

UNIVERSIDADE DE PASSO FUNDO
INSTITUTO DE CIÊNCIAS EXATAS E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM
COMPUTAÇÃO APLICADA

CPTEC TOOLKIT: um conjunto de
ferramentas para acesso e
disponibilização de dados climáticos

Felipe Borella

Passo Fundo

2018

UNIVERSIDADE DE PASSO FUNDO
INSTITUTO DE CIÊNCIAS EXATAS E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

**CPTEC TOOLKIT: UM CONJUNTO
DE FERRAMENTAS PARA ACESSO
E DISPONIBILIZAÇÃO DE DADOS
CLIMÁTICOS**

Felipe Borella

Dissertação apresentada como requisito parcial
à obtenção do grau de Mestre em Computação
Aplicada na Universidade de Passo Fundo.

Orientador: Prof. Dr. Carlos Amaral Holbig

Coorientador: Prof. Dr. Willingthon Pavan

Passo Fundo

2018

CIP – Catalogação na Publicação

B731c Borella, Felipe
CPTEC TOOLKIT : um conjunto de ferramentas para
acesso e disponibilização de dados climáticos / Felipe
Borella. – 2018.
74 f. : il. color. ; 30 cm.

Orientador: Prof. Dr. Carlos Amaral Holbig.
Coorientador: Prof. Dr. Willingthon Pavan.
Dissertação (Mestrado em Computação Aplicada) –
Universidade de Passo Fundo, 2018.

1. Computação. 2. Mudanças climáticas. 3. Plataforma
PROJETA. 4. Sistemas de Informações Geográficas.
I. Holbig, Carlos Amaral, orientador. II. Pavan, Willingthon,
coorientador. III. Título.

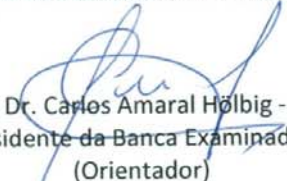
CDU: 004

Catalogação: Bibliotecária Jucelei Rodrigues Domingues - CRB 10/1569

**ATA DE DEFESA DO
TRABALHO DE CONCLUSÃO DE CURSO DO ACADÊMICO**

FELIPE BORELLA

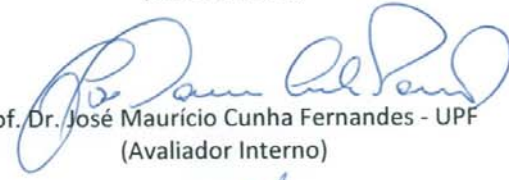
Aos dezoito dias do mês de dezembro do ano de dois mil e dezessete, às 15 horas e 30 minutos, realizou-se, no Instituto de Ciências Exatas e Geociências, prédio B5, da Universidade de Passo Fundo, a sessão pública de defesa do Trabalho de Conclusão de Curso **“CPTec TOOLKIT: Um conjunto de ferramentas para acesso e disponibilização de dados climáticos”**, de autoria de Felipe Borella, acadêmico do Curso de Mestrado em Computação Aplicada do Programa de Pós-Graduação em Computação Aplicada – PPGCA/UPF. Segundo as informações prestadas pelo Conselho de Pós-Graduação e constantes nos arquivos da Secretaria do PPGCA, o aluno preencheu os requisitos necessários para submeter seu trabalho à avaliação. A banca examinadora foi composta pelos doutores Carlos Amaral Hölbig, Willingthon Pavan, José Maurício Cunha Fernandes e Alexandre Tagliari Lazzaretti. Concluídos os trabalhos de apresentação e arguição, a banca examinadora considerou o candidato APROVADO. Foi concedido o prazo de até quarenta e cinco (45) dias, conforme Regimento do PPGCA, para o acadêmico apresentar ao Conselho de Pós-Graduação o trabalho em sua redação definitiva, a fim de que sejam feitos os encaminhamentos necessários à emissão do Diploma de Mestre em Computação Aplicada. Para constar, foi lavrada a presente ata, que vai assinada pelos membros da banca examinadora e pela Coordenação do PPGCA.



Prof. Dr. Carlos Amaral Hölbig - UPF
Presidente da Banca Examinadora
(Orientador)




Prof. Dr. Willingthon Pavan - UPF
(Coorientador)



Prof. Dr. José Maurício Cunha Fernandes - UPF
(Avaliador Interno)



Prof. Dr. Alexandre Tagliari Lazzaretti - IFSul
(Avaliador Externo)



Prof. Dr. Rafael Rieder
Coordenador do PPGCA

AGRADECIMENTOS

Agradecer primeiramente a Deus, pela saúde e por me nortear pelos melhores caminhos.

Ao meu orientador Carlos Amaral Holbig e coorientador Willingthon Pavan, pelos ensinamentos, sugestões e apoio nesta jornada.

À minha noiva Angelise Longo, pelo amor, carinho, paciência e ausência as vezes. Pelas palavras de incentivo nos momentos difíceis.

Aos meu pais Moacir Borella (in memoriam), Delva Botezini Borella que sempre me apoiaram em tudo. Esta conquista é para vocês.

Ao meu irmão André Borella e sua esposa Adriana Rutkoski pelo apoio e companheirismo.

Aos meus colegas de trabalho José Henrique Debastiani Andreis e Marcos Godoy Debastiani, pela troca de idéias e auxílio durante o desenvolvimento.

À equipe do grupo de pesquisa Mosaico da Universidade de Passo Fundo, em especial para Angela Mazzonetto, Guilherme Tocchetto, Jônatas Casiraghi Vaz e Pablo Chitolina, pela amizade, auxílio e companheirismo.

Ao professor Alexandre Tagliari Lazzaretti, por apresentar uma arquitetura semelhante a que precisávamos.

Ao pessoal do INPE por ter disponibilizado este trabalho, em especial a Chou Sin Chan, Luiz Henrique Coura, Jorge Luís Gomes, do qual eu me orgulho e muito por ter concluído.

Aos professores do PPGCA, por todo conhecimento passado.

Fica aqui meu muito obrigado a todos.

CPTEC TOOLKIT: UM CONJUNTO DE FERRAMENTAS PARA ACESSO E DISPONIBILIZAÇÃO DE DADOS CLIMÁTICOS

RESUMO

A previsão do tempo e do clima produzem dados importantes para tomada de decisão em diversas áreas, especialmente na agricultura, na hidrologia e na geração de energia. A combinação de tecnologias modernas e a disponibilização de dados meteorológicos fornecem um mecanismo para aumentar a produtividade e minimizar riscos relacionados ao clima. Dentro deste contexto, este trabalho foi elaborado visando o desenvolvimento de um kit de ferramentas para automatizar o processo de requisição, de disponibilização e de visualização de dados gerados por modelos de mudanças climáticas que são executados pelo CPTEC/INPE. O objetivo principal deste trabalho foi o estudo e desenvolvimento de uma plataforma computacional capaz de disponibilizar os dados de forma direta ao usuário. Este processo, normalmente, era realizado de forma manual no CPTEC/INPE. Com o desenvolvimento da primeira versão do CPTEC Toolkit, obteve-se como resultado a disponibilização de vários serviços, tanto via web como via API, que facilitaram o processo de requisição de dados e análise de mudanças climáticas e que puderam ser integrados ao PROJETA (Projeções de mudanças climáticas para a América do Sul regionalizadas pelo Modelo Eta), um projeto do INPE, UPF, Ministério do Meio Ambiente e agência de fomento GIZ da Alemanha.

Palavras-Chave: API, Big Data, Mudanças climáticas, Plataforma PROJETA, Sistemas de Informações Geográficas.

CPTEC TOOLKIT: A SET OF TOOLS TO ALLOW ACCESS AND TO PROVIDE CLIMATE DATA

ABSTRACT

Weather and climate forecast are major instruments for decision making in several areas, especially in agriculture, hydrology and power generation. The combination of modern technologies and readily available environmental data provides a valuable mechanism to increase productivity and minimize climate risks. Therefore, this study was carried out with the goal of building a toolkit to automate data requests, providing access and visualization of data generated by the weather forecast and climate models from CPTEC/INPE. The main objective was to develop a computational platform to provide data to users and API services. This Toolkit should replace manually querying and emailing the data. The first version of CPTEC Toolkit delivers several services via API, furthering the process of data request and climate change analysis on PROJETA (Projections of climatic changes for South America regionalized by Eta Model), a project of INPE, UPF, Ministry of the Environment and GIZ development agency of Germany.

Keywords: API, Big Data, Climate changes, PROJETA Platform, Geographic Information System.

LISTA DE FIGURAS

Figura 1.	Dimensões disponíveis e que devem ser consideradas ao avaliar a viabilidade de uma solução Big Data [18].	21
Figura 2.	Diversos sensores espalhados que podem enviar suas posições, serem acessados via web, controlados via web remotamente. Adaptado de Esri [46].	27
Figura 3.	Previsão de precipitação total acumulada de 24 h válida para 12 UTC 10 junho 1995 [8].	38
Figura 4.	Precipitação total observada no dia 10 de junho de 1995.[8].	39
Figura 5.	Diagrama de classes que descreve a estrutura e relação entre as classes do sistema.	46
Figura 6.	Fluxo da plataforma até a disponibilização para os clientes (aplicativos, sistemas web, etc).	49
Figura 7.	Diferentes produtos que podem ser criados a partir de uma integração com a API.	49
Figura 8.	E-mail enviado pela plataforma ao concluir o processamento da requisição. .	51
Figura 9.	Tabela raster_data, com os arquivos no formato TIFF, que foram importados via comandos disponibilizados pelo PostGIS.	51
Figura 10.	Primeiro passo da requisição, onde o usuário informa dados referentes ao cenário climático, frequência, localização, formato e período.	52
Figura 11.	Tela de acesso da plataforma CPTEC. O ambiente é restrito e somente é liberado se o usuário possuir um usuário e senha válidos.	52
Figura 12.	Área administrativa visualizada após o login do usuário, com a manutenção de Modelos em destaque.	53
Figura 13.	Passo 1 - Escolha dos filtros iniciais para a requisição.	53
Figura 14.	Passo 2 - Visualização de amostra do dado, onde é apresentado no mapa. .	54
Figura 15.	Passo 3 - Usuário informa os dados pessoais, qual a utilização dos dados e instituição a qual ele pertence.	54
Figura 16.	Passo 4 - Conclusão da requisição, que pode ser visualizada no último passo, apresentando um resumo da requisição.	54
Figura 17.	Passo 4 - Mensagem de sucesso da requisição, apresentada após concluir a mesma.	55
Figura 18.	Resumo e download da requisição.	55
Figura 19.	Arquivo package.json com o metadados da plataforma, que são configurações do nome da aplicação, equipe e pacotes utilizados no desenvolvimento.	56
Figura 20.	Mapeamento objeto relacional da entidade país.	57

Figura 21.	Tabela e gráfico dos testes das requisições - Máquina do CPTEC/INPE.	63
Figura 22.	Tabela e gráfico dos testes das requisições - Máquina local.	63
Figura 23.	Arquivo CSV gerado pela plataforma.	64
Figura 24.	Montagem da URL que será disponibilizada através de layout apresentado, conforme foi desenvolvido na API.	65

LISTA DE TABELAS

Tabela 1.	Tecnologias que permitem trabalhar com GIS.	26
Tabela 2.	Publicação de dados - Possibilita a recuperação e o envio de dados.	28
Tabela 3.	Acesso a Dados - Pacotes que fazem interface com repositórios de dados . .	28
Tabela 4.	Literatura - Pacotes que fazem interface com jornais ou metadados.	31
Tabela 5.	Pacotes de computação escalável.	31
Tabela 6.	Pacotes de interface com as bases de dados.	32
Tabela 7.	Pacotes para visualização de dados.	32
Tabela 8.	Pacotes para manipulação de dados.	33
Tabela 9.	Pacotes para manipulação de dados Geospaciais.	33
Tabela 10.	Variáveis Climáticas.	40
Tabela 11.	Índices de Extremos Climáticos.	43
Tabela 12.	Esquema de tabelas para armazenamento dos dados.	46

LISTA DE ABREVIATURAS

GIS – Geographic Information System

GPS – Global Positioning System

SGDB – Sistema de gerenciamento de banco de dados

TIFF – Tagged Image File Format

API – Application Programming Interface

GB – Gigabyte

TB – Terabyte

MHz – Megahertz

TI – Tecnologia da Informação

NPM – Node Package Manager

SUMÁRIO

1	INTRODUÇÃO	19
2	REVISÃO DA LITERATURA	21
2.1	BIG DATA	21
2.1.1	Os 5 V's do Big Data	21
2.2	FERRAMENTAS PARA BIG DATA	23
2.3	SISTEMA DE INFORMAÇÃO GEOGRÁFICA (SIG)	25
2.4	BIG DATA E GIS	25
2.5	PACOTES EM R PARA MANIPULAÇÃO DE DADOS CLIMÁTICOS	27
2.6	VISUALIZAÇÃO DE DADOS CLIMÁTICOS EM R	34
3	MODELOS DE MUDANÇAS CLIMÁTICAS	37
3.1	MODELO ETA	37
3.2	DADOS CLIMÁTICOS E METEOROLÓGICOS DOS MODELOS DE PREVISÃO	39
4	CPTEC TOOLKIT	45
4.1	ARMAZENAMENTO DOS DADOS CLIMÁTICOS	45
4.1.1	Descrição das tabelas	46
4.2	APPLICATION PROGRAMMING INTERFACE (API)	48
4.3	INTERFACE WEB - FRONT-END	50
4.3.1	Características do Kit de Ferramentas	50
4.3.2	Requisição	52
4.3.3	Modularização e carga de módulos sob demanda	55
4.4	APLICAÇÃO BACK-END	55
4.4.1	Recursos da API	57
5	TESTES E RESULTADOS	61
5.1	ANÁLISE DA METODOLOGIA PROPOSTA	61
5.2	ANÁLISE DOS RESULTADOS	62
5.3	USABILIDADE	64
5.4	INSTALAÇÃO NO CPTEC/INPE	64
6	CONCLUSÕES E TRABALHOS FUTUROS	67

1. INTRODUÇÃO

Os sistemas de informações geográficas (GIS) tem ganhado destaque na comunidade de geociências pelo significativo papel que representam nas ciências da terra. Uma grande variedade de ferramentas tem sido desenvolvidas para fornecimento e manipulação de Big Data [1]. Existem diversas maneiras de coletar, analisar e visualizar estes dados, sendo que o avanço tecnológico, as formas tradicionais estão perdendo seu espaço, e a todo momento surgem novas soluções. Dentre elas podemos destacar a Internet das coisas (IoT), Web Services, Sistemas de posicionamento global (GPS), sensoriamento remoto e análise de Big Data [2, 3, 4, 5].

Big Data, é um paradigma emergente aplicado em um conjunto de dados, cujo tamanho é superior à capacidade de ferramentas comumente usadas para capturar, gerenciar e processar os dados dentro de um tempo aceitável [6] e pode ser manipulado com várias tecnologias, utilizando o que há de melhor no mercado para auxiliar na capacidade de transformar o negócio, mudando a maneira de como é feita a análise dos dados obtendo informações úteis para tomadas de decisão.

Aliado a este fato, existe uma crescente demanda por serviços, conteúdos e produtos diversificados, que atendem segmentos e necessidades cada vez mais específicas. Lidar com esse fluxo de informações constantes, torna-se uma tarefa difícil para as empresas e instituições, que muitas vezes precisam tomar decisões rápidas baseadas em análises de dados.

A agricultura tem uma relação muito forte com o clima, o qual interfere diretamente na produção. A quantidade de chuvas, temperaturas e outros elementos, tem ligação direta no resultado da colheita.

Diversos modelos climáticos globais vem sendo utilizados para gerar projeções de mudanças climáticas, por meio de diferentes cenários. O modelo de área limitada Eta, foi desenvolvido pela Universidade de Belgrado em conjunto com o Instituto de Hidrometeorologia da Iugoslávia e se tornou operacional no National Centers for Environmental Prediction (NCEP) [7]. Esse modelo foi instalado no CPTEC em 1996 com o objetivo de complementar a previsão numérica de tempo que era realizada desde o início de 1995 com o modelo de circulação geral atmosférica [8].

Os dados gerados pelo modelo Eta são armazenados no CPTEC em arquivos em formato binário. Entretanto, grande parte dos usuários destes dados solicitam as informações em formato texto, principalmente valores diários e mensais das variáveis bidimensionais disponibilizadas pelo modelo.

A criação do CPTEC Toolkit teve como premissa automatizar o processo de filtragem dos dados binários disponibilizados pelo CPTEC/INPE, integrando tecnologias de acesso aos dados e armazenamento dos mesmos. Esse kit de ferramentas tem o objetivo de fornecer facilidade e respostas rápidas para o usuário, e ao mesmo tempo eliminar a criação de scripts manuais feitos pelos técnicos do INPE, diminuindo a espera dos usuários pelos dados solicitados e liberando o técnico para desempenhar outras atividades.

Com o CPTEC Toolkit também é possível escolher um formato pré-definido para download da requisição. Essa funcionalidade trouxe benefícios para os usuários na usabilidade dos dados, facilitando a importação dos dados em sistemas legados das empresas e instituições. Com isso espera-se que o CPTEC/INPE promova o aumento no número de requisições.

Outra característica importante do CPTEC Toolkit, é a visualização dos dados que podem ser expressados por meio de mapas ou imagens, o que é muito comum em ferramentas sobre a perspectiva GIS.

O desenvolvimento deste trabalho foi dividido em algumas etapas independentes. No capítulo dois, é apresentada a revisão da literatura. No capítulo três, será descrito o modelo ETA, que se trata da origem dos dados binários fornecidos pelo CPTEC/INPE. Nos demais capítulos, são focados no CPTEC Toolkit, tecnologias utilizadas, padrões da API, banco de dados, testes, conclusões e trabalhos futuros.

2. REVISÃO DA LITERATURA

Estamos vivendo em uma era com a geração de uma grande massa de dados. O termo Big Data está aparecendo em muitos contextos, como por exemplo: meteorologia, simulações, redes sociais, vendas e etc, e está presente nas mais diversas áreas como: biologia, meio ambiente, finanças, saúde e etc. Geralmente os dados são semi-estruturados e heterogêneos, o que torna extremamente complexo trabalhar com abordagens tradicionais em tempo real [9].

2.1 BIG DATA

O termo Big Data, é um paradigma emergente aplicado em um conjunto de dados, cujo tamanho é superior à capacidade de ferramentas comumente usadas para capturar, gerenciar e processar os dados dentro de um tempo aceitável [6]. Várias tecnologias estão sendo discutidas para apoiar o manipulação de Big Data, tais como banco de dados [10], sistemas de armazenamento escaláveis [11, 12], plataformas de computação em nuvem [13, 14] e MapReduce [15, 16].

2.1.1 Os 5 V's do Big Data

Muitos artigos tratam o Big Data com três dimensões: volume, variedade, velocidade. É seguro dizer que mais duas começam ganhar destaque: valor e veracidade. Artigos atuais tratam o Big Data com cinco dimensões [17] como apresentado na Figura 1.



