



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

FERRAMENTAS COMPUTACIONAIS PARA TRATAR A QUALIDADE PARA
DADOS OBTIDOS ATRAVÉS DE SENSORES

Guilherme Alvarenga Dias

Relatório de Iniciação Científica do
Programa PIBIC, orientada pelo Dr.
Alan James Peixoto Calheiros e pelo.
Dr. Pedro Luiz Pizzigatti Corrêa.

INPE

São José dos Campos

2021

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3945-6923/6921

Fax: (012) 3945-6919

E-mail: pubtc@sid.inpe.br

**COMISSÃO DO CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO
DA PRODUÇÃO INTELECTUAL DO INPE (DE/DIR-544):****Presidente:**

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Membros:

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

Dr. Amauri Silva Montes - Coordenação Engenharia e Tecnologia Espaciais (ETE)

Dr. André de Castro Milone - Coordenação Ciências Espaciais e Atmosféricas
(CEA)

Dr. Joaquim José Barroso de Castro - Centro de Tecnologias Espaciais (CTE)

Dr. Manoel Alonso Gan - Centro de Previsão de Tempo e Estudos Climáticos
(CPT)

Dr^a Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr. Plínio Carlos Alvalá - Centro de Ciência do Sistema Terrestre (CST)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

Clayton Martins Pereira - Serviço de Informação e Documentação (SID)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Simone Angélica Del Ducca Barbedo - Serviço de Informação e Documentação
(SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

EDITORAÇÃO ELETRÔNICA:

Marcelo de Castro Pazos - Serviço de Informação e Documentação (SID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SID)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

FERRAMENTAS COMPUTACIONAIS PARA TRATAR A QUALIDADE PARA
DADOS OBTIDOS ATRAVÉS DE SENSORES

Guilherme Alvarenga Dias

Relatório de Iniciação Científica do
Programa PIBIC, orientada pelo Dr.
Alan James Peixoto Calheiros e pelo
Dr. Pedro Luiz Pizzigatti Corrêa.

INPE

São José dos Campos

2021

Dados Internacionais de Catalogação na
Publicação (CIP) Sobrenome, Nomes.

Cutter Ferramentas Computacionais Para Tratar A Qualidade Para Dados Obtidos Através De Sensores / Nome Completo do Autor1; Nome Completo do Autor2. – São José dos Campos : INPE, 2021.

xiii + 27 p. ; (aa/bb/cc/dd)

Dissertação ou Tese (Mestrado ou Doutorado em Nome do Curso) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, AAAA.

Orientador : José da Silva.

1. Palavra chave. 2. Palavra chave 3. Palavra chave. 4. Palavrachave. 5. Palavra chave I. Título.

CDU 000.000



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 NãoAdaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).

**ATENÇÃO! A FOLHA DE
APROVAÇÃO SERÁ IN-
CLUIDA**

POSTERIORMENTE.

Mestrado ou Doutorado em
Nome do Curso

RESUMO

Este projeto de iniciação científica teve como objetivo o desenvolvimento de algoritmos computacionais capazes de tratar dados brutos provenientes de sensores meteorológicos mantidos pelo Instituto Nacional de Pesquisas Espaciais (INPE), processá-los para um padrão de dados aceito internacionalmente e validá-los. Também foram desenvolvidos algoritmos para a interpretação gráfica das grandezas primárias medidas e para verificar a consistência dos dados de modo amigável ao usuário final. Dentre os sensores de chuva do INPE, foram analisados os dados provenientes dos disdrômetros de impacto, Joss-Waldvogel, e do disdrômetro a laser, PARSIVEL, ambos importantes na análise microfísica da precipitação. Os algoritmos de processamento de dados foram construídos para interpretar dados brutos gerados pelos sensores, levando em consideração dados faltantes, e converter o formato proprietário em *Network Common Data Form* (NetCDF4), amplamente utilizados pela comunidade científica. Esses arquivos seguem o padrão do *Atmospheric Radiation Measurement* (ARM/EUA), que normaliza a estrutura dos dados para algo aceito internacionalmente e sob alto rigor de qualidade. A avaliação dos dados processados por esses algoritmos foi feita com o uso de métricas bem difundidas na comunidade científica. De modo a avaliar a qualidade das informações dos equipamentos aqui relacionados, foi realizada inicialmente uma intercomparação entre sensores colocalizados para um experimento de campo liderado pelo INPE, o SOSCHUVA. Foi também utilizado um instrumento de referência para chuva, o pluviômetro. Os disdrômetros colocalizados no sítio de Jaquariúna-SP durante a campanha mostraram alta correlação, 0,9 entre as taxas de chuva medidas pelo pluviômetro e pelo disdrômetro Joss-Waldvogel e de 0,92 entre o pluviômetro e o disdrômetro PARSIVEL, indicando uma alta confiabilidade nas medidas realizadas por esses instrumentos e nos algoritmos desenvolvidos. Assim, visamos executar esses algoritmos em outras campanhas do INPE, como o CHUVA, GoAmazon e ATTO, facilitando o seu processamento e distribuição. Espera-se com essas ferramentas otimizar os processamentos dos dados e acelerar as pesquisas realizadas com estes instrumentos, trazendo visibilidade à instituição por meio da qualidade das informações.

Palavras-chave: Ciência dos dados. Qualidade de Dados. Dados meteorológicos.

ABSTRACT

This undergraduate research project aimed to develop computational algorithms capable of treating raw data originating from meteorological sensors maintained by the Instituto Nacional de Pesquisas Espaciais (National Institute for Space Research - INPE), to process that data to an internationally recognized data standard and to validate it. Algorithms for the graphical interpretation of the primary variables of the data were also developed, along with additional algorithms designed to verify its consistency in a user-friendly way. Among INPE's climate sensors, the data analyzed came from Joss-Waldvogel impact disdrometers and from PARSIVEL laser disdrometers, both of which are important in the microphysical analysis of rain. The data processing algorithms were developed to interpret raw data generated by the sensors, taking missing data into account, and to convert their proprietary file format into a Network Common Data Form (NetCDF4) format, which is widely used by the international scientific community. These files follow the Atmospheric Radiation Measurement (ARM/USA) data standards, which standardize the structure of the data to a high degree of quality. The data processed by these algorithms was evaluated with commonly used metrics in the wider scientific community. In order to evaluate the quality of the information obtained from the previously mentioned sensors, data from colocalized sensors belonging to the INPE maintained SOSCHUVA campaign was intercompared. A pluviometer was also used as a reference instrument for rain measurements. The colocalized disdrometers located at the Jaquariúna-SP site showed high correlations, namely, of 0.9 between the Joss-Waldvogel and pluviometer measurements and of 0.92 between the PARSIVEL and pluviometer sensors, which indicates a high degree of reliability in the measurements made by those sensors and in the data processing algorithms. We wish to use these algorithms in other INPE campaigns, such as CHUVA, GoAmazon and ATTO, facilitating the processing and distribution of data. With these tools, we hope to optimize the processing of meteorological data and to accelerate research developed with these instruments, bringing attention to the institution by its high data quality.

Keywords: Data Science. Data Quality. Meteorological Data.

