATMOSPHERIC MODEL (BAM) ATRAVÉS DE BIBLIOTECAS PYTHON

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

Lucas Ribeiro Mallmann (FATEC, Bolsista PIBIC/CNPq)

E-mail: <u>lucas.mallmann@inpe.br</u>

Eugênio Sper de Almeida (INPE /CPTEC, Orientador)

E-mail: eugenio.almeida@cptec.inpe.br

Ivo Kenji Koga (INPE /CPTEC, Coorientador)

E-mail: ivo.koga@cptec.inpe.br

Julho de 2017

DADOS DE IDENTIFICAÇÃO

Nome: Lucas Ribeiro Mallmann

Local de Trabalho: INPE/ CPTEC

Título do Projeto: Manipulação e visualização de dados gerado pelo modelo atmosférico Brazilian Global Atmospheric Model (BAM) através de bibliotecas Python

Tipo de bolsa: Bolsa PIBIC-IC

Período: abril/2017 a julho/2017

MANIPULAÇÃO E VISUALIZAÇÃO DE DADOS GERADOS PELO MODELO ATMOSFÉRICOS BRAZILIAN GLOBAL ATMOSPHERIC MODEL (BAM) ATRAVÉS DE BIBLIOTECAS PYTHON

Lucas Ribeiro Mallmann¹ (FATEC, Bolsista PIBIC/CNPq) Eugenio Sper de Almeida¹ (INPE/CPTEC, Orientador) Ivo Kenji Koga¹ (INPE/CPTEC, Coorientador)

RESUMO

A previsão global do tempo e clima do Centro de Previsão do Tempo e Estudos Climáticos (CPTEC) é gerada pelo Brazilian Global Atmospheric Model (BAM), modelo numérico atmosférico desenvolvido pelo INPE/ CPTEC que entrou em execução em 2016. Este modelo gera como saídas campos meteorológicos representados por matrizes multidimensionais, na forma de arquivos no formato GRIB. Este trabalho tem como objetivo a manipulação e visualização dos dados meteorológicos gerados pelo BAM, utilizando as bibliotecas iris, gribapi e pygrib, desenvolvidas utilizando a linguagem Python. No sistema atual, a manipulação e visualização dos arquivos são realizadas principalmente através de shell scripts e do software Grid Analysis and Display System (GrADS). Nesta fase inicial, utilizou-se a saída do BAM referente ao dia 02/02/2017 e a biblioteca iris para visualizar informações relativas ao campo meteorológico temperatura do ar. Gerou-se gráficos referentes a distribuição da temperatura do ar sobre o globo terrestre, a variação da temperatura sobre uma determinada longitude e interpolação de dados. Esses resultados preliminares demonstram o potencial da biblioteca iris para manipulação e visualização de dados de saída de modelos meteorológicos. Os próximos passos incluem a exploração de outras funcionalidades da biblioteca iris. As mesmas atividades desenvolvidas com a iris também deverão ser desenvolvidas para a gribapi e pygrib, visando um estudo comparativo dessas bibliotecas.

¹ Centro de Previsão do Tempo e Estudos Climáticos – E-mail: {lucas.mallmann, eugenio.almeida, ivo.koga}@inpe.br

SUMÁRIO

1 – INTRODUÇÃO	5
2 - OBJETIVOS	6
2.1 OBJETIVO GERAL	6
2.2 OBJETIVOS ESPECÍFICOS	6
3 – FUNDAMENTAÇÃO TEÓRICA	7
3.1 – MODELOS ATMOSFÉRICOS	7
3.2 – BRAZILIAN GLOBAL ATMOSPHERIC MODEL (BAM)	8
3.3 – ARQUIVO GRIB	8
4 – MATERIAIS E MÉTODOS	11
5 – RESULTADOS PRELIMINARES	13
6 – CONCLUSÃO	15
7 - REFERÊNCIAS	16

1 – INTRODUÇÃO

Os resultados gerados em uma previsão numérica de tempo e clima por um modelo atmosférico, consiste em um conjunto de matrizes multidimensionais. A previsão global do INPE/CPTEC é gerada pelo *Brazilian Global Atmospheric Model* (BAM), um modelo numérico atmosférico global que entrou em vigor a partir de 2016 e que gera previsões do tempo para a América do Sul.

A geração da previsão consiste em um *workflow* que coleta informações meteorológicas, altera sua resolução espacial, executa o modelo de previsão numérica no supercomputador Cray-XE6 e gera figuras correspondentes aos campos meteorológicos para a análise por meteorologistas e visualização na página do INPE/CPTEC².