Figura 1. Dimensões disponíveis e que devem ser consideradas ao avaliar a viabilidade de uma solução Big Data [18].

O primeiro dos 5V's se refere ao volume de dados gerados pela população na internet (amostras de dados estruturados, mensagens de email, tweets, facebook, etc.) [9]. Possuir um bom modelo de análise é uma parte importante, mas estar apto a processar um volume grande de dados trará mais

confiança e precisão ao resultado final. O MapReduce¹ tem comprovado ser bem completo quando se trata de larga escala em estruturas de processamento de dados. Um exemplo de ferramenta que faz uso do MapReduce é o Hadoop² [9].

Na sua grande maioria, os dados em geral estão prontos para serem utilizados pelas aplicações. Os mesmos são gerados nos mais diversos formatos e estruturas. Além disso, as fontes de dados são diversificadas, o que gera ainda mais variedade. Os dados podem aparecer na forma de texto, áudio, vídeo, gerados por pessoas, gerados por máquinas, sensores e etc. Os dados podem aparecer em diferentes níveis estruturais [9], contendo uma ou mais divisões que se relacionam na forma de pai-filho, até mesmo a ausência completa de estruturas, ou seja, não há um padrão definido. Quando feita a organização das fontes de dados, invariavelmente podemos dispensar algumas informações. Isso pode acontecer no processo de análise bruta dos dados. Um dos princípios do Big Data é, quando permitido e possível, manter todos os dados.

Possuir uma informação precisa no momento errado não é uma grande vantagem, porém possuir uma informação precisa no momento certo é, de fato, uma vantagem que o dono dos dados procura. Hoje em dia, grandes empresas anseiam por rápida movimentação de informações para o seu próprio benefício. O tráfego de dados cresceu exponencialmente, devido a existência de vários dispositivos que permitem o acesso à rede, como tablets, smartphones e outros dispositivos. A importância da velocidade está no grau de rapidez em que o fluxo de informações é realizado, desde a entrada dos dados, até o ponto onde estes se tornam úteis e são utilizados na tomada de decisões [9].

Geralmente o armazenamento é feito na forma de pares chave e valor, armazenado em banco de dados orientado a colunas. Os bancos de dados que seguem essas estratégias fazem parte de uma categoria conhecida como NoSQL³.

Para que o resultado do processo de análise seja satisfatório, é necessário obter dados verídicos, de acordo com a realidade. O conceito de velocidade está alinhado ao conceito de veracidade, pela necessidade constante de análise em tempo real, isso significa que os dados condizem com a realidade do momento. A verificação dos dados coletados juntamente com o propósito da análise, são pontos chave para a obtenção das informações que agreguem valor ao processo [19].

O uso do Big Data torna-se imprescindível quando todos os esforços são direcionados para gerar um valor, que é a informação passível de análise e a sua conversão em material para ser aplicado nas decisões de uma empresa [20].

É altamente recomendável que o resultado final faça sentido, bem como a autenticidade da informação gerada. O valor gasto para desenvolver e manter uma infraestrutura de Big Data é muito alta, precisando assim retornar o investimento.

¹Modelo de programação que suporta computação paralela em grandes coleções de dados.

²Plataforma de computação distribuída voltada para clusters e processamento de grandes massas de dados.

³Termo usado para descrever bancos de dados não relacionais de alto desempenho

2.2 FERRAMENTAS PARA BIG DATA

No últimos anos, um número cada vez maior de organizações, de diversos portes e segmentos, vem empregando a filosofia de Big Data no que diz respeito a análise de dados e como ferramenta de apoio estratégico, com o intuito de melhorar os seus processos de trabalho e adquirir informações valiosas, como tendências de mercado, comportamento dos consumidores e etc.

Existem muitas ferramentas para se trabalhar com Big Data. Abaixo são listadas algumas delas, juntamente com suas características.

- a) **Apache Spark Sql:** O Apache Spark torna possível o processamento de grandes quantidades de dados. Ele foi construído com o intuito de ser extremamente veloz, fácil de usar e fornece recursos avançados para se lidar com análises de dados. Ele foi desenvolvido em 2009, na Universidade de Berkeley, tornando-se open source em 2010, como um projeto da Apache. Ele fornece uma certa vantagem quando comparado a outros frameworks de processamento Big Data, primeiramente por fornecer uma ferramenta completa para o processamento e manipulação de diversos tipos de dados (texto, imagem, gráficos, etc.) e também para situações diversas como o processamento de dados previamente armazenados e dados que chegam em um fluxo contínuo (streaming de dados). O Framework do Spark fornece uma série de facilidades, com APIs de alto nível em Java, Scala, Python, permitindo que aplicações executem em um cluster Hadoop até 100 vezes mais rápida em memória e cerca de 10 vezes mais rápida em disco quando comparada a outras ferramentas [21].
- b) **Apache Hadoop:** O Hadoop foi originalmente criado como parte do projeto Apache Nutch, mas logo se transformou em um projeto da Apache. Hadoop é um framework para desenvolver aplicações que processam grandes volumes de dados em paralelo sobre vários nós (milhares de nós) de hardware. Este projeto pertence à Apache, sendo uma solução open source para trabalhar com Big Data. O Hadoop pode ser utilizado para realizar uma consulta em um grande volume de dados e obter uma resposta rápida, utilizando a plataforma de computação distribuída fornecida por ele [22]. Após a criação do Hadoop, uma variedade de projetos foi criada dentro da fundação Apache. Esses projetos utilizam o Hadoop como mecanismo central, fornecendo funcionalidades adicionais. Dentre eles podemos destacar: Cassandra, HBase, Hive, Pig e o ZooKeeper.
- c) **Cassandra:** É um sistema de banco de dados distribuído, uma solução que se encaixa na categoria NoSQL. Armazena as informações sob a estrutura de chave e valor. O projeto foi desenvolvido pelo Facebook, e em 2008 seu código fonte foi liberado para a comunidade [23]. Atualmente é mantido pela Apache.
- d) **MongoDB:** MongoDB é uma ferramenta diferente do modelo relacional tradicional. Possui um conceito de NoSQL (Banco de dados que não utiliza SQL). Ele é altamente escalável e utilizado em soluções que demandam uma alta performance. Possui uma poderosa linguagem de consulta baseada em documento, tornando a mudança de paradigma mais suave, ou seja, é possível converter comandos SQL utilizando funções de uma API de consultas. Os dados são armazenados usando BSON, uma versão binária do JSON, utilizando pares de chave e valor para mapear os dados [24].

- e) **Linguagem R:** R é uma linguagem open source desenvolvida por Ross Ihaka e Robert Gentleman no departamento de Estatística da universidade de Auckland, na Nova Zelândia. É uma linguagem muito utilizada por estatísticos, cientistas e pessoas que precisam realizar análise estatística de dados [25]. A linguagem R fornece aos seus usuários uma grande variedade de pacotes, que podem ser utilizados para cálculos estatísticos, machine learning⁴, modelagem e simulação, extração de dados, limpeza de dados, carregamento de dados, transformação em diversos formatos e opções de visualização. Além de todas as facilidades que o R fornece aos seus usuários, ele é capaz de conectar com vários bancos de dados, como por exemplo: MySQL, SQLite, MongoDB. Isso pode ser muito útil na hora de trabalhar com o grande número de dados, oriundos de diferentes fontes [26].
- f) **Linguagem HTML:** Com o HTML é possível criar aplicações ricas para a internet, devido ao grande número de componentes e frameworks disponíveis para mostrar os resultados. O AngularJS é um framework escrito em JavaScript, de código aberto, mantido pelo Google, que auxilia na execução de single-page applications e permite estender o vocabulário HTML. O ambiente resultante é expressivo, legível e rápido para o desenvolvimento [27]. Existem alguns recursos visuais quando se trabalha com HTML e JavaScript. Várias bibliotecas de mapas e gráficos são disponibilizadas gratuitamente para a comunidade como é o caso do leafletjs para interação com mapas, chartjs, flotcharts e jquery.sparkline para gráficos.
- g) **NodeJS:** NodeJS é um interpretador de código JavaScript que funciona do lado do servidor. O seu principal objetivo é auxiliar na criação de aplicações de alta escalabilidade, atuando como back-end como servidor, com códigos capazes de manipular dezenas de milhares de conexões simultâneas, numa única máquina física [28].
- h) **PostgreSQL:** PostgreSQL é um sistema gerenciador de banco de dados objeto relacional. É um dos bancos de código aberto mais avançados, contando com recursos como: consultas complexas, chaveamento estrangeiro, integridade transacional, controle de concorrência, suporte ao mapeamento objeto relacional, facilidade de acesso através do pgAdmin, criação de triggers, views, suporta as linguagens PL/pgSQL, PL/Python, PL/Java, PL/Perl e uma das mais importantes é a estrutura para guardar dados Georeferenciados através do PostGIS [29].
- i) **PostGIS:** O PostGIS é uma extensão espacial gratuita e de código fonte livre para o PostgreSQL. O mesmo suporta dados geográficos e permite o acesso aos dados através de SQL(Structured Query Language). O PostGIS ainda oferece uma série de funções para se trabalhar com esses dados, algumas muito próximas as oferecidas em banco de dados como Oracle e SQL Server da Microsoft [30].

⁴Sub-campo da inteligência artificial dedicado ao desenvolvimento de algoritmos e técnicas que permitam ao computador aprender e aperfeiçoar as técnicas.

2.3 SISTEMA DE INFORMAÇÃO GEOGRÁFICA (SIG)

SIG é uma coleção de ferramentas e informações geográficas, que são aplicadas em diferentes áreas, podendo ser utilizadas em diversos softwares e serviços, auxiliando na tomada de decisões [31].

Algumas tecnologias têm crescido rapidamente no ramo das ciências agrícolas, tais como a agricultura de precisão, gestão dos recursos naturais, economia agrícola e a gestão ambiental. Equipamentos avançados que fazem o uso do GIS, GPS, entre outros, os quais auxiliam a maneira como os produtores agrícolas administram os seus negócios [32].

As alterações climáticas como mudanças de temperaturas, chuvas e outros fenômenos, tem impactos sobre a agricultura. Determinar que as mudanças climáticas afetam a agricultura é uma tarefa delicada e muitas vezes é necessária uma análise mais profunda, devido à variedade de efeitos que podem ocorrer na natureza [33].

Existem algumas técnicas que podem ajudar na análise dos impactos, como os modelos de simulação, que são ferramentas importantes para avaliar os potenciais de uma safra. A partir desses modelos, é possível analisar toda a gestão do crescimento e desenvolvimento de uma cultura, bem como associar isso aos dados climáticos, para eventuais tomadas de decisões. Os modelos tiveram muito sucesso no passado, ajudando na produção, principalmente nos quesitos de otimização do uso da água [34] e compreensão dos impactos efetuados pelas mudanças climáticas sobre as culturas [33]. Diversos estudos têm demonstrado uma forte relação de sucesso entre modelos de culturas e GIS. Muitos pesquisadores têm utilizado esses modelos para estudar e descobrir qual o real efeito que as mudanças climáticas fazem sobre as mais variadas culturas existentes [35].

2.4 BIG DATA E GIS

O sucesso de qualquer aplicação que tem dependência direta com GIS, depende da seleção, coleta e manipulação dos dados. Com a enorme quantidade de sensores espalhados, os dados crescem de forma exponencial dia a dia, ganhando proporções muitas vezes inimagináveis, demandando de técnicas de avançadas de processamento, dando espaço para os sistemas inteligentes.

- a) **GIServices - Sistemas Inteligentes:** Sistemas inteligentes GIS concentram-se no armazenamento, consulta, visualização e análise de uma grande quantidade de dados georeferenciadas. Serviços como WebServices [36], APIs, pacotes R, possibilitam as interações entre aplicações, melhoria dos modelos, e disponibilizações diversas através da rede.
- b) **Tecnologias disponíveis:** São vários os benefícios que os sistemas inteligentes oferecem. Isso se dá pelo fato de existirem diversas integrações de tecnologias no mercado. Essas tecnologias estão agrupadas em várias categorias, conforme a Tabela 1, e são de extrema importância para os sistemas GIS. A infra-estrutura de computação na nuvem, possibilita a redefinição das capacidades da geociência [37], tornando possível um serviço integrado para ajudar as pessoas em diversas

tomadas de decisão em tempo real, o que por muitas vezes era praticamente impossível de conseguir. A análise de Big Data, tal como as tecnologias MapReduce e NoSQL [38, 39], têm ajudado na extração de informações e descobertas de conhecimento, incluindo algoritmos de mineração de dados, técnicas de *machine learning*, inteligência artificial, modelos estatísticos, podendo ser acessados via GIServices, com intuito de transformar dados em informações.

Tabela 1: Tecnologias que permitem trabalhar com GIS.

Categorias	Tecnologias
Web e serviços	AJAX, HTML5, JQuery, SSL, HTTP, XML, WSDL, JSON, SOAP, REST, NodeJS, R, Adobe Flex, AngularJS, ReactJS, openCPU.
Web semântica	RDF, OWL, SAWSDL, OWL-S, WSMO, SWSF.
Fluxo de Trabalho	BPEL, BPMN, XPDL, YAWL.
Sensoriamento	Sensores, Internet das coisas, crowdsourcing ⁵ , Sistemas de observação da Terra.
Banco de dados Escaláveis	NoSQLDB, NewSQLDB, ArrayDB, in-memory DBMS, CloudDB, MongoDB, Cassandra, HBase, Hive, PIG, ZooKeeper, PostgreSQL, PostGIS.
Dados intensivos	Hadoop, Apache Spark SQL, Computação na nuvem, processamento de fluxo, processamento paralelo, Google Cloud Dataflow, HPC.
Análises avançadas	Análise de big data, textos, áudios, imagens, dados da Web, tag cloud, computação social.

- c) **Infraestrutura:** Existem modelos que são implementados como serviços. Esses modelos são heterogêneos e requerem infraestruturas ricas capazes de integrar sistemas computacionais, dados, informações, redes, sensores, instalações experimentais, middleware⁶, e etc, integrando os fluxos de trabalho [40, 41, 42].
- d) **Sistemas Inteligentes na Nuvem:** Com o grande procura de espaço e processamento, grandes empresas iniciaram suas trajetórias em tecnologias de computação na nuvem, reduzindo custos e manutenções internas, bem como preocupações com backup e mais uma série de problemas enfrentados por gestores de TI. Hoje é possível ter uma estrutura inteira na nuvem, ou seja, sistemas e bancos de dados ficam hospedados fora da empresa [43]. A maioria das empresas que fornecem o serviço na nuvem, possibilitam seus clientes acessarem estruturas escaláveis, pré-firmadas em

⁵Contribuição Colaborativa ou Colaboração Coletiva, terceirização de serviços.

⁶No campo da computação, é um programa de computador que faz a mediação entre software e demais aplicações.

um contrato, ou seja, necessitando mais espaço físico, memória, é possível fazer tudo isso online, sem interromper o processo. Esse tipo de serviço tem se tornado cada vez mais presente no meio da Tecnologia da Informação [44].

- e) **Sensores:** Existem vários tipos de sensores, que podem ser de espaço, ar, terrestres, oceânicos, e etc. No entanto é muito custoso extrair informações úteis, devido a infinidade tipos de observações que são armazenadas nesses sensores, bem como a variedade de aparelhos existentes no mercado, criando assim um desafio enorme para os pesquisadores e interessados [45]. A Figura 2 demonstra algumas formas de coletar dados por meio de diferentes sensores.



Figura 2. Diversos sensores espalhados que podem enviar suas posições, serem acessados via web, controlados via web remotamente. Adaptado de Esri [46].

- f) **Sistemas Inteligentes Semânticos:** Um das grandes promessas da ciência colaborativa e interdisciplinar sem dúvidas é o avanço da Web Semântica, onde a colaboração e compartilhamento de conhecimento, tem sido de grande importância para os pesquisadores na comunidade. Com isso é possível fornecer informações através de serviços web, pré-processados, auxiliando interessados em suas tomadas de decisões e na disseminação do conhecimento [47, 48].

2.5 PACOTES EM R PARA MANIPULAÇÃO DE DADOS CLIMÁTICOS

Existem vários pacotes R que disponibilizam acesso, manipulação e visualização de dados climáticos armazenados em grandes repositórios ou centros de pesquisa. Neste trabalho são explorados, inicialmente, os seguintes pacotes: rOpenSci, rnoaa, rclifro, rncep [25].

- a) **rOpenSci:** O rOpenSci é uma lista completa de pacotes R. Os mesmos fornecem acesso a dados científicos, de texto completo de artigos de revistas a repositórios de dados, incluindo métricas de

impacto em tempo real. Nas tabelas abaixo são apresentados os pacotes integrantes do conjunto de pacotes rOpenSci.

- **Publicação de dados:** Este conjunto de pacotes permite a recuperação e também podem ser capazes de submeter arquivos e textos para pesquisas avançadas. Os pacotes disponíveis estão ilustrados na Tabela 2.

Tabela 2: Publicação de dados - Possibilita a recuperação e o envio de dados.

Pacote	Descrição
dataone	Pesquisa em repositórios. Inclui repositórios como o KNB e Dryad .
dvn	Interface para a Rede DataVerse.
EML	Usado para ler, escrever, integração de dados usando o Ecological Metadata Language.
rfigshare	Envio de dados, figuras e texto, consulta e recuperação de dados no Figshare.
RNeXML	I/O em R - próxima geração XML para dados filogenética.

- **Acesso a dados:** Estes pacotes têm acesso a diferentes bases de dados, podendo acessar diretamente API das mais diversas áreas. Os pacotes disponíveis estão ilustrados na Tabela 3.

