



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS  
DEPARTAMENTO DE ENGENHARIA CARTOGRÁFICA  
CURSO DE ENGENHARIA CARTOGRÁFICA E DE AGRIMENSURA

MARCELLA FERNANDES DE OLIVEIRA MELO

**AUTOMAÇÃO DE DOWNLOADS DIÁRIOS DAS OBSERVAÇÕES GNSS DAS  
ESTAÇÕES DA RBMC E DAS EFEMÉRIDES PRECISAS DO IGS**

Recife

2021

MARCELLA FERNANDES DE OLIVEIRA MELO

**AUTOMAÇÃO DE DOWNLOADS DIÁRIOS DAS OBSERVAÇÕES GNSS DAS  
ESTAÇÕES DA RBMC E DAS EFEMÉRIDES PRECISAS DO IGS**

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia Cartográfica e de Agrimensura da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do grau de Bacharel em Engenharia Cartográfica e de Agrimensura.

Orientador: Prof. Cezario de Oliveira Lima Junior.

Recife

2021

Catálogo na fonte  
Bibliotecária Margareth Malta, CRB-4 / 1198

M528a      Melo, Marcella Fernandes de Oliveira.  
Automação de downloads diários das observações GNSS das estações da RBMC e das efemérides precisas do IGS / Marcella Fernandes de Oliveira Melo. - 2021.  
45 folhas, il., gráfs., tabs.

Orientador: Prof. Dr. Cezário de Oliveira Lima Junior.

TCC (Graduação) – Universidade Federal de Pernambuco. CTG. Departamento de Engenharia Cartográfica, 2021.  
Inclui Referências.

1. Engenharia cartográfica. 2. Aquisição de dados. 3. Efemérides precisas. 4. IGS. 5. Dados GNSS. 6. RBMC. 7. *Python*. I. Lima Junior, Cezário de Oliveira (Orientador). II. Título.

UFPE

526.1 CDD (22. ed.)

BCTG/2021-69

MARCELLA FERNANDES DE OLIVEIRA MELO

**AUTOMAÇÃO DE DOWNLOADS DIÁRIOS DAS OBSERVAÇÕES GNSS DAS  
ESTAÇÕES DA RBMC E DAS EFEMÉRIDES PRECISAS DO IGS**

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia Cartográfica e de Agrimensura da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do grau de Bacharel em Engenharia Cartográfica e de Agrimensura.

Aprovado em: 12 / 04 / 2021.

**BANCA EXAMINADORA**

---

Prof. Cezário de Oliveira Lima Junior (Orientador)

Universidade Federal de Pernambuco

---

Prof. Dr. Silvio Jacks dos Anjos Garnés (Examinador Interno)

Universidade Federal de Pernambuco

---

Engenheiro Cartógrafo Edgar dos Reis (Examinador Externo)

Companhia Hidrelétrica do São Francisco

## **AGRADECIMENTOS**

Primeiramente a Deus, por ter me dado saúde todos os dias, possibilitando as condições de vencer os desafios.

Aos familiares, em especial minha mãe Edilene Silva por todo o amor, companheirismo, dedicação, cuidado, paciência, conselhos e incentivos, sendo de fundamental importância para a realização do presente trabalho e em toda a graduação.

Ao Prof. Cezário Lima pelas orientações e contribuições para elaboração deste TCC.

Ao Prof. Dr. Silvio Jacks pela ideia central do trabalho, por suas valiosas contribuições e aconselhamentos, por ter acreditado em mim, por me incentivar e ajudar na seleção do mestrado, seus conselhos foram fundamentais para participar de todo o processo seletivo.

A Prof. Dr. Andréa de Seixas por sua dedicação durante a graduação, por seu incentivo e apoio, por ter contribuído infinitamente com a minha formação, pela sua orientação durante mais de dois anos somando iniciação científica, monitoria, disciplinas e projeto de pesquisa.

Ao Engenheiro Cartógrafo Edgar dos Reis por todos os ensinamentos durante o estágio na CHESF que contribuíram bastante para a minha formação, me ensinando a ter uma nova visão sobre a profissão e potencializando o interesse pela área da Geodésia.

Ao Engenheiro Eletricista Lucas Gusman e ao Engenheiro Mecânico Ricardo Santos pela força, leitura do trabalho, opiniões, contribuições com o abstract e por testarem o programa. Além do aprendizado diário e suporte para alcançar os objetivos.

Ao MSc. Geógrafo Tiago Holanda por suas leituras, apoio e auxílio com as normas ABNT, suas correções foram fundamentais para atingir um bom resultado.

A Aline Damaso, Dyego Farias, Vanessa Nunes e Isabella Gomes por tá comigo desde o começo dessa caminhada, dando força, apoio e carinho, por acreditarem em mim quando nem eu acreditava mais, sem eles a graduação teria sido muito mais difícil.

Aos alunos da graduação do curso de Engenharia Cartográfica da UFPE (Júlia Pontes, Alice Soares, Beatriz Leopoldina, Lucas Israel e Bruno Barros).

## RESUMO

Nas últimas décadas o interesse em se realizar o posicionamento de feições terrestres com alta acurácia tem sido cada vez maior. A obtenção dos dados das estações GNSS da Rede Brasileira de Monitoramento Contínuo (RBMC), providas pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e das efemérides precisas, providas pelo *International GNSS Service* (IGS), representam, respectivamente, parte integrante do procedimento padrão para realização dos métodos de posicionamento relativo e por ponto preciso. Embora o acesso aos bancos de dados do IBGE e do IGS seja livre, em algumas situações específicas, principalmente acadêmicas e científicas, há uma dificuldade em se obter esses dados em grande quantidade ou até mesmo dados mais antigos que já não existem mais no acervo das duas instituições. Sendo assim, o trabalho objetiva a construção de uma rotina para automatizar a aquisição dos dados diários das observações GNSS das estações da RBMC e das efemérides precisas, com intuito de criar uma infraestrutura de acervo geodésico para consulta em processamentos de dados e de séries temporais. Para isso, foi desenvolvido um programa em linguagem *Python* com o auxílio das bibliotecas adicionais *Wget* e *Selenium*, podendo ser executado tanto em sistema *Linux* quanto em sistema *Windows* por meio de um compilador, desde que tenha o *Mozilla Firefox* instalado e conexão com a Internet. Como resultado foi obtido um programa com interface e funcionamento intuitivo, mostrando-se uma alternativa de grande importância principalmente aos usuários que tenham necessidade de realizar o *download* de grande quantidade de dados, evitando o envolvimento excessivo com tarefas repetitivas, impedindo possíveis falhas humanas. Além disso, houve êxito em adquirir grandes quantidades de dados organizados em uma estruturação de armazenamento, sendo estes de grande contribuição para a criação da infraestrutura de acervo geodésico para futuras consultas em processamentos de dados e de séries temporais. Ao final, são listados os benefícios atingidos pelo usuário com a utilização do programa em comparação com o método tradicional.

Palavras-chave: Aquisição de dados. Efemérides precisas. IGS. Dados GNSS. RBMC. *Python*.

## **ABSTRACT**

On last decades the interest in realize the position of lands fictions with high accuracy has been each time greater. The obtention of stations GNSS's data of Rede Brasileira de Monitoramento Contínuo (RBMC), sources by Instituto Brasileiro de Geografia e Estatística (IBGE) and of precise ephemerides, sources by International GNSS Service (IGS), represent, respectively, part integrant of standard procedure to realization of relative position methods and by precise point. Although the acess the banks of IBGE and IGS's data be free, on some situations specific, mainly academic and scientific, there is a difficult on obtain this data on wide amount or until data older than yet do not exist on the two institutions's acquis. Therefore, the job aim the build of a routine to automatize the RBMC and precise ephemerides's acquisition of day data of GNSS observations, with purpose of create an infrastructure of the geodesic acquis for consult on processing of dados and of temporal series. For this, was developed a program on Python language with support of library addictions Wget and Selenium, can be executed as Linux system as Windows system, since that has Mozilla Firefox installed and with Internet connection. As result was obtained a program with intuitive interface and operation, showing an alternative of wide importance mainly the users who have need of realize the download of great amount of data organized, avoid the involvement excessive with tasks repetitive, preventing possible humans failures. In addition, there was success on acquire great amounts of organized data in an storage structuration, being of great contribution to the creation of infrastructure of geodesic acquis for future consults on operation of data and of temporal series. The end, are listed the benefits achieved by user with the utilization of the program on comparation with the traditional method.

Keywords: Acquisition of data. Precise ephemerides. IGS. GNSS data. RBMC. Phyton.

## LISTA DE ILUSTRAÇÕES

|             |                                                                                                         |    |
|-------------|---------------------------------------------------------------------------------------------------------|----|
| Figura 1 -  | Exemplo de uma estrutura de um documento HTML.....                                                      | 25 |
| Figura 2 -  | Estrutura em árvore de um documento HTML.....                                                           | 26 |
| Figura 3 -  | Interação entre usuário e servidor.....                                                                 | 27 |
| Figura 4 -  | Modelo de uma mensagem HTTP.....                                                                        | 27 |
| Quadro 1-   | URLs dos diretórios iniciais das soluções processadas pelo IGS.....                                     | 34 |
| Figura 5 -  | Exibição da interface inicial do “ <i>menu.py</i> ”.....                                                | 36 |
| Figura 6 -  | Interface de <i>login</i> do <i>Earthdata</i> usada para ter acesso aos produtos do IGS.....            | 37 |
| Figura 7 -  | Página de redirecionamento do <i>Earthdata</i> .....                                                    | 37 |
| Figura 8 -  | Interface do diretório principal com as pastas das semanas GPS.....                                     | 37 |
| Figura 9 -  | Execução da opção [1] do “ <i>menu.py</i> ”.....                                                        | 38 |
| Figura 10 - | Execução da opção [2] do “ <i>menu.py</i> ”.....                                                        | 38 |
| Figura 11 - | Execução do módulo “ <i>looper.py</i> ”.....                                                            | 39 |
| Quadro 2 -  | Comparativo de obtenção dos dados alvo do trabalho entre o método convencional e o uso do programa..... | 40 |



## LISTA DE ABREVIATURAS E SIGLAS

|          |                                                                   |
|----------|-------------------------------------------------------------------|
| API      | Application Programming Interface                                 |
| BDS      | BeiDou System                                                     |
| CDDIS    | Crustal Dynamics Data Center                                      |
| CDMA     | Code Division Multiple Access                                     |
| DGPS     | Differential Global Positioning System                            |
| DNS      | Domain Name System                                                |
| EOSDIS   | Earth Observing System Data and Information System                |
| FAGS     | Federation of Astronomical and Geophysical Data Analysis Services |
| FDMA     | Frequency Division Multiple Access                                |
| FTP      | File Transfer Protocol                                            |
| GLONASS  | Global Orbiting Navigation Satellite System                       |
| GNSS     | Global Navigation Satellite System                                |
| GPS      | Global Positioning System                                         |
| HTML     | Hypertext Markup Language                                         |
| HTTP     | Hypertext Transfer Protocol                                       |
| IAG      | International Association of Geodesy                              |
| IBGE     | Instituto Brasileiro de Geografia e Estatística                   |
| ICAO     | International Civil Aviation Organization                         |
| IERS     | International Earth Rotation and Reference System Service         |
| IGS      | International GNSS Service                                        |
| IPS      | Internet Protocol Suite                                           |
| ITRF2005 | International Terrestrial Reference Frame 2005                    |
| MEO      | Medium Earth Orbit                                                |
| NASA     | National Aeronautics and Space Administration                     |
| NGS      | National Geodetic Survey                                          |
| P        | Precise or Protected                                              |
| PPP      | Posicionamento por Ponto Preciso                                  |
| PPS      | Posicionamento por Ponto Simples                                  |
| PRN      | Pseudo Random Noise                                               |
| PZ       | Parametry Zemli                                                   |
| RBMC     | Rede Brasileira de Monitoramento Contínuo                         |
| RINEX2   | Receiver INdependent EXchange format 2                            |

|      |                                             |
|------|---------------------------------------------|
| SCA  | Sistemas de Controle Ativos                 |
| SGB  | Sistema Geodésico Brasileiro                |
| SPP  | Single Point Positioning                    |
| sp3  | Standard Product 3                          |
| TAI  | International Atomic Time                   |
| TCP  | Transmission Control Protocol               |
| URL  | Uniform Resource Locator                    |
| URSS | União das Repúblicas Socialistas Soviéticas |
| UTC  | Coordinated Universal Time                  |
| WGS  | World Geodetic System                       |