Atualmente esse *workflow* consiste em um conjunto de *shell scripts*, sendo um deles responsável pela execução do software de visualização científica *Grid Analysis and Display System* (GrADS³) para a geração de figuras. Neste processo, a interação ocorre através de arquivos e também é necessária a segmentação dos processos, o que influencia no seu desempenho.

Em meteorologia e climatologia, a adoção de Python é observada nos grandes centros de meteorologia mundiais, que desenvolveram bibliotecas para a manipulação e visualização de dados meteorológicos: NOAA (pygrib)⁴, UK Met Office (íris⁵) e ECMWF (gripapi⁶). No contexto do CPTEC, são utilizados diversos artefatos, dentre eles: IDL, Matlab, *shell script*, GrADS. A utilização de Python se mostra adequada em nosso ambiente pois poderemos integrar todas essas funcionalidades em um ambiente único.

Este trabalho apresenta a manipulação e visualização dos dados gerados pelo BAM através da linguagem Python, como alternativa a programação utilizando *shell scripts* e GrADS.

³ Disponível em: http://cola.gmu.edu/grads/

⁵ Disponível em: https://scitools.org.uk/iris

² Disponível em: www.cptec.inpe.br

⁴ Disponível em: https://pypi.python.org/pypi/pygrib

⁶ Disponível em: https://software.ecmwf.int/wiki/display/GRIB/Python/packag/gribapi.

2 - OBJETIVOS

OBJETIVO GERAL 2.1

O presente trabalho tem por objetivo analisar e manipular dados gerados pelo BAM utilizando bibliotecas Python.

2.2 OBJETIVOS ESPECÍFICOS

- Analisar e avaliar bibliotecas para a manipulação e visualização de dados atmosféricos: pygrib ⁷, iris ⁸ e gribapi ⁹.
- Definir e adaptar a biblioteca adequada
- Manipular e visualizar as saídas do modelo BAM (formato GRIB)

⁷ Disponível em: (https://pypi.python.org/pypi/pygrib).

B Disponível em: (https://scitools.org.uk/iris).
 Disponível em: (https://scitools.org.uk/iris).
 Disponível em: (https://scitools.org.uk/iris).

3 – FUNDAMENTAÇÃO TEÓRICA

3.1 – MODELOS ATMOSFÉRICOS

Modelos numéricos atmosféricos são modelos matemáticos construídos utilizando um conjunto completo de equações dinâmicas primitivas que regem movimentos atmosféricos, e leis físicas para prever as condições futuras da atmosfera.

O conceito de equações primitivas baseia-se em identificar a origem de uma função, ou seja, em saber de qual função provém a equação que estamos a trabalhar. Os modelos numéricos atmosféricos utilizam as equações dinâmicas, ou seja, que trabalham com a origem dos movimentos (LMA, 2015).

Eles podem representar de uma maneira simples apenas o estado dinâmico da atmosfera, ou conter os processos físicos que atuam sobre ela. Podem prever fenômenos de grande escala, como tornados, precipitação e fluxos de vento. Os modelos numéricos atmosféricos geram as saídas em forma de matrizes de cinco dimensões, sendo que as três primeiras dimensões correspondem aos dados, a quarta corresponde à variável e quinta ao tempo.

A atmosfera é complexa, porém os modelos atmosféricos devem ser capazes de representar todas as relações possíveis entre seus componentes, ou seja, a relação entre os ambientes e suas variáveis. Os modelos de previsão de tempo e clima descrevem o comportamento da troposfera e estratosfera, a partir de condições iniciais do estado da atmosfera e de condições de contorno da superfície (LMA, 2015).

As previsões são computadas utilizando equações matemáticas, que atuam sobre a física e a dinâmica da atmosfera, porém essas equações não são lineares e são impossíveis de serem resolvidas com total exatidão.

Alguns processos físicos não podem ser calculados explicitamente pelos modelos numéricos. Isso acontece devido ao fato de o processo acorrer em uma escala menor que a do modelo. Portanto, eles precisam ser representados de uma maneira que o modelo consiga compreender. Essa maneira se chama parametrização, ou seja, é a representação dos processos físicos que ocorrem em uma escala espacial menor que a resolução do modelo, por variáveis que podem ser descritas por ele. Cada processo físico é representado por equações que indicam as parametrizações.

A parte dinâmica é representada pelas equações primitivas em que são utilizadas.