LISTA DE TABELAS

	<u>Pág.</u>
1 Variáveis principais do disdrômetro de impacto Joss-Waldvogel geradas pelo algoritmo de processamento de dados	23
2 Variáveis principais do disdrômetro a laser PARSIVEL geradas pelo algoritmo de processamento de dados	24
3 Diâmetros e velocidades de queda brutas para cada classe de gota. Valores relativos ao instrumento PARSIVEL.....	24
4 Diâmetros e velocidades de queda brutas para cada classe de gota. Valores relativos ao instrumento Joss-Waldvogel	25

LISTA DE FIGURAS

	<u>Pág.</u>
1 Cabeçalho (header) do arquivo NetCDF de um disdrômetro a laser seguindo o padrão ARM e aberto utilizando-se a biblioteca NetCDF4 em Python.	4
2 Primeiras colunas de um arquivo de dados brutos obtidos pelo disdrômetro de impacto Joss-Waldvogel.	7
3 Primeiras linhas de um arquivo típico do disdrômetro a laser PARSIVEL. Medições referentes ao instante 15:09:00 na data 03/03/2014.	9
4 Gráfico de dispersão de medições de taxa de chuva (mm.h^{-1}) obtidas de um pluviômetro de báscula e de um disdrômetro de impacto na campanha SOSCHUVA.	14
5 Gráfico de dispersão de medições de taxa de chuva (mm.h^{-1}) obtidas de um pluviômetro de báscula e de um disdrômetro a laser na campanha SOSCHUVA.	15
6 Gráfico de dispersão de medições de taxa de chuva (mm.h^{-1}) obtidas de um pluviômetro de báscula e de um disdrômetro a laser na campanha SOSCHUVA.	16
7 Gráfico da precipitação medida por um disdrômetro Joss-Waldvogel na campanha GoAmazon, em Manacapuru-AM, durante o dia 04/02/2014, gerada pelo algoritmo desenvolvido.	17
8 Gráfico da taxa de chuva medida por um disdrômetro Joss-Waldvogel na campanha GoAmazon, em Manacapuru-AM, durante o dia 04/02/2014, gerada pelo algoritmo desenvolvido.	17
9 Gráfico da refletividade radar medida por um disdrômetro PARSIVEL na campanha GoAmazon, em Manacapuru-AM, durante o dia 04/03/2014, gerada pelo algoritmo desenvolvido.	18
10 Gráfico da taxa de chuva medida por um disdrômetro PARSIVEL na campanha GoAmazon, em Manacapuru-AM, durante o dia 04/03/2014, gerada pelo algoritmo desenvolvido.	18

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
1.1 Objetivos.....	2
1.2 Atividades desenvolvidas durante vigência da bolsa	2
2 DESENVOLVIMENTO	3
2.1 Padrão de dados ARM.....	3
2.2 Padrão de nomenclatura de arquivos.....	4
2.3 Representação de <i>flags</i> de qualidade de dados	5
2.3 Características dos sensores meteorológicos	5
2.5 Formato dos dados brutos – Disdrômetro Joss-Waldvogel.....	6
2.6 Formato dos dados brutos – Disdrômetro a Laser PARSIVEL.....	8
2.7 Algoritmos de leitura e de geração de arquivos netcdf	10
2.8 Disdrômetro Joss-Waldvogel.....	10
2.9 Disdrômetro PARSIVEL.....	12
2.10 Validação dos dados.....	13
2.11 Visualização dos dados processados	16
3 CONCLUSÕES	19
3.1 Trabalhos Futuros	19
REFERÊNCIAS BIBLIOGRÁFICAS.....	21
APÊNDICE A – TABELAS DAS VARIÁVEIS PRINCIPAIS DOS INSTRUMENTOS	23

1 INTRODUÇÃO

O processamento otimizado e a verificação da qualidade de dados provenientes de sensores meteorológicos são essenciais para a aceleração de pesquisas realizadas com esses instrumentos, da mesma forma, a conversão desses dados para padrões aceitos internacionalmente, sob alto rigor de qualidade, facilita em sua distribuição. Nesse sentido, durante o período de desenvolvimento do projeto, foram criados algoritmos computacionais na linguagem Python capazes de ler arquivos brutos, em formato de texto, de dados gerados por dois tipos de sensores meteorológicos mantidos pelo Instituto Nacional de Pesquisas Espaciais (INPE), o disdrômetro de impacto Joss-Waldvogel e o disdrômetro a laser PARSIVEL, e convertê-los para um desses padrões.

Os algoritmos desenvolvidos convertem os formatos proprietários dos dados desses dois instrumentos em *Network Common Data Form* (NetCDF4), um formato de arquivo amplamente utilizado para dados meteorológicos. Em particular, foi escolhido o padrão do *Atmospheric Radiation Measurement* (ARM/EUA) para os dados, devido principalmente ao seu alto grau de aceitação na comunidade acadêmica internacional. O ARM é um programa do governo norte-americano que promove estudos relacionados a fatores que afetam o clima. O ARM distribui sensores em *sites* ao redor do mundo, que coletam dados e os transmitem para um sistema de *Data Publishing* (ARM, 2020) com um volume total em torno de 10 PB (*petabytes*).

Tal padrão especifica, dentre outras características, o formato dos dados para cada grandeza principal medida pelos instrumentos, o formato das variáveis de qualidade de dados e a nomenclatura de variáveis.

Esse processo de conversão leva em consideração dados faltantes no período de amostragem dos arquivos processados e, uma vez gerados os arquivos no formato NetCDF4, foram realizadas análises para verificar a consistência e confiabilidade dos dados processados. Foram utilizadas medições feitas durante as campanhas GoAmazon, em Manacapuru, Amazonas e SOSCHUVA em Campinas, São Paulo.

Por fim, foram também desenvolvidos *scripts* simples e amigáveis ao usuário que criam gráficos para as grandezas principais dos sensores utilizados a partir dos arquivos NetCDF4, facilitando a visualização dessas informações.

1.1 Objetivos

De acordo com o plano de trabalho deste projeto de pesquisa, o objetivo geral deste estudo foi desenvolver ferramentas que tratem da qualidade de dados obtidos por sensores que coletam dados sobre precipitação e nuvens em experimentos de campo com participação do INPE

Os objetivos específicos são apresentados a seguir e dizem respeito à criação de ferramentas adaptadas para os diferentes sensores de chuva mantidos pelo INPE, levando em consideração a estrutura de seus dados brutos e outras especificidades.

- a) Desenvolver uma ferramenta de leitura de e processamento para dados brutos do disdrômetro de impacto Joss-Waldvogel (JOSS;WALDVOGEL, 1967);
- b) Desenvolver uma ferramenta de leitura e de processamento para dados brutos do disdrômetro a laser PARSIVEL (ARM, 2020);
- c) Desenvolver ferramentas para avaliar a qualidade dos dados processados por meio de intercomparações entre instrumentos colocalizados e para gerar gráficos das suas grandezas principais.

1.2 Atividades desenvolvidas durante vigência da bolsa

Durante o período de vigência da bolsa, foram desenvolvidas as seguintes atividades.

- a) Desenvolvido algoritmo em Python para processar dados do disdrômetro de impacto Joss-Waldvogel, gerando arquivos NetCDF4 diários seguindo o padrão ARM;
- b) Desenvolvido algoritmo em Python para processar dados do disdrômetro a laser PARSIVEL, gerando arquivos NetCDF4 diários seguindo o padrão ARM;
- c) Desenvolvido algoritmo em Python para ler os dados no formato NetCDF4 e gerar gráficos das variáveis;
- d) Desenvolvido algoritmo em Python para verificar a qualidade dos dados processados, utilizando métricas de correlação e bias, gerando-se também um gráfico de dispersão com os dados da intercomparação entre instrumentos.