## SUMÁRIO

|              |                                                              |           |
|--------------|--------------------------------------------------------------|-----------|
| <b>1</b>     | <b>INTRODUÇÃO .....</b>                                      | <b>12</b> |
| <b>2</b>     | <b>JUSTIFICATIVA.....</b>                                    | <b>14</b> |
| <b>3</b>     | <b>OBJETIVOS .....</b>                                       | <b>15</b> |
| 3.1          | OBJETIVO GERAL.....                                          | 15        |
| 3.2          | OBJETIVOS ESPECÍFICOS .....                                  | 15        |
| <b>4</b>     | <b>REFERÊNCIAL TEÓRICO .....</b>                             | <b>16</b> |
| 4.1          | GLOBAL NAVIGATION SATELLITE SYSTEMS (GNSS) .....             | 16        |
| <b>4.1.1</b> | <b>GPS .....</b>                                             | <b>17</b> |
| <b>4.1.2</b> | <b>GLONASS.....</b>                                          | <b>18</b> |
| 4.2          | OBSERVÁVEIS BÁSICAS DO GNSS .....                            | 18        |
| 4.3          | MÉTODOS DE POSICIONAMENTO.....                               | 19        |
| <b>4.3.1</b> | <b>Posicionamento Absoluto.....</b>                          | <b>19</b> |
| <b>4.3.2</b> | <b>Posicionamento Relativo.....</b>                          | <b>20</b> |
| 4.4          | FONTES PÚBLICAS DE DADOS GNSS .....                          | 21        |
| <b>4.4.1</b> | <b>RBMC (Rede Brasileira de Monitoramento Contínuo).....</b> | <b>21</b> |
| <b>4.4.2</b> | <b>IGS (International GNSS Service).....</b>                 | <b>22</b> |
| 4.4.2.1      | Tempo GPS .....                                              | 23        |
| 4.5          | PYTHON.....                                                  | 24        |
| 4.6          | INTEGRATED DEVELOPMENT ENVIRONMENT (IDE) .....               | 24        |
| 4.7          | WEB SCRAPING .....                                           | 24        |
| <b>4.7.1</b> | <b>Hypertext Markup Language (HTML).....</b>                 | <b>25</b> |
| <b>4.7.2</b> | <b>Hypertext Transfer Protocol (HTTP).....</b>               | <b>26</b> |
| <b>4.7.3</b> | <b>File Transfer Protocol (FTP).....</b>                     | <b>28</b> |
| 4.8          | GIT E GITHUB.....                                            | 29        |
| <b>5</b>     | <b>METODOLOGIA .....</b>                                     | <b>30</b> |
| 5.1          | DOWNLOAD E INSTALAÇÃO DO INTERPRETADOR PYTHON.....           | 30        |
| 5.2          | BAIXANDO E INSTALANDO A IDE PYCHARM .....                    | 30        |
| 5.3          | INSTALAÇÃO DE BIBLIOTECAS PYTHON.....                        | 31        |
| 5.4          | MENU.PY .....                                                | 31        |
| 5.5          | LOOPER.PY .....                                              | 31        |
| 5.6          | DOWNLOADER.PY .....                                          | 32        |
| <b>5.6.1</b> | <b>Download IBGE.....</b>                                    | <b>32</b> |
| <b>5.6.2</b> | <b>Download NASA.....</b>                                    | <b>33</b> |

|          |                                        |           |
|----------|----------------------------------------|-----------|
| 5.7      | OBSERVAÇÕES RELEVANTES AO LEITOR ..... | 35        |
| <b>6</b> | <b>RESULTADOS E DISCUSSÕES .....</b>   | <b>36</b> |
| <b>7</b> | <b>CONSIDERAÇÕES FINAIS.....</b>       | <b>41</b> |
|          | <b>REFERÊNCIAS .....</b>               | <b>42</b> |

## 1 INTRODUÇÃO

Nas últimas décadas o interesse em se realizar o posicionamento de feições terrestres com alta acurácia tem sido cada vez maior. Nesse sentido, as tecnologias espaciais vêm sendo amplamente empregadas e seu pleno domínio é fundamental para que se tenha um melhor aproveitamento do sistema. O GNSS (*Global Navigation Satellite System*), uma das tecnologias espaciais de posicionamento mais avançadas, tem revolucionado as atividades relacionadas com posicionamento. Devido a isso, a quantidade de aplicações é bastante ampla e continua aumentando, indo desde Geodésia, Geodinâmica, Agricultura de Precisão, Meteorologia, Navegação, até as atividades de lazer (ALVES, 2013).

Neste âmbito, o Brasil dispõe de uma rede de estações GNSS de operação permanente, a Rede Brasileira de Monitoramento Contínuo (RBMC), sendo a estrutura geodésica de referência mais precisa do país, cujas informações atendem tanto a comunidade científica quanto a prática. As observações advindas das estações da RBMC desempenham um papel muito importante, principalmente para a realização de processamento GNSS no posicionamento utilizando o método relativo (COSTA, 2008). Em algumas situações específicas, principalmente acadêmicas e científicas, é necessário obter grande quantidade de dados, de meses a anos de observações para uma ou mais estações da RBMC. No entanto, todo o procedimento padrão para obtenção dos dados, torna esse procedimento moroso e susceptível a erros, demandando a elaboração de soluções para suprir esta dificuldade.

De maneira similar, a obtenção de efemérides precisas, representa parte integrante do procedimento padrão para tratamento de observações pelo método do Posicionamento por Ponto Preciso, método bastante utilizado na geodésia e em outras áreas que requerem alta acurácia. De acordo com Seeber (2003), as efemérides precisas possuem acurácia superior às efemérides transmitidas, pois são estimadas com base em observações de uma rede global de confiança e alta qualidade, como, por exemplo, a rede do IGS (*International GNSS Service*). De modo análogo à obtenção de dados da RBMC, todo o procedimento padrão para obtenção dos dados, requer elevado grau de interação entre o usuário e o portal, tornando esse procedimento moroso e susceptível a erros na obtenção de grande quantidade de arquivos, requerem esforços para proporcionar soluções mais práticas.

Partindo das dificuldades supracitadas, este trabalho tem como objetivo apresentar soluções para automatizar *downloads* diários de observações GNSS das estações da RBMC e efemérides precisas do IGS, visando à criação de uma infraestrutura de acervo geodésico para consulta em processamento de dados e de séries temporais. Para tanto, foi elaborada uma

rotina em linguagem *Python*, usando diversas bibliotecas para concluir o objetivo final além de técnicas de *Web Scraping*, demandando apenas conexão com a internet e a instalação das bibliotecas adicionais *Wget* e *Selenium*.

## 2 JUSTIFICATIVA

Embora o acesso aos bancos de dados do IBGE (Instituto Brasileiro de Geografia e Estatística) e do IGS seja livre, em algumas situações específicas, principalmente acadêmicas e científicas, há uma dificuldade em se obter esses dados em grande quantidade ou até mesmo dados mais antigos que já não existem mais no acervo das duas instituições. A motivação de se obter uma rotina de automatização de *downloads* para criação de acervo geodésico ocorreu por uma constante demanda de alunos do Laboratório de Geodésia do Departamento de Engenharia Cartográfica da Universidade Federal de Pernambuco, sobre estudos relacionados com séries temporais de dados posicionais. Como atividade rotineira, os alunos desempenham a árdua tarefa de obter e processar grande quantidade de dados GNSS da RBMC e das efemérides precisas. Apesar do processamento possibilitar uma automatização por meio de *softwares* para processamento GNSS, a obtenção dos dados ainda é uma tarefa altamente desgastante e trabalhosa, podendo consumir semanas de trabalho para realização das atividades repetitivas.

Envolvidos nesse contexto, este trabalho se justifica por estar inserido dentro de uma temática bastante atual e relevante, abrangendo um problema comum a todos os usuários de dados GNSS que precisam obtê-los em grande escala ou dados mais antigos que já não existem mais no acervo das duas instituições. Dessa forma, a solução foi a idealização de uma infraestrutura de acervo geodésico para consulta em processamentos de dados e de séries temporais.

### 3 OBJETIVOS

Para a estruturação deste trabalho foram traçadas metas sobre o tema pretendido, e observados alguns pressupostos e de acordo com eles os objetivos desta pesquisa que serão apresentados a seguir.

#### 3.1 OBJETIVO GERAL

Automatizar *downloads* diários de dados GNSS com intuito de criar uma infraestrutura de acervo geodésico para consulta em processamentos de dados e de séries temporais.

#### 3.2 OBJETIVOS ESPECÍFICOS

Construção de uma rotina em *python* para:

- a) Automatizar *downloads* diários das observações GNSS brutas das estações da Rede Brasileira de Monitoramento Contínuo - RBMC providas pelo Instituto Brasileiro de Geografia e Estatística – IBGE;
- b) Automatizar *downloads* diários das efemérides precisas providas pelo *International GNSS Service* – IGS;
- c) Realizar *downloads* de grandes quantidades de dados para séries temporais de forma rápida;
- d) Elaboração de uma estrutura de armazenamento dos dados organizada em diretórios para criação de uma infraestrutura de acervo geodésico.



## 4 REFERÊNCIAL TEÓRICO

Este tópico tem por objetivo apresentar uma revisão e o estado da arte das principais características dos conhecimentos utilizados para o entendimento total do trabalho, a saber: (i) a descrição dos dois principais sistemas globais de navegação; (ii) modelo das observáveis GNSS; (iii) descrição dos diferentes métodos de posicionamento, dando destaque para os que utilizam as efemérides precisas e as estações da RBMC; (iv) as principais fontes públicas de dados GNSS; (v) uma breve descrição sobre python e ambiente integral de desenvolvimento; (vi) os conceitos sobre *Web Scraping* e (vii) e a utilização do git e github.

### 4.1 GLOBAL NAVIGATION SATELLITE SYSTEMS (GNSS)

O surgimento das técnicas espaciais na Geodésia, na década de 1970, provocou uma revolução nas atividades de navegação e posicionamento, proporcionando ao longo do tempo uma melhoria de precisão significativa. Durante este período, ocorreu o advento dos sistemas globais americano GPS (*Global Positioning System*) e soviético GLONASS (*Global Orbiting Navigation Satellite System*), desenvolvidos de maneira concomitante e independente, seguidos pelos sistemas Galileo e BDS (*BeiDou System*). Além desses sistemas globais, existem até o momento, os sistemas regionais, em desenvolvimento em países como a Índia e o Japão (MONICO, 2008).

Ainda segundo a referência supracitada, a junção dos sistemas supracitados é denominada de GNSS (*Global navigation satellite system* - Sistema Global de Navegação por Satélite). Tal nome foi concebido em 1991, durante a 10ª Conferência de Navegação Aérea, quando a Associação Internacional de Aviação Civil (*International Civil Aviation Organization* - ICAO) reconheceu que a fonte primária para a navegação seria o GNSS.