- Segunda Lei de Newton (Equação do Movimento)
- Equação do gás ideal (Equação de Estado)

- Primeira Lei da Termodinâmica
- Conservação de Massa

Contudo, existem diversos tipos de modelos numéricos com os quais é possível obter um valor aproximado dessas equações.

Os modelos globais e regionais são um dos tipos de modelos numéricos atmosféricos, sendo que os modelos regionais dependem dos globais. Cada modelo atmosférico pode utilizar um diferente modelo numérico com o objetivo de se utilizar o modelo mais apropriado.

3.2 – BRAZILIAN GLOBAL ATMOSPHERIC MODEL (BAM)

O BAM é um modelo atmosférico de circulação global que substituiu o antigo Modelo de Circulação Geral Atmosférico (*Atmospheric Global Circulation Model* – AGCM) do INPE/CPTEC. Ele entrou em vigor a partir de 2016, substituindo o AGCM que estava em operação desde 2010, e possui uma resolução de 20 quilômetros e 96 camadas na vertical.

O modelo (AGCM) era uma versão modificada e melhorada do modelo global *Center for Ocean-Land-Atmosphere Studies* (COLA) dos Estados Unidos, que entrou em vigor na inauguração do CPTEC, em 1994. Porém, devido à necessidade de melhorar as previsões de tempo e clima em altas resoluções na América Latina, foi desenvolvido um novo modelo global próprio mais adequado.

A partir de sua adoção, houve uma maior confiabilidade e qualidade nas previsões geradas pelo CPTEC. Esse modelo gera quatro saídas diárias nos períodos 00, 06, 12 e 18 UTC na forma de campos meteorológicos representados como matrizes bidimensionais e tridimensionais com informações sobre o vento, temperatura, entre outros.

3.3 – ARQUIVO GRIB

O GRIB é um formato de arquivo padrão do *World Meteorological Organization* (WMO) utilizado para o armazenamento e troca de dados meteorológicos em grade. Além dos campos meteorológicos, cada arquivo contém informações sobre: resolução, data, variáveis, nível e centro originário. Ele é um formato criado para ser compacto e portável

através de arquiteturas computacionais. Com o passar dos anos, a WMO lançou três versões diferentes do padrão GRIB.

- GRIB 0 Hoje em dia é obsoleta, não tem mais suporte e é raramente usada.
- GRIB 1 Não é a versão mais atual, e também não há mais aprimoramentos.
 Porém devido ao seu vasto uso, ainda é aceita pela WMO.
- GRIB 2 Representa uma grande e significante modernização do padrão GRIB, sendo amplamente utilizada ultimamente nos grandes centros meteorológicos, como por exemplo: *Uk Met Office* (Serviço Nacional de Meteorologia do Reino Unido).

No formato GRIB os dados são descritos na própria mensagem, pois é um formato de arquivo auto descritivo. Um arquivo pode conter uma ou mais mensagens, em que cada uma contém diversas seções. As seções irão definir quais dados serão transmitidos em suas mensagens. Após a leitura das seções a parte seguinte da mensagem, que contém os dados, poderá ser lida. As características dos parâmetros já devem estar definidas na tabela padrão do WMO e cada arquivo conta com uma série de números que deve ser descrita no início, que correspondem às entradas na Tabela do Formato GRIB do manual do WMO.

O formato GRIB 1 apresenta algumas limitações. A transmissão e armazenamento de previsões de um longo prazo não é muito eficiente, não possui conversor para os dados que forem requisitados, além de problemas na transmissão de dados multidimensionais. Porém, apesar de não ter mais suporte, ainda é aceito devido a sua vasta utilização em vários centros meteorológicos. A tabela 1 apresenta o formato do arquivo contento as suas diferentes seções.

Tabela 1. Formato do arquivo GRIB 1

SECTION 0	Indicator section
SECTION 1	Product definition section
SECTION 2	Grid description section
SECTION 3	Bit-map Section

SECTION 4	Binary Data Section
SECTION 5	7777 (End Section)

A seção 0 do GRIB 1apenas identifica a mensagem. A seção 1 define o horário ou os períodos de tempo, os parâmetros, os níveis e a grade utilizada. Pode ser uma grade pré-definida, referenciada por um número. Caso a grade não esteja previamente definida, é necessário que o arquivo contenha a seção 2, que irá definir a grade ideal para o arquivo de acordo com o manual padrão da WMO. A seção 3 contém o mapa de bits, ou "bit-map", que apesar de ser opcional é de extrema importância. Ele tem uma relação direta com os dados que estão em grade, pois caso o dado esteja presente em determinado ponto da grade, o bit tem valor atribuído a 1. Caso contrário o valor é 0. A seção 4 começa pela descrição do sistema de encapsulamento utilizado, pois o arquivo GRIB trabalha com os dados comprimidos, sendo essa uma de suas vantagens. A seção número 5 apresenta uma indicação que o arquivo chegou ao fim.