Tabela 3: Acesso a Dados - Pacotes que fazem interface com repositórios de dados

Pacote	Descrição
AntWeb	Acesso ao maior banco de dados de formigas do mundo.
bold	Pacote R da base de dados da vida.
ckanr	Cliente R para CKAN API RESTful.
chromer	Interações com o banco de dados contagem de cromossomos.
clifro	Um cliente R para acesso a base de dados de clima da Nova Zelândia.
Continuação na próxima página	

Tabela 3 – Continuação

Pacote	Descrição
ecoengine	Um cliente R para acesso a base de dados de clima da Nova Zelândia.
ecoretriever	Cliente R envolvendo o Python software Ecoretriever.
gender	Pacote de R para codificar gênero com base em nomes e datas de nascimento.
geonames	Pacote R com diversos nomes de estabelecimentos por todo o mundo.
historydata	Pacote R com conjuntos de dados de interesse para os historiadores.
internetarchive	API para o Internet Archive.
musemeta	Pacote R para Museus.
neotoma	Pesquisas usando o base de dados do Neotoma Paleocological.
paleobioDB	Acesso ao dados do base de dados do Paleobiology.
pangaear	Acesso ao base de dados do Pangaea.
helminthR	Acesso ao base de dados de parasitas do Museu de História Natural de Londres
pleiades	Recuperar dados de metadados e de localização espacial em lugares arqueológicos.
rAvis	Dados sobre a ocorrência de espécies de aves na base de dados do ProyectoAvis
rbison	Pacote R para de informações sobre biodiversidade servindo a nossa nação.
RCryptsy	Permite acessar a base de dados do Cryptsy.
rdpla	Pacote R para interagir com Digital Public Library of America.
rdryad	Obter dados do Dryad, um repositório de trabalhos acadêmicos publicados
rebird	Pesquisar ocorrências de avistamentos de aves recolhidas na base de dados eBird
rentrez	Permite acessar a base de dados do NCBI.
reol	Pacote R para a Enciclopédia da Vida.
rerddap	Pacote R para trabalhar com servidores ERDDAP.
reuropeana	Pacote R para interagir com a base de dados Europeia.
rfishbase	Acessar quaisquer dados de peixes na base de dados do Fishbase.org
Continuação na próxima página	

Tabela 3 – Continuação

Pacote	Descrição
rfisheries	Pesquisar e recuperar dados do OpenFisheries.org
rgbif	Pesquisar espécies na base de dados da Global Biodiversity Information Facility.
rglobi	Pacote R para os dados de interações bióticas
rif	Pacote R para acesso a base de dados da Neuroscience Information Network.
rinat	Pacote R para a API fornecida pelo iNaturalist.
rnoaa	Acesso aos dados climáticos de acesso da NOAA.
rnbn	Acesso a base de dados do Reino Unido Rede Nacional de Biodiversidade.
rnpn	Acesso aos dados fenológicos da US National Phenology Network.
rotl	Interações com a API do Open Tree of Life.
rredlist	Interações com a API IUCN Red List.
RSelenium	Pacote R para o Selenium WebDriver remoto.
rsnps	Pesquisar dados de Nucleotide Polymorphism no openSNP.
RStars	Acesso aos dados do Digital Universe.
rusda	Interface para a base de dados do USDA.
rvertnet	Acesso aos registros de vertebrados no VertNet.
rWBClimate	Interface com os dados climáticos utilizados no Banco Mundial
spocc	Coleta dados do GBIF, iDigBio, Ecoengine, AntWeb, eBird, and USGS's BISON.
spoccutils	Pacote R para limpeza e visualização de dados.
taxize	Obter identificadores taxonômicos de muitas bases de dados diferentes.
traits	Espécies de traças, incluindo Polytraits, BETTYdb, NCBI, Phylomatic.
treebase	Programação para Treebase, banco de dados de árvores filogenéticas.
USAboundaries	Dados de fronteiras dos Estados Unidos.
webchem	Recuperar informações químicas de muitas bases de dados.

- **Literatura** - Estes conjuntos interagem com textos e artigos acadêmicos, incluindo PLOS, PMC, PeerJ, Elife, arXiv, IEEE entre outros. Os pacotes disponíveis estão ilustrados na Tabela 4.

Tabela 4: Literatura - Pacotes que fazem interface com jornais ou metadados.

Pacote	Descrição
aRxiv	Acesso ao repositório da arXiv.
bmc	Uma interface para a base de dados do Biomed Central (BMC).
csl	Explorar estilos CSL e localidades.
fulltext	Textos acadêmicos, incluindo PLOS, PMC, PeerJ, Elife, arXiv, e muitos mais.
IEEEER	Interface para o IEEE Xplore Gateway.
oai	Trabalhar com OAI-PMH (Protocolo Open Archives Initiative for Metadata Harvesting).
pdftools	Pacote rápido e portátil para extrair texto e dados a partir de documentos PDF.
rbhl	Acesso ao texto integral e metadados em texto digitalizado e OCR.
rcrossref	Uma interface R para Crossref para pesquisa para API de metadados.
rdatacite	Wrapper para DataCite.
rebi	Acesso à base de dados europeia PubMed Central.
rorcid	A interface de programação para a API do Orcid.org.
rplos	Conecta-se à API de texto completo da Public Library of Science (PLOS).

- **Pacotes de computação escalável** - Estes conjuntos colaboram com a escalabilidade da computação. Conta com pacotes importantes para análise de dados, reuso de códigos, expressões regulares e etc. Os pacotes disponíveis estão ilustrados na Tabela 5.

Tabela 5: Pacotes de computação escalável.

Pacote	Descrição
gistr	É uma interface para GIST do GitHub e R.
Continuação na próxima página	

Tabela 5 – Continuação

Pacote	Descrição
git2r	Ligações Git e R.
textreuse	Pacote de R para detectar a reutilização de texto e semelhança.

- **Pacotes de interface com as bases de dados:** Estes conjuntos colaboram com interfaces para base de dados. Conta com pacotes importantes para pesquisas Elasticsearch⁷. Os pacotes disponíveis estão ilustrados na Tabela 6.

Tabela 6: Pacotes de interface com as bases de dados.

Pacote	Descrição
elastic	Pacote R para o Elasticsearch.
elasticsearchdsl	Um DSL para Elasticsearch.
etseed	Pacote R para a ETCD, um armazenamento de chave-valor dos povos em CoreOS.
nodbi	Fornecer uma interface única para interagir com muitos bancos de dados NoSQL.
sofa	Interface fácil de CouchDB em R.
solrium	Um cliente de R para Apache Solr.

- **Pacotes para visualização de dados:** Estes pacotes importantes principalmente para a visualização de dados. É possível construir gráficos para facilitar no entendimento das informações. Os pacotes disponíveis estão ilustrados na Tabela 7.

Tabela 7: Pacotes para visualização de dados.

Pacote	Descrição
plotly	Interface de programação para a API Plot.ly - construir gráficos com ggplot2.
magick	Processamento de imagem.

⁷Armazenamento de dados de forma não relacional, muito bom para retornar buscas muito pesadas. Por ser um motor de pesquisa textual altamente escalável, permite armazenar e analisar grandes volumes de informações praticamente em tempo real

- **Pacotes para manipulação de dados:** Pacotes para manipulação de dados. Servem para manipular os dados JSON, por exemplo. Usado na validação de alguns formatos. Os pacotes disponíveis estão ilustrados na Tabela 8.

Tabela 8: Pacotes para manipulação de dados.

Pacote	Descrição
finch	Leitor de Darwin Core Archive.
jq	Pacote R para jq, um processador JSON.

- **Pacotes para manipulação de dados Geospaciais:** Pacotes para manipulação de dados geoespaciais. Conta com vários pacotes para manipular especialmente dados do tipo GeoJSON. Os pacotes disponíveis estão ilustrados na Tabela 8.

Tabela 9: Pacotes para manipulação de dados Geospaciais.

Pacote	Descrição
cartographer	Fornecer mapas interativos em documentos RMarkdown ou no console R.
geoaxe	Corte polígonos em pedaços.
geojsonio	Pacote R que converte muitos formatos de dados geoJSON para topoJSON.
lawn	Pacote R para turf.js, uma biblioteca JavaScript para análise geoespacial avançada.
proj	Pacote R para proj4js, uma biblioteca JavaScript para coordenadas.
wellknown	Converter WKT para GeoJSON e vice-versa. Inspirado por Python GEOMET.

Alguns pacotes de destaque serão estudados neste trabalho, pois se encaixam com o grande volumes de dados e principalmente com dados geográficos.

- b) **RNOAA:** O pacote *rnoaa* é um pacote para acesso de dados climáticos da NOAA (*National Oceanic and Atmospheric Administration*), incluindo temperatura e precipitação, bem como dados de cobertura de gelo do mar e eventos climáticos extremos. Com este pacote, é possível o acesso de clima, dados meteorológicos graves, dados de gelo do mar, dados das bóias espalhadas em alto

mar, tornados, tempestades, previsão do tempo e etc [49]. Para acessar a API, é necessário uma autenticação, que vai se dar através de um token, cadastrado no site da NOAA. Não é permitido em hipótese alguma o uso da API sem esta autenticação [49, 50, 51, 52].

- c) **rNOMADS**: Trata-se de um pacote para os modelos NOAA (*National Oceanic and Atmospheric Administration*), que permitem aos usuários R baixar e processar os modelos globais. O pacote ainda suporta uma grande variedade de modelos, que vão desde dados meteorológicos até climáticos e regionais em alta resolução. O rNOMADS permite trabalhar com formatos do tipo grib.
- d) **CLIFRO**: O pacote clifro foi projetado para facilitar o acesso aos dados de clima da Nova Zelândia via CliFlo (sistemas web de acesso aos dados). Ele fornece funções para o usuário fazer login, escolher os tipos de dados e estações apropriadas e, em seguida, consultar o banco de dados. Uma vez que os dados tenham sido transferidos, eles são armazenados como objetos específicos em R com o objetivo principal de garantir a visualização de dados. A exploração é feita com o mínimo esforço e máxima eficiência. Para usar o pacote CliFlo é necessário uma assinatura, caso contrário, os dados de apenas uma estação está disponível. O cadastro é gratuito e tem a duração de 2 anos ou 2.000.000 linhas, sem renovação, que permite o acesso a cerca de 6.500 estações climáticas em torno de Nova Zelândia e Pacífico [53].
- e) **RNCEP**: RNCEP é um pacote de funções na linguagem R para computação estatística e gráficos. Destina-se a simplificar o acesso, organização e visualização de dados de tempo do NCEP/NCAR e a reanálise dos conjuntos de dados do NCEP/DOE. Estes são dois conjuntos de dados de alta qualidade com cobertura global de muitas variáveis atmosféricas relevantes. Este pacote de funções destina-se principalmente para os ecologistas para ajudar a integrar mais facilmente dados atmosféricos em pesquisas ecológicas. As funções do pacote, no entanto, são robustas e é provável que seja útil em muitas disciplinas.

2.6 VISUALIZAÇÃO DE DADOS CLIMÁTICOS EM R

A aplicação de técnicas de visualização podem ajudar a investigar dados climáticos sob vários aspectos simultaneamente para uma ampla gama de questões de pesquisa. Os pesquisadores, têm como objetivo melhorar a compreensão dos processos naturais, para avaliar a qualidade dos resultados de modelos climáticos e identificar problemas para a tomada de decisão [54]. Alguns pacotes R disponíveis no CRAN, podem ajudar a perceber instabilidades de acordo com a visualização dos diversos conjuntos de dados.

- a) **RNetCDF**: NetCDF é uma abstração que suporta uma exibição de dados como uma coleção de objetos, portáteis auto-descritivos que podem ser acessados por meio de uma interface simples. Valores de matrizes podem ser acessado diretamente, sem saber detalhes de como os dados são armazenados [55].

- b) **maps**: Pacote para desenhar mapas geográficos. A área de cada região correspondente no mapa é calculado, e as regiões que correspondem ao mesmo elemento de regiões têm suas áreas combinadas, através do X e Y no mapa [56].
- c) **mapdata**: Trata-se de um pacote extra para mapas. Alguns exemplos de mapas que o pacote possui: mapa da China, Japão, Nova Zelândia, rios pelo mundo, pacífico e mundi [57].
- d) **mapproj**: Este pacote tem por objetivo converter latitudes e longitudes em coordenadas projetadas. É possível plotar diferentes mapas [58].
- e) **verification**: Pacote utilitário para verificar previsões discretas, contínuas e probabilística e previsões expressas como distribuições paramétricas [59].
- f) **ensembleBMA**: Pacote de ajuste do modelo BMA de previsões. Permite a especificação de um modelo para criar previsões probabilísticas através de um conjunto de observações meteorológicas [60].

Esse capítulo teve como objetivo verificar o que está sendo utilizado para trabalhar com Big Data e GIS. O estudo compreende desde tecnologias, técnicas e importância da análise de grandes volumes de dados. No desenvolvimento do kit de ferramentas do CPTEC/INPE foram utilizadas técnicas e materiais abordados nesse capítulo, trazendo uma maior segurança e confiabilidade na arquitetura proposta bem como para os usuários finais. Além disto, pode-se perceber onde é possível escalar com o projeto, como, por exemplo, na criação de pacotes que podem ser distribuídos livremente pela comunidade auxiliando nas pesquisas.

3. MODELOS DE MUDANÇAS CLIMÁTICAS

Muitos modelos climáticos globais vem sendo utilizados para gerar projeções de mudanças climáticas, por meio de diferentes cenários [8]. O modelo de área limitada Eta, foi desenvolvido pela Universidade de Belgrado em conjunto com o Instituto de Hidrometeorologia da Iugoslávia e se tornou operacional no National Centers for Environmental Prediction (NCEP) [7]. Esse modelo foi instalado no CPTEC em 1996 com o objetivo de complementar a previsão numérica de tempo que era realizada desde o início de 1995 com o modelo de circulação geral atmosférica [8].

3.1 MODELO ETA

O modelo Eta prevê com maiores detalhes todos os fenômenos associados a frentes, orografia, brisa marítima, tempestades severas, ou seja, sistemas que se caracterizam com organização em mesoscala. Como trata-se de um modelo não linear, a previsibilidade é menor e as previsões são de curto prazo. Elas se estendem até 48 horas e cobrem a maior parte da América do Sul. Estas previsões são fornecidas duas vezes ao dia, dispostas em uma grade regular com resolução de 40km [8].

A grade horizontal possui resolução disponível nas versões 40km e 80km. Trata-se de uma grade regular com coordenadas esféricas, tendo o ponto de interseção entre o equador e o meridiano de 0o transladado para o centro do domínio. O domínio do modelo cobre aproximadamente a região compreendida entre as longitudes de 25oW a 90oW e as latitudes de 12oN a 45oS. Outra divisão é a coordenada vertical cujo a principal característica deste modelo é a coordenada vertical definida por Mesinger no ano de 1984. As versões de 40km e 80km possuem 38 camadas, sendo a resolução maior nos baixos níveis e diminuindo com a altura; a primeira camada tem 20m de espessura. Por último temos a orografia que é representada na forma de degraus cujo topo é parecido com a interface das camadas. A altura de cada degrau é obtida a partir do método de silhueta. Cada degrau possui um ponto de massa no centro e os quatro pontos de velocidade são localizados nos vértices [8].

A condição inicial do modelo é obtida através de uma análise estática, ajustada de acordo com observações do horário da análise. A previsão de 12h do modelo global do CPTEC é utilizada como estimativa inicial. Os erros podem ser minimizados de acordo com o método de interpolação ótima. A dinâmica utiliza um esquema forward-backward modificado por Janjic no ano de 1979, que trata dos termos responsáveis pelo ajuste, enquanto o esquema Euler-backward modificado trata dos termos de advecção horizontal e vertical [8].

A Figura 3 mostra uma previsão de precipitação acumulada em 24h válida para 12UTC de 10 de junho de 1995. A resolução de 40km foi utilizada. Esta foi uma previsão de 36h que teve como condição inicial a análise do modelo global do NCEP na resolução T126L28 [8].

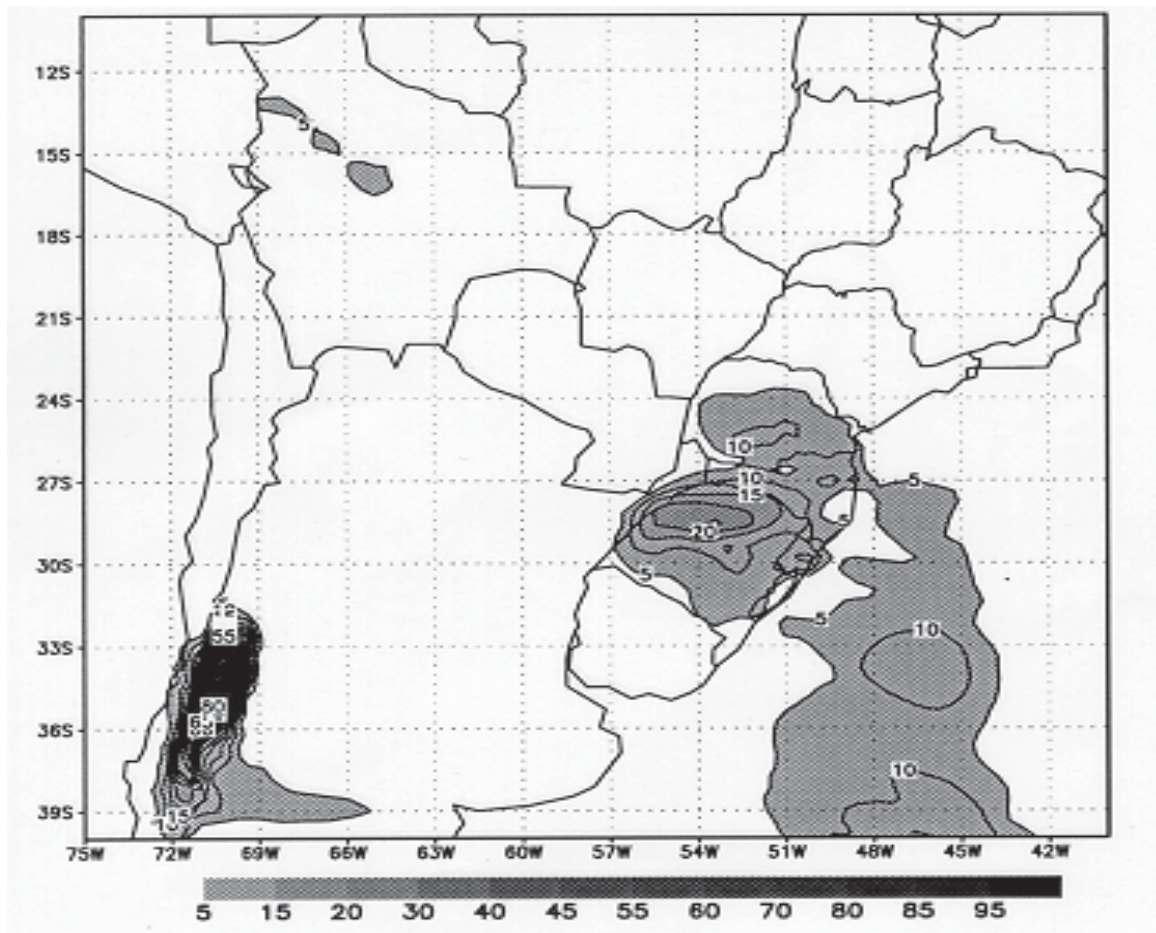


Figura 3. Previsão de precipitação total acumulada de 24 h válida para 12 UTC 10 junho 1995 [8].

Os núcleos de máxima precipitação, que não foram previstos pelo mesmo modelo global, apresentaram valores um pouco abaixo daqueles observados como apresenta na Figura 4. Este foi um caso de intensa precipitação sobre a região Sul que ocorreu devido à passagem de um cavado⁸ de onda curta de acordo com a Climanálise em 1995 [8].

⁸Região alongada de relativa pressão atmosférica baixa, frequentemente associados com frentes.