2 DESENVOLVIMENTO

2.1 Padrão de dados ARM

Como mencionado acima, o padrão escolhido para armazenar os dados processados a partir das medições brutas dos instrumentos foi aquele utilizado pelo *Atmospheric Radiation Measurement* (ARM, 2020), no que segue é feita uma descrição geral de suas exigências.

Em termos do formato de arquivo esperado para os dados processados, é preferido o formato NetCDF3 ou NetCDF4, um tipo de formato de arquivo utilizado amplamente para armazenamento de dados científicos graças à sua capacidade de lidar eficientemente com dados em vetores (*arrays*) uni ou multidimensionais e, adicionalmente, por ser suportado por uma variedade de bibliotecas em diferentes linguagens de programação, como C, C++, Python (Unidata, 2020), além de oferecer a possibilidade de adicionar os metadados diretamente nos arquivos.

Internamente, o arquivo em formato NetCDF deve apresentar uma série de *flags* que descrevam, por exemplo, o instrumento do qual foram retirados os dados, o intervalo de amostragem utilizado, o sítio no qual o instrumento se encontra, dentre outras características dos dados. Em especial, devem também ser especificadas as dimensões dos dados armazenados, que serão usadas para indexar essas informações. Como exemplo, no padrão de dados ARM, um disdrômetro a laser deve apresentar o tempo (*time*), a classe de gotas (*drop class*) e as velocidades de queda (*fall velocity*) como dimensões para indexar vetores multidimensionais como o espectro bruto de gotas, *raw_spectrum*, seguindo a nomenclatura ARM.

Por fim, também existem exigências quanto à nomenclatura das variáveis, bem como o tipo de dados para cada uma. A maior parte das variáveis devem ser tratadas como números de ponto flutuante de 32 bits, *float32*, enquanto outras como as variáveis temporais, por exemplo, devem ser dadas como números de ponto flutuante de 64 bits, *float64*.

Tabelas com descrições completas das *flags* e variáveis para os instrumentos Joss-Waldvogel e PARSIVEL estão disponíveis no Apêndice A deste documento. Abaixo é

apresentado um exemplo do cabeçalho de um arquivo NetCDF seguindo o padrão descrito.

Figura 1 - Cabeçalho (header) do arquivo NetCDF de um disdrômetro a laser seguindo o padrão ARM e aberto utilizando-se a biblioteca NetCDF4 em Python.

```
<class 'netCDF4._netCDF4.Dataset'>
root group (NETCDF4 data model, file format HDF5):
  description: Data collected from a laser disdrometer in the ATTO-Campina site
  site_id: att
  platform_id: parsivel
  facility_id: cam
  data_level: b0
  location_description: Amazon Tall Tower Observatory (ATTO), Amazonia, Brazil
  datastream: attlaserdisdcam.bo
  serial_number: N/A
  calib_date: N/A
  sampling_interval: 60 seconds
  averaging_interval: N/A
  dimensions(sizes): time(1440), drop_class(32), fall_vel(32)
  variables(dimensions): uint32 base_time(), float64 time_offset(time), float64 time(time), float32 rain_rate(time), float32
zdb(time), float32 drop_class(drop_class), float32 fall_vel(fall_vel), float32 delta_diam(drop_class), float32 raw_spectrum(time),
e,drop_class,fall_vel), float32 interval_sample(time), float32 err_code(time), float32 F_laser_area(), float32 lat(), float32 lon(),
float32 Alt()
  groups:
```

Fonte: Próprio autor (2021).

2.2 Padrão de nomenclatura de arquivos

O padrão de dados ARM também versa quanto aos nomes dos arquivos em formato NetCDF gerados. Estes devem ter o seguinte formato geral: (sss)(inst)(qualifier)(temporal)(Fn).(dl).(yyyymmdd).(hhmmss).nc. A primeira posição “sss” refere-se ao identificador do sítio onde o instrumento se encontra, já “inst” identifica o sensor meteorológico que produziu o dado, “yyyymmdd” informa o ano, mês e dia quando se iniciou a amostragem dos dados e “hhmmss” informa o horário desse instante em horas, minutos e segundos. Por fim, “dl” refere-se ao nível dos dados armazenados, por exemplo: 00, a0, b1, c1, c2, indicando níveis crescentes de processamento e de cálculos realizados sobre os dados brutos, maiores detalhes podem ser encontrados no manual de dados do ARM (ARM, 2020).

Os campos restantes, como “qualifier”, “temporal”, “Fn” são opcionais ou, no caso do último, utilizados apenas para sítios do ARM com suas próprias designações.

Como exemplo, o título de um arquivo gerado durante este projeto foi o seguinte: attimpactdisdcam.b0.20140201.000000.nc. Note que “att” refere-se ao sítio da ATTO (*Amazon Tall Tower Observatory*) de onde vieram os dados processados, “impactdis” identifica o instrumento utilizado, o disdrômetro de impacto e “cam” é o identificador do sítio da campina, a 10 km da torre ATTO, na região amazônica. O nível de dados deste

arquivo é “b0”, ou seja, apenas foram processados os dados brutos do sensor, sem serem realizados cálculos para adicionar outras variáveis além daquelas obtidas pelo instrumento. Por fim, os dados foram gerados à meia noite do dia 01/02/2014.

2.3 Representação de *flags* de qualidade de dados

O padrão de dados ARM também especifica o formato que as *flags* de qualidade de dados devem tomar nos arquivos NetCDF final, especificamente, eles devem contar com uma série de flags de qualidade de dados (*Quality Control – QC*) para todas suas medições principais, que representam na forma de bits valores de verdadeiro ou falso para testes que possam ter sido aplicados aos dados, sendo representados da seguinte maneira:

- a) $\text{bit}_1 = 00000001 = 0x01 = 2^0 = 1$, $\text{qc_bit}_1_description = \text{Valor igual a } -999$ (*missing_value*);
- b) $\text{bit}_2 = 00000010 = 0x02 = 2^1 = 2$, $\text{qc_bit}_2_description = \text{Valor é menor do que } valid_min$;
- c) $\text{bit}_3 = 00000100 = 0x04 = 2^2 = 4$, $\text{qc_bit}_3_description = \text{Valor é maior do que } valid_max$;
- d) $\text{bit}_4 = 00001000 = 0x08 = 2^3 = 8$, $\text{qc_bit}_4_description = \text{Diferença entre valor atual e anterior excede } valid_delta$.

Dessa forma, um valor nulo para todos os bits de uma variável $qc_{\langle\text{nome da variável}\rangle}$ indica que todos os testes de qualidade foram bem-sucedidos. Por outro lado, se essa mesma variável tiver valor 10, por exemplo, isso indicaria que essa medição falhou tanto o teste de valor mínimo quanto no de delta máximo. Por fim, vale notar ainda que operações *bitwise* podem ser utilizadas para selecionar erros específicos nos dados para serem analisados.

Os valores máximos, mínimos e de diferença mencionados acima estão disponíveis, para os instrumentos trabalhados neste projeto, no Apêndice A deste relatório.

2.3 Características dos sensores meteorológicos

Os sensores meteorológicos utilizados neste projeto, como mencionado anteriormente, foram o disdrômetro de impacto Joss-Waldvogel e o disdrômetro a laser PARSIVEL,

além de pluviômetros de balança utilizados nas etapas de validação dos dados processados.