O GRIB 2 foi uma evolução do GRIB 1, na qual houveram grandes melhorias. Além da autodescrição ter sofrido um grande aprimoramento, houve também melhorias quanto a sua flexibilidade e expansão, as quais facilitam a sua manipulação. Com a autodescrição, os dados são descritos na mensagem do próprio arquivo. A sua flexibilidade permite a troca dos arquivos em várias arquiteturas diferentes de computador sem problema, e a expansão permite fazer previsões de períodos maiores que o GRIB 1. O GRIB 2 é mais eficiente quanto à compressão do arquivo e também na extração de dados, podendo guardar informações sobre previsões de um período maior. A tabela 2 apresenta a estrutura do arquivo GRIB 2.

Tabela 2. Formato do arquivo GRIB 2

SECTION 0	Indicator Section
SECTION 1	Identification Section
SECTION 2	Identification Section

SECTION 3	Grid Definition Section
SECTION 4	Product Definition Section
SECTION 5	Data Representation
	Section
SECTION 6	Bit-map Section
SECTION 7	Binary Data Section
SECTION 8	End Section

A seção 0 GRIB 2, assim como a do GRIB 1, é uma seção curta que apenas identifica a mensagem. A seção 2 no GRIB 2 é opcional, e pode ser usada para transmitir qualquer tipo de informação ou parâmetros. A seção 3 define qual grade será utilizada. A seção 4 irá definir o parâmetro. Na seção 5, é definido o modo de encapsulamento utilizado, assim como a seção 4 do GRIB 1. A seção 6 define a presença de dados em cada ponto da grade, porém o dado está presente na seção 7.

4 – MATERIAIS E MÉTODOS

Este projeto teve como base o estudo dos arquivos em formato GRIB, gerados pelo BAM, utilizando as bibliotecas: iris, gribapi e pygrib, da linguagem Python. Os arquivos foram obtidos através de acesso aos dados de saída do modelo pelo acesso ao supercomputador do CPTEC, e foi dado o início da instalação das bibliotecas necessárias à manipulação destes arquivos.

Para a instalação das bibliotecas foi utilizado o software Anaconda¹⁰. Ele é uma ferramenta de gerenciamento de bibliotecas e ambientes virtuais, que tem como principal vantagem o fato de vários pacotes científicos já estarem inclusos em sua instalação, o que facilita e agiliza o desenvolvimento. Outra grande vantagem é que o Anaconda permite trabalhar com diversos ambientes virtuais, e cada um deles com um interpretador Python diferente, para que seja possível desenvolver trabalhos diversos sem precisar mexer em um outro interpretador. Para o desenvolvimento deste trabalho, a versão Python 3.6 foi instalada em um computador, porém para a instalação das bibliotecas foi utilizada a versão 2.7, devido a sua maior compatibilidade com os pacotes necessários.

_

¹⁰ Disponível em.: https://www.continuum.io/downloads

Foram obtidos 2 arquivos no formato GRIB, gerados pelo BAM. Um dos arquivos está no formato GRIB 1, e o outro está no formato GRIB 2, com tamanhos de 368 MB e 97.5 MB respectivamente. Os dois arquivos contêm informações meteorológicas referentes à temperatura relativa, umidade do ar, entre outras. Ambos os arquivos são referentes à data de 02/02/2017, com saídas no horário 12 UTC.

Para o fluxo de análise dos dados meteorológicos, atualmente é realizada a aquisição dos dados utilizando scripts em *shell script* que acessam a web, ou modificar os parâmetros no código fonte de um modelo atmosférico, e rodar esse modelo para obter os dados desejados. Com os dados, ou os arquivos dos modelos, é possível escrever um programa para esses dados, utilizando IDL ou MATLAB a fim de conduzir uma análise estatística dos dados. Finalmente é possível visualizar os dados através de uma plotagem utilizando o GrADS. Em geral são utilizadas diversas ferramentas para a análise dos dados, sendo cada uma independente entre si e fazendo as comunicações através de arquivos.