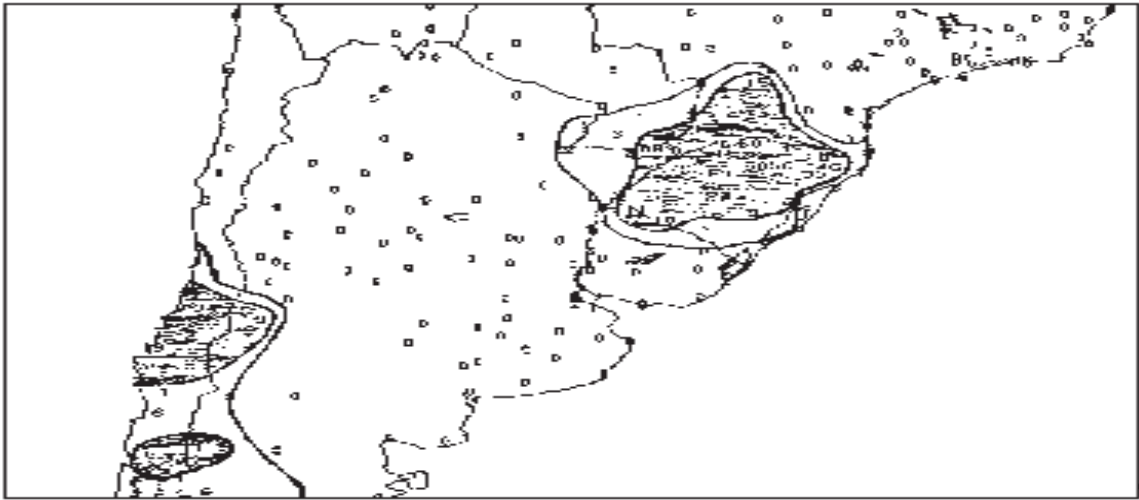


Figura 4. Precipitação total observada no dia 10 de junho de 1995.[8].

3.2 DADOS CLIMÁTICOS E METEOROLÓGICOS DOS MODELOS DE PREVISÃO

A principal ferramenta para estudos das mudanças climáticas são os modelos globais do sistema terrestre. Por outro lado, os impactos em diferentes setores de atividades são geralmente de caráter local. A resolução dos modelos globais, de cerca de 200×200 km, é considerada grosseira para estes estudos. Modelos regionais climáticos, com tamanhos de grade de cerca de 50 km a 20 km, buscam atender esta necessidade de detalhamento. Entretanto, esse tamanho de grade ainda é considerado grosseiro para a maioria dos estudos relacionados com os recursos hídricos, em particular com aqueles relacionados com a disponibilidade hídrica para abastecimento humano e para a segurança alimentar. Assim, a principal inovação desta plataforma é o emprego de vários cenários de mudanças climáticas em alta resolução espacial. Uma segunda inovação é o emprego do modelo em altíssima resolução, cerca de 5 km, em uma escala compatível com a tomada de decisões relacionados com o gerenciamento dos recursos hídricos, como é o caso dos mananciais de abastecimento afetados pelos conflitos no uso da água. O modelo Eta pode operar em modo não-hidrostático, enquanto que a maioria dos modelos climáticos simplificam em modo hidrostático.

O INPE, a partir de 1997, tem fornecido previsões numéricas de tempo com maior detalhamento para grande parte da América do Sul utilizando o modelo Eta. Desde 2002, o modelo passou a gerar previsões climáticas sazonais na resolução de 40 km sobre a região [61]. O modelo também foi utilizado para produzir estudos de mudanças climáticas [62, 63, 64] e apoiar na elaboração da Segunda Comunicação Nacional do Brasil à Convenção-Quadro das Nações Unidas sobre Mudanças do Clima⁹. Novos cenários de emissão foram propostos pelo IPCC [65] e geraram novas projeções de mudanças climáticas globais. Projeções dos modelos HadGEM2-ES e MIROC5, em cenários de emissão RCP4.5 e RCP8.5, foram regionalizadas pelo modelo Eta do INPE [66]. RCP4.5 e RCP8.5 se referem aos cenários de emissão de gases do efeito estufa e derivam do inglês “*Representative Concentration Pathway*”, correspondendo às forçantes radiativas de $4.5 Wm^2$ e $8.5 Wm^2$, respectiva-

⁹Disponível em http://www.mct.gov.br/upd_blob/0215/215070.pdf.

mente. Tais simulações regionalizadas apoiaram a elaboração da Terceira Comunicação Nacional¹⁰ e o estudo denominado “Brasil 2040: cenários e alternativas de adaptação à mudança do clima” da Secretaria de Assuntos Estratégicos da Presidência da República.

Os dados das previsões de mudanças climáticas apresentam um total de 59 variáveis para um período histórico e futuro entre os anos de 1960 e 2100. Esses dados são gerados em formato binário e disponibilizados manualmente aos usuários. As variáveis climáticas e de extremos climáticos geradas pelos modelos e que foram armazenadas no banco de dados do CPTEC/INPE são descritas nas Tabelas 10 e 11, respectivamente.

Tabela 10: Variáveis Climáticas.

Nome	Sigla	Descrição	Unidade
Cobertura de Nuvens Altas	HINV	Fração de cobertura de nuvens altas.	0-1
Cobertura de Nuvens Baixas	LWNV	Fração de cobertura de nuvens baixas.	0-1
Cobertura de Nuvens Médias	MDNV	Fração de cobertura de nuvens médias.	0-1
Direção do vento a 10 m	D10M	Pode ser comparado à direção do vento medido na estação meteorológica.	graus meteorológicos
Direção do vento a 100 m	D100	Pode ser comparado à direção do vento medido a 100 metros de altura de uma torre.	graus meteorológicos
Energia Potencial Disponível para Convecção	CAPE	Energia disponível na atmosfera para produzir nuvens originadas do aquecimento nos baixos níveis da troposfera. Também conhecido como CAPE. Quando o valor é alto, acima de cerca de 1500 J/kg, indica disponibilidade de energia para gerar nuvens tipo Cumulonimbus. Quando o valor é baixo, menor que cerca de 500 J/kg, indica estabilidade atmosférica.	J/kg
Escoamento Subsuperficial	RNSG	Runoff dentro do solo.	mm/3h, mm/dia, mm/mês, mm/ano
Escoamento Superficial	RNOF	Equivale a Runoff da hidrologia.	mm/3h, mm/dia, mm/mês, mm/ano

Continuação na próxima página

¹⁰Disponível em <https://goo.gl/Qe1g2y>

Tabela 10 – Continuação

Nome	Sigla	Descrição	Unidade
Evapotranspiração	EVTP	Equivale a evaporação. Em áreas com cobertura vegetal, é considerada a transpiração da planta e evaporação da água interceptada da chuva sobre as folhas	mm/3h, mm/- dia, mm/mês, mm/ano
Evapotranspiração Potencial	EVPP	Equivale a evapotranspiração com o solo em capacidade de campo, isto é, em ótimas condições hídricas.	mm/3h, mm/- dia, mm/mês, mm/ano
Fluxo de Calor Latente à Superfície	CLSF	Equivale à energia transmitida da superfície da terra para atmosfera devido à evaporação. A conversão para evaporação da superfície é fornecida na variável evapotranspiração.	W/m^2
Fluxo de Calor no Solo	GHFL	Equivale a energia transmitida dentro das camadas do solo.	W/m^2
Fluxo de Calor Sensível à Superfície	CSSF	Equivale a energia transmitida da superfície da terra para atmosfera devido ao calor da superfície.	W/m^2
Neve	NEVE	Neve acumulada dentro de um período de tempo.	mm/3h, mm/- dia, mm/mês, mm/ano
Precipitação Total	PREC	Precipitação acumulada dentro de um período de tempo.	mm/3h, mm/- dia, mm/mês, mm/ano
Pressão à Superfície	PSLC	Pressão atmosférica no nível da localidade.	hPa
Pressão ao Nível Médio do Mar	PSLM	É pressão atmosférica reduzida ao nível médio do mar. É usada para comparação da pressão atmosférica entre localidades, independente das diferenças de altitude entre elas.	hPa
Radiação de Onda Curta Emergente à Superfície	OCES	Fluxo de radiação solar refletida pela superfície terrestre.	W/m^2
Radiação de Onda Curta Incidente à Superfície	OCIS	Fluxo de radiação solar que atinge a superfície terrestre.	W/m^2
Radiação de Onda Longa Emergente à Superfície	OLES	Fluxo de radiação emitida pela superfície terrestre.	W/m^2

Continuação na próxima página

Tabela 10 – Continuação

Nome	Sigla	Descrição	Unidade
Radiação de Onda Longa Incidente à Superfície	OLIS	Fluxo de radiação terrestre que atinge a superfície terrestre.	W/m^2
Temperatura a 2 m	TP2M	Pode ser comparada à temperatura medida na estação meteorológica.	°C
Temperatura à Superfície (continente e oceano)	TSFC	Temperatura à superfície da terra, quando medida sobre a terra, ou temperatura da superfície do mar quando medida sobre o oceano.	°C
Temperatura do Ponto de Orvalho a 2 m	DP2M	Temperatura em que se forma o orvalho, estimada na altura da estação meteorológica. Alta temperatura do ponto de orvalho indica alta quantidade de umidade no ar.	°C
Temperatura do Solo na Camada de 0-10 cm	TGSC	Temperatura da primeira camada do solo. Recebe valor indefinido sobre oceano.	°C
Temperatura do Solo na Camada de 10-40 cm	TGRZ	Temperatura da camada de 10 a 40 cm do solo. Recebe valor indefinido sobre oceano.	°C
Temperatura Máxima a 2 m	MXTP	É a temperatura máxima dentro de um período de tempo.	°C
Temperatura Mínima a 2 m	MNTP	É a temperatura mínima dentro de um período de tempo.	°C
Umidade do Solo na Camada de 0-10 cm	USSL	Umidade da primeira camada do solo. Recebe valor indefinido sobre oceano.	0-1
Umidade do Solo na Camada de 10-40 cm	UZRS	Umidade da camada de 10 a 40 cm do solo. Recebe valor indefinido sobre oceano.	0-1
Umidade Relativa a 2 m	UR2M	Pode ser comparada à umidade relativa estimada na estação meteorológica.	%
Vento a 10 m	W10M	Pode ser comparado à força do vento medida na estação meteorológica.	m/s
Vento a 100 m	W100	Pode ser comparado à força do vento medido a 100 metros de altura de uma torre.	m/s

Tabela 11: Índices de Extremos Climáticos

Nome	Sigla	Descrição	Unidade
Número máximo de dias consecutivos sem chuva no ano	CDD	Este índice reflete a duração de períodos de estiagem.	dias
Número máximo de dias consecutivos no ano com temperatura mínima diária menor que percentil 10	CSDI	Este índice reflete a duração das ondas de frio.	dias
Número máximo de dias consecutivos com chuva no ano	CWD	Este índice reflete a duração de períodos chuvosos.	dias
Amplitude anual média da temperatura diurna	DTR	Este índice reflete a diferença entre a temperatura mínima e a máxima do dia. A maior diferença indica maior amplitude térmica diária.	°C
Total de dias com temperatura abaixo de 0oC	FD	Este índice reflete a frequência de dias de possível ocorrência de geada e condições desfavoráveis para culturas sensíveis à geada.	dias
Total de dias com temperatura máxima abaixo de 0oC	ID	Este índice reflete a frequência de dias congelantes.	dias
Precipitação anual total dos dias úmidos	PRCPTOT	Fornecer a precipitação acumulada durante o ano.	mm
Número de dias no ano com RR >= 10 mm	R10mm	Este índice reflete frequência de eventos de chuva >10mm/dia.	dias
Número de dias no ano com RR >= 20 mm	R20mm	Este índice reflete frequência de eventos de chuva >20mm/dia.	dias
Número de dias no ano com RR >= 25 mm	R25mm	Este índice reflete frequência de eventos de chuva >25mm/dia.	dias
Máxima precipitação anual em 1 dia	RX1day	Precipitação diária mais intensa ocorrida no ano. O índice reflete a intensidade de chuva que pode causar inundações bruscas.	mm
Máxima precipitação anual em 5 dias consecutivos	RX5day	Precipitação acumulada em 5 dias, máxima ocorrida em um ano. É um indicador para possibilidade de deslizamento.	mm
Precipitação anual total dos dias em que RR > percentil 95	R95p	Indica precipitação de intensidade extrema.	mm

Continuação na próxima página

Tabela 11 – Continuação

Nome	Sigla	Descrição	Unidade
Precipitação anual total dos dias em que RR > percentil 99	R99p	Indica precipitação de intensidade muito extrema.	mm
Precipitação anual dividida pelo número de dias com chuva	SDII	Indica a intensidade média da precipitação durante o ano.	mm
Número de dias no ano em que TX > 25°C	SU25	Indica a frequência de dias no ano em que a temperatura máxima excede 25oC. É indicativo de dias de verão para latitudes médias onde as temperaturas máximas geralmente são menores que 25oC.	dias
Porcentagem anual de dias em que TN < percentil 10	TN10p	Indica a frequência das noites muito frias durante o ano.	%
Porcentagem anual de dias em que TN > percentil 90	TN90p	Indica a frequência dos dias muito quentes durante o ano.	%
Menor temperatura mínima anual	TNn	Menor temperatura mínima anual	°C
Maior temperatura mínima anual	TNx	Maior temperatura mínima anual	°C
Número de dias no ano em que TN > 20°C	TR	Indica a frequência de noites com temperatura acima de 20oC. É um indicador empregado para latitudes médias onde as temperaturas mínimas geralmente não excedem 20oC.	dias
Porcentagem anual de dias em que TX < percentil 10	TX10p	Indica a frequência de dias muito frios no ano.	%
Porcentagem anual de dias em que TX > percentil 90	TX90p	Indica a frequência de dias muito quentes no ano.	%
Menor temperatura máxima anual	TXn	Menor temperatura máxima anual	°C
Maior temperatura máxima anual	TXx	Maior temperatura máxima anual	°C
Número máximo de dias consecutivos no ano com TX > percentil 90	WSDI	Indica a duração das ondas de calor no ano.	dias

Este capítulo teve como objetivo apresentar uma breve introdução sobre o Modelo Eta, suas características e as variáveis que são geradas com a sua execução.

4. CPTEC TOOLKIT

Para desenvolver o kit de ferramentas para o CPTEC/INPE foram estudadas tecnologias, com o objetivo de verificar o comportamento e a possibilidade de integração entre cada uma delas. O desenvolvimento do CPTEC Toolkit proporcionou facilidade nas requisições dos usuários, podendo interagir de diversas maneiras nos processos de requisição, de disponibilização e de visualização dos dados climáticos. Neste capítulo a plataforma CPTEC Toolkit¹¹ é apresentada, bem como uma descrição das ferramentas utilizadas para a sua construção.

4.1 ARMAZENAMENTO DOS DADOS CLIMÁTICOS

Para efetuar o armazenamento dos dados foi escolhido o SGDB PostgreSQL, que é um banco de dados com uma gama considerável de recursos. As principais características que levaram a esta escolha foram: transações, triggers, segurança de acesso, necessidade de trabalhar com objeto relacional, tipos de dados, funções e a integridade dos dados. Uma extensão do PostgreSQL chamada PostGIS também foi utilizada, por possuir inúmeras funções úteis nas consultas, bem como armazenamento de dados georeferenciados, os quais foram fundamentais no decorrer do desenvolvimento do toolkit.

Os arquivos binários disponibilizados pelo CPTEC/INPE foram lidos e armazenados no PostgreSQL. Para que fosse possível essa leitura, foi necessário a criação de um script em R, com determinadas funções de leitura, manipulação e gravação dos dados em formato raster.

O PostgreSQL armazena os dados em formatos de tabelas, ou seja, os scripts gerados pelo pacote devem estar neste formato SQL ou TIFF, para que seja possível realizar o processo de conversão e inserção dos dados na base de dados do projeto.

A base de dados contém simulações de mudanças climáticas em escala reduzida a 20 km pelo Modelo Eta, para toda América do Sul e parte da América Central e escala de 5 km para a região sudeste do Brasil [62, 63, 64]. Para a resolução de 20 km são disponibilizadas as regionalizações do modelo Eta, acoplado com os modelos globais HadGEM2-ES e MIROC CanESM, neste primeiro momento. Para a resolução de 5 km foi disponibilizado a regionalização do modelo global HadGEM2-ES. O período de disponibilização dos dados é de 1961 a 2100, em cenários de alta e baixa emissão dos gases de efeito estufa (RCP8.5 e RCP4.5). Serão disponibilizados 33 variáveis climáticas e 26 índices de extremos climáticos derivados das projeções do modelo Eta. A lista das variáveis climáticas e dos índices de extremos estão listados nas Tabelas 10 e 11, respectivamente.

Atualmente a base de dados está com o tamanho total de 30TB. Com a inserção de todos os dados horários, bem como variáveis que ainda não foram incluídas, a estimativa é que chegue próximo aos 80TB.

¹¹Disponível em: <https://projeta.cptec.inpe.br>.

4.1.1 Descrição das tabelas

Neste item são detalhadas as estruturas das tabelas criadas no banco de dados do projeto. A Figura 5, apresenta o diagrama de classes que é uma representação da estrutura e relações entre as classes.

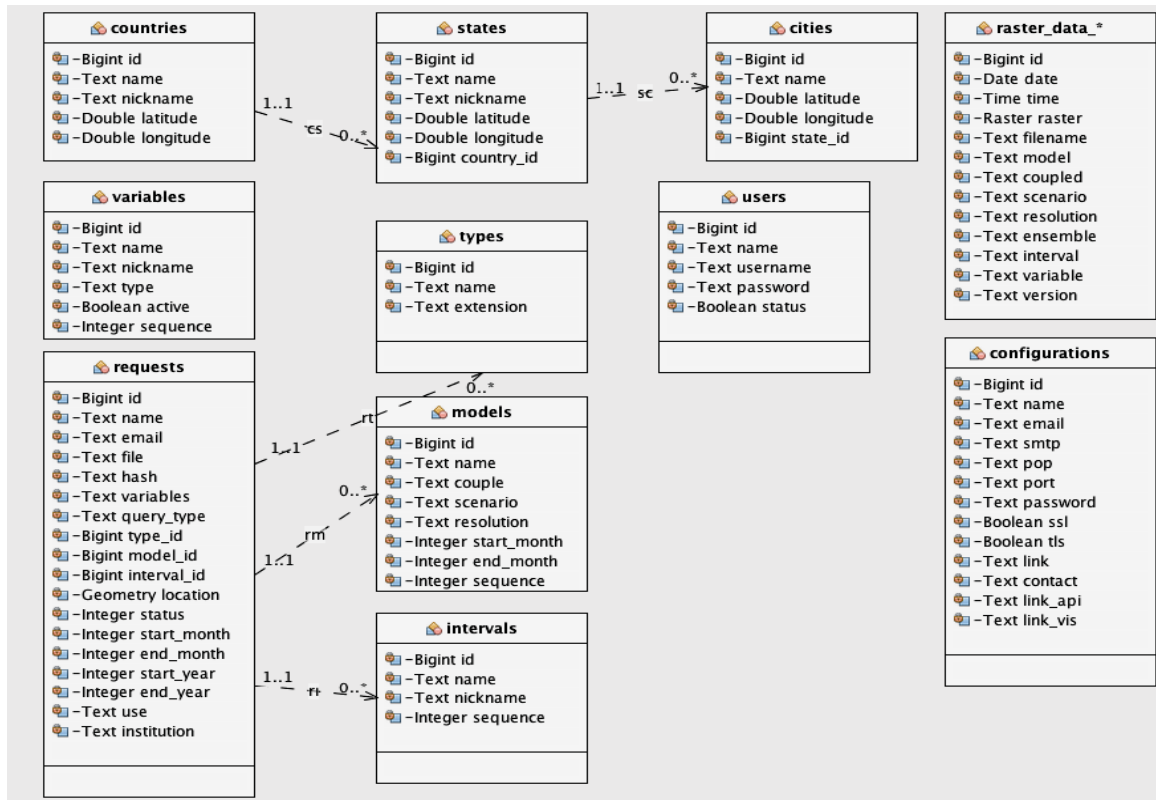


Figura 5. Diagrama de classes que descreve a estrutura e relação entre as classes do sistema.