O primeiro disdrômetro trata-se de um instrumento capaz de determinar a distribuição de tamanho de gotas da precipitação em um certo local ao transformar o impacto mecânico da queda de uma gota de água no sensor do dispositivo em um pulso elétrico (JOSS;WALDVOGEL, 1967). A amplitude desse pulso relaciona-se diretamente ao diâmetro da gota, permitindo que a distribuição seja obtida.

Além da distribuição de tamanho de gotas, o disdrômetro também é capaz de analisar uma série de dados relacionados à precipitação, como a velocidade de queda de cada gota, a taxa de precipitação e a refletividade radar.

O disdrômetro a laser, por sua vez, tem como função principal medir a distribuição de tamanhos de gotas da precipitação em um local, bem como suas velocidades, além de dados adicionais como a refletividade radar, visibilidade e precipitação nessa região. A medição das gotas é feita quando elas interrompem o feixe de um laser horizontal projetado pelo disdrômetro, quando isso ocorre, a intensidade óptica do laser medida pelo instrumento cai, de tal forma que o diâmetro da gota é proporcional a essa redução (LÖFFLER-MANG;JOSS, 2000).

Com essas informações, o instrumento é capaz de determinar a distribuição de classes de gotas, além de caracterizar o tipo de precipitação, como chuvisco, granizo, neve, dentre outros.

Como esses sensores foram desenvolvidos por diferentes empresas, o formato de seus dados brutos é distinto e proprietário, assim, desenvolvemos algoritmos independentes para processar os dados de cada um deles, levando em consideração os formatos como os dados são organizados internamente nos seus arquivos, que serão descritos a seguir.

2.5 Formato dos dados brutos – Disdrômetro Joss-Waldvogel

Os dados brutos do disdrômetro de impacto Joss-Waldvogel são armazenados em arquivos texto (ASCII) e, internamente, separados em colunas, que representam suas medidas, e temporalmente separados em linhas, cada nova linha sendo uma medição realizada após um dado período. Os algoritmos desenvolvidos neste projeto consideram

sempre uma resolução temporal de um minuto entre medições, ou seja, nos arquivos trabalhados observa-se uma nova linha a cada minuto desde o início das medições.

Na figura abaixo observam-se as primeiras linhas de um dos arquivos utilizados para testar o funcionamento do algoritmo desenvolvido para o disdrômetro de impacto. É possível observar que as primeiras duas colunas armazenam a data e o horário em que a medição foi feita. A terceira coluna é utilizada para armazenar informações de *status* do instrumento e a quarta informa a diferença temporal entre a medição atual e anterior que, como mencionado anteriormente, será sempre de 60 segundos nos arquivos trabalhados neste projeto.

Segundo o manual do instrumento, as colunas “n1” até “n20” contabilizam o número de gotas de uma dada classe detectadas pelo sensor, cujos diâmetros e velocidades de queda são documentadas, também, no manual do instrumento. As últimas três colunas, denominadas “RI [mm/h]”, “RA [mm]” e “RAT [mm]” armazenam, respectivamente a taxa de chuva, quantidade de chuva e precipitação total durante o período de amostragem.

Figura 2 – Amostra de dados de um arquivo bruto obtida pelo disdrômetro de impacto Joss-Waldvogel para o dia 02 de fevereiro de 2014 durante a campanha GoAmazon.

YYYY-MM-DD	hh:mm:ss	Status	Interval [s]	n1	n2	n3	n4	n5	n6	n7	n8	n9	n10
2014-02-02	13:55:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	13:56:00	60	1	0	0	0	0	0	0	0	0	0	0
2014-02-02	13:57:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	13:58:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	13:59:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:00:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:01:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:02:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:03:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:04:00	60	1	1	8	0	0	0	0	0	0	0	0
2014-02-02	14:05:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:06:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:07:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:08:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:09:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:10:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:11:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:12:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:13:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:14:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:15:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:16:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:17:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:18:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:19:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:20:00	60	0	0	0	0	0	0	0	0	0	0	0
2014-02-02	14:21:00	60	0	0	0	0	0	0	0	0	0	0	0

Fonte: Próprio autor (2021)

Esse formato permite que o algoritmo de processamento de dados leia cada arquivo do instrumento com auxílio da função “*read_csv*”, parte da biblioteca *Pandas* para manipulação de dados em Python. Como será detalhado adiante, cada linha é lida

considerando-se espaços em branco como separadores e então os dados numéricos são organizados em vetores dentro de um *Data Frame* com os dados de todos os arquivos processados até então.

2.6 Formato dos dados brutos – Disdrômetro a Laser PARSIVEL

O disdrômetro a laser PARSIVEL também armazena suas medições em um arquivo ASCII, contudo, sua formatação é distinta daquela vista para o disdrômetro de impacto. Para cada medida desde o início da amostragem (considerando, novamente, uma taxa de amostragem de um minuto) são escritas 98 linhas com todas as medições realizadas pelo instrumento, *flags* descrevendo a operação do sensor e outras informações relevantes ao instrumento .

Dentre essas 98 linhas, são de interesse para este trabalho apenas dez, que contêm as informações brutas necessárias para gerar o arquivo NetCDF, sendo as seguintes:

- a) Linha 1 – Taxa de chuva (mm.h^{-1});
- b) Linha 7 – Refletividade radar (dB);
- c) Linha 9 – Intervalo de amostragem (s);
- d) Linha 20 – Horário da medição (hh:mm:ss);
- e) Linha 21 – Data da medição (dd.mm.yyyy);
- f) Linha 25 – Código de erro (adimensional);
- g) Linha 93 – Espectro de gotas (*Raw Spectrum*, adimensional).

A linha 93 tem um formato diferente das outras, pois não é composta de um número inteiro ou em ponto flutuante, mas sim de uma matriz de dimensão 32x32 que apresenta o número de gotas de uma certa classe de diâmetro e de velocidade de queda em cada uma de suas posições. A posição nas colunas indica a classe de diâmetro das gotas detectadas, da classe 0 até 31, enquanto a posição nas linhas indica a velocidade de queda, novamente da classe 0 até a classe 31. Tabelas com os valores numéricos de cada classe estão disponíveis no Apêndice A deste relatório.

Abaixo é apresentado um exemplo das primeiras 28 linhas de um arquivo gerado por um instrumento PARSIVEL utilizado na campanha GoAmazon.

Como será detalhado adiante, este formato permite uma leitura do arquivo do PARSIVEL feita linha por linha, onde as linhas buscadas são identificadas pelos primeiros números em cada uma, como “01:” e “91:” para taxa de chuva e espectro de gotas respectivamente.

2.7 Algoritmos de leitura e de geração de arquivos netcdf

Os algoritmos de processamento dos arquivos brutos dos sensores estudados, bem como aqueles utilizados para gerar gráficos a partir das suas variáveis primárias e para realizar a validação dos dados foram desenvolvidos inteiramente na linguagem Python, especificamente em Python 3.9.5, com a utilização de bibliotecas auxiliares para manipulação de dados e operações no formato NetCDF.

O ambiente de programação majoritariamente utilizado para desenvolver os algoritmos foi o *Jupyter Notebook*, plataforma escolhida inicialmente por facilitar o projeto inicial e compartilhamento de códigos no formato de *Notebooks*. Durante o período final de testes e verificação dos algoritmos, foi utilizado também o *Visual Studio Code*.