O objetivo principal do GNSS é a determinação instantânea de posição, velocidade e tempo de um usuário, quaisquer fossem as condições atmosféricas e sua localização geográfica (SEEBER, 2003). Devido a isso, o GNSS tem facilitado amplamente todas as atividades que envolvem posicionamento, surgindo assim, inúmeras aplicações. De acordo com Monico (2008) pode-se destacar: atividades ligadas à geodinâmica e a sistemas de observação da superfície terrestre (e.g. Zanutta et al. 2017; Sośnica et al. 2018; Farolfi et al. 2019); estudos relacionados à atmosfera, abrangendo tanto ionosfera quanto troposfera (e.g. FENG, 2008; Guerova et al. 2016; Bianchi et al. 2016; Rahmani et al. 2020); e o

estabelecimento de redes geodésicas locais, regionais, continentais e globais, tanto ativas quanto passivas (e.g. Kaminskis et al. 2020).

Atualmente, o GPS e o GLONASS são os principais sistemas de navegação por satélite, sendo os únicos a dispor de constelações completas. Por este motivo, tais sistemas serão descritos em seguida.

#### 4.1.1 GPS

O GPS é um sistema de radionavegação desenvolvido e controlado pelos Estados Unidos da América, onde foi declarado operacional no ano de 1995, inicialmente criado para fins militares, expandido posteriormente para uso civil. É composto por uma constelação nominal de 24 satélites em órbita terrestre média (*Medium Earth Orbit* - MEO) e por estações terrestres distribuídas pelo globo com o intuito de monitorar e controlar os satélites, sendo suas coordenadas vinculadas ao sistema de referência WGS-84 (*World Geodetic System*) (SEEBER, 2003). Em fevereiro de 2021, 31 satélites estão operacionais (GPS, 2021).

Ao longo dos anos, o GPS passou por diversos projetos de modernização tanto no segmento espacial quanto no de controle. Em relação às frequências, é utilizada a técnica CDMA (*Code Division Multiple Access* - divisão do código para múltiplo acesso), no qual todos os satélites transmitem na mesma frequência e a identificação é dada por um código exclusivo de cada satélite, o PRN (*Pseudo Random Noise*) (SPILKER, 1996; MONICO, 2008).

A princípio o sistema possuía apenas duas ondas portadoras: L1 e L2, contudo apenas a portadora L2 transmitia o código P (*Precise or Protected* – Preciso ou Protegido), reservado para uso dos militares norte-americanos e usuários autorizados (MONICO, 2008). Em decorrência dos projetos de modernização do GPS, a partir do bloco IIR-M a portadora L2 passou a transmitir seu segundo sinal civil, denominado de L2C. Entre outras mudanças, uma nova portadora, L5, foi desenvolvida na geração do bloco IIF, com o terceiro sinal civil L5C. Com quatro satélites em órbita, uma nova geração denominada de GPS III conta com um quarto sinal civil em L1 (L1C) (GPS, 2021).

#### 4.1.2 GLONASS

O GLONASS é o sistema global desenvolvido pela antiga União das Repúblicas Socialistas Soviéticas (URSS) e atualmente é mantido pelo governo russo. De forma similar ao GPS, foi criado inicialmente para fins militares, expandido posteriormente para uso civil (SEEBER, 2003). Sobre seu segmento espacial, o GLONASS é constituído de 24 satélites operacionais e 03 de reserva, distribuídos em três planos orbitais inclinados aproximadamente 64,8° em relação ao equador. A altitude média é de 19.100 km e o período orbital é de 11h15min (Tempo Sideral) (MONICO, 2008).

Ainda segundo a referência supracitada, embora a constelação completa tenha sido alcançada em 1995, em razão do curto período de vida útil dos satélites, bem como da não reposição destes, o número de satélites passou por um longo período abaixo do necessário. Em 2002, o governo russo estabeleceu um plano de reestabelecimento da cobertura global e modernização do sistema, hoje a constelação conta com 28 satélites em órbita, dos quais apenas 23 estão ativos (fevereiro de 2021) (IAC, 2021).

Em relação às frequências, é utilizada a técnica FDMA (*Frequency Division Multiple Access* - Múltiplo Acesso pela Divisão da Frequência), na qual cada satélite transmite em uma frequência, sendo o satélite identificado de acordo com a mesma. No que concerne ao referencial geodésico, é utilizado o PZ 90 (*Parametry Zemli* 1990) (MONICO, 2008).

Assim como no GPS, inicialmente, a portadora L1 possuía sinal para a comunidade civil e usuários autorizados, em contrapartida, a segunda portadora era designada somente para usuários autorizados. A partir da geração GLONASS-M em 2004, um sinal civil também foi adicionado na portadora L2 (HOFMANN-WELLENHOF et al., 2008). Os satélites da geração GLONASS-K1 transmitem, além dos sinais FDMA, uma nova portadora L3, CDMA. A nova geração GLONASS-K2 contará com a tecnologia CDMA também nas portadoras L1 e L2 (REVNIVYKH et al., 2017).

#### 4.2 OBSERVÁVEIS BÁSICAS DO GNSS

As observáveis básicas do GNSS são a pseudodistância e a fase da onda portadora. A pseudodistância é uma medida do tempo de propagação do sinal entre o satélite e o receptor, convertida para distância e obtida sem considerar o erro do não sincronismo entre os relógios, sendo a precisão dessa medida na ordem métrica. A equação da pseudodistância de um satélite S rastreado por um receptor r é dado pela Equação (1) (MONICO, 2008):

$$PD_R^S = \rho_r^S + c[dt_r - dt^S] + I_r^S + T_r^S + dm_r^S + \varepsilon_{PD_r}^S \quad (1)$$

em que  $PD_R^S$  refere-se à observação da pseudodistância (m),  $\rho_r^S$  é a distância geométrica dada por  $\rho_r^S = c\tau_r^S$  onde  $\tau_r^S$  é o intervalo de propagação do sinal (s) e  $c$  é a velocidade da luz no vácuo (m/s),  $dt^S$  é o erro do relógio do satélite (s),  $dt_r$  é o erro do relógio do receptor (s),  $\varepsilon_{PD_r}^S$  é o erro da medida de pseudodistância (m),  $I_r^S$  é o atraso ionosférico (m),  $T_r^S$  é o atraso troposférico (m) e  $dm_r^S$  são os efeitos do multicaminho (m).

A fase da onda portadora é uma observável muito mais precisa que a pseudodistância, sua medida tem precisão da ordem de até 1/1000 do ciclo. A equação da fase da onda portadora é dada pela Equação (2) (MONICO, 2008):

$$\phi_r^S(t) = f \left( \frac{\rho_r^S - I_r^S + T_r^S + dm}{c} \right) + f[dt_r - dt^S] + [\phi_t^S(t_0) - \phi_r(t_0)] + N_r^S + \varepsilon\phi_r^S \quad (2)$$

em que  $\phi_r^S$  refere-se à observação da fase da onda portadora (ciclos),  $f$  é a frequência da onda portadora (Hz),  $\phi_r(t_0)$  é a fase gerada no receptor em uma época de referência (ciclos),  $\phi_t^S(t_0)$  é a fase transmitida pelo satélite em uma época de referência (ciclos),  $N_r^S$  é a ambiguidade da fase (ciclos) e  $\varepsilon\phi_r^S$  é o erro da fase da onda portadora (ciclos).

### 4.3 MÉTODOS DE POSICIONAMENTO

De acordo com Monico (2008) o posicionamento refere-se à determinação da posição de objetos com relação a um referencial específico. Os métodos de posicionamento GNSS podem ser divididos em Posicionamento Absoluto, também conhecido como Posicionamento por Ponto; Posicionamento Relativo; e Posicionamento Diferencial, também conhecido pela sigla DGPS (que não será explorado neste trabalho).

#### 4.3.1 Posicionamento Absoluto

No posicionamento absoluto as coordenadas do receptor são estimadas diretamente em relação ao geocentro. É subdividido em Posicionamento por Ponto Simples (*Single Point Positioning* - SPP) e Posicionamento por Ponto Preciso (*Precise Point Positioning* - PPP). No PPS, método bastante utilizado pela comunidade civil em atividades de navegação,

observações de pseudodistância normalmente em simples frequência medidos por um receptor são processadas para resolver sua posição, dadas as posições e erros dos relógios dos satélites GNSS, calculados a partir das mensagens de navegação (*Broadcast Ephemerides* - Efemérides Transmitidas). Por outro lado, no PPP, o método requer uso de observações de fase e pseudodistância em dupla frequência GPS, órbitas e relógios precisos (*Precise Ephemerides*- Efemérides Precisas) (MONICO, 2008; ODIJK, 2017; KOUBA et al., 2017).

As efemérides precisas são oriundas de estimativas feitas pelos centros de análise subordinados ao Serviço GNSS Internacional (*International GNSS Service* - IGS). Os arquivos disponibilizados à comunidade civil contêm as coordenadas das órbitas e as correções dos relógios precisos. Devido aos esforços do IGS, desde 1990 até os dias atuais, o PPP teve bastante notoriedade para geodésia e muitas outras aplicações que requerem alta acurácia, como, por exemplo, em geodinâmica, meteorologia, metrologia e assim por diante (KOUBA et al., 2017). No que diz respeito à acurácia das efemérides, pesquisas realizadas nas duas últimas décadas (HÉROUX; KOUBA, 2001; SEEPERSAD; BISNATH, 2014) têm mostrado potencial na ordem milimétrica a centimétrica para um período longo de coleta de dados.

#### **4.3.2 Posicionamento Relativo**

O posicionamento relativo tem como objetivo determinar as coordenadas de uma estação desconhecida em relação à outra conhecida, isto é, visa à determinação do vetor entre dois pontos que frequentemente são chamados de linha de base (HOFMANN-WELLENHOF, 2008). Para realizar o posicionamento relativo o usuário deve dispor de no mínimo dois receptores. Entretanto, atualmente o usuário pode ter no mínimo um receptor, devidos aos Sistemas de Controle Ativos (SCA), nos quais os receptores rastreiam continuamente os satélites visíveis e os dados podem ser acessados via internet. No Brasil, tem-se a Rede Brasileira de Monitoramento Contínuo (RBMC), sob responsabilidade do IBGE, além de outras estações contínuas (MONICO, 2008).

O posicionamento relativo pode ser realizado de diversas maneiras, dentre as quais têm-se: relativo estático, semicinemático e cinemático. No posicionamento estático, os dados são coletados por receptores estacionários e para que a posição seja obtida com alta precisão, é necessário um longo período de observação. Já no posicionamento relativo semicinemático não é necessário um longo período de permanência na estação, porém o receptor permanece ligado durante todo o período do levantamento. Quando o receptor ocupa a estação de

referência enquanto o segundo receptor se desloca sobre as feições de interesse, o método é denominado de relativo cinemático. No que diz respeito à precisão, para o estático a precisão varia de 0,01 a 1 ppm enquanto para os demais a precisão varia de 1 a 10 ppm (SEEBER, 2003; MONICO, 2008).

#### 4.4 FONTES PÚBLICAS DE DADOS GNSS

Nesta seção são descritos a RBMC (Rede Brasileira de Monitoramento Contínuo) e o IGS (Serviço GNSS Internacional), os quais são redes geodésicas ativas gerenciadas por instituições públicas responsáveis por manter os sistemas e divulgar os dados via rede.

##### 4.4.1 RBMC (Rede Brasileira de Monitoramento Contínuo)

O Brasil possui uma rede geodésica ativa de referência equipada com receptores GNSS, que proporcionam, uma vez por dia ou em tempo real, observações para a determinação de coordenadas. Tal rede de estações GNSS é denominada de Rede Brasileira de Monitoramento Contínuo (RBMC), suas estações desempenham o papel do ponto de coordenadas conhecidas pertencentes ao Sistema Geodésico Brasileiro (SGB) na utilização do método relativo. Conforme IBGE (2021a), instituto responsável pela implantação e manutenção da RBMC, sua infraestrutura geodésica é constituída atualmente por 150 estações que se encontram distribuídas ao longo do território nacional.