Devido a quantidade de modelos, cenários, intervalos e resoluções, foi necessário fazer a divisão das tabelas raster data para o armazenamento. A estruturação destas tabelas está descrita na Tabela 12.

Tabela 12: Esquema de tabelas para armazenamento dos dados.

Modelo	Cenário	Intervalo	Resolução	Tabela
MIROC5	RCP4.5	3 HORAS	20KM	raster_data_1
MIROC5	RCP4.5	DIÁRIO	20KM	raster_data_2
MIROC5	RCP4.5	MENSAL	20KM	raster_data_3
MIROC5	RCP4.5	ANUAL	20KM	raster_data_4
MIROC5	RCP8.5	3 HORAS	20KM	raster_data_5
MIROC5	RCP8.5	DIÁRIO	20KM	raster_data_6

Continuação na próxima página

Tabela 12 – Continuação

Modelo	Cenário	Intervalo	Resolução	Tabela
MIROC5	RCP8.5	MENSAL	20KM	raster_data_7
MIROC5	RCP8.5	ANUAL	20KM	raster_data_8
MIROC5	HISTÓRICO	3 HORAS	20KM	raster_data_9
MIROC5	HISTÓRICO	DIÁRIO	20KM	raster_data_10
MIROC5	HISTÓRICO	MENSAL	20KM	raster_data_11
MIROC5	HISTÓRICO	ANUAL	20KM	raster_data_12
CanESM2	RCP4.5	3 HORAS	20KM	raster_data_13
CanESM2	RCP4.5	DIÁRIO	20KM	raster_data_14
CanESM2	RCP4.5	MENSAL	20KM	raster_data_15
CanESM2	RCP4.5	ANUAL	20KM	raster_data_16
CanESM2	RCP8.5	3 HORAS	20KM	raster_data_17
CanESM2	RCP8.5	DIÁRIO	20KM	raster_data_18
CanESM2	RCP8.5	MENSAL	20KM	raster_data_19
CanESM2	RCP8.5	ANUAL	20KM	raster_data_20
CanESM2	HISTÓRICO	3 HORAS	20KM	raster_data_21
CanESM2	HISTÓRICO	DIÁRIO	20KM	raster_data_22
CanESM2	HISTÓRICO	MENSAL	20KM	raster_data_23
CanESM2	HISTÓRICO	ANUAL	20KM	raster_data_24
HadGEM2-ES	RCP4.5	3 HORAS	20KM	raster_data_25
HadGEM2-ES	RCP4.5	DIÁRIO	20KM	raster_data_26
HadGEM2-ES	RCP4.5	MENSAL	20KM	raster_data_27
HadGEM2-ES	RCP4.5	ANUAL	20KM	raster_data_28
HadGEM2-ES	RCP8.5	3 HORAS	20KM	raster_data_29
HadGEM2-ES	RCP8.5	DIÁRIO	20KM	raster_data_30
HadGEM2-ES	RCP8.5	MENSAL	20KM	raster_data_31
HadGEM2-ES	RCP8.5	ANUAL	20KM	raster_data_32
HadGEM2-ES	HISTÓRICO	3 HORAS	20KM	raster_data_33
HadGEM2-ES	HISTÓRICO	DIÁRIO	20KM	raster_data_34
HadGEM2-ES	HISTÓRICO	MENSAL	20KM	raster_data_35
HadGEM2-ES	HISTÓRICO	ANUAL	20KM	raster_data_36
HadGEM2-ES	RCP4.5	3 HORAS	5KM	raster_data_37
HadGEM2-ES	RCP4.5	DIÁRIO	5KM	raster_data_38
HadGEM2-ES	RCP4.5	MENSAL	5KM	raster_data_39
HadGEM2-ES	RCP4.5	ANUAL	5KM	raster_data_40
HadGEM2-ES	RCP8.5	3 HORAS	5KM	raster_data_41
HadGEM2-ES	RCP8.5	DIÁRIO	5KM	raster_data_42
Continuação na próxima página				

Tabela 12 – Continuação

Modelo	Cenário	Intervalo	Resolução	Tabela
HadGEM2-ES	RCP8.5	MENSAL	5KM	raster_data_43
HadGEM2-ES	RCP8.5	ANUAL	5KM	raster_data_44
HadGEM2-ES	HISTÓRICO	3 HORAS	5KM	raster_data_45
HadGEM2-ES	HISTÓRICO	DIÁRIO	5KM	raster_data_46
HadGEM2-ES	HISTÓRICO	MENSAL	5KM	raster_data_47
HadGEM2-ES	HISTÓRICO	ANUAL	5KM	raster_data_48

4.2 APPLICATION PROGRAMMING INTERFACE (API)

O uso das APIs estão cada vez mais sendo utilizados, já que as funcionalidades e suas vantagens são inúmeras como: alto poder de integração, viabilização de parcerias, customização de serviços, entre outros.

Com o banco de dados disponível, os arquivos binários convertidos e importados, foi possível iniciar o projeto back-end, que é responsável por todo o processo de inclusão, alteração, remoção, listagens e processamento dos dados, bem como a disponibilização de uma API que foi utilizada por toda a plataforma.

As conversões dos arquivos binários foram feitas através de um programa desenvolvido em R, ondem são gerados arquivos do tipo TIF. Estes arquivos são importados através do comando `raster2pgsql`, que possibilita o uso de imagens diretamente no banco de dados, sem a ajuda de linguagens extras.

Os dados são disponibilizados por meio de links fornecidos pela API. Cada caso de uso foi analisado e projetado conforme necessidades da plataforma. Foram definidos parâmetros padrões e específicos, para que o serviço conseguisse efetuar as consultas de maneira correta, fornecendo os dados de acordo com o desejado pelos usuários e de acordo com as orientações recebidas dos pesquisadores do CPTEC. O fluxo consiste no usuário acessar via HTTP, através de um navegador de sua preferência, uma aplicação angular. A aplicação por sua vez, através de uma API que acessa a camada servidor NodeJS. A camada faz o acesso ao banco de dados, onde são feitas requisições do tipo GET, POST, PUT e DELETE. A representação do fluxo é apresentado na Figura 6, onde o usuário acessa através de navegador de sua preferência, via HTTP, uma aplicação AngularJS, que por sua vez acessa a API. A API acessa a camada de servidor que foi desenvolvido em NodeJS, acessando recursos do banco de dados, retornando à requisição que pode ser GET, POST, PUT e DELETE. O pacote R apresentado na figura, foi usado inicialmente para fazer a transformação dos dados binários para tif. Posteriormente o CPTEC/INPE passou a fornece-los no formato correto para inclusão no banco.

A troca de informações por meio de API tem papel fundamental, padronizando essas integrações e fornecendo uma certa autonomia para as empresas e instituições usufruírem dos serviços disponíveis.

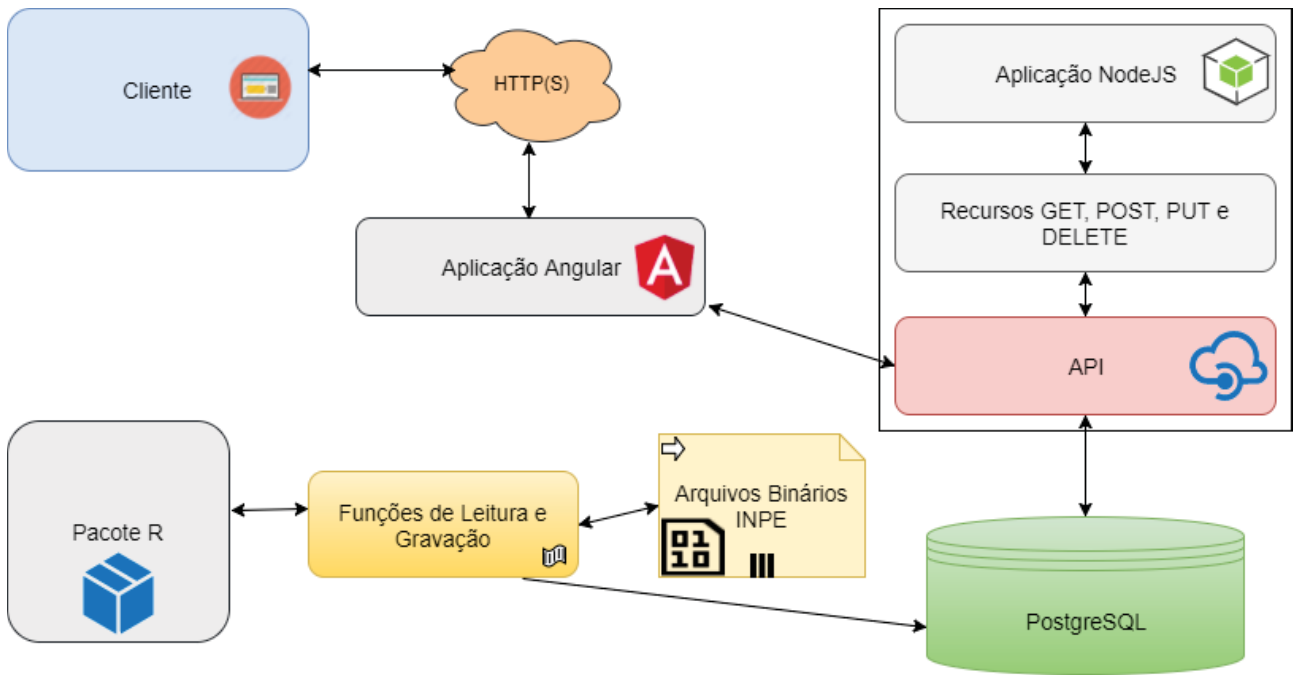


Figura 6. Fluxo da plataforma até a disponibilização para os clientes (aplicativos, sistemas web, etc).

A plataforma geral do kit de ferramentas foi desenvolvida utilizando recursos que possibilitam a criação de novos produtos e serviços, conforme a Figura 7, que apresenta a aplicabilidade de uma API em vários produtos que com ela podem ser desenvolvidos. Com isso foi possível dar maior flexibilidade no tráfego dos dados, podendo ser distribuída a terceiros, sem intervenção dos técnicos responsáveis pela plataforma.

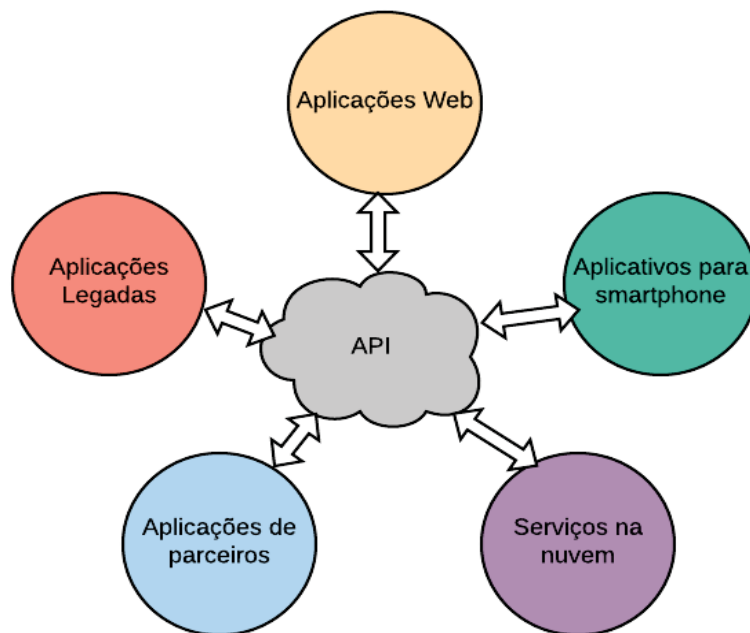


Figura 7. Diferentes produtos que podem ser criados a partir de uma integração com a API.

4.3 INTERFACE WEB - FRONT-END

A interface de um software ou aplicativo deve ser tratada com muita atenção e cuidado, podendo muitas vezes proporcionar uma má experiência do usuário, o que pode acarretar negativamente na avaliação das funcionalidades do software ou do aplicativo. Contudo, se a interface for bem projetada e desenvolvida obedecendo os padrões web, pode garantir aos usuários a realização das tarefas de forma simples e eficiente.

O kit de ferramentas deste projeto conta com aplicações web e aplicativos, desenvolvidos com padrões atuais e a utilização de HTML5, com o intuito de deixar o usuário seguro das tarefas que ele está desempenhando, tanto nas requisições de dados como na visualização, trazendo a clareza das informações requisitadas.

4.3.1 Características do Kit de Ferramentas

Foram utilizadas soluções de desenvolvimento baseadas na web, utilizando o conceito Cliente/Servidor. Dessa forma, visou-se o desenvolvimento de uma infraestrutura capaz de operar e tratar informações oriundas de requisições web por meio de serviços API¹². Alguns detalhes imprescindíveis foram relatados pelos técnicos e responsáveis do CPTEC/INPE durante o desenvolvimento e foram essenciais para o desenvolvimento:

- a) **O envio de um e-mail para o solicitante, com o link para o download:** Após o usuário finalizar a solicitação, é enviado um e-mail com os dados da requisição e um link para download. No conteúdo do e-mail são enviados o link, um conjunto de referências bibliográficas para ser citados quando utilizados em publicações técnicas ou científicas, como é mostrado na Figura 8. Uma chave foi gerada para cada requisição, facilitando a busca quando o usuário utilizar o link enviado. Esta chave é única e faz parte do endereço enviado no e-mail, facilitando a busca da requisição no momento que o usuário abrir o link.
- b) **O limite de espaço de armazenamento de dados:** Os arquivos foram convertidos e armazenados no banco de dados PostgreSQL, respeitando o limite de armazenamento definido pelo TI do CPTEC/INPE. Atualmente, tem-se armazenado aproximadamente 30 TB de dados.
- c) **Os arquivos binários são armazenados em formato TIFF no banco de dados PostgreSQL:** Os arquivos binários foram convertidos para objetos TIFF pela equipe do CPTEC/INPE, que nos passaram os mesmos para armazenar no banco de dados conforme apresentados na Figura 9. Os arquivos TIFF foram inseridos através do comando raster2pgsql, disponibilizado pelo PostGIS.

¹²É um conjunto de rotinas e padrões de programação para acesso a um aplicativo de software ou plataforma baseado na Web.

projeta@inpe.br
para mim

21:21 (Há 0 minutos)

Olá Felipe Borella

Informamos que a sua requisição está disponível.
Clique no link abaixo para ser direcionado até a área de download

<http://192.168.1.50/cptec/#/download/dbe6d4c69b5134a08f87aefdc026e66abb6aa109>

Quando os dados forem utilizados para trabalhos acadêmicos e publicações utilizar nas referências bibliográficas a fonte dos dados CPTEC/INPE da seguinte forma "Dados gerados pelo CPTEC/INPE" e as referências bibliográficas:

1. Chou, S.C., Lyra, A., Mourão, C., Dereczynski, C., Pilotto, I., Gomes, J., Bustamante, J., Tavares, P., Silva, A., Rodrigues, D., Campos, D., Chagas, D., Sueiro, G., Siqueira, G., Nobre, P. and Marengo, J. (2014) Evaluation of the Eta Simulations Nested in Three Global Climate Models. American Journal of Climate Change, 3, 438-454. doi:10.4236/ajcc.2014.35039. http://www.scirp.org/journal/PaperInformation.aspx?PaperID=52887#VakHq_IVko
2. Chou, S.C., Lyra, A., Mourão, C., Dereczynski, C., Pilotto, I., Gomes, J., Bustamante, J., Tavares, P., Silva, A., Rodrigues, D., Campos, D., Chagas, D., Sueiro, G., Siqueira, G. and Marengo, J. (2014) Assessment of Climate Change over South America under RCP 4.5 and 8.5 Downscaling Scenarios. American Journal of Climate Change, 3, 512-527. doi: 10.4236/ajcc.2014.35043. http://www.scirp.org/journal/PaperInformation.aspx?PaperID=52877#Vaklh_IVko
3. Lyra, A., Tavares, P., Chou, S.C., Sueiro, G., Dereczynski, C.P., Sondermann, M., Silva, A., Marengo, J., Giarolla, A. 2017. Climate change projections over three metropolitan regions in Southeast Brazil using the non-hydrostatic Eta regional climate model at 5-km resolution Theor Appl Climatol. doi:10.1007/s00704-017-2067-z. <https://link.springer.com/article/10.1007/s00704-017-2067-z>
4. Terceira Comunicação Nacional do Brasil à Convenção-Quadro das Nações Unidas sobre Mudança do Clima

Executive Summary: <http://unfccc.int/resource/docs/natc/branc3es.pdf>

Volume 1: <http://unfccc.int/resource/docs/natc/branc3v1.pdf>

Volume 2: <http://unfccc.int/resource/docs/natc/branc3v2.pdf>

Volume 3: <http://unfccc.int/resource/docs/natc/branc3v3.pdf>

Figura 8. E-mail enviado pela plataforma ao concluir o processamento da requisição.

date	time	rast	filename	model	model_coupled	scenario	model_resolution	variable
date	time withou..	raster	text	characte...	character varyi...	character ...	character varying...	character var
1998-02-27	00:00:00	01000001009A9999999999C93F9...	Eta_MIROC5_20_Historical_climat...	Eta	MIROC5	Historical	20	PREC
1998-02-28	00:00:00	01000001009A9999999999C93F9...	Eta_MIROC5_20_Historical_climat...	Eta	MIROC5	Historical	20	PREC
1998-03-01	00:00:00	01000001009A9999999999C93F9...	Eta_MIROC5_20_Historical_climat...	Eta	MIROC5	Historical	20	PREC
1998-03-02	00:00:00	01000001009A9999999999C93F9...	Eta_MIROC5_20_Historical_climat...	Eta	MIROC5	Historical	20	PREC
1998-03-03	00:00:00	01000001009A9999999999C93F9...	Eta_MIROC5_20_Historical_climat...	Eta	MIROC5	Historical	20	PREC

Figura 9. Tabela raster_data, com os arquivos no formato TIFF, que foram importados via comandos disponibilizados pelo PostGIS.