Até o término do projeto, foram desenvolvidos cinco algoritmos, dois sendo os *scripts* principais de leitura para os arquivos de cada instrumento, outros dois sendo *scripts* contendo funções auxiliares para a escrita dos arquivos em NetCDF e um para realizar a validação dos dados processados.

Para que os *scripts* sejam executados corretamente, eles devem estar no mesmo diretório onde se encontram os arquivos brutos em ASCII dos instrumentos. Uma vez executados, serão gerados os arquivos “.nc” (NetCDF) seguindo o padrão ARM descrito anteriormente.

2.8 Disdrômetro Joss-Waldvogel

Foram escritos dois algoritmos para realizar a leitura e o processamento do instrumento Joss-Waldvogel, intitulados “Joss_DataProcess” e “Joss_Uutilities”.

O algoritmo “Joss_DataProcess” é responsável por realizar a leitura de todos os arquivos de texto brutos do disdrômetro. Tais arquivos não precisam necessariamente estar ordenados por data dentro do diretório dos dados, pois o código é capaz de organizar os dados cronologicamente, ademais, são aceitos eventuais dados faltantes, que são contabilizados e recebem o valor numérico de -99,0 no arquivo NetCDF final. O resultado

da execução do programa são arquivos em NetCDF diários com eventuais dados faltantes inseridos caso o arquivo dos dados brutos não seja diário, no caso de falha de medidas.

Vale notar ainda que esse algoritmo não realiza cálculos sobre os dados brutos, apenas os organiza seguindo os padrões de dados ARM, ou seja, chegando apenas ao nível de dados “b0” deste padrão.

Em termos de desenvolvimento, foram utilizadas diversas bibliotecas auxiliares disponíveis para Python, notadamente: *Numpy* para realizar operações numéricas simples, *Pandas* por oferecer estruturas de dados úteis para organizar os dados brutos, como *DataFrames*, *netCDF4* para realizar a escrita dos arquivos em formato “.nc” e outras padrões do Python para escrita de arquivos, como *pathlib*, e *datetime* para manipular formatos de datas.

Durante sua execução, o algoritmo “Joss_DataProcess” primeiramente detecta todos os arquivos de texto na pasta, que devem ser os arquivos brutos do disdrômetro, e na sequência os organiza em um vetor de arquivos, ordenando-os por data de criação. Em seguida, os dados dos arquivos nesse vetor são lidos com auxílio do leitor “read_csv” da biblioteca *Pandas* utilizando os espaços em branco no arquivo dos dados brutos como separadores.

Após a leitura, todos os dados são copiados para *DataFrames*, estruturas de dados também da biblioteca *Pandas* que são de grande auxílio na manipulação de dados multidimensionais e de diferentes tipos. Os dados são então ordenados por data e, caso eles não sejam diários, *flags* de valor “-99.0” são utilizadas para extendê-los, também sendo utilizadas em caso de dados faltantes, o que incrementa a qualidade desses dados em comparação aos brutos.

Após a construção dos *DataFrames*, é chamada a função “Joss_GenerateCDF”, que é declarada no arquivo “Joss_Uilities” mencionado anteriormente. Essa função é executada para cada *DataFrame* com os dados diários e é responsável por escrever os arquivos diários NetCDF4, com as dimensões, variáveis e tipos de dados exigidos pelo padrão ARM e por salvá-los no mesmo diretório onde se encontram os arquivos brutos originais.

Por fim, com a contração da função “Joss_GenerateGraph” foi possível gerar gráficos de todas as variáveis principais armazenadas nos arquivos em NetCDF4, com o auxílio da biblioteca gráfica *Matplotlib*, que serão apresentados adiante.

2.9 Disdrômetro PARSIVEL

Os algoritmos desenvolvidos para o instrumento PARSIVEL foram construídos de maneira semelhante àqueles do instrumento Joss-Waldvogel, o *script* “Pars_DataProcess” é responsável por realizar a leitura inicial do arquivo bruto do PARSIVEL e converter o formato dos dados de um padrão proprietário para dados armazenados em *DataFrames* de leitura simples, enquanto o *script* “Pars_Utilities” contém funções adicionais para realizar a escrita de arquivos em formato NetCDF e realizar a plotagem de variáveis principais. São utilizadas as mesmas bibliotecas dos algoritmos para o sensor Joss-Waldvogel.

Ao ser executado no diretório onde se encontram os arquivos brutos do sensor PARSIVEL, o *script* “Pars_DataProcess” detecta todos os arquivos e em seguida realiza uma leitura linha por linha de cada um deles, preenchendo vetores apenas com os dados nas linhas procuradas, mencionadas anteriormente na seção “Formato dos dados brutos – Disdrômetro a Laser PARSIVEL”. Vale notar que para as linhas que armazenam os 1024 valores das matrizes *Raw Spectrum* (ou espectro bruto em português) é realizada uma formatação para que esses dados possam ser lidos como matrizes de dimensões 32x32, através da função *reshape*, nativa do Python.

Em seguida, as informações de horário e data das medições são transformadas em objetos *datetime* de maneira a facilitar manipulações com dados temporais e, por fim, *DataFrames* com todos os dados extraídos são criados. Contudo, diferentemente do algoritmo “Joss_DataProcess”, o ordenamento por data, preenchimento de dados faltantes e geração de dados diários é realizado por uma função externa que é declarada no arquivo “Pars_Utilities”, chamada “PARS_CDFPerDay”.

Uma vez ordenados, os dados nos *DataFrames* são utilizados pela função “createCDF” para gerar os arquivos NetCDF diários de forma semelhante ao que foi feito para o disdrômetro de impacto. A plotagem das variáveis principais do sensor também é possível pela função “PARS_GenerateGraph”, declarada no arquivo “Pars_Utilities”. Ao final da

execução, serão obtidos todos os arquivos NetCDF diários na pasta onde se encontram os arquivos brutos.

2.10 Validação dos dados

Após a criação dos arquivos NetCDF, realizou-se uma etapa de validação dos dados nestes arquivos, de maneira a verificar a qualidade destes, fator essencial para garantir a confiabilidade do processamento realizado pelos algoritmos desenvolvidos e dos dados em si. Especificamente, foram intercomparadas variáveis comuns aos dois sensores, utilizando-se arquivos brutos colocalizados, ou seja, nos mesmos locais, de modo que seja razoável supor que ambos tenham sido submetidos às mesmas condições meteorológicas e ambientais.

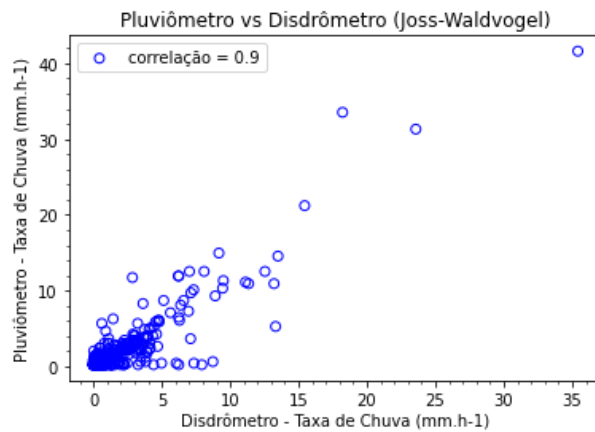
Para esse processo, também foram utilizados dados vindos de pluviômetros de balança, que são tidos como os mais confiáveis dentre os instrumentos estudados, seguidos do disdrômetro de impacto Joss-Waldvogel e, com menor grau de robustez dentre os instrumentos, o disdrômetro PARSIVEL.