De acordo com IBGE (2021b), as estações da RBMC são compostas por receptores que coletam e armazenam continuamente as observações do código e da fase das ondas portadoras transmitidas pelos satélites das constelações GPS ou GLONASS. As observações são organizadas, ainda na memória do receptor, em arquivos diários, correspondendo a sessões iniciando às 00h01min e encerrando às 24h00min (tempo universal), com intervalo de rastreamento de 15 seg. Após o encerramento de uma sessão, os arquivos com as respectivas observações são transferidos do receptor para o Centro de Controle da Rede Brasileira de Monitoramento Contínuo dos Sistemas GNSS - RBMC, na Coordenação de Geodésia (Rio de Janeiro-RJ).

A partir deste ponto, são criados novos arquivos em formato padrão RINEX2 (*Receiver INdependent EXchange format 2*), nos quais é realizado um controle de qualidade das observações. Em seguida os arquivos de dados RINEX2 e as órbitas transmitidas são

compactados e disponibilizados na área de *download* do portal do IBGE (IBGE, 2021c) via FTP (*File Transfer Protocol*).

#### 4.4.2 IGS (*International GNSS Service*)

O Serviço GNSS Internacional (*International GNSS Service* - IGS) trata-se de um Centro Técnico do IERS (*International Earth Rotation and Reference System Service* - Serviço Internacional de Rotação da Terra e de Sistema de Referência) em assuntos relacionados com o GPS, o GLONASS e outros sistemas de navegação por satélite planejados para a próxima década. O IGS é um Serviço da IAG (*International Association of Geodesy*) e também da FAGS (*Federation of Astronomical and Geophysical Data Analysis Services*) (MONICO, 2008).

O serviço IGS está baseado em uma rede global de monitoramento e estações de rastreamento GNSS onde disponibiliza inúmeros produtos, dentre os quais efemérides precisas dos satélites GPS e GLONASS. O Brasil está integrado à rede IGS através da RBMC, e os dados das estações são repassados para um centro global da rede, situado no *Crustal Dynamics Data Center* (CDDIS) da NASA (*National Aeronautics and Space Administration*) (IGS, 2021).

O IGS produz três tipos de efemérides e correções para o relógio dos satélites, denominadas de efemérides IGS (resultante da combinação das órbitas de vários centros de análises disponíveis a partir do 13º dia após a coleta dos dados), IGR (órbitas IGS rápidas, disponíveis com latência de 17 horas) e IGU (órbitas IGS ultrarrápidas disponível com latência de 3 horas) (IGS, 2021).

O arquivo das efemérides precisas é composto pelas coordenadas X, Y e Z dos satélites, em quilômetros, atualmente referenciadas ao ITRF2005 (*International Terrestrial Reference Frame* 2005), além das correções dos relógios dos satélites, em microssegundos, os quais são dados, em épocas equidistantes, a cada 15 minutos (MONICO, 2008). Conforme Spofford e Remondi (1996), a identificação das efemérides precisas se dá com base na sigla do órgão que a produziu, na semana GPS correspondente e no dia da semana GPS (começa com 0 no domingo e vai até 6 no sábado). A extensão utilizada é o sp3 (*Standard Product* 3), representação ASCII definido pelo NGS (*National Geodetic Survey*).

As acurácias, latências e instantes de atualizações das efemérides precisas estão demonstradas na Tabela 1. Tais informações foram retiradas do site oficial do IGS (IGS, 2021).

Tabela 1 - Acurácias de órbitas e relógios dos satélites, latências e atualizações

| <b>Efemérides dos satélites GPS /<br/>Satélite e Relógios das Estações</b> |                           | <b>Acurácia</b>            | <b>Latência</b> | <b>Atualizações</b>      |
|----------------------------------------------------------------------------|---------------------------|----------------------------|-----------------|--------------------------|
| Ultrarrápida<br>(observada)                                                | Órbita                    | ~3 cm                      | 3 - 9 horas     | às 03, 09, 15,<br>21 UTC |
|                                                                            | Relógios dos<br>satélites | ~150 ps RMS<br>~50 ps SDev |                 |                          |
| Rápida                                                                     | Órbita                    | ~2.5 cm                    | 17 - 41 horas   | às 17 UTC<br>diária      |
|                                                                            | Relógios dos<br>satélites | ~75 ps RMS<br>~25 ps SDev  |                 |                          |
| Final                                                                      | Órbita                    | ~2.5 cm                    | 12 - 18 dias    | Toda quinta-<br>feira    |
|                                                                            | Relógios dos<br>satélites | ~75 ps RMS<br>~20 ps SDev  |                 |                          |

Fonte: Adaptado de IGS (2021).

#### 4.4.2.1 Tempo GPS

Conforme Achanta (2015), o tempo GPS é sincronizado com o TAI (*International Atomic Time*). A época de referência inicial do tempo GPS é a 00h00min UTC (*Coordinated Universal Time*) de 6 de janeiro de 1980 (domingo). Os segundos intercalados não são adicionados ao tempo do GPS. Portanto, a diferença entre o tempo UTC e GPS varia em incrementos de segundos cada vez que um segundo intercalado é adicionado à escala de tempo UTC.

O tempo GPS é representado pelo número da semana GPS, começando a partir do tempo da época do GPS (6 de janeiro de 1980). As semanas são numeradas desde 0 e vão até 1.023 e então retornam para 0. Este ciclo da contagem (“rollover cycle”) tem 1.024 semanas ou 7.168 dias, que representam aproximadamente 19,6 anos civis. É importante para os dispositivos de determinação do tempo que usam GPS (tais como receptores e relógios GPS) lidar com o ciclo de contagem do número da semana GPS para garantir que o tempo exato seja reportado (MONICO, 2008; ACHANTA, 2015). No entanto, para nomenclatura das pastas dos arquivos CDDIS não é considerado o ciclo de contagem.



## 4.5 PYTHON

Inventado por Guido van Rossum em 1989, *Python* está entre as linguagens de programação mais utilizadas ao redor do mundo. Desde sua estreia no mercado em 1991, ela vem se destacando e evoluindo de acordo com seu tempo. Essa linguagem possui tipagem forte e dinâmica como sua principal característica. Deixando o código mais limpo e eficiente (YEDA et al, 2018).

O *Python* é muito conhecido por sua consistência e transparência. Segundo Piotrowski (2016), essa linguagem é interpretada, todavia programadores podem utilizá-la como *script*. Suporta os mais comuns paradigmas, como orientação a objetos, até aqueles pouco utilizados, como a funcional. Com uma variedade de bibliotecas e uma comunidade *opensource* sempre ativa, o *Python* se mostra uma linguagem muito popular e de fácil acesso. Além disso, os interpretadores de *Python* estão disponíveis para os principais sistemas operacionais (*Linux*, *Mac OS*, *Windows*, *Android*, *iOS*, *BSD*, etc.).

## 4.6 INTEGRATED DEVELOPMENT ENVIRONMENT (IDE)

A forma mais comum para programar em *Python* é através do uso de um ambiente de programação, também conhecido como ambiente integral de desenvolvimento (*Integrated Development Environment – IDE*). De acordo com Python Brasil (2021), existem inúmeras IDE's para se programar em Python. Dentre essa ampla gama de possibilidades, tem-se: *Idle*, *PyCharm community*, *Komodo-Edit*, *NetBeans*, *Spyder 2*, *Eclipse*, *IPython* e *PyScripter*. Para o presente trabalho será abordado o *PyCharm community*, o qual é uma multiplataforma desenvolvida pela companhia *JetBrains*. Esta edição é liberada sob a licença da Apache. Essa IDE fornece análise de código, um depurador gráfico, um testador de unidade integrado, integração com sistemas de controle de versão (VCSes), e suporta desenvolvimento de *web* com *Django*.

## 4.7 WEB SCRAPING

Embora *Web Scraping* não seja uma terminologia nova, há alguns anos esta prática tem sido mais comumente conhecida como *screen scraping*, *data mining*, *web harvesting* ou variações similares. Em tese, *Web Scraping* é a prática de coletar dados de maneira automatizada, fazendo requisições HTTP (*Hypertext Transfer Protocol* - Protocolo de

Transferência de Hipertexto) a um servidor da *Web*, na finalidade de extrair páginas HTML (*HyperText Markup Language - Linguagem de Marcação de Hipertexto*) ou outros tipos de arquivos encontrados em um site (MITCHELL, 2016; VARGIU; URRU, 2013).

#### 4.7.1 Hypertext Markup Language (HTML)

O HTML é uma linguagem utilizada para representação visual de conteúdo na *Web*, inicialmente proposta por Tim Berners-Lee (1989). Embora esteja em constante evolução desde o seu início, a gramática básica dos documentos HTML tem se mantido, fazendo com que esta linguagem seja um dos padrões mais importantes para o trabalho na *Web* (MUNZERT, 2015).

Ainda segundo a referência supracitada, um arquivo HTML é essencialmente um arquivo de texto, o que torna a linguagem poderosa é a sua estrutura em marcação, que permite definir quais partes do documento devem ser projetadas como cabeçalho, links, tabelas e afins. Estas definições em marcação acontecem por meio de sequências de caracteres predefinidas, conhecidas como *tags*, que envolvem uma determinada parte do texto. A marcação informa ao navegador como o documento está estruturado. A Figura 1 ilustra um exemplo simples de um código-fonte com as *tags* de abertura e fechamento.

Figura 1 - Exemplo de uma estrutura de um documento HTML

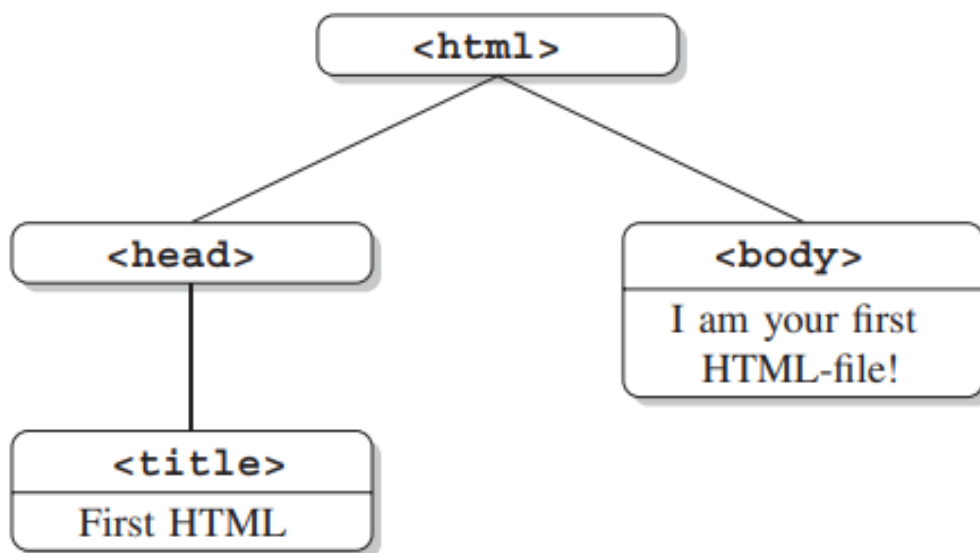
```
1  <!DOCTYPE html>
2  <html>
3    <head>
4      <title>First HTML</title>
5    </head>
6    <body>
7      I am your first HTML file!
8    </body>
9  </html>
```

Fonte: Munzert (2015).

A estrutura de um documento HTML pode ser interpretada como uma árvore. Como ilustrado na Figura 2, seus elementos internos seguem uma regra rígida de aninhamento: cada

*tag* tem seu início e fim. A infração desta regra implica em uma má formação do documento (MUNZERT, 2015).