Na linguagem Python, todas as ferramentas são utilizadas dentro de um único interpretador, e o fluxo dos dados não precisa acontecer através de diversos arquivos. É possível acessar uma variável em qualquer ferramenta, tornando o fluxo de trabalho mais robusto e flexível (LIN, 2012b).

Para a leitura dos arquivos, era necessária a instalação da biblioteca gribapi, a qual possui como dependência o compilador Fortran 90 e o g++, pois não era possível decodificar os arquivos do tipo GRIB 2 sem elas.

Após a instalação das bibliotecas, deu-se início aos estudos da documentação da biblioteca íris, que é uma ferramenta de fácil uso e sustentada por uma comunidade para a análise e visualização de dados meteorológicos e oceanográficos. A biblioteca providencia uma estrutura na qual as informações são armazenadas em cubos. Essa estrutura de cubo é uma representação da Convenção de Metadados CF (CF *Metadata Convention*¹¹). Um cubo descreve apenas um único fenômeno, e possui um nome, uma unidade, e um vetor que representa o dado da variável. Possui também coordenadas para representação do cubo no mundo real. O arquivo GRIB contém diversas informações de diversas variáveis, o que faz com que uma lista de cubos seja retornada de sua leitura, e a partir dessa lista de cubos, é possível manipulá-los individualmente. A partir da leitura,

-

¹¹ Disponível em: http://cfconventions.org/

pudemos obter gráficos utilizando a biblioteca Matplotlib, que já vem inclusa na biblioteca iris.

5 – RESULTADOS PRELIMINARES

Utilizando a biblioteca iris, foi possível fazer a leitura do arquivo GRIB 2, analisando seus campos e variáveis, e identificar os fenômenos meteorológicos nele. Após a leitura dos campos disponíveis, foi possível extrair campos específicos para a análise.

De acordo com a biblioteca iris, quando um arquivo é carregado, é retornada uma lista com todos os cubos possíveis. Analisando esses cubos, foi possível perceber que todos possuem coordenadas de latitude e longitude, correspondendo sempre a 1000 valores de latitude, e 2000 valores de longitude. A latitude apresenta seu valor máximo em 89.91° e seu valor mínimo com valor de –89.91°, enquanto a longitude apresenta seu valor máximo em 0.0° e seu valor máximo em 360°.

Manipulando o arquivo GRIB 2, os dados de temperatura do ar foram extraídos para serem analisados. Foram identificados 3 campos contendo informações sobre a temperatura do ar, porém possuíam coordenadas diferentes.

Na leitura do arquivo, foi possível obter a sua data de saída, que foi totalmente compatível com o esperado. Através de testes de verificação dos dados, utilizando a biblioteca *matplotlib*, foi possível visualizar graficamente os dados em relação a latitude ou longitude desejada.

A figura 1 apresenta um mapa com a distribuição da temperatura no globo terrestre a uma altitude de 2 metros. As temperaturas ficam mais baixas quando o azul se torna mais escuro e as temperaturas ficam mais altas quando o vermelho se torna mais escuro.

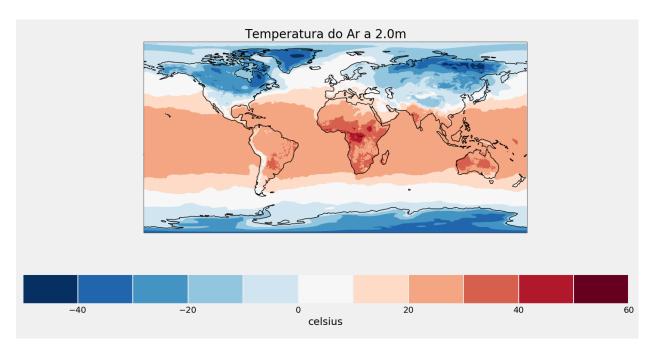


Figura 1 – Distribuição da temperatura do ar sobre o globo a 2.0 metros da superfície.

O gráfico da figura 2 apresenta a variação da temperatura do ar na latitude -72.27° em relação a longitude, que varia de 0° a 360° . Os valores são representados em graus Celsius, e apresentam apenas valores negativos para a latitude -72.27° .