As métricas estatísticas utilizadas para realizar essa validação foram a correlação, que mede o grau com o qual duas variáveis estão linearmente relacionadas e o *bias* ou viés, que representa a diferença entre o valor esperado de uma medição e o valor real obtido (DEVORE, 2006). Também foram criados gráficos de dispersão para visualmente indicar a relação entre as variáveis comparadas. O *script* utilizado para realizar essa análise utilizou da biblioteca “scipy” e “sklearn” para obter funções estatísticas para o cálculo da correlação e do viés.

Inicialmente, foram processados dados do programa SOSCHUVA (MACHADO, 2015), vindos de instrumentos colocalizados em Jaquariúna (SP). Foram comparadas medições de um disdrômetro a laser PARSIVEL e de um disdrômetro de impacto Joss-Waldvogel, após serem processadas pelos algoritmos, com medições de um pluviômetro a balança no mesmo período, o período de amostragem dos instrumentos em questão foi de aproximadamente um ano, entre setembro de 2016 e outubro de 2017

A análise dos dados processados pelos algoritmos, utilizando as métricas supracitadas, resultou nos seguintes gráficos, onde foram comparadas as medições da taxa de chuva entre instrumentos.

Figura 4 - Gráfico de dispersão de medições de taxa de chuva (mm.h^{-1}) obtidas de um pluviômetro de bscula e de um disdrmetro de impacto na campanha SOSCHUVA.

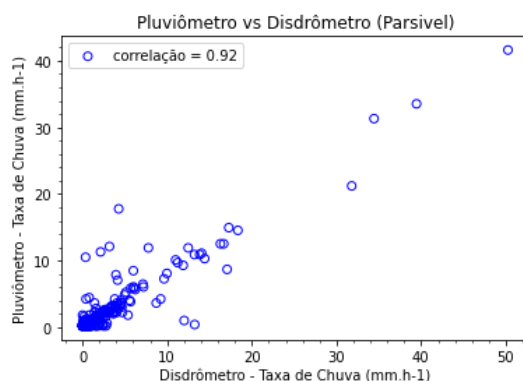


Fonte: Prprio autor (2021).

Pelo grfico apresentado acima,  possvel verificar visualmente e numericamente pelo valor de 0,9 de correlao que os dados apresentam alta relao entre eles. Alm disso, foi possvel verificar que os algoritmos de processamento de dados desenvolvidos apresentam os mesmos resultados que aqueles apresentados por Calheiros (CALHEIROS, 2018), pois mostraram os mesmos resultados ali obtidos para os dados do instrumento Joss-Waldvogel aps converso para o formato NetCDF4. Alm da correlao, tambm foi obtido um valor de *bias* de $-0,2 \text{ mm.h}^{-1}$, ou seja, h uma variao pequena entre os valores esperados da taxa de chuva, vindos do pluvimetro de bscula, e os valores obtidos de fato dos dados processados do disdrmetro de impacto.

A mesma anlise foi realizada entre o pluvimetro e o instrumento PARSIVEL, gerando o seguinte grfico.

Figura 5 - Gráfico de dispersão de medições de taxa de chuva (mm.h^{-1}) obtidas de um pluviômetro de balança e de um disdrômetro a laser na campanha SOSCHUVA.

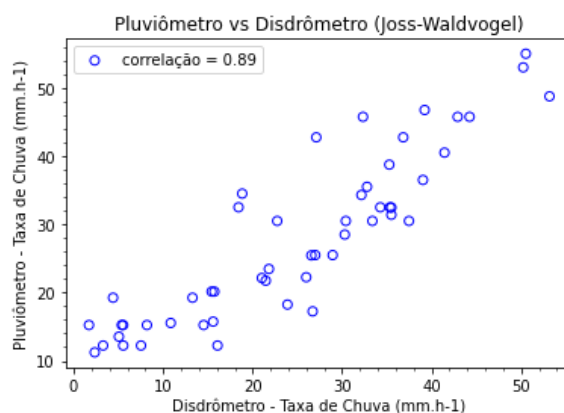


Fonte: Próprio autor (2021).

Novamente, foi obtido um valor alto de correlação (0,92) ao serem comparadas as medições de taxa de chuva (mm.h^{-1}) do pluviômetro com as medições vindas de arquivos processados em formato NetCDF do disdrômetro a laser PARSIVEL. O valor do *bias* calculado entre esses dados foi de 0.4 mm.h^{-1} , indicando pequena variação entre os dados do pluviômetro e do PARSIVEL. Esses fatos indicam a confiabilidade dos algoritmos desenvolvidos para o processamento de dados também para este instrumento.

Uma vez testados os dados para uma campanha cujo dados foram validados, a próxima etapa do trabalho foi realizar a execução dos *scripts* também para dados vindos de instrumentos da campanha GoAmazon, em Manacapuru, Amazonas. Em seguida apresenta-se o gráfico de dispersão com medidas de taxa de chuva feitas por um pluviômetro de balança e por um disdrômetro de impacto colocalizados, durante os dias 31 de janeiro e 2 de fevereiro de 2014.

Figura 6 - Gráfico de dispersão de medições de taxa de chuva (mm.h^{-1}) obtidas de um pluviômetro de balança e de um disdrômetro a laser na campanha SOSCHUVA.



Fonte: Próprio autor (2021).

De forma semelhante aos dados da campanha SOSCHUVA, observa-se que os dados dos sensores de Manacapuru estão fortemente correlacionados, com taxa de correlação de 0,89. O valor do *bias* entre essas medições, por sua vez, foi de $-3,05 \text{ mm.h}^{-1}$, ou seja, um pouco maior do que aqueles vistos antes, mas ainda pequeno frente aos valores das medições. Assim, é possível afirmar que os algoritmos são igualmente confiáveis no processamento de dados provenientes dos sensores localizados para outras campanhas do INPE.

2.11 Visualização dos dados processados

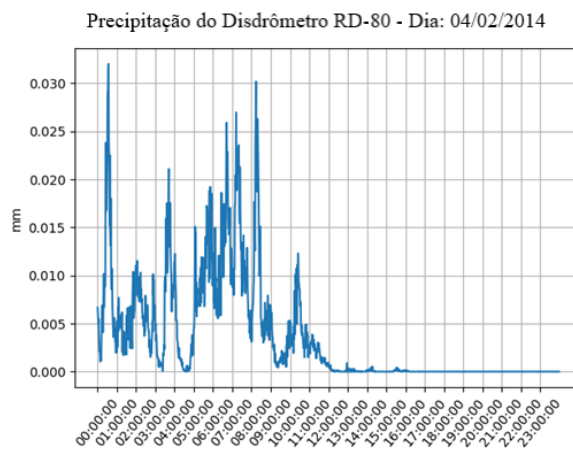
Além dos algoritmos de processamento de dados desenvolvidos para os disdrômetros, foram criados paralelamente *scripts* mais simples e amigáveis ao usuário final responsáveis por visualizar graficamente as variáveis primárias destes sensores a partir da leitura dos arquivos NetCDF gerados. Estes algoritmos podem ser utilizados diretamente após a criação dos arquivos NetCDF, internamente aos algoritmos “Pars_DataProcess” e “Joss_DataProcess”, ou individualmente.