Figura 2 - Estrutura em árvore de um documento HTML



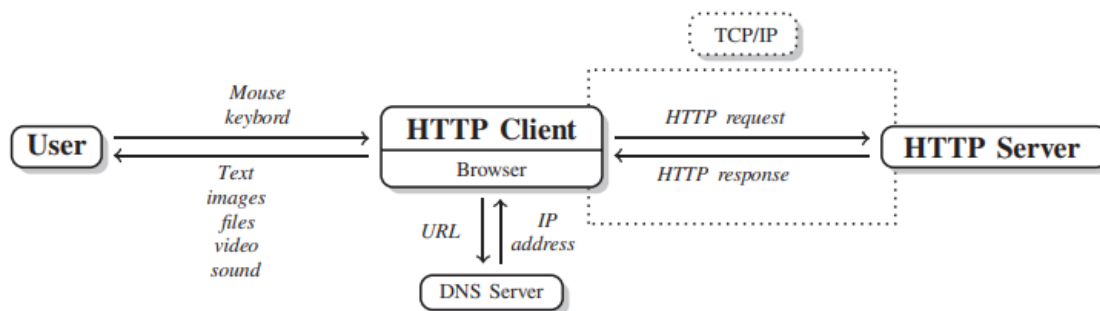
Fonte: Munzert (2015).

O acesso visual a um documento HTML possibilita o conhecimento de sua estrutura, o que implica que é possível então realizar o *parsing* do mesmo. O *parsing* é realizado por meio de expressões regulares, percorrendo sua árvore para encontrar informações que correspondam a alguns padrões e armazenando suas informações de maneira estruturada, na finalidade de formar uma base de dados (PENMAN, 2009; VARGIU; URRU, 2013).

#### 4.7.2 Hypertext Transfer Protocol (HTTP)

Para retirar dados da internet, é necessário permitir que o software se comunique com servidores e serviços *Web*. O HTTP, Protocolo de Transferência de Hipertexto, é o protocolo mais comum para a comunicação entre um cliente *Web* (navegadores, por exemplo) e servidores. Virtualmente, qualquer página HTML, imagem ou vídeo que é visto por meio de um navegador é entregue por este protocolo. O protocolo HTTP pode requisitar praticamente qualquer tipo de recurso disponível em um servidor e também pode ser usado para o envio de dados para o servidor (MUNZERT, 2015). A Figura 3 ilustra como ocorre a interação entre usuário e servidor.

Figura 3 - Interação entre usuário e servidor

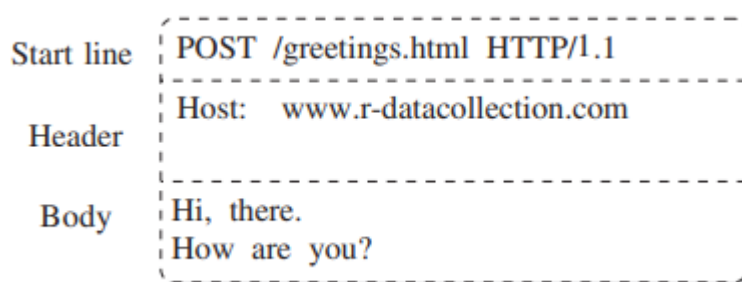


Fonte: Munzert (2015).

Para acessar o conteúdo da Web, costuma-se digitar URLs no navegador ou simplesmente clicar em links. Por trás dessa camada projetada para interação com o usuário, existem várias outras camadas (técnicas, padrões e protocolos) que fazem tudo funcionar. Juntos, são chamados de IPS (*Internet Protocol Suite*). Os protocolos mais proeminentes desse conjunto são TCP (*Transmission Control Protocol*) e IP (*Internet Protocol*), no qual representam a camada de Internet (IP) e a camada de transporte (TCP), sendo responsáveis pela transferência confiável de dados entre computadores na rede (FOROUZAN, 2002; FALL, 2011; MUNZERT, 2015).

Quando um *website* é acessado, o navegador se torna cliente HTTP. O servidor pergunta ao servidor DNS (*Domain Name System*) qual endereço IP está associado com o domínio URL alvo. Após receber esta resposta, o navegador estabelece uma conexão com o servidor HTTP requisitada via TCP/IP. Uma vez que a conexão é estabelecida, cliente e servidor podem trocar informações (MUNZERT, 2015). No caso do *Web Scraper*, estas informações serão requisições HTTP. Como ilustrado na Figura 4, tais mensagens possuem uma estrutura definida em três partes: linha inicial, cabeçalho e corpo.

Figura 4 - Modelo de uma mensagem HTTP



Fonte: Munzert (2015).

No caso de uma requisição a linha inicial define o método utilizado, seguida pelo caminho para o recurso e a versão HTTP mais alta que o cliente pode manipular. O cabeçalho contém metadados e o corpo contém os dados carregados pela mensagem.

Existem vários métodos de requisições HTTP, mas para um *Web Scraper* os métodos mais relevantes são GET e POST. Enquanto o método GET solicita documentos ao servidor, o método POST pode ser utilizado para enviar arquivos (MUNZERT, 2015).

### 4.7.3 File Transfer Protocol (FTP)

O *File Transfer Protocol* (FTP), é um dos protocolos mais antigos e mais amplamente utilizados na Internet, onde foi desenvolvido, como o próprio nome sugere, para transferir arquivos de um local para outro. O FTP é bidirecional, o que significa que pode ser usado para transferir arquivos do cliente para o servidor (*upload*) e do servidor para o cliente (*download*). Além disso, vem com um conjunto de comandos que especificam quais arquivos transferir, quais diretórios criar, excluir e muitas outras operações (MUNZERT, 2015).

Ao contrário do HTTP que permite conexões persistentes, *keepalive*, ou seja, conexões entre cliente e servidor que são mantidas por várias transferências, no FTP a conexão deve ser restabelecida após cada transferência. Além disso, o FTP não oferece suporte nativo a *proxies* e *pipelining*, ou seja, vários pedidos simultâneos antes de receber uma resposta. Por outro lado, o FTP pode ser mais rápido em certas circunstâncias, já que não vem com um monte de campos de cabeçalho como HTTP, apenas os arquivos binários ou ASCII são transferidos (MUNZERT, 2015).

Ainda segundo a referência supracitada, o FTP é um protocolo cliente-servidor, ou seja, o cliente solicita os arquivos e o servidor os fornece. Como tal, o FTP requer dois canais básicos para estabelecer uma conexão: o canal de comando (inicia a instrução e carrega informações básicas) e o canal de dados (transfere os dados do arquivo entre os dois dispositivos). Para estabelecer uma conexão, os usuários precisam fornecer credenciais para o servidor FTP, que normalmente usa a porta número 21 como seu modo padrão de comunicação. Além disso, as conexões FTP podem ser estabelecidas em dois modos diferentes: o modo ativo e o modo passivo.

No FTP ativo, o servidor assume uma função ativa aprovando uma solicitação de dados. No entanto, o modo ativo às vezes pode ter problemas com *firewalls*, que bloqueiam sessões não autorizadas. Este problema é resolvido com o modo passivo onde o servidor não

mantém ativamente a conexão, o cliente inicia o comando e a conexão de dados (MUNZERT, 2015).

#### 4.8 GIT E GITHUB

Para disponibilizar os códigos referentes ao trabalho, é utilizada uma ferramenta de versionamento muito popular entre os programadores, o *Git*. Com diversas funções, o *Git* é mais utilizado por pesquisadores com intuito de compartilhar seus projetos. Com a principal característica de ser descentralizado, seu maior atributo é funcionar como servidor ou cliente. Analisando esses pontos, os programadores têm total liberdade, dependendo de seu nível de acesso, de fazer alterações de forma simultânea sem a ocorrência de erros e remotamente (RAM, 2013).

Ainda segundo a referência supracitada, por ser uma ferramenta com diferentes usabilidades, o *Git* mostra-se como a melhor opção para facilitar a colaboração. Com o controle de *commits*, o autor do projeto no qual está sendo trabalhado tem fácil acesso a todas as alterações feitas por outros colaboradores interessados em participar desse determinado projeto, podendo ser desde a criação de aplicativos para celular até grandes sistemas de análises de dados. Dessa forma, caso haja alguma alteração contrária ao ideal perseguido pelo autor, pode-se localizar esse *commit* e analisá-lo junto com o colaborador, a fim de chegar a um ponto de interesse entre os dois.

Como plataforma de codificação, é utilizado o *GitHub* – site colaborativo especializado em guardar códigos fontes baseado no sistema *Git*. Semelhante a uma rede social, essa plataforma tem opções para “seguir” e acompanhar projetos, além de adicionar a sua rede de amigos, colaboradores alinhados ao tema de seu trabalho. Ajudando a criar uma ligação mais forte com o projeto em desenvolvimento e ter maior visibilidade para futuras colaborações (KALLIAMVAKOU et. al., 2014).

## 5 METODOLOGIA

O projeto foi realizado exclusivamente na linguagem *Python*, usando diversas bibliotecas para concluir o objetivo final. A linguagem foi selecionada para este trabalho devido a sua flexibilidade e sintaxe limpa, dispondo de um grande conjunto de bibliotecas com recursos adicionais permitindo, portanto, construir programas sofisticados. Além disso, pode executar os códigos de *Web Scraping* exigidos, usando as bibliotecas *Wget* e *Selenium*.

Para melhor organização, controle, reutilização de linhas de comando e o entendimento do todo, de modo a facilitar a expansão organizada e a manutenção do programa, foi realizada uma modularização. Dessa forma, o código foi dividido em três módulos, a saber: “*downloader.py*”, “*looper.py*” e “*menu.py*”.

Como ambiente integral de desenvolvimento, foi utilizada a IDE da *Jetbrains*, o *PyCharm Community*. Ademais, para compilar os módulos na IDE é necessário que seja instalado na máquina o interpretador do *Python* na versão a qual se deseja programar. Diante disso, nesta seção é apresentado brevemente como baixar e instalar tanto o interpretador *Python* quanto a IDE *Pycharm*, o qual será utilizado para compilar os códigos, além de uma breve descrição dos módulos “*downloader.py*”, “*looper.py*” e “*menu.py*”, bem como algumas observações relevantes ao leitor.

### 5.1 DOWNLOAD E INSTALAÇÃO DO INTERPRETADOR PYTHON

Para realizar o *download* do interpretador *Python* basta acessar o site do *Python*, na seção de *downloads*, ou através do link: <https://www.python.org/downloads/>. Dessa forma, é possível baixar a versão do interpretador *Python* de preferência do usuário, para este trabalho foi utilizada a versão 3.6.0. Após essa etapa, o *download* de um executável será iniciado, ao fim do *download*, basta executar o arquivo com dois cliques. A instalação recomendável segue-se clicando em *Install Now*, onde será instalado em um diretório padrão. Em seguida, uma tela de progresso irá aparecer. Após a finalização, basta clicar em *close* para concluir a instalação.

### 5.2 BAIXANDO E INSTALANDO A IDE PYCHARM

Para instalar o *PyCharm*, basta acessar o site da empresa *JeBrains*, na aba *Tools* e na seção IDEs, selecionando a opção *PyCharm*, ou acessar diretamente através do link:

<https://www.jetbrains.com/pt-br/pycharm/download/#section=windows>. Ressalta-se que a opção de download que deve ser selecionada é a Community por ser a opção gratuita deste IDE. A seguir, é realizado o download de um executável, sendo necessário executar o instalador e aceitar as permissões de segurança. Dessa forma, uma janela será aberta e para concluir a instalação, basta ir clicando em “Next >”. Após esses passos, basta selecionar se deseja reiniciar agora (*Reboot now*) ou reiniciar depois manualmente (*I want to manually reboot later*) e clicar em “Finish”. Por fim, o PyCharm estará instalado e pronto para uso.

### 5.3 INSTALAÇÃO DE BIBLIOTECAS PYTHON

Uma das grandes vantagens de se programar em Python é o grande número de bibliotecas disponíveis para inúmeras operações diferentes. Muitas bibliotecas vêm instaladas no interpretador do Python, porém caso haja interesse, é possível baixar outras bibliotecas que disponibilizam outras funções, classes, entre outros.

Para realizar a instalação das bibliotecas adicionais através da IDE utilizada, ou seja, utilizando o *PyCharm*, basta seguir o caminho: *File* → *Settings* → *Project* → *Project Interpreter* → *Install* (ícone de “+” no canto superior direito da tela). Nesta etapa, basta digitar o nome da biblioteca desejada, no caso do presente trabalho, a biblioteca *Wget* e *Selenium* e seguir sua instalação de forma intuitiva.