- d) **O solicitante pode escolher o domínio e o período do cenário, através de parâmetros em tela:** Foram desenvolvidas interfaces com parametrizações flexíveis para que o usuário consiga completar sua requisição de forma correta, conforme apresentado na Figura 10.
- e) **As requisições devem ser geradas em um arquivo compactado, juntamente com uma cópia do pedido (PDF ou imagem), para que o CPTEC/INPE possa garantir que o pedido foi atendido corretamente:** O resultado da requisição foi comprimido em formato ZIP, juntamente com o PDF da requisição do usuário.
- f) **Na plataforma é possível efetuar cadastro de países, estados, cidades, cenários climáticos e outros cadastros que forem necessários, facilitando algumas requisições:** Na área administrativa, é possível fazer a manutenção destes cadastros, necessários para o andamento do processo, bem como facilitar as requisições dos usuários. Esta área é utilizada apenas para usuários autenticados, para garantir maior segurança dos dados. As telas de login e da área administrativa são mostradas nas Figuras 11 e 12. Para efetuar o login, o usuário deve informar o e-mail e senha, validar por meio da API, a qual libera um token, caso o usuário seja válido, acessando assim a área administrativa.

Figura 10. Primeiro passo da requisição, onde o usuário informa dados referentes ao cenário climático, frequência, localização, formato e período.

Figura 11. Tela de acesso da plataforma CPTEC. O ambiente é restrito e somente é liberado se o usuário possuir um usuário e senha válidos.

- g) **A possibilidade de fazer upload de arquivos de formato shapefile facilitando na parametrização dos dados solicitados:** Para facilitar algumas requisições, o usuário pode fazer upload de arquivos em formato shapefile, que tem por objetivo importar dados geoespaciais ajudando na filtragem dos dados requisitados.

De acordo com as características e necessidades dos usuários do CPTEC/INPE foram definidos alguns padrões de sistema, os quais são descritos nas seções seguintes.

4.3.2 Requisição

A requisição dos dados foi dividida em quatro etapas. O primeiro passo refere-se a escolha dos dados desejados, conforme apresentado na Figura 13. O usuário deve selecionar um cenário climático, frequência, localização (que pode ser por ponto, por cidade, desenho no mapa ou importação

PROJETA Projeções de mudanças climáticas para a América do Sul regionalizadas pelo Modelo Eta

MENU PRINCIPAL

- Manutenções
- Cadastros
- Modelos**
- Frequências
- Variáveis
- Modelos/Frequências
- Tipos de Saída
- Cidades
- Estados
- Países
- Configuração

Modelos

Novo Nome Pesquisar

Nome	Modelo	Modelo acoplado	Cenário	Resolução	
Eta - 20 km, Historical, BESM, continental. - 1961-2005	ETA	BESM	Historical	20	Editar
Eta - 20 km, Historical, CANESM2, continental. - 1961-2005	ETA	CANESM2	Historical	20	Editar
Eta - 20 km, Historical, HADGEM2-ES, continental. - 1961-2005	ETA	HADGEM2-ES	Historical	20	Editar
Eta - 20 km, Historical, MIROC5, continental. - 1961-2005	ETA	MIROC5	Historical	20	Editar
Eta - 20 km, RCP4.5, BESM, continental. - 2006-2099	ETA	BESM	RCP4.5	20	Editar
Eta - 20 km, RCP4.5, CANESM2, continental. - 2006-2099	ETA	CANESM2	RCP4.5	20	Editar
Eta - 20 km, RCP4.5, HADGEM2-ES, continental. - 2006-2099	ETA	HADGEM2-ES	RCP4.5	20	Editar
Eta - 20 km, RCP4.5, MIROC5, continental. - 2006-2099	ETA	MIROC5	RCP4.5	20	Editar
Eta - 20 km, RCP8.5, BESM, continental. - 2006-2099	ETA	BESM	RCP8.5	20	Editar
Eta - 20 km, RCP8.5, CANESM2, continental. - 2006-2099	ETA	CANESM2	RCP8.5	20	Editar

Primeira Anterior 1 2 Próxima Última

Figura 12. Área administrativa visualizada após o login do usuário, com a manutenção de Modelos em destaque.

de arquivos shapefile), no mínimo uma variável (para frequência anual ou mensal é possível selecionar até cinco variáveis) e o período que é mês/ano inicial e mês/ano final.

1º Passo Escolha dos dados

2º Passo Visualização de amostra do dado

3º Passo Dados e formato dos dados

4º Passo Concluir requisição

Escolha o cenário climático, frequência e demais dados necessários

Cenário Climático: Eta - 20 km, RCP4.5, MIROC5, continental.

Frequência: ANUAL

Localização: Por cidade

Formato do arquivo dos dados: JSON

Variáveis: TP2M - Temperatura a 2 m da Superfície

Mês Inicial: Janeiro

Ano Inicial: 2006

Mês Final: Dezembro

Ano Final: 2006

Cidade: MARAU

Prezado usuário, recomenda-se realizar a requisição de períodos maiores que 10 anos.

Próximo

Figura 13. Passo 1 - Escolha dos filtros iniciais para a requisição.

No passo dois é apresentado o mapa com a área (ou ponto ou cidade) de abrangência da amostra requisitada pelo usuário. A Figura 14 apresenta, como exemplo, a escolha por cidade que foi realizada pelo usuário no passo anterior, que trata-se de Marau, Rio Grande do Sul, cidade onde resido.

O formulário apresentado no passo três requer informações pessoais do usuário: nome, e-mail, motivo para utilização do dado e instituição que está requerendo os dados (Figura 15).

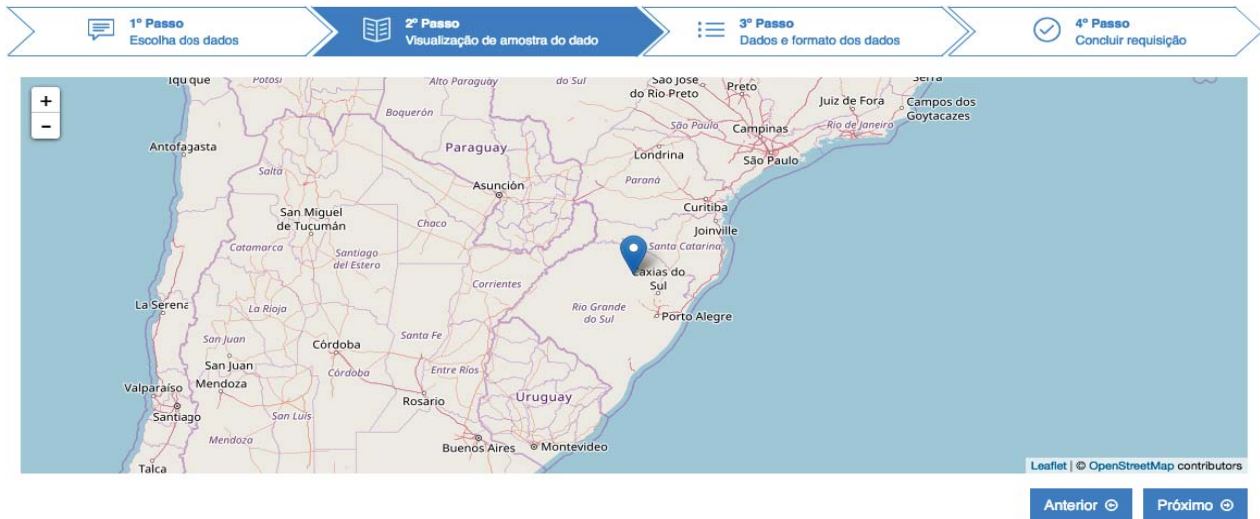


Figura 14. Passo 2 - Visualização de amostra do dado, onde é apresentado no mapa.

Informar dados pessoais e escolha do formato dos arquivos

Nome	E-mail
Felipe Borella	borellaster@gmail.com
Utilização do dado	Instituição
Estudo	Universidade de Passo Fundo

Figura 15. Passo 3 - Usuário informa os dados pessoais, qual a utilização dos dados e instituição a qual ele pertence.

O passo quatro é um resumo da requisição com o objetivo que o usuário possa fazer uma última conferência antes de concluir a mesma, conforme mostrado na Figura 16. Após concluir este passo, é apresentado uma mensagem informando sobre o processamento da requisição e sobre o envio do e-mail para download. A mensagem é apresentada na Figura 17.

Confirmar requisição e finalizar

<p>Resumo da requisição: Cenário Climático: Eta - 20 km, RCP4.5, MIROC5, continental. Frequência: ANUAL Localização: Por cidade Cidade: MARAU (-28.449, -52.2) Variáveis: TP2M Período: 1/2006 até 12/2006 Utilização do dado: Estudo Instituição: Estudo</p>	<p>Dados pessoais: Nome: Felipe Borella E-mail: borellaster@gmail.com</p> <p>Formato do arquivo dos dados: Arquivo: JSON</p>
--	---

Figura 16. Passo 4 - Conclusão da requisição, que pode ser visualizada no último passo, apresentando um resumo da requisição.

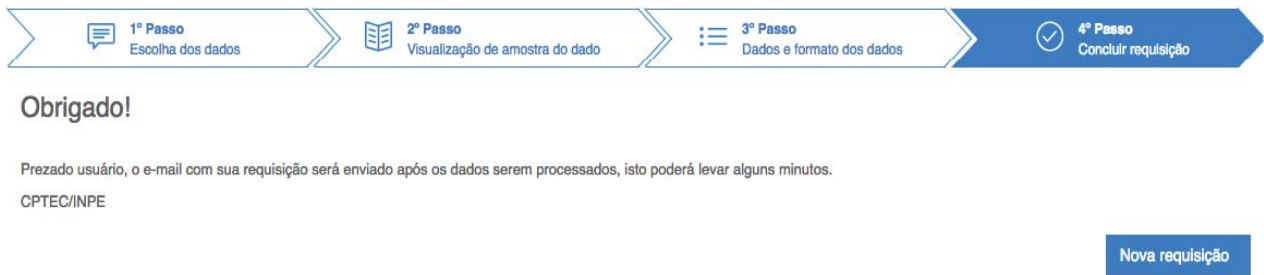


Figura 17. Passo 4 - Mensagem de sucesso da requisição, apresentada após concluir a mesma.

Depois que o usuário acessar o link que foi identificado no e-mail enviado ao usuário informando o local para download dos dados da requisição, será apresentado um breve resumo da requisição feita, parecido com o passo quatro, porém com opção de download da requisição, como pode ser visualizado na Figura 18.



Figura 18. Resumo e download da requisição.

4.3.3 Modularização e carga de módulos sob demanda

Toda a camada de visualização, pública e privada, e acesso a API, foi desenvolvida utilizando recursos do HTML5, CSS e JavaScript com o auxílio do AngularJS [27]. São vários os arquivos que precisam ser carregados quando abre-se a plataforma via navegador. Para solucionar este problema, buscou-se o conceito de modularização, que trata da separação de funcionalidades, redução de complexidade e principalmente, o reuso de código. Utilizou-se o RequireJS [67] para facilitar tanto o desenvolvimento quanto a rapidez da versão de distribuição final da plataforma. Outra biblioteca, chamada ocLazyLoad [68], permite o carregamento de arquivos adicionais usando um terceiro módulo que registra no Angular, deixando os mesmos disponíveis para uso na aplicação.

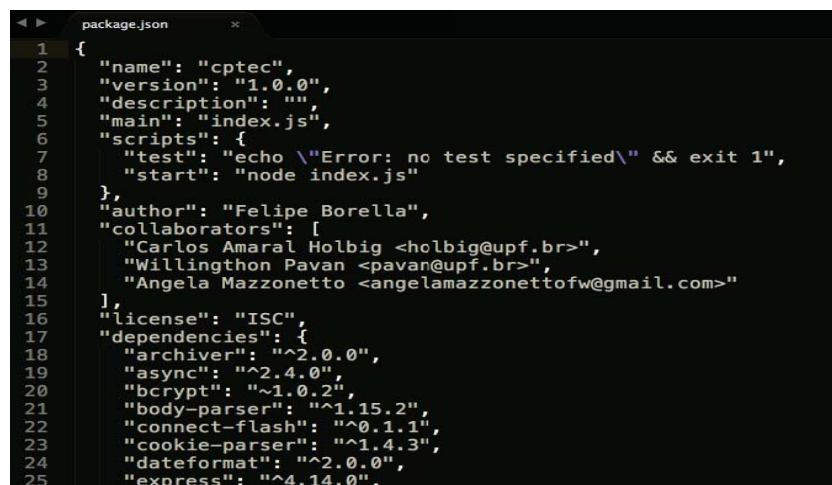
4.4 APLICAÇÃO BACK-END

Para a aplicação do back-end foi utilizado o NodeJS [28] que forneceu todo o suporte para a criação de uma API para grande volume de dados e requisições. Além de bibliotecas tradicionais, foi

utilizado o NPM¹³, Gerenciador de Pacotes do Node, que é um repositório online onde são publicados projetos de código livre para o Node.js. Trata-se de um utilitário de linha de comando que interage com este repositório, ajudando na instalação de pacotes, gerenciamento de versões e de dependências dos mesmos.

A seguir serão listadas algumas das principais bibliotecas utilizadas para que fosse possível trabalhar com este grande volume de dados.

- a) **ExpressJS**: Com o ExpressJS, foi possível armazenar as informações do metadados do projeto no arquivo `package.json`, como é mostrado na Figura 19, onde são armazenadas as características das bibliotecas, versões, arquivo principal da aplicação, entre outros dados. Toda biblioteca que foi baixada do NPM, cria uma linha no `package.json`. Isto serve para que outros integrantes do time de desenvolvimento possam rodar o projeto, com as mesmas versões, evitando problemas. Toda a API foi desenvolvida utilizando camadas, seguindo o padrão MVC (*Model/View/Controller*)¹⁴, permitindo, assim, a reutilização de códigos e a integração de novos módulos no futuro. O ExpressJS possibilitou o *setup* básico da aplicação e ainda serviu de alicerce para alguns padrões que foram utilizados durante o desenvolvimento.



```
1 {
2   "name": "cptec",
3   "version": "1.0.0",
4   "description": "",
5   "main": "index.js",
6   "scripts": {
7     "test": "echo \\\"Error: no test specified\\\" && exit 1",
8     "start": "node index.js"
9   },
10  "author": "Felipe Borella",
11  "collaborators": [
12    "Carlos Amaral Holbig <holbig@upf.br>",
13    "Willingthon Pavan <pavan@upf.br>",
14    "Angela Mazzone <angelamazzonefw@gmail.com>"
15  ],
16  "license": "ISC",
17  "dependencies": {
18    "archiver": "^2.0.0",
19    "async": "^2.4.0",
20    "bcrypt": "~1.0.2",
21    "body-parser": "^1.15.2",
22    "connect-flash": "^0.1.1",
23    "cookie-parser": "^1.4.3",
24    "dateformat": "^2.0.0",
25    "express": "^4.14.0",
```

Figura 19. Arquivo `package.json` com o metadados da plataforma, que são configurações do nome da aplicação, equipe e pacotes utilizados no desenvolvimento.

- b) **SequelizeJS**: Para acesso ao banco de dados foi utilizado o SequelizeJS [69], uma biblioteca do NodeJS, que oferece uma solução baseada em esquemas para modelar os dados da uma aplicação. Possui ainda um sistema de conversão de tipos de dados, validação, criação de consultas, que cobrem uma série de técnicas utilizadas para modificar ou melhorar comportamentos, essencial para lógica de negócios. O SequelizeJS fornece um mapeamento *Object Relational Mapping* (ORM). Na Figura 20 é mostrado o mapeamento da entidade país. O framework permite que sejam mapeados os campos das tabelas, tipo de dado, nome da tabela, associações entre tabelas, enfim,

¹³Node Package Manager

¹⁴Padrão de arquitetura de software que separa a representação da informação da interação do usuário com ele.

todo o suporte necessário para que a ferramenta consiga identificar e montar algumas consultas automáticas.



```

country.js
1  'use strict';
2  module.exports = function(sequelize, DataTypes) {
3    var country = sequelize.define('country', {
4      name: DataTypes.STRING,
5      nickname: DataTypes.STRING,
6      latitude: DataTypes.DOUBLE,
7      longitude: DataTypes.DOUBLE,
8      zoom: DataTypes.INTEGER
9    }, {
10     underscored: true,
11     classMethods: {
12       associate: function(models) {
13       }
14     }
15   });
16   return country;
17 };
18

```

Figura 20. Mapeamento objeto relacional da entidade país.

- c) **Archiver**: Biblioteca distribuída pela comunidade do NPM, que serve para criar arquivos. Para este projeto em específico foi utilizada para criar os arquivos no formato zip, necessários para geração do retorno das requisições.
- d) **Async**: Pacote utilizado para efetuar o processamento das requisições de modo assíncrono, dando maior facilidade e robustez nas funções utilizadas pelo Node.js.
- e) **Bcrypt**: Pacote utilizado para garantir que as senhas dos usuários fossem armazenadas com segurança, evitando, assim, problemas futuros. Este pacote possui um recurso chamado salt, que é um recurso que garante que a mesma senha nunca será igualmente criptografada duas vezes..
- f) **GeoJSON, Js2xmlparser e Json2csv**: Pacotes utilizado para transformar os dados da requisição em o formato geojson, xml e csv, respectivamente, disponibilizando os mesmos em formato zip para download.
- g) **Passport**: O Passport foi utilizado na implementação da autenticação e para proteger alguns links da API, onde somente usuários autorizados tem acesso para manipulação.
- h) **Jwt-simple**: Utilizado na autenticação, distribuindo aos usuários que fazem o login. O usuário recebe um token, que será utilizado nas sessões, dando a liberdade necessária para as manutenções no uso interno.

4.4.1 Recursos da API

Os recursos normalmente são dados, atribuindo o nome das entidades, para fazer sentido do ponto de vista de quem vai consumir a API. Utilizamos estes nomes no plural, por exemplo: uma

entidade chamada *country*, terá na URL */countries*. A grande vantagem de uma API, é o aproveitamento de métodos HTTP existentes para implementar funcionalidades significativas em apenas um único endpoint¹⁵. Não existem convenções de nomenclatura a seguir e a estrutura de URL é limpa e clara.

a) **Informações Gerais** - Trata-se da documentação básica de funcionamento da API, padrões adotados e nomenclaturas, bem como as especificações de cada um dos itens. É necessário identificar quais ações se aplicam e como serão mapeados para a API. Princípios RESTful fornecem estratégias para lidar com ações CRUD usando métodos HTTP.

- **:app** - Refere-se ao nome da aplicação;
- **:rest** - Parte da URL que identifica a comunicação em REST;
- **:version** - Versionamento da API;
- **:resource** - Nome da entidade;

b) **GET, UPDATE, DELETE:** <http://:app/:rest/:version/:resource/:id>

- **:id** - identificador do recurso;
- **GET** - Recupera o recurso do respectivo identificador;
- **UPDATE** - Atualiza o recurso do respectivo identificador. Enviado no corpo da requisição;
- **DELETE** - Solicita a remoção do respectivo recurso pelo seu identificador;

c) **GET, POST:** <http://:app/:rest/:version/:resource>

- **GET** - Recupera uma listagem dos recursos;
- **POST** - Novo recurso a ser inserido. Enviado no corpo da requisição;

d) **GET:** Suporte a paginação - <http://:app/:rest/:version/:resource?page=1&size=10>

- **page** - Identifica o índice inicial para listagem (valor default 1)
- **size** - Quantidade de itens que irão retornar na requisição (valor default 10).