Quando executados, os algoritmos irão detectar os arquivos NetCDF presentes na pasta onde se encontram e então salvarão as imagens dos gráficos em pastas com títulos no formato: (inst)Plots(dd)(mm)(yy), onde “inst” é um identificador do instrumento do qual os dados foram retirados e “dd”, “mm” e “yy” dizem respeito ao dia, mês e ano em que a medição foi realizada. As imagens em si seguem um padrão de nomenclatura semelhante:

(var)(dd)(mm)(yy), onde “var” indica qual a variável plotada na figura. Foi utilizada a biblioteca *matplotlib* para a geração das imagens.

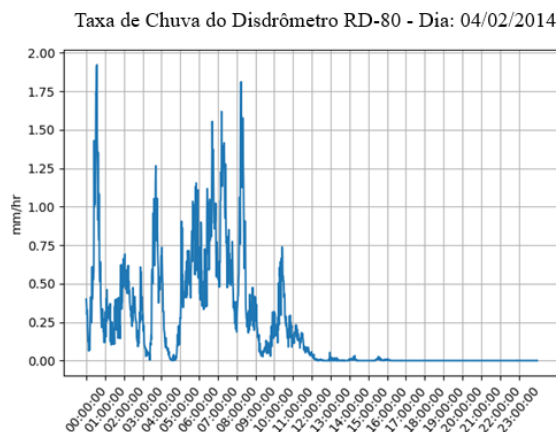
Abaixo são apresentados gráficos com dados diários das variáveis analisadas de um disdrômetro de impacto Joss-Waldvogel retirados de instrumentos utilizados durante a campanha GoAmazon.

Figura 7 - Gráfico da precipitação medida por um disdrômetro Joss-Waldvogel durante a campanha GoAmazon, em Manacapuru-AM, no dia 04/02/2014.



Fonte: Próprio autor (2021).

Figura 8 - Gráfico da taxa de chuva ($\text{mm}\cdot\text{h}^{-1}$) medida por um disdrômetro Joss-Waldvogel durante a campanha GoAmazon, em Manacapuru-AM, no dia 04/02/2014.

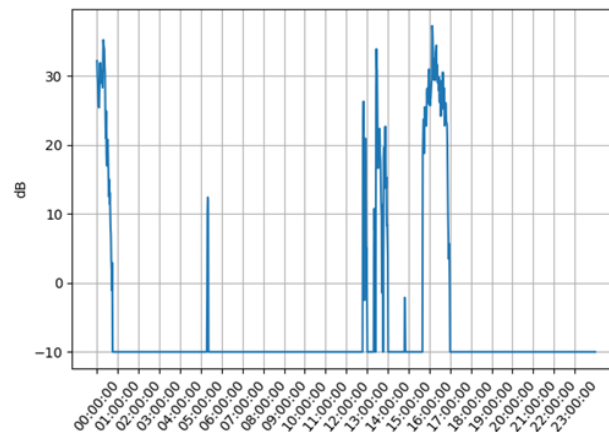


Fonte: Próprio autor (2021).

Adicionalmente, são apresentados também os gráficos das variáveis principais do instrumento PARSIVEL, também utilizado durante a campanha GoAmazon, na mesma localização dos sensores anteriores.

Figura 9 - Gráfico da refletividade radar (dBZ) medida por um disdrômetro PARSIVEL durante a campanha GoAmazon, em Manacapuru-AM, no dia 04/03/2014.

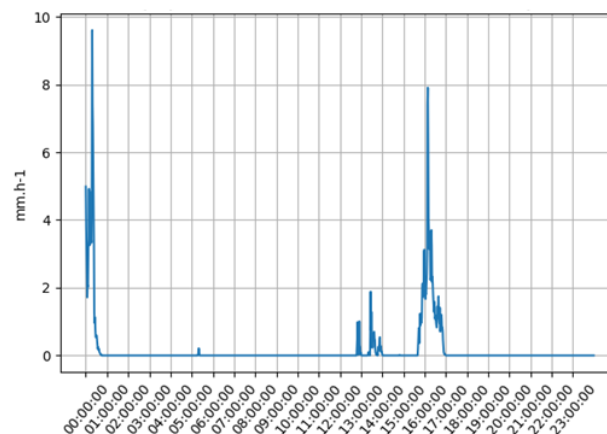
Refletividade radar pelo software Ott do Disdrômetro a Laser - Dia: 04/03/2014



Fonte: Próprio autor (2021).

Figura 10 - Gráfico da taxa de chuva (mm.h⁻¹) medida por um disdrômetro PARSIVEL durante a campanha GoAmazon, em Manacapuru-AM, no dia 04/03/2014.

Taxa de chuva pelo software Ott do Disdrômetro a Laser - Dia: 04/03/2014



Fonte: Próprio autor (2021).

3 Conclusões

Este trabalho teve o objetivo de desenvolver algoritmos computacionais capazes de converter em formatos proprietários para NetCDF, organizá-los, definir controles de qualidade e visualizar os dados provenientes de sensores meteorológicos mantidos pelo Instituto Nacional de Pesquisas Espaciais (INPE), os disdrômetros Joss-Waldvogel e PARSIVEL, de acordo com os padrões de dados do *Atmospheric Radiation Measurement* (ARM, órgão do Departamento de Energia dos EUA), de tal forma a torná-los mais acessíveis à comunidade científica e mais facilmente compartilhados nacional e internacionalmente.

Considerando os resultados obtidos a partir dos algoritmos criados, que foram capazes de criar arquivos NetCDF diários a partir de arquivos brutos vindos principalmente das campanhas SOS-CHUVA e GoAmazon, levando em consideração as validações estatísticas realizadas, que mostraram um alto grau de confiabilidade dos arquivos processados, é possível afirmar que essas ferramentas computacionais foram bem-sucedidas e são, portanto, apropriadas para uso em campanhas mantidas pelo INPE que façam uso de dados meteorológicos.

Igualmente, os algoritmos para visualização das variáveis principais corretamente plotaram os dados dos arquivos NetCDF gerados permitindo evidenciar que os dados convertidos estão dentro de um padrão esperado e contam com as *flags* de dados faltantes. Dada a relativa facilidade de uso e simplicidade desses *scripts*, espera-se que possam ser utilizados pelos mentores destes instrumentos e por outros usuários interessados nos seus dados.

Apenas um ponto restante não foi desenvolvido por questão do tempo dedicado aos algoritmos dos sensores restantes, os algoritmos para o radiômetro MP3000A, que seguiriam uma estrutura e maneira de funcionamento semelhante aos códigos “Pars_DataProcess” e “Joss_DataProcess” apresentados.

3.1 Trabalhos Futuros

Na proposta de bolsa, apenas um ponto não foi desenvolvido, que devido ao tempo dedicado aos algoritmos dos sensores aqui analisados não foi possível avançar de modo apropriado. Este objetivo não alcançado, tratava dos algoritmos para o radiômetro

MP3000A. Os dados deste instrumento são mais complexos e volumosos o que requeria mais tempo do que disposto. Contudo, os algoritmos aqui desenvolvidos podem ser adaptados e aplicados a este equipamento em trabalhos futuros.

REFERÊNCIAS BIBLIOGRÁFICAS

Atmospheric Radiation Measurement. ARM Research Facility, 2020. Página explicativa. Disponível em: <https://www.arm.gov/about>. Acesso em: 15/08/21.