### 5.4 MENU.PY

O “*menu.py*” é o módulo para interatividade com o usuário, onde cada opção solicita uma função do módulo “*downloader.py*”. No que diz respeito aos dados de entrada necessários para seu funcionamento, o programa demanda apenas o fornecimento das datas do atual calendário Gregoriano desejadas.

### 5.5 LOOPER.PY

O “*looper.py*” é o módulo que automatiza o *download* dos dados diários, ou seja, possui uma estrutura de repetição *while* baseada em contagem regressiva, a qual é responsável por executar o módulo “*downloader.py*” periodicamente, no caso, toda 01h00min. Para que os dados sejam baixados diariamente o “*looper.py*” deve estar executando no computador. Além disso, o módulo também informa quanto tempo falta para



baixar os dados do IBGE e da NASA de modo que o usuário consiga visualizar o programa funcionando.

## 5.6 DOWNLOADER.PY

O "*downloader.py*" é o módulo que possui as funções para realizar os *downloads* das observações GNSS das estações da RBMC e das efemérides precisas, localizadas, respectivamente, no site do IBGE e no site da NASA. Tal módulo é utilizado como uma biblioteca pelos módulos "*menu.py*" e "*looper.py*", não sendo necessário executá-lo. O "*downloader.py*" foi dividido em funções com a finalidade de organizar o programa em blocos de código para que algumas instruções não precisassem ser repetidas. É necessário ressaltar que a forma de *downloads* de ambos os sites são diferentes, o do IBGE é realizado via FTP e o da NASA via HTTP, sendo este último mais complexo por possuir *login* e senha, dificultando assim, a automação.

### 5.6.1 Download IBGE

O *download* via IBGE, como descrito anteriormente, é voltado para obtenção das observações GNSS das estações da RBMC. Seu funcionamento baseia-se em acesso à base de dados via FTP para obtenção dos arquivos desejados. Para o *download* dos dados foi utilizada a biblioteca adicional *wget*, a qual é uma ferramenta, não interativa, para realizar transferências de arquivos através de comunicação com o banco de dados do IBGE via protocolo FTP, utilizando poucas linhas de código, sendo necessário apenas importar a biblioteca, declarar a URL desejada em uma variável e utilizar a função *download()*.

As observações GNSS das estações da RBMC estão disponíveis em estrutura de subdiretórios sendo estes anuais e dias corridos do ano, o diretório inicial para esses arquivos é: [https://geoftp.ibge.gov.br/informacoes\\_sobre\\_posicionamento\\_geodesico/rbmc/dados/](https://geoftp.ibge.gov.br/informacoes_sobre_posicionamento_geodesico/rbmc/dados/)

A URL para o *download* de cada arquivo segue a seguinte lógica: URL do diretório inicial/ano/dias corridos do ano/nome do arquivo. O nome do arquivo se dá com base na sigla da estação RBMC, os dias do ano contados sequencialmente desde o dia 01 de janeiro e a sua ordem de gravação, por exemplo, o arquivo "perc0701.zip" contém em seu nome sua respectiva estação RBMC [RECIFE], o dia corrido do ano [070] e sua ordem de gravação [1]. Dessa forma, a URL de *download* para o arquivo citado anteriormente tem a forma:

[https://geoftp.ibge.gov.br/informacoes\\_sobre\\_posicionamento\\_geodesico/rbmc/dados/2021/070/perc0701.zip](https://geoftp.ibge.gov.br/informacoes_sobre_posicionamento_geodesico/rbmc/dados/2021/070/perc0701.zip).

Para facilitar a organização dos arquivos obtidos, o programa distribui os arquivos baixados em um sistema de pastas, utilizando para isso, a biblioteca *os*, a qual fornece funções relacionadas a funcionalidades que são dependentes do sistema operacional. A partir dessa biblioteca foi possível fazer manipulações de estruturas de diretórios, tais como verificar se o diretório existe, criar diretórios e tratar exceções. Os dados obtidos são separados por ano, mês e o dia dentro de uma pasta denominada IBGE, no qual a nomenclatura do arquivo é a mesma utilizada pela instituição.

### 5.6.2 Download NASA

O *download* via NASA é realizado por meio do CDDIS, como já mencionado, é voltado para obtenção das efemérides precisas. Seu funcionamento baseia-se em acesso à base de dados via HTTP para obtenção dos arquivos desejados. Anteriormente, os arquivos CDDIS eram acessados via protocolo FTP anônimo que permitia os usuários automatizarem facilmente os downloads de arquivos, porém havia problemas do ponto de vista do sistema de segurança. Por causa das restrições de segurança da NASA, após o dia 31 de outubro de 2020, o CDDIS extinguiu o uso do FTP não seguro para *uploads* de arquivos de provedores de dados e passou a utilizar um sistema novo e atualizado, o qual foi projetado para usar protocolo HTTP para *upload* de arquivos (MICHAEL; NOLL, 2019; MICHAEL; NOLL, 2020).

Atualmente, as operações de *upload* de arquivo CDDIS via HTTP usa o processo de *login* EOSDIS *Earthdata*, requer que todos os provedores de dados se registrem no EOSDIS para uma conta de usuário, têm disponibilidade tanto na *web* quanto em linha de comando, permite a criação de *scripts* com URL e outras opções do usuário, além de eliminar o problema de duas portas com FTP e *firewalls* (NOLL; MICHAEL, 2018).

Para realizar o *download* dos dados foi utilizado o *framework Selenium* (MUTHUKADAN, 2021). O *Selenium Python* possibilita a automação de um navegador ou testes em aplicações *web*, além de fornecer uma API simples para escrever testes funcionais ou de aceitação usando o *WebDriver*. O *WebDriver* é o programa que intermedia a comunicação entre um *script* e um navegador. Por meio das funcionalidades do *WebDriver* é possível simular ações do usuário dentro do navegador, como, por exemplo, realizar *login* em um site, rolar a barra da página do site, clicar em botões e redirecionamento.

Cada navegador possui seu próprio *WebDriver* e é geralmente disponibilizado pelos próprios desenvolvedores dos navegadores, sendo necessário, além de ter a biblioteca devidamente instalada em seu ambiente de desenvolvimento, baixar um *WebDriver*. No caso do presente trabalho foi utilizado o *GeckoDriver* (MOZILLA, 2021) que é o *WebDriver* para o *Mozilla Firefox*, pois não é necessário nenhuma configuração adicional, basta que ele esteja instalado. Além disso, para manipulação da página *web* com o *WebDriver* foi necessário possuir conhecimento em HTML, pois o programa percorre a página *web* localizando *tags* para executar alguma ação.

As soluções processadas pelos centros de análise IGS estão disponíveis em uma estrutura de diretórios com as semanas GPS no acervo do CDDIS, onde as URLs para acessar os diretórios iniciais são:

Quadro 1 - URLs dos diretórios iniciais das soluções processadas pelo IGS

| Soluções            | URL                                                                                                             |
|---------------------|-----------------------------------------------------------------------------------------------------------------|
| GPS e GPS + GLONASS | <a href="https://cddis.nasa.gov/archive/gnss/products/">https://cddis.nasa.gov/archive/gnss/products/</a>       |
| Somente GLONASS     | <a href="https://cddis.nasa.gov/archive/glonass/products/">https://cddis.nasa.gov/archive/glonass/products/</a> |

Fonte: A Autora (2021).

A URL para o *download* de cada arquivo segue a seguinte lógica: URL do diretório inicial/semana GPS/nome do arquivo. Onde o nome do arquivo se dá com base na sigla do órgão que a produziu, na semana GPS correspondente e no dia da semana GPS (começa com 0 no domingo e vai até 6 no sábado), por exemplo, o arquivo “igs21480.sp3.z” contém as efemérides precisas finais [IGS] do domingo [0] da semana GPS [2148] e sua extensão [sp3]. Dessa forma, a URL de *download* para o arquivo citado anteriormente tem a forma: <https://cddis.nasa.gov/archive/gnss/products/2148/igs21480.sp3.Z>.

Com a finalidade de facilitar a entrada dos dados no programa, foi criada uma função que calcula a semana GPS, fazendo com que o usuário trabalhe apenas com datas do atual calendário Gregoriano ao invés de entrar com a semana GPS. Tal cálculo foi realizado com o embasamento do tópico 4.4.2.1, onde tem como data de referência o dia 6 de janeiro de 1980 (domingo). Para tanto, foi utilizado o módulo *datetime* do *Python*, o qual fornece ferramentas para manipulação de datas e horas.

Assim como no *download* via IBGE, para facilitar a organização dos arquivos obtidos, o *download* via NASA também distribui os arquivos baixados em um sistema de pastas. Tendo como arquivo de saída as efemérides IGR, IGS e IGL, sendo a última relativa à

constelação GLONASS. Os dados obtidos são separados por ano, mês, dia e sigla do órgão que a produziu dentro de uma pasta denominada NASA, no qual a nomenclatura do arquivo é a mesma utilizada pela instituição.

## 5.7 OBSERVAÇÕES RELEVANTES AO LEITOR

O funcionamento do programa no geral só é possível havendo conexão com a internet durante todo o período de *download* e ter instalado no computador o *Mozilla FireFox*, podendo ser executado tanto em sistema *Linux* quanto em *Windows*. Além disso, os módulos devem estar localizados na mesma pasta juntamente com o *GeckoDriver*.

No que se refere a sua interface, seguindo um propósito de minimização de custo computacional, optou-se por mantê-lo em linha de comando. Adicionalmente, ressalta-se que foram realizados todos os testes de consistência nas possíveis entradas do usuário. Para fins de brevidade, os códigos são omitidos do presente trabalho, estando estes disponíveis para livre acesso à comunidade em Melo (2021), um repositório virtual do GitHub. Neste item constam apenas algumas observações e indicações relevantes ao usuário.

Além disso, é importante salientar que na pasta do programa no GitHub encontra-se o *GeckoDriver* para o Windows na versão 64, caso o sistema do usuário seja diferente, é necessário fazer o download para o sistema adequado em Mozilla (2021).

## 6 RESULTADOS E DISCUSSÕES

Como mencionado anteriormente, o programa é um conjunto de três módulos denominados de “*downloader.py*”, “*looper.py*” e “*menu.py*”. Ambos os módulos estão disponíveis em <https://github.com/MarcellaFMelo/Program>, repositório virtual do GitHub, em uma pasta sob o nome de “*program*”, com seus respectivos códigos abertos.

A interface inicial do “*menu.py*” é exibida na Figura 5, onde as opções do usuário são: [1] - Baixar dados desde 01/01/2010 (realiza o *download* dos dados desde o dia 01/01/2010, data determinada como mínima pelo desenvolvedor); [2] - Baixar dados desde uma data específica (realiza o *download* dos dados em um intervalo de tempo, variando desde o dia anterior ao inserido pelo usuário até a data atual); [3] - Baixar dados de ontem (realiza o *download* dos dados do dia anterior) e [4] – Sair (Finaliza o programa).