Com o ferramental descrito acima, foi possível desenvolver o conjunto de ferramentas proposto, dando uma maior liberdade para os usuários no quesito facilidade de acesso e inclusão da requisição e, por outro lado, eliminando o serviço manual realizado no CPTEC/INPE.

Os estudos realizados desde o início do projeto também foram de grande valia, pois determinaram alguns padrões importantes no desenvolvimento da plataforma. O uso de todas estas ferramentas computacionais desempenharam bem o seu papel, dando o corpo necessário na arquitetura da plataforma.

¹⁵É o nome para uma entidade em um terminal de uma conexão da camada de transporte.

O projeto se comportou muito bem trabalhando em separado da API, assim como a camada de visualização, atingindo o objetivo de fazer com que a plataforma não fique fechada para uso exclusivo do CPTEC/INPE e sim que possa se utilizada por terceiros.

A requisição facilitou para que os usuários possam ter acesso aos dados de maneira fácil e intuitiva, de uma forma muito rápida, de acordo com as suas necessidades.

5. TESTES E RESULTADOS

Este capítulo é dedicado a análise da metodologia, das ferramentas e dos resultados esperados pela plataforma. São discutidos os principais aspectos técnicos que viabilizaram a pesquisa bem como sua concretização. Testes no ambiente de desenvolvimento e de produção também foram realizados.

5.1 ANÁLISE DA METODOLOGIA PROPOSTA

O JavaScript é uma linguagem interpretada para *browsers*, para atualizar conteúdo dinamicamente, permitindo assim que ações sejam realizadas enquanto outras ainda estão sendo processadas. O contexto dos *browsers* contribuiu para que o JavaScript evoluísse de uma forma diferente das demais linguagens, focando em performance e em possibilitar a criação de interfaces com uma melhor experiência para o usuário [70]. Foi com este intuito que desde o início optou-se por utilizar o JavaScript tanto no servidor quanto no cliente.

O NodeJS se mostrou uma linguagem confiável no decorrer do desenvolvimento, de fácil manutenção e de rapidez na execução de tarefas. É uma linguagem que dificilmente consome muita memória no seu desenvolvimento, liberando o sistema operacional para executar outras tarefas. Possui características de processamento assíncrono, facilitando, assim, a liberação da aplicação até a conclusão do processamento.

O NPM é o repositório mais importante da comunidade de desenvolvedores NodeJS. Atualmente conta com mais de 475.000 (quatrocentos e setenta e cinco mil) pacotes, prontos para serem baixados e utilizados gratuitamente entre os desenvolvedores.

Testes no processamento das requisições foram realizados com períodos de mais de 100 (cem) anos combinados com mais de uma variável, realizando a geração do arquivo de saída em um tempo aceitável. Os pacotes que foram utilizados para o processamento foram baixados do NPM.

Algumas boas práticas de desenvolvimento foram utilizadas para melhorar o desempenho de algumas consultas. A paginação mostrou-se necessária em alguns casos de uso da área administrativa, logo se obteve como um padrão, mesmo para casos menores.

Mesmo com uma gama de recursos disponibilizados pelo SequelizeJS, para determinadas consultas, fez-se o uso de consultas nativas, por questões de desempenho. Algumas consultas se fizeram necessárias pois, com o auxílio de bibliotecas do PostGIS, não foi possível o uso de algum facilitador para efetuar as consultas de forma automatizada.

5.2 ANÁLISE DOS RESULTADOS

A plataforma passou por algumas medições, que foram constantemente testadas, para verificar a operabilidade, quando a carga de dados fosse maior que a máquina de testes.

Para os testes locais, o seguinte equipamento foi utilizado: MacBook Air, sistema operacional macOS Sierra versão 10.12.6, processador 1.6 GHZ Intel Core i5 e equipado com 8 GB DDR3¹⁶ com velocidade de 1600 MHz e um disco de estado sólido com Interface PCI Express de 256 GB de capacidade e a configuração da máquina do CPTEC/INPE utilizada foi: processador Intel 32 cores, equipado com 64 GB de memória com um *storage*¹⁷ de 66 TB de disco, conforme informações repassadas pela equipe de TI do CPTEC/INPE.

O banco de dados de testes local, possuía apenas 2 (duas) variáveis inseridas, com um tamanho total de 18 (dezoito) GB, com dados do ano de 2006 a 2009, diferente da base utilizada no CPTEC, a qual continha todas as variáveis possíveis. Para cada caso de teste foram realizadas cinco execuções do programa, onde as médias dos tempos de execução foram consideradas minimizando, assim, influências externas que pudessem interferir em resultados individuais. Os filtros utilizados para os testes foram:

- a) **Cenário Climático:** ETA MIROC5 RCP4.5 20KM;
- b) **Frequência:** DIÁRIA;
- c) **Localização (por cidade):** Marau-RS;
- d) **Formato de arquivos de dados:** CSV;
- e) **Variável:** TP2M Temperatura 2 metros da superfície;

O processo anterior acontecia da seguinte forma: o usuário informava as coordenadas geográficas (latitude e longitude), de uma área retangular ou um ponto isolado, os quais abrangem uma região de filtragem do conjunto de dados. Todo o processo de filtragem dos dados era feito utilizando scripts em Fortran 90, elaborados especificamente para cada requisição. Para agilizar o processo, era criado um arquivo de configuração com padrões pré-estabelecidos, onde são escritas as variáveis, área abrangida ou pontos isolados da requisição. O script em Fortran fazia a leitura da configuração, e gerava um arquivo texto com o resultado, que era disponibilizado para o usuário.

De conhecimento do processo anterior, verifica-se que em ambos os testes o tempo foi aceitável, visto a quantidade de dados armazenadas em banco. O tempo das requisições está dentro do esperado, uma vez que quando o processo era feito de forma manual, e o mesmo podia levar mais de um dia útil. O processamento acontece de forma automática sem intervenção de pessoas internas do CPTEC/INPE.

¹⁶taxa dupla de transferência nível 3 de memória síncrona dinâmica de acesso aleatório

¹⁷Dispositivos projetados especificamente para armazenamento de dados, onde através de uma conexão via rede, é possível conectar servidores à um storage, facilitando assim a expansão da capacidade de armazenamento sem impacto na produção, garantindo maior flexibilidade e confiabilidade no armazenamento.

A tabela dos valores e o gráfico estão apresentados na Figuras 21 e 22. A Figura 23 mostra o arquivo gerado no formato do arquivo dos dados CSV.

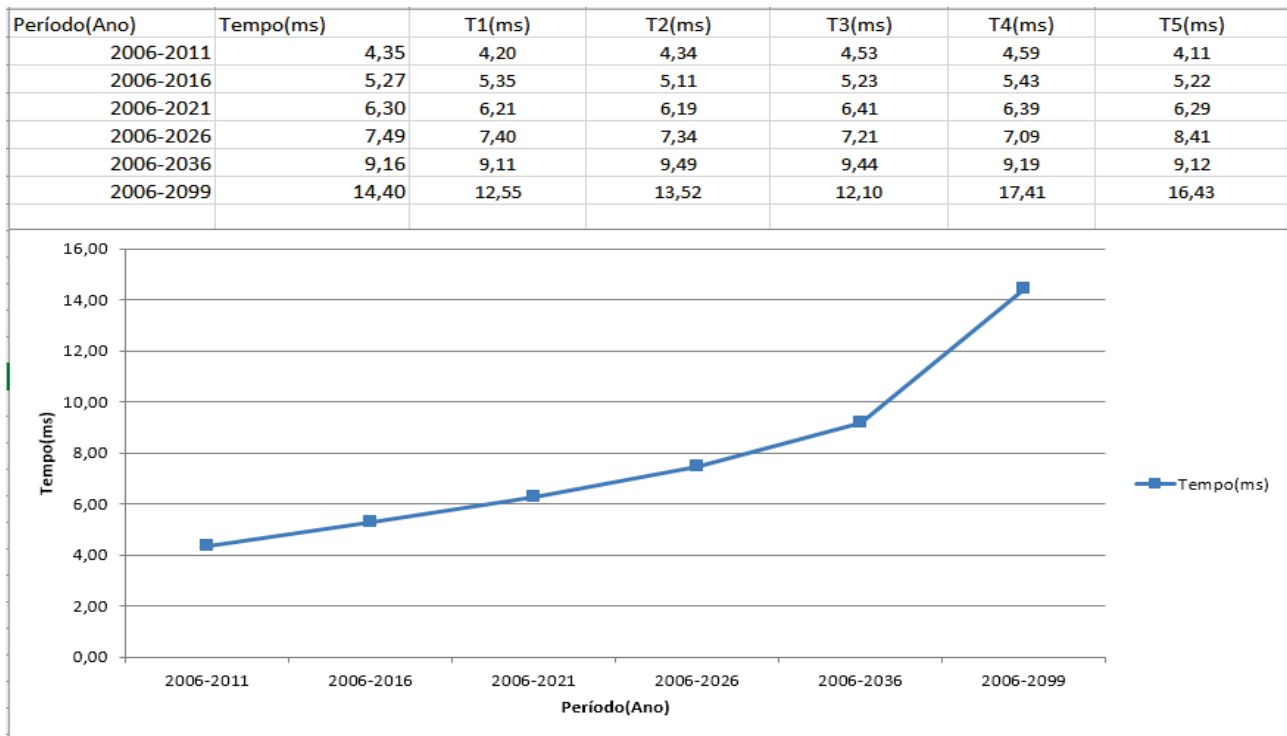


Figura 21. Tabela e gráfico dos testes das requisições - Máquina do CPTEC/INPE.

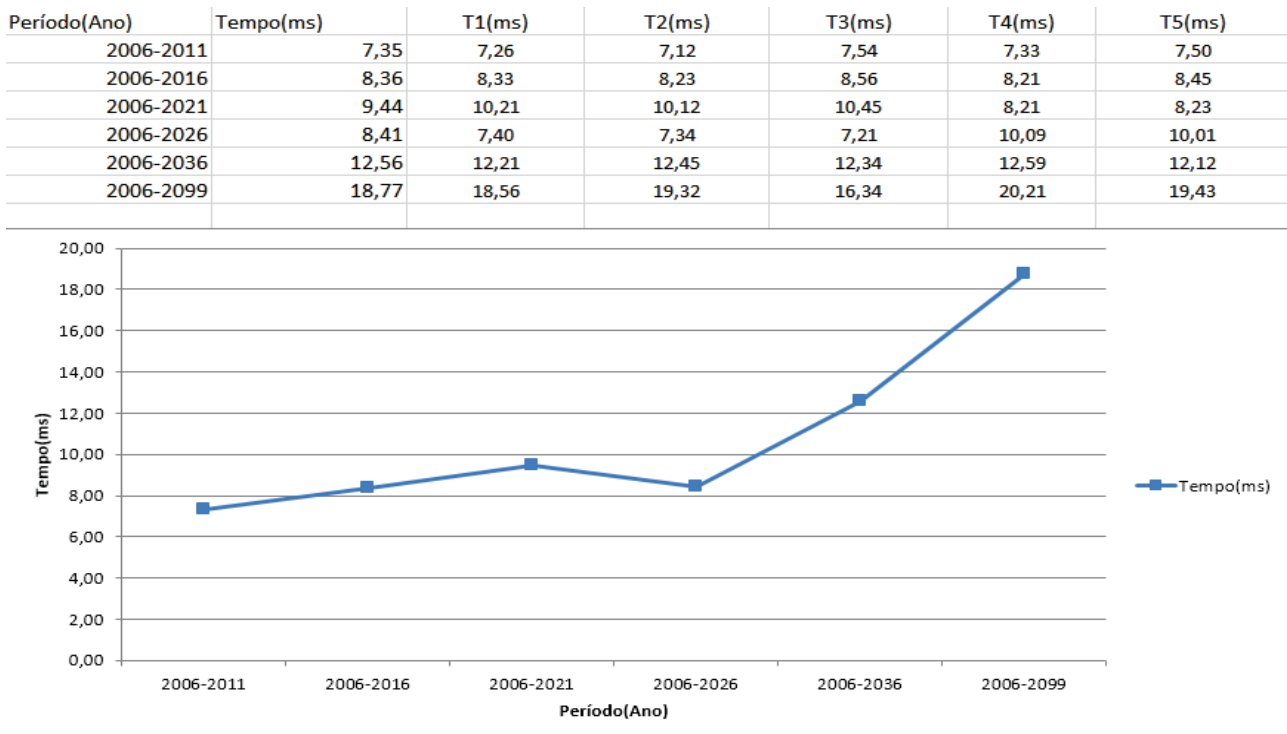


Figura 22. Tabela e gráfico dos testes das requisições - Máquina local.


```

"Valor";"Data";"Hora";"Variavel";"Latitude";"Longitude"
17.7917785644531;"2004-01-01";"00:00:00";"TP2M";"-28.37";"-52.175"
19.0158081054688;"2004-01-02";"00:00:00";"TP2M";"-28.37";"-52.175"
18.7553405761719;"2004-01-03";"00:00:00";"TP2M";"-28.37";"-52.175"
17.1433410644531;"2004-01-04";"00:00:00";"TP2M";"-28.37";"-52.175"
18.3220825195312;"2004-01-05";"00:00:00";"TP2M";"-28.37";"-52.175"
19.1419677734375;"2004-01-06";"00:00:00";"TP2M";"-28.37";"-52.175"
17.9074401855469;"2004-01-07";"00:00:00";"TP2M";"-28.37";"-52.175"
20.0035705566406;"2004-01-08";"00:00:00";"TP2M";"-28.37";"-52.175"
19.093017578125;"2004-01-09";"00:00:00";"TP2M";"-28.37";"-52.175"
19.1392211914062;"2004-01-10";"00:00:00";"TP2M";"-28.37";"-52.175"
19.3292236328125;"2004-01-11";"00:00:00";"TP2M";"-28.37";"-52.175"
16.5270385742188;"2004-01-12";"00:00:00";"TP2M";"-28.37";"-52.175"
16.3938598632812;"2004-01-13";"00:00:00";"TP2M";"-28.37";"-52.175"
18.6234741210938;"2004-01-14";"00:00:00";"TP2M";"-28.37";"-52.175"
18.4501647949219;"2004-01-15";"00:00:00";"TP2M";"-28.37";"-52.175"
18.8007507324219;"2004-01-16";"00:00:00";"TP2M";"-28.37";"-52.175"
20.2192993164062;"2004-01-17";"00:00:00";"TP2M";"-28.37";"-52.175"
22.0781555175781;"2004-01-18";"00:00:00";"TP2M";"-28.37";"-52.175"
18.5213623046875;"2004-01-19";"00:00:00";"TP2M";"-28.37";"-52.175"
19.3392028808594;"2004-01-20";"00:00:00";"TP2M";"-28.37";"-52.175"
20.1052856445312;"2004-01-21";"00:00:00";"TP2M";"-28.37";"-52.175"

```

Figura 23. Arquivo CSV gerado pela plataforma.

5.3 USABILIDADE

Nos dias atuais é importante se preocupar com a usabilidade de um software, *website* entre outros. Deve-se levar em consideração alguns pontos: capacidade de o usuário realizar tarefas básicas, rapidez na execução, relatar aos usuários sobre validações e campos obrigatórios, erros que o usuário pode cometer e, não menos importante, apresentar um visual agradável [71].

A plataforma foi pensada de uma maneira que o usuário se sinta seguro das atividades executadas e de acordo com as necessidades relatadas pelo pessoal do CPTEC/INPE. Foi adquirido um tema HTML para facilitar na modernização do visual. O tema ajudou muito no desenvolvimento pois, além de possuir um visual moderno, o mesmo continha vários componentes reutilizáveis, os quais se encaixaram perfeitamente para alguns casos de uso. O modelo utilizado na requisição, em formato de wizard, foi feito com base nos exemplos disponibilizados pelo mesmo, por exemplo.

5.4 INSTALAÇÃO NO CPTEC/INPE

O projeto ficou dividido em duas etapas, na primeira o banco de dados roda apenas em uma rede interna, por questões de segurança. A API acessa esta base de dados e pode efetuar as consultas necessárias. O banco precisou de algumas alterações no local de armazenamento, devido ao grande volume de dados que foi incluído.

A API disponibilizada possui os filtros necessários bem como a documentação da mesma para uso. Esta funcionalidade serve de auxílio para gerar a URL válida. De posse desta URL, é

possível integrar com softwares existentes. Na Figura 24 é apresentado como é realizada a montagem da URL.

Cenário Climático	Frequência	Mês Inicial	Ano Inicial	Mês Final	Ano Final
20 km, RCP4.5, continental, MIROC5.	ANUAL	Janeiro	2006	Dezembro	2099
Variável	Latitude	Longitude			
PSLM - Pressão ao Nível Médio do Mar	-12	-49	Gerar URL		

<https://projeta.cptec.inpe.br/api/v1/public/ETA/1/YEARLY/4/1/2006/12/2099/PSLM/-12/-49/>

- Cenário climático

Exemplo:
<https://projeta.cptec.inpe.br/api/v1/public/ETA/1/DAILY/2/1/2005/12/2005/TP2M/-12/-49/>
 Atributos: **model** e **id**

Link para os cenários climáticos
<https://projeta.cptec.inpe.br/api/v1/public/models>

```
{
  "data": [
    {
      "id": "1",
      "name": "20 km, RCP4.5, continental, MIROC5.",
      "model": "ETA",
      "couple": "MIROC5",
      "scenario": "RCP4.5",
      "resolution": "20",
      "start_year": "2006",
      "end_year": "2099",
      "lat": "-12",
      "lon": "-49"
    }
  ]
}
```

+ Frequência

+ Período

Figura 24. Montagem da URL que será disponibilizada através de layout apresentado, conforme foi desenvolvido na API.

A equipe da AdaptaClima, projeto fruto da parceria do Ministério do Meio Ambiente, Fundo Newton operado pelo Conselho Britânico, Centro de Estudos em Sustentabilidade da FGV-EAESP (GVces) e o International Institute for Environment and Development (IIED), iniciou a utilização da API no mês de outubro de 2017. O seu objetivo é sistematizar e disponibilizar informações em áreas temáticas prioritárias relacionadas à Adaptação à Mudança do Clima, promover a comunicação e a articulação entre quem produz e quem utiliza o conhecimento e estimular a produção de conhecimento em relação às lacunas identificadas [72].