JOSS, J.; WALDVOGEL, A. *Ein spektrograph für niederschlagstropfen mit automatischer auswertung. pure and applied geophysics*, v. 68, n. 1, p. 240–246, 12 1967. Disponível em: <https://doi.org/10.1007/BF00874898>.

Löffler-Mang, M., Joss, J., 2000. An optical disdrometer for measuring size and velocity of hydrometeors. *J. Atmos. Ocean. Technol.* 17 (2), 130–139.

Atmospheric Radiation Measurement. Microwave Radiometer Profiler (MWRP) Instrument Handbook, 2020. Manual. Disponível em: https://www.arm.gov/publications/tech_reports/handbooks/mwrp_handbook.pdf. Acesso em: 15/08/21.

Atmospheric Radiation Measurement. ARM Data File Standards Version: 1.3, 2020. Manual. Disponível em: <https://www.arm.gov/publications/programdocs/doe-sc-arm-15-004.pdf>. Acesso em: 15/08/21.

netCDF4 API *documentation*, 2020. Documentação. Disponível em: <https://unidata.github.io/netcdf4-python/>. Acesso em: 15/08/21.

Instituto Nacional de Pesquisas Espaciais. *Disdrometer - RD-80 User Guide for DISDRODATA 2.0*, 2009. Manual. Disponível em: http://ftp.cptec.inpe.br/chuva/read_me/alcantara/eq_disdrometer/esp_joss/Manual_Joss.pdf. Acesso em: 15/08/21.

DEVORE, Jay L. **Probabilidade e estatística**: para engenharia e ciências. São Paulo: Pioneira Thomson Learning, 2006.

MACHADO, L. A. T. Previsão imediata de tempestades intensas e entendimento dos processos físicos no interior das nuvens: O sos- chuva (sistema de observação e previsão de tempo severo). 2015. Disponível em: http://chuvaproject.cptec.inpe.br/portal/pdf/relatorios/Rel_dados_sos_chuva.pdf.

CALHEIROS, A. J. P. Relatório sobre os dados do projeto sos-chuva (pluviômetros, disdrômetros, mrr e mp3000a). 2018. Disponível em: http://chuvaproject.cptec.inpe.br/portal/pdf/relatorios/Rel_dados_sos_chuva.pdf. 18, 30

APÊNDICE A – TABELAS DAS VARIÁVEIS PRINCIPAIS DOS INSTRUMENTOS

Tabela A.1 – Variáveis principais do disdrômetro de impacto Joss-Waldvogel geradas pelo algoritmo de processamento de dados.

Tipo	Variável	Tipo de dado (<i>Data Type</i>)	Unidade
Tempo base em Epoch	base_time	Ponto flutuante de 32 sem sinal	segundos desde YYYY-mm-dd XX:XX:XX X:XX
Desvio de tempo em relação a base_time	time_offset	Ponto flutuante de 64 bits	segundos desde YYYY-mm-dd XX:XX:XX X:XX
Desvio de tempo em Relação à meia noite	time	Ponto flutuante de 64 bits	segundos desde YYYY-mm-dd XX:XX:XX X:XX
Latitude norte	lat	Ponto flutuante de 32 bits	graus
Longitude leste	lon	Ponto flutuante de 32 bits	graus
Altitude	alt	Ponto flutuante de 32 bits	metros acima do nível do mar
Número de gotas	num_drop	Ponto flutuante de 32 bits	inteiro
Diâmetro médio da Classe de gotas	drop_class	Ponto flutuante de 32 bits	mm
Taxa de chuva	rain_rate	Ponto flutuante de 32 bits	mm/h
Velocidade de queda	fall_vel	Ponto flutuante de 32 bits	m/s
Diâmetro médio da classe de partícula	drop_class	Ponto flutuante de 32 bits	mm

Fonte: Próprio autor (2021).

Tabela A.2 – Variáveis principais do disdrômetro a laser PARSIVEL geradas pelo algoritmo de processamento de dados.

Tipo	Variável	Tipo de dado (<i>Data Type</i>)	Unidade
Tempo base em Epoch	base_time	Ponto flutuante de 32 sem sinal	segundos desde YYYY-mm-dd XX:XX:XX X:XX
Desvio de tempo em relação a base_time	time_offset	Ponto flutuante de 64 bits	segundos desde YYYY-mm-dd XX:XX:XX X:XX
Desvio de tempo em Relação à meia noite	time	Ponto flutuante de 64 bits	segundos desde YYYY-mm-dd XX:XX:XX X:XX
Latitude norte	lat	Ponto flutuante de 32 bits	graus
Longitude leste	lon	Ponto flutuante de 32 bits	graus
Altitude	alt	Ponto flutuante de 32 bits	metros acima do nível do mar
Diâmetro médio da classe de partícula	drop_class	Ponto flutuante de 32 bits	mm
Taxa de precipitação	rain_rate	Ponto flutuante de 32 bits	mm/h
Velocidade de queda bruta	fall_vel	Ponto flutuante de 32 bits	m/s
Distribuição bruta de tamanho de partículas	raw_spectrum	Ponto flutuante de 32 bits	N/A
Fator de refletividade radar	radar_reflectivity_factor	Ponto flutuante de 32 bits	dB
Intervalo de amostragem	interval_sample	Ponto flutuante de 32 bits	s
Código de erro	err_code	Ponto flutuante de 32 bits	N/A
Área de laser sensível	F_laser_area	Ponto flutuante de 32 bits	m ²

Fonte: Próprio autor (2021).

Tabela A.3 – Diâmetros e velocidades de queda brutas para cada classe de gota. Valores relativos ao instrumento PARSIVEL.

Diâmetro (mm)	Velocidade de queda (m/s)
0.062	0.05

continua

Tabela A.3 - Continuação

0.1870	0.15
0.3120	0.25
0.4370	0.35
0.5690	0.45
0.6870	0.55
0.8120	0.65
0.9370	0.75
1.0620	0.85
1.1870	0.95
1.3750	1.10
1.6250	1.30
1.8750	1.50
2.1250	1.70
2.3750	1.90
2.7500	2.20
3.2500	2.60
3.7500	3.00
4.2500	3.40
4.7500	3.80
5.5000	4.40
6.5000	5.20
7.5000	6.00
8.5000	6.80
9.5000	7.60
11.000	8.80
13.000	10.40
15.000	12.00
17.000	13.60
19.000	15.20
21.500	17.60
24.500	20.80

Fonte: Próprio autor (2021).

Tabela A.4 – Diâmetros e velocidades de queda brutas para cada classe de gota. Valores relativos ao instrumento Joss-Waldvogel.

Diâmetro (mm)	Velocidade de queda (m/s)
0.359	1.435
0.455	1.862
0.551	2.267

continua

Tabela A.4 – Continuação

0.656	2.692
0.771	3.154
0.913	3.717
1.116	4.382
1.331	4.986
1.506	5.423
1.665	5.793
1.912	6.315
2.259	7.009
2.584	7.546
2.869	7.903
3.198	8.258
3.544	8.556
3.916	8.784
4.350	8.965
4.859	9.076
5.373	9.137

Fonte: Próprio autor (2021).

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas e documentação de projetos de engenharia.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Constam destas publicações o International Standard Serial Number (ISSN) que é Um código único e definitivo para identificação de títulos de seriados

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.

Manuais Técnicos (MAN)

São publicações de caráter técnico Que incluem normas, procedimentos, instruções e orientações.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Programas de Computador (PDC)

São a sequência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.