Figura 5 – Exibição da interface inicial do “*menu.py*”

```

=== Downloads das observações GNSS das estações da RBMC e das efemérides precisas do IGS ===

[1]-Baixar dados desde 01/01/2010
[2]-Baixar dados desde uma data específica
[3]-Baixar dados de ontem
[4]-Sair

Favor digitar o número da opção desejada:

```

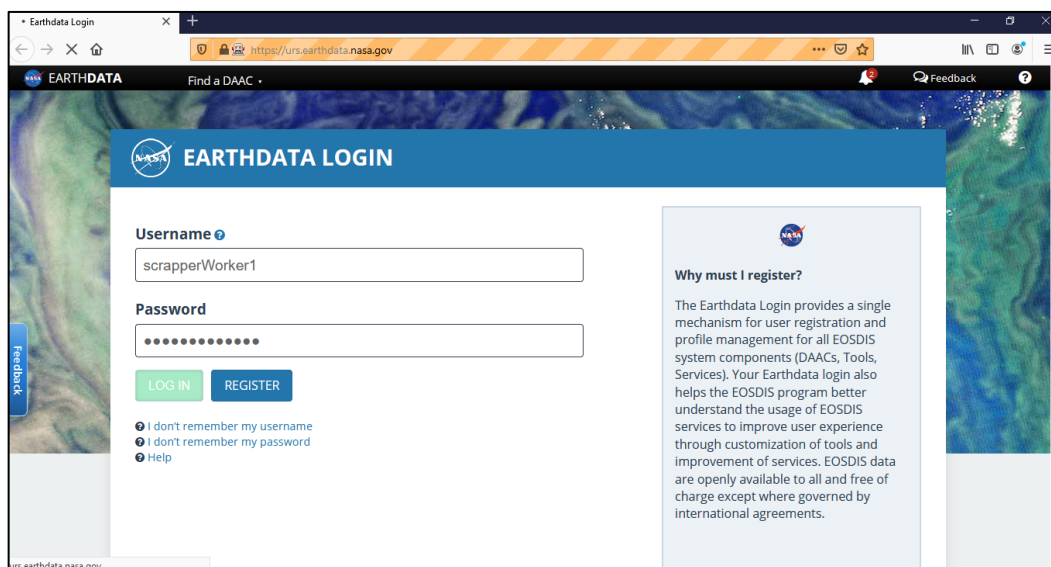
Fonte: A Autora (2021).

Todas as opções do “*menu.py*” e o “*looper.py*” demandam o *download* das efemérides precisas do IGS via NASA. Por consequência, o *GeckoDriver*, *WebDriver* do *Selenium*, realiza a automação do navegador *Mozilla Firefox*, simulando ações do usuário dentro do navegador. Assim, o *GeckoDriver* abre a janela de navegação do *Mozilla Firefox* com a URL de *login* do *Earthdata*, insere o *username* e o *password* já cadastrados nos campos apropriados e clica no botão de *login* automaticamente, como demonstrado na Figura 6.

Posteriormente, o programa aguarda 60s antes de iniciar o processo de *download*, pois após o *login*, o site realiza o redirecionamento da página, como ilustrado na Figura 7.

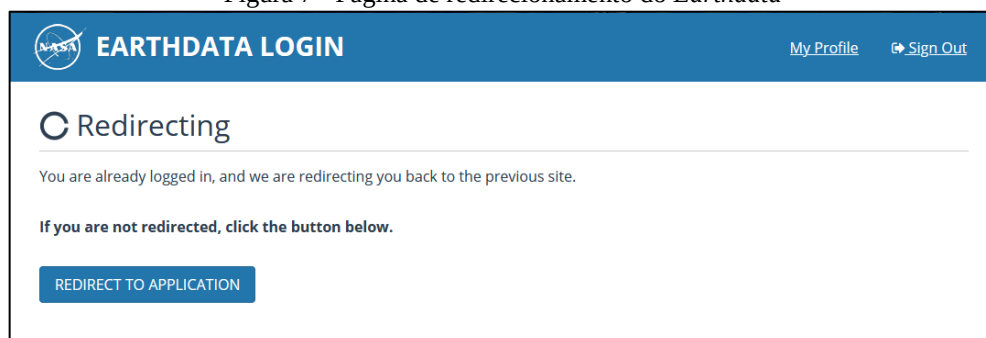
Em seguida, o *GeckoDriver* inicia o processo de *download* das efemérides precisas trabalhando com alternância de janelas do navegador. O programa inicia o *Firefox* com o diretório principal (ilustrado na Figura 8), posteriormente, abre o diretório secundário (pasta da semana GPS desejada) em outra janela, baixa o arquivo pretendido e depois fecha a janela, voltando para o diretório principal.

Figura 6 - Interface de *login* do *Earthdata* usada para ter acesso aos produtos do IGS



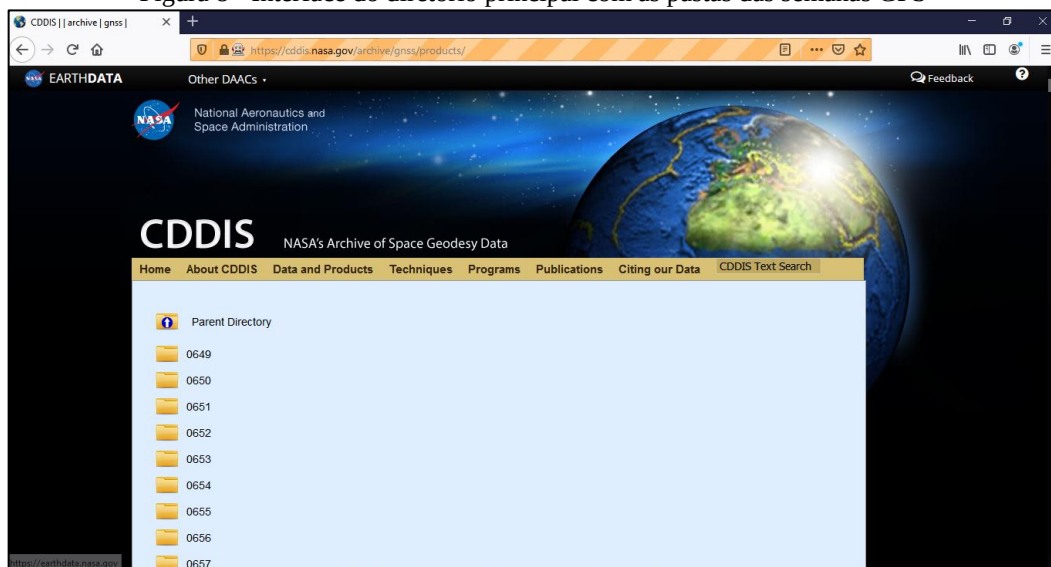
Fonte: A Autora (2021).

Figura 7 - Página de redirecionamento do *Earthdata*



Fonte: A Autora (2021).

Figura 8 - Interface do diretório principal com as pastas das semanas GPS



Fonte: A Autora (2021).

A opção [1] do “*menu.py*” realiza o *download* das observações GNSS das estações da RBMC e das efemérides precisas do IGS desde o dia 01/01/2010, tal data foi definida como mínima devido ao fato de ser a data mínima de *download* no servidor FTP do IBGE. Além disso, o programa pergunta se o usuário tem certeza de que quer baixar todos os dados desde essa data, caso a resposta seja sim, o programa realiza o *download*, como demonstrado na Figura 9. Essa opção se torna relevante na criação da infraestrutura para acervo geodésico, podendo realizar o *download* dos dados disponíveis dos servidores de ambas as instituições.

Figura 9 – Execução da opção [1] do “*menu.py*”

```

=== Downloads das observações GNSS das estações da RBMC e das efemérides precisas do IGS ===

[1]-Baixar dados desde 01/01/2010
[2]-Baixar dados desde uma data específica
[3]-Baixar dados de ontem
[4]-Sair

Favor digitar o número da opção desejada: 1
Certeza que voce deseja baixar todos os dados desde 01/01/2010?
Sim(S) ou Não(N): S
downloadMinimalYear
01/01/2010 - 07/04/2021

```

Fonte: A Autora (2021).

A opção [2] do “*menu.py*” realiza o *download* dos dados em um intervalo de tempo, variando desde o dia anterior ao inserido pelo usuário até o atual momento, onde o programa solicita o fornecimento da data desejada, indicando o formato exigido (dd/mm/aaaa). Tal procedimento é ilustrado na Figura 10.

Figura 10 – Execução da opção [2] do “*menu.py*”

```

=== Downloads das observações GNSS das estações da RBMC e das efemérides precisas do IGS ===

[1]-Baixar dados desde 01/01/2010
[2]-Baixar dados desde uma data específica
[3]-Baixar dados de ontem
[4]-Sair

Favor digitar o número da opção desejada: 2
Favor informar a data no formato dd/mm/aaaa. Ex: 31/01/2020
05/04/2021
Download de Período: 04/04/2021 - 07/04/2021

```

Fonte: A Autora (2021).

O “*looper.py*” é o módulo que realiza a automação dos *downloads* dos dados diários, devendo este permanecer executando no computador para o seu funcionamento. Os *downloads* diários são programados para que ocorram todos os dias à 01h00min. Ademais, com o intuito de o usuário conseguir visualizar o programa funcionando, o módulo faz uma contagem regressiva do tempo restante para baixar os dados, como demonstrado na Figura 11.

Figura 11 – Execução do módulo “*looper.py*”

```
Download de ontem: 06/04/2021
10:13:49|para baixar dados do IBGE e NASA
10:13:48|para baixar dados do IBGE e NASA
10:13:47|para baixar dados do IBGE e NASA
10:13:46|para baixar dados do IBGE e NASA
10:13:45|para baixar dados do IBGE e NASA
10:13:44|para baixar dados do IBGE e NASA
10:13:43|para baixar dados do IBGE e NASA
10:13:42|para baixar dados do IBGE e NASA
10:13:41|para baixar dados do IBGE e NASA
10:13:40|para baixar dados do IBGE e NASA
10:13:39|para baixar dados do IBGE e NASA
10:13:38|para baixar dados do IBGE e NASA
10:13:37|para baixar dados do IBGE e NASA
10:13:36|para baixar dados do IBGE e NASA
```

Fonte: A Autora (2021).

A fim de facilitar a estruturação de armazenamento dos dados para criação de uma infraestrutura de acervo geodésico, o programa distribui os arquivos baixados em um sistema de diretórios dentro da pasta em que os módulos se encontram. A distribuição dos sistemas de pastas dos dados das observações GNSS das estações da RBMC e das efemérides precisas são dadas, respectivamente, por:

- (i) Program → Data → ano → mês → dia → IBGE → Estações da RBMC;
- (ii) Program → Data → ano → mês → dia → NASA → Efemérides precisas (IGL, IGR E IGS).

Com o intuito de dimensionar as melhorias impostas pelo programa desenvolvido em relação aos métodos tradicionais de obtenção dos dados das observações GNSS das estações



da RBMC e das efemérides precisas, foi utilizada uma série de parâmetros comparativos informais entre ambos, conforme disposto no Quadro 2.

Quadro 2 – Comparativo de obtenção dos dados alvo do trabalho entre o método convencional e o uso do programa

| <b><i>Downloads dos dados pelo método convencional</i></b>                       | <b><i>Downloads dos dados com o uso do programa</i></b>                                                                         |
|----------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|
| Demanda conexão com a internet                                                   | Demanda conexão com a internet                                                                                                  |
| Processo manual, exigindo tempo e atenção do usuário a cada passo.               | Processo automático, permitindo ao usuário a execução de outras atividades enquanto os dados são adquiridos.                    |
| Possibilidade de falha humana durante a aquisição.                               | Não há possibilidade de falha humana durante a aquisição, desde que os dados de entrada estejam corretos.                       |
| Demanda organização do usuário ao realizar os downloads dos dados.               | Facilita a estruturação de armazenamento dos dados, pois o programa distribui os arquivos baixados em um sistema de diretórios. |
| Maior tempo para aquisição                                                       | Menor tempo para aquisição (dependendo da velocidade da conexão com a internet).                                                |
| Processo mecânico e repetitivo, sem aproveitamento entre as etapas consecutivas. | Código livre com aproveitamento de funções intermediárias.                                                                      |
| Os sites de ambas as instituições podem ficar fora do ar.                        | Os usuários terão acesso aos dados através do servidor.                                                                         |

Fonte: A Autora (2021).

Neste segmento, todos os elementos citados no Quadro 1 e todos os demais pontos abordados ao longo do presente trabalho atuam em prol do programa desenvolvido, ficando perceptível os benefícios relacionados ao uso do programa em comparação com o método tradicional.