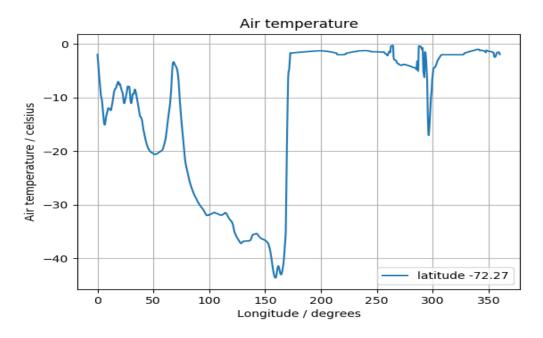


Figura 2. Variação da temperatura do ar de acordo com a longitude

Foi utilizado o método da interpolação linear para criar um novo conjunto de dados, partindo dos dados previamente conhecidos. Este método realiza a interpolação ligando cada um dos dados interpolados utilizando uma reta. Na figura 3 é possível observar os dados interpolados e originais. Os valores originais estão representados pelos pontos pretos, e os pontos vermelhos são os valores interpolados. Foi feita uma extração do valor original para analisar apenas as latitudes de 80° a 90°, representadas pelo eixo Y. No eixo X estão os valores da temperatura, em Celsius, referentes à latitude 0.0°. É possível perceber que os dados diferem muito pouco, e a interpolação se mostra aparentemente precisa.

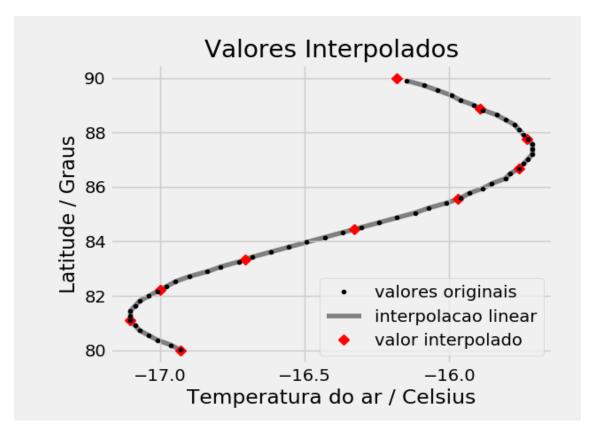


Figura 3. Interpolação linear da temperatura do ar

6 – CONCLUSÃO

Este trabalho, iniciado em abril de 2017, teve como objetivo a manipulação e visualização de dados gerados pelo modelo atmosférico BAM utilizando a biblioteca iris da linguagem Python. Foi possível visualizar a distribuição do campo meteorológico

relativo a temperatura do ar, além de visualizar a variação da temperatura sobre a latitude, e realizar a comparação dos valores interpolados pela biblioteca com os valores originais do arquivo.

Cada arquivo GRIB possui tabelas específicas para definir as suas variáveis. Cada centro meteorológico possui uma tabela própria que caracterizam os seus dados. No INPE/CPTEC, a tabela que define essas informações é a de número 46. Apesar disso, ela não está catalogada nas bibliotecas Python do ECMWF, UK Met Office e NOAA. Isso acarreta em dados desconhecidos na leitura dos arquivos, mostrando a necessidade da inclusão da tabela 46 nas bibliotecas.

Em trabalhos futuros iremos avaliar as outras bibliotecas disponíveis, como a pygrib e griapi pois os resultados presentes foram obtidos apenas utilizando a biblioteca íris. Serão realizados os testes para a análise dos dados, verificando também a possibilidade de incluir a tabela do CPTEC nas bibliotecas.

7 - REFERÊNCIAS

Dey, C. (2007).Guide to the WMO Table Driven Code Form Used for the Representation and Exchange of Regularly Spaced Data In Binary Form: FM 92 GRIBEdition2.

Disponível em [https://www.wmo.int/pages/prog/www/WMOCodes/Guides/GRIB/GRIB2 062006.pdf].

WMO. (2003). **Introduction to GRIB Edition 1 and GRIB Edition 2**. Disponível em [https://www.wmo.int/pages/prog/www/WMOCodes/Guides/GRIB/Introduction GRIB 1-GRIB2.pdf].

LIN,J. W. B (2012a). A Hands-On Introduction to Using Python in the Atmospheric and Oceanic Sciences. Disponível em [http://home.chpc.utah.edu/~u0035056/python/a hands on introduction to using pyth on in the atmospheric and oceanic sciences.pdf].

LIN,J. W. B (2012b). **Why Python Is the Next Wave in Earth Sciences Computing**. Disponível em [http://journals.ametsoc.org/doi/pdf/10.1175/BAMS-D-12-00148.1].

Laboratório de Modelagem Atmosférica (LMA). **NÍVEIS**. 2015. Disponível em: [http://lma.cptec.inpe.br/niveis/nivel1.php]. Acesso em: jul. 2017.