## 7 CONSIDERAÇÕES FINAIS

O presente trabalho teve como objetivo automatizar *downloads* diários de dados GNSS com intuito de criar uma infraestrutura de acervo geodésico para consulta em processamentos de dados e de séries temporais. Para isso, foi desenvolvido um programa em linguagem *Python* com o auxílio das bibliotecas adicionais *Wget* e *Selenium*, podendo ser executado tanto em sistema *Linux* quanto em *Windows* por meio de um compilador, desde que tenha o *Mozilla Firefox* instalado. O programa encontra-se disponível para livre *download* à comunidade em um repositório virtual do *GitHub*, podendo este ser utilizado e implementado desde que devidamente referenciado.

Os módulos do programa possuem uma interface e funcionamentos intuitivos, apresentando-se como alternativa de grande importância principalmente aos usuários que tenham necessidade de realizar o *download* de grande quantidade de dados, evitando o envolvimento excessivo com tarefas repetitivas. Além disso, o programa impede possíveis falhas humanas, sendo capaz de proporcionar mais tempo para a execução de outras tarefas.

Como resultados obteve-se a automatização de *downloads* diários das observações GNSS das estações da RBMC e das efemérides providas, respectivamente, pelo IBGE e pelo IGS, além do êxito em adquirir grandes quantidades de dados organizados em uma estruturação de armazenamento. Tais dados contribuem para a criação da infraestrutura de acervo geodésico para futuras consultas em processamentos de dados e de séries temporais.

E por fim, a criação de uma infraestrutura de acervo geodésico vai auxiliar os alunos do Laboratório de Geodésia do Departamento de Engenharia Cartográfica e Agrimensura da Universidade Federal de Pernambuco em suas pesquisas, facilitando o tratamento dos dados e permitindo utilizar o tempo gasto de *download* manual para se dedicar à análise e à produção científica. Além disso, futuramente os alunos terão à disposição todos os dados em um servidor apresentando vantagens a usuários que precisem realizar o *download* e se deparem com o inconveniente dos sites fora do ar.

## REFERÊNCIAS

- ACHANTA, S.; WATT, S. T.; SAGEN, E.. Mitigating GPS vulnerabilities. *In: Proceedings of the 17th Annual Power and Energy Automation Conference*. 2015.
- ALVES, D. B. M.; ABREU, P. A. G.; SOUZA, J. S. S. GNSS: status, modelagem atmosférica e métodos de posicionamento. **Revista Brasileira de Geomática**, v. 1, n. 1, 2013.
- BERNERS-LEE, T. J. **Information management**: a proposal. 1989.
- BIANCHI, C. E., MENDOZA, L. P. O., FERNÁNDEZ, L. I., NATALI, M. P., MEZA, A. M.; MOIRANO, J. F. Multi-year GNSS monitoring of atmospheric IWV over Central and South America for climate studies. *In: Annales Geophysicae*. Copernicus GmbH, 2016. p. 623-639.
- COSTA, S. M. A., LIMA, M. A. A., JÚNIOR, N., ABREU, M. A., SILVA, A., FORTES, L. P. S. RBMC em tempo real, via NTRIP, e seus benefícios nos levantamentos RTK e DGPS. **Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação**, v. 2, p. 8-11, 2008.
- FALL, K. R.; STEVENS, W. R. **TCP/IP illustrated**: the protocols. addison-Wesley, 2011 , v. 1.
- FAROLFI, G.; PIOMBINO, A.; CATANI, F. Fusion of GNSS and Satellite Radar Interferometry: Determination of 3D Fine-Scale Map of Present-Day Surface Displacements in Italy as Expressions of Geodynamic Processes. **Remote Sensing**, v. 11, n. 4, p. 394, 2019.
- FENG, Y. GNSS three carrier ambiguity resolution using ionosphere-reduced virtual signals. **Journal of Geodesy**, v. 82, n. 12, p. 847-862, 2008.
- FOROUZAN, B. A. **TCP/IP protocol suite**. McGraw-Hill Higher Education, 2002.
- GPS. **Global Positioning System**. Space segment. Disponível em : <https://www.gps.gov/systems/gps/space/>. Acesso em: 20 de fev. de 2021.
- GUEROVA, G., JONES, J., DOUŠA, J., DICK, G., HAAN, S. D., POTTIAUX, E., ... & BENDER, M. Review of the state of the art and future prospects of the ground-based GNSS meteorology in Europe. **Atmospheric Measurement Techniques**, v. 9, n. 11, p. 5385-5406, 2016.
- HEROUX, P, KOUBA, J. GPS precise point positioning using IGS orbit products. **Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy** .26(6–8), p. 573–578, 2001. DOI. 10.1016/S1464-1895(01)00103-X.
- HOFMANN-WELLENHOF, B.; LICHTENEGGER, H.; COLLINS, J. **GNSS – Global Navigation Satellite Systems GPS, GLONASS, Galileo, and more**. New York: Springer-Verlag, 2008. DOI: 10.1007/978-3-211-73017-1.

IAC. **Information and Analysis Center: GLONASS**. Disponível em: <https://www.glonass-iac.ru/en/GLONASS/>. Acesso em: 24 de fev. de 2021.

IBGE. **Banco de Dados Geodésicos - BDG**. Disponível em: <http://www.bdg.ibge.gov.br/appbdg/>. Acesso em: 10 de fev. de 2021. 2021a.

IBGE. **Download de Produtos - RBMC**. Disponível em: <https://www.ibge.gov.br/geociencias/informacoes-sobre-posicionamento-geodesico/rede-geodesica/16258-rede-brasileira-de-monitoramento-continuo-dos-sistemas-gnss-rbmc.html?=&t=downloads>. Acesso em: 10 de fev. de 2021. 2021c.

IBGE. **Rede Brasileira de Monitoramento Contínuo dos Sistemas GNSS - RBMC**. Disponível em: <https://www.ibge.gov.br/geociencias/informacoes-sobre-posicionamento-geodesico/rede-geodesica/16258-rede-brasileira-de-monitoramento-continuo-dos-sistemas-gnss-rbmc.html?=&t=sobre>. Acesso em: 10 de fev. de 2021. 2021b.

IGS. **International GNSS Service**. Disponível em: <https://www.igs.org/>. Acesso: 12 de fev. de 2021.

KALLIAMVAKOU, E., GOUSIOS, G., BLINCOE, K., SINGER, L., GERMAN, D. M., DAMIAN, D.. The promises and perils of mining github. *In: Proceedings of the 11th working conference on mining software repositories*. 2014. p. 92-101.

KAMINSKIS, J., SULAKOVA, L., SALMIS, K., KAULINS, J., & GOLDBERGS, L. SLR and GNSS Test Field for Global Geodetic Network Assessment in Riga. *In: Environmental Engineering. Proceedings of the International Conference on Environmental Engineering. ICEE*. Vilnius Gediminas Technical University, Department of Construction Economics & Property, 2020. p. 1-6.

KOUBA, J.; LAHAYE, F.; TÉTREAULT, P. Precise Point Positioning. *In: TEUNISSEN, P. J. G.; MONTENBRUCK, O. (Ed.) Springer Handbook of Global Navigation Satellite Systems*. Berlin: Springer, 2017. p. 723-752.

MELO, M. F. O. **Program**. Repositório. Disponível em: <https://github.com/MarcellaFMelo/Program>. Acesso em: 10 de abr. de 2021.

MICHAEL, B. P.; NOLL, C. E. NASA CDDIS: Supporting Global Geodetic and Geophysical Research and Applications. *In: AGU Fall Meeting Abstracts*. 2019. p. IN31B-0794.

MICHAEL, B. P.; NOLL, C. E. Technology Changes and User Community Reaction—An Example from CDDIS. *In: AGU Fall Meeting 2020*. AGU, 2020.

MITCHELL, R. **Web Scraping com Python**. São Paulo: Novatec Editora Ltda, 2016.

MONICO, J. F. G. **Posicionamento pelo GNSS – Descrição, fundamentos e aplicações**. São Paulo: Editora UNESP, 2008. 476p.

MOZILLA. **GeckoDriver**. Repositório. Disponível em: <https://github.com/mozilla/gecko-driver/releases>. Acesso em: 22 de mar. de 2021.

MUNZERT, S. **Automated data collection with r a practical guide to web scraping and text mining**. John Wiley & Sons Inc., 2015.

MUTHUKADAN, B. **Selenium with Python**. Disponível em: <https://selenium-python.readthedocs.io/index.html>. Acesso em: 18 de mar. de 2021.

NOLL, C. E.; MICHAEL, B. P. Important Changes to User Access at the NASA CDDIS. In: **AGU Fall Meeting Abstracts**. 2018. p. G31B-0675.

ODIJK, D. Positioning Model. In: TEUNISSEN, P. J. G.; MONTENBRUCK, O. (Ed.) **Springer Handbook of Global Navigation Satellite Systems**. Springer, Berlin, 2017. p. 605-638.

PENMAN, R. B.; BALDWIN, T.; MARTINEZ, D. Web scraping made simple with sitescraper. **Retrieved from the internet**, 2009.

PIOTROWSKI, A. An Analysis of the use of the Python Language in Robot Applications. **Applied Computer Science**, v. 12, n. 2, 2016.

PYTHON BRASIL. **Ides para python**. Disponível em: <https://python.org.br/ferramentas/>. Acesso em: 18 de mar. de 2021.

RAHMANI, Y., ALIZADEH, M. M., SCHUH, H., WICKERT, J., & TSAI, L. C. Probing vertical coupling effects of thunderstorms on lower ionosphere using GNSS data. **Advances in Space Research**, v. 66, n. 8, p. 1967-1976, 2020.

RAM, K. Git can facilitate greater reproducibility and increased transparency in science. **Source code for biology and medicine**, v. 8, n. 1, p. 1-8, 2013.

REVNIVYKH, S.; BOLKUNOV, A.; SERDYUKOV, A.; MONTENBRUCK, O. GLONASS. In: TEUNISSEN, P. J. G.; MONTENBRUCK, O. (Ed.) **Springer Handbook of Global Navigation Satellite Systems**. Springer, Berlin, 2017. p. 219-246.

SEEBER, G. **Satellite Geodesy**. 2. ed. New York: Walter de Gruyter, 2003.

SEEPERSAD, G.; BISNATH, S. Challenges in assessing PPP performance. **Journal of Applied Geodesy** v. 8, n. 3, p. 205–222, 2014. DOI. 10.1515/jag-2014-0008.

SOŚNICA, K.; BURY, G.; ZAJDEL, R.. Contribution of Multi-GNSS Constellation to SLR-Derived Terrestrial Reference Frame. **Geophysical Research Letters**, v. 45, n. 5, p. 2339-2348, 2018.

SPIPKER, J. J. Tropospheric Effects on GPS. PARKINSON, B. W.; SPIPKER, J.J. **Global positioning system: Theory and applications**, v.1. Cambridge, American Institute of Aeronautics and Astronautics, p.517-46, 1996.

SPOFFORD, P.R.; REMONDI, B.W. **The National Geodetic Survey Standard GPS Format SP3**, 1996. Disponível em: [https://www.ngs.noaa.gov/orbits/sp3\\_docu.txt](https://www.ngs.noaa.gov/orbits/sp3_docu.txt). Acesso em: 20 de fev. de 2021.

VARGIU, E.; URRU, M. Exploiting web scraping in a collaborative filtering- based approach to web advertising. **Artif. Intell. Research**, v. 2, n. 1, p. 44 – 54, 2013.

YEDA, L.; BING, Z.; RENQIANG, W. Application of Python language in UOE molding simulation of pipeline steel. *In: MATEC Web of Conferences*. EDP Sciences, 2018. p. 01018.

ZANUTTA, A. N. T. O. N. I. O., NEGUSINI, M. O. N. I. A., VITTUARI, L. U. C. A., CIANFARRA, P., SALVINI, F., MANCINI, F., ... & CAPRA, A. Monitoring geodynamic activity in the Victoria Land, East Antarctica: Evidence from GNSS measurements. **Journal of Geodynamics**, v. 110, p. 31-42, 2017.