Devido a estrutura do TI do CPTEC/INPE, fomos incumbidos da tarefa de utilizar o SVN como ferramenta de versionamento da plataforma, tanto na parte HTML, quanto na parte de servidor. Assim que testado pela equipe de desenvolvimento e demais colaboradores, os códigos fontes foram enviados para uma base. O CPTEC nos forneceu usuário e senha para que pudéssemos enviar os arquivos. A estrutura do projeto no SVN foi estruturada da seguinte forma: foram criadas duas pastas, API e VIEW, para que pudesse ser enviada os códigos separadamente. A instalação está sendo feita de forma automática, separadas por um intervalo de tempo, definido pela equipe de TI do CPTEC. Quando se submete o código, o *deploy* vai verificar se tem alterações no código, e fará de forma automática a atualização, executando os comandos necessários, caso necessite de alguma biblioteca nova gerenciada pelo NPM, principalmente do lado da API.

Devido algumas políticas internas do CPTEC, no início teve-se algumas restrições para rodar a plataforma mas, com auxílio sempre prestativo da equipe do CPTEC, estas restrições foram sanadas no transcorrer do projeto.

Foi de fundamental importância os testes realizados desde o início do desenvolvimento da plataforma. Com isso, foi possível testar várias hipóteses de arquiteturas, até a decisão de optar pela que foi descrita no capítulo 4.

6. CONCLUSÕES E TRABALHOS FUTUROS

O CPTEC Toolkit, por meio da plataforma PROJETA, está disponível no endereço <https://projeta.cptec.inpe.br>, onde é possível realizar as requisições aos dados dos modelos de mudanças climáticas, de acordo com os parâmetros necessários para o usuário. Também se encontra no link a descrição para a utilização da API, onde é apresentado um layout com a documentação necessária para fazer a integração com websites e ferramentas de terceiros.

A arquitetura utilizada no desenvolvimento da plataforma possibilitou a identificação de alguns pontos positivos durante o desenvolvimento, como a oportunidade de escalar a utilização dos dados disponibilizados pelo CPTEC/INPE através da API. A estrutura permite, ainda, que sejam incluídos novos módulos e que sejam adaptados novos cenários existentes de maneira fácil e de forma organizada.

O processo de requisição, de disponibilização e de visualização dos dados de mudanças climáticas foi automatizado. Desta forma, o usuário pode fazer uso da interface gráfica e filtrar conforme sua necessidade. Após a conclusão da requisição, em poucos minutos, ele irá receber um e-mail com os passos para download dos dados solicitados. Algumas configurações internas foram feitas para que o usuário solicite uma quantidade muito grande de dados, principalmente para frequência de 3 em 3 horas. Para providenciar esta restrição, limitações foram definidas em relação a quantidade máxima de variáveis e o intervalo de período que pode ser solicitado em uma mesma requisição, de acordo com o tipo de frequência realizada.

O uso da API permitiu não só a integração da necessidade do CPTEC/INPE, como também possibilitou o uso externo da mesma, para que outros órgãos fizessem o uso, conforme especificações internas do CPTEC. A comunicação é feita por URL, ou seja, qualquer linguagem de programação que dê suporte ao HTTP, poderá integrar de maneira fácil via o uso desta API.

Como trabalhos futuros, sugere-se uma exploração no desenvolvimento de gráficos sobre os dados dos modelos climáticos. A disponibilização de recursos de visualização e a manipulação de dados em tela seriam muito importantes para os usuários. Este tipo de funcionalidade, no estilo *dashboard*, agregaria muito na plataforma. Este trabalho está em desenvolvimento no grupo de pesquisa Mosaico da UPF pelo aluno do curso de Ciência da Computação, Guilherme Tocchetto. Aspectos a respeito da visualização destes dados já estão sendo tratados com o CPTEC/INPE e serão trabalhados em novos projetos já a partir de 2018.

A disponibilização de um pacote R para acesso aos dados contidos na base de dados do projeto no CPTEC é outra funcionalidade a ser disponibilizada pelo toolkit. Esta integração possibilitaria aos usuários da linguagem R acessar de forma fácil e eficiente estes dados e utilizá-los em suas aplicações. Este pacote está sendo desenvolvido pela aluna do curso de Ciência da Computação, Marina Dezordi Lopes.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] HEY, T.; TREFETHEN, A. E. Cyberinfrastructure for e-science. *Science*, American Association for the Advancement of Science, v. 308, n. 5723, p. 817–821, 2005. ISSN 0036-8075. Disponível em: <<http://science.sciencemag.org/content/308/5723/817>>.
- [2] WANG, S. et al. Cybergis software: a synthetic review and integration roadmap. *International Journal of Geographical Information Science*, v. 27, p. 2122–2145, 2013. Disponível em: <<http://dx.doi.org/10.1080/13658816.2013.776049>>.
- [3] YUAN, J. et al. A linked data approach for geospatial data provenance. *IEEE Transactions on Geoscience and Remote Sensing*, v. 51, n. 11, p. 5105–5112, Nov 2013. ISSN 0196-2892.
- [4] BRÖRING, A. et al. New generation sensor web enablement. *Sensors*, v. 11, n. 3, p. 2652, 2011. ISSN 1424-8220. Disponível em: <<http://www.mdpi.com/1424-8220/11/3/2652>>.
- [5] MANYIKA, J. et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. [S.l.], 2011.
- [6] WIGAN, M. R.; CLARKE, R. Big data's big unintended consequences. *Computer*, IEEE Computer Society, Los Alamitos, CA, USA, v. 46, n. 6, p. 46–53, 2013. ISSN 0018-9162.
- [7] BLACK, T. L. The new nmc mesoscale eta model: Description and forecast examples. *Weather and Forecasting*, v. 9, n. 2, p. 265–278, 1994. Disponível em: <[http://dx.doi.org/10.1175/1520-0434\(1994\)009<0265:TNNMEM>2.0.CO;2](http://dx.doi.org/10.1175/1520-0434(1994)009<0265:TNNMEM>2.0.CO;2)>.
- [8] CHOU, S.-C. *Modelo Regional ETA*. 1999. Disponível em: <<http://climanalise.cptec.inpe.br/rclimanl/boletim/cliente>>. Acesso em: Oct. 10, 2016.
- [9] YIN, S.; KAYNAK, O. Big Data for Modern Industry: Challenges and Trends. *Proceedings of the IEEE*, v. 103, n. 2, p. 143–146, 2015. ISSN 00189219.
- [10] YUAN, D. et al. A highly practical approach toward achieving minimum data sets storage cost in the cloud. *IEEE Transactions on Parallel and Distributed Systems*, v. 24, n. 6, p. 1234–1244, June 2013. ISSN 1045-9219.
- [11] ZHANG, X. et al. A privacy leakage upper bound constraint-based approach for cost-effective privacy preserving of intermediate data sets in cloud. *IEEE Transactions on Parallel and Distributed Systems*, v. 24, n. 6, p. 1192–1202, June 2013. ISSN 1045-9219.
- [12] WANG, L.; KHAN, S. U. Review of performance metrics for green data centers: a taxonomy study. *The Journal of Supercomputing*, v. 63, n. 3, p. 639–656, 2013. ISSN 1573-0484. Disponível em: <<http://dx.doi.org/10.1007/s11227-011-0704-3>>.

- [13] LIU, X. et al. Do we need to handle every temporal violation in scientific workflow systems? *ACM Trans. Softw. Eng. Methodol.*, ACM, New York, NY, USA, v. 23, n. 1, p. 5:1–5:34, fev. 2014. ISSN 1049-331X. Disponível em: <<http://doi.acm.org/10.1145/2559938>>.
- [14] WANG, L. et al. Towards enabling cyberinfrastructure as a service in clouds. *Computers and Electrical Engineering*, v. 39, n. 1, p. 3 – 14, 2013. ISSN 0045-7906. Special issue on Recent Advanced Technologies and Theories for Grid and Cloud Computing and Bio-engineering. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0045790612000882>>.
- [15] ZHANG, X. et al. A scalable two-phase top-down specialization approach for data anonymization using mapreduce on cloud. *IEEE Transactions on Parallel and Distributed Systems*, v. 25, n. 2, p. 363–373, Feb 2014. ISSN 1045-9219.
- [16] WANG, L. et al. G-hadoop: Mapreduce across distributed data centers for data-intensive computing. *Future Generation Computer Systems*, v. 29, n. 3, p. 739 – 750, 2013. ISSN 0167-739X. Special Section: Recent Developments in High Performance Computing and Security. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167739X12001744>>.
- [17] LIM, E.-P.; CHEN, H.; CHEN, G. Business intelligence and analytics: Research directions. *ACM Trans. Manage. Inf. Syst.*, ACM, New York, NY, USA, v. 3, n. 4, p. 17:1–17:10, jan. 2013. ISSN 2158-656X. Disponível em: <<http://doi.acm.org/10.1145/2407740.2407741>>.
- [18] SOFTWARE., E. *Eclipse Blog*. Disponível em: <<http://blog.eclipse.com.br/blog/2015/9/8/a-internet-das-coisas-e-seu-papel-no-aumento-da-productividade-da-industria>>. Acesso em: Jun. 29, 2016.
- [19] AMAN, S.; CHELMIS, C.; PRASANNA, V. Addressing data veracity in big data applications. *Proceedings - 2014 IEEE International Conference on Big Data, IEEE Big Data 2014*, p. 1–3, 2015.
- [20] ISHWARAPPA; ANURADHA, J. A brief introduction on big data 5vs characteristics and hadoop technology. *Procedia Computer Science*, v. 48, p. 319 – 324, 2015. ISSN 1877-0509. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1877050915006973>>.
- [21] APACHE. *Apache Spark*. Disponível em: <<http://spark.apache.org/>>. Acesso em: Jun. 29, 2016.
- [22] APACHE. *Apache Hadoop*. Disponível em: <<http://hadoop.apache.org/>>. Acesso em: Jun. 29, 2016.
- [23] APACHE. *Apache Cassandra*. Disponível em: <<http://cassandra.apache.org/>>. Acesso em: Jun. 29, 2016.
- [24] INC., M. *MongoDB*. Disponível em: <<https://www.mongodb.com>>. Acesso em: Jun. 29, 2016.
- [25] R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2016. Disponível em: <<https://www.R-project.org/>>.

- [26] PRAJAPATI, V. *Big Data Analytics with R and Hadoop*. [S.l.]: Packt Publishing, 2013. ISBN 178216328X, 9781782163282.
- [27] Google. *AngularJS: Superheroic JavaScript MVW Framework*. Vienna, Austria, 2010. Disponível em: <<https://angularjs.org>>.
- [28] FOUNDATION, N. *Node.js*. 2015. Disponível em: <<https://nodejs.org/en/>>. Acesso em: Oct. 31, 2016.
- [29] STONEBRAKER, M. *PostgreSQL*. 1996. Disponível em: <<https://www.postgresql.org/>>. Acesso em: Oct. 31, 2016.
- [30] RAMSEY, P. *PostGIS*. 2001. Disponível em: <<http://postgis.net/>>. Acesso em: Oct. 31, 2016.
- [31] RUI-SHENG, J. et al. Research on distributed GIS process modeling and integration. *Proceedings of 2008 IEEE International Symposium on IT in Medicine and Education, ITME 2008*, p. 33–38, 2008.
- [32] SRIHARAN, S.; EVERITT, J. H.; FLETCHER, R. Geographic information system (gis) and remote sensing (rs): Undergraduate academic curriculum and precollege training program. In: *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*. [S.l.: s.n.], 2008. v. 3, p. III – 1421–III – 1424. ISSN 2153-6996.
- [33] KNOX, J. et al. Climate change impacts on crop productivity in africa and south asia. *Environmental Research Letters*, v. 7, n. 3, p. 034032, 2012. Disponível em: <<http://stacks.iop.org/1748-9326/7/i=3/a=034032>>.
- [34] FORTES, P.; PLATONOV, A.; PEREIRA, L. Gisareg—a {GIS} based irrigation scheduling simulation model to support improved water use. *Agricultural Water Management*, v. 77, n. 1–3, p. 159 – 179, 2005. ISSN 0378-3774. Special Issue on Land and Water Use: Environmental Management Tools and Practices. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378377405000806>>.
- [35] JAGTAP, S. S.; JONES, J. W. Adaptation and evaluation of the cropgro-soybean model to predict regional yield and production. *Agriculture, Ecosystems and Environment*, v. 93, n. 1–3, p. 73 – 85, 2002. ISSN 0167-8809. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167880901003589>>.
- [36] D, B. *Web Services Architecture*. 2004. Disponível em: <<http://www.w3.org/TR/ws-arch/>>. Acesso em: Jun. 10, 2016.
- [37] YANG, C.; XU, Y.; NEBERT, D. Redefining the possibility of digital earth and geosciences with spatial cloud computing. *International Journal of Digital Earth*, v. 6, n. 4, p. 297–312, 2013. Disponível em: <<http://dx.doi.org/10.1080/17538947.2013.769783>>.

- [38] IBM; ZIKOPOULOS, P.; EATON, C. *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. 1st. ed. [S.l.]: McGraw-Hill Osborne Media, 2011. ISBN 0071790535, 9780071790536.
- [39] INSTITUTE, M. G. et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, 2011. ISBN 9780983179696. Disponível em: <<https://books.google.com.br/books?id=vN1CYAAACAAJ>>.
- [40] LANIAK, G. F. et al. Integrated environmental modeling a vision and roadmap for the future. *Environmental Modelling Software*, v. 39, p. 3 – 23, 2013. ISSN 1364-8152. Thematic Issue on the Future of Integrated Modeling Science and Technology. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1364815212002381>>.
- [41] BASTIN, L. et al. Managing uncertainty in integrated environmental modelling the uncertweb framework. *Environmental Modelling Software*, v. 39, p. 116 – 134, 2013. ISSN 1364-8152. Thematic Issue on the Future of Integrated Modeling Science and Technology. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1364815212000564>>.
- [42] NATIVI, S.; MAZZETTI, P.; GELLER, G. N. Environmental model access and interoperability the {GEO} model web initiative. *Environmental Modelling Software*, v. 39, p. 214 – 228, 2013. ISSN 1364-8152. Thematic Issue on the Future of Integrated Modeling Science and Technology. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1364815212000898>>.
- [43] EVANGELIDIS, K. et al. Geospatial services in the cloud. *Computers Geosciences*, v. 63, p. 116 – 122, 2014. ISSN 0098-3004. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0098300413002719>>.
- [44] YUE, P. et al. Geoprocessing in cloud computing platforms – a comparative analysis. *International Journal of Digital Earth*, v. 6, n. 4, p. 404–425, 2013. Disponível em: <<http://dx.doi.org/10.1080/17538947.2012.748847>>.
- [45] BRÖRING, A. et al. New generation sensor web enablement. *Sensors*, v. 11, n. 3, p. 2652, 2011. ISSN 1424-8220. Disponível em: <<http://www.mdpi.com/1424-8220/11/3/2652>>.
- [46] ESRI. *Esri*. Disponível em: <<http://www.esri.com/news/arcwatch/0312/gis-in-education-the-web-and-beyond.html>>. Acesso em: Jun. 20, 2016.
- [47] HENDLER, J. Science and the semantic web. *Science*, American Association for the Advancement of Science, v. 299, n. 5606, p. 520–521, 2003. ISSN 0036-8075. Disponível em: <<http://science.sciencemag.org/content/299/5606/520>>.
- [48] EGENHOFER, M. J. Toward the semantic geospatial web. In: *Proceedings of the 10th ACM International Symposium on Advances in Geographic Information Systems*. New York, NY, USA: ACM, 2002. (GIS '02), p. 1–4. ISBN 1-58113-591-2. Disponível em: <<http://doi.acm.org/10.1145/585147.585148>>.

- [49] ROPENSCI. *R interface to many NOAA data APIs*. 2016. Disponível em: <<https://github.com/ropensci/rnoaa>>. Acesso em: May. 27, 2016.
- [50] NOAA. *Request Web Services Token*. 2016. Disponível em: <<http://www.ncdc.noaa.gov/cdo-web/token>>. Acesso em: May. 27, 2016.
- [51] CHAMBERLAIN NICHOLAS POTTER, K. R. S.; EDMUND, H. *rnoaa: NOAA climate data from R. R package version 0.5.2*. 2016. Disponível em: <<https://github.com/ropensci/rnoaa>>. Acesso em: May. 27, 2016.
- [52] CHAMBERLAIN, S. *rnoaa: 'NOAA' Weather Data from R*. [S.l.], 2016. R package version 0.5.6. Disponível em: <<http://CRAN.R-project.org/package=rnoaa>>.
- [53] SEERS, B.; SHEARS, N. *New Zealand's Climate Data in R — An Introduction to clifro*. Auckland, New Zealand, 2015. Disponível em: <<http://stattech.wordpress.fos.auckland.ac.nz/2015-02-new-zealands-climate-data-in-r-an-introduction-to-clifro/>>.
- [54] NOCKE, T.; SCHUMANN, H.; BÖHM, U. Methods for the visualization of clustered climate data. *Computational Statistics*, v. 19, n. 1, p. 75–94, 2004. ISSN 1613-9658. Disponível em: <<http://dx.doi.org/10.1007/BF02915277>>.
- [55] MICHNA, P.; WOODS with contributions from M. *RNetCDF: Interface to NetCDF Datasets*. [S.l.], 2016. R package version 1.8-2. Disponível em: <<http://CRAN.R-project.org/package=RNetCDF>>.
- [56] BECKER, O. S. code by R. A.; MINKA, A. R. W. R. version by Ray Brownrigg. Enhancements by T. P.; DECKMYN., A. *maps: Draw Geographical Maps*. [S.l.], 2016. R package version 3.1.0. Disponível em: <<https://CRAN.R-project.org/package=maps>>.
- [57] BECKER, O. S. code by R. A.; BROWNRIGG., A. R. W. R. version by R. *mapdata: Extra Map Databases*. [S.l.], 2016. R package version 2.2-6. Disponível em: <<https://CRAN.R-project.org/package=mapdata>>.
- [58] BROWNRIGG, D. M. P. for R by R.; MINKA, T. P.; BIVAND. transition to Plan 9 codebase by R. *mapproj: Map Projections*. [S.l.], 2015. R package version 1.2-4. Disponível em: <<https://CRAN.R-project.org/package=mapproj>>.
- [59] LABORATORY, N. R. A. *verification: Weather Forecast Verification Utilities*. [S.l.], 2015. R package version 1.42. Disponível em: <<https://CRAN.R-project.org/package=verification>>.
- [60] FRALEY, C. et al. *ensembleBMA: Probabilistic Forecasting using Ensembles and Bayesian Model Averaging*. [S.l.], 2015. R package version 5.1.2. Disponível em: <<https://CRAN.R-project.org/package=ensembleBMA>>.
- [61] CHOU, S. C.; BUSTAMANTE, J. F.; GOMES, J. L. Evaluation of Eta model seasonal precipitation forecasts over south america, nonlin. *Processes Geophys.*, p. 537–555, 2005.

- [62] PESQUERO, J. F. et al. Climate downscaling over South America for 1961- 1970 using the Eta model. *Theoretical and Applied Climatology*, p. 75–93, 2010.
- [63] CHOU, S. C. et al. Downscaling of South America present climate driven by 4-member HadCM3 runs. *Climate Dynamics*, p. 635–653, 2012.
- [64] MARENGO, J. A. et al. Development of regional future climate change scenarios in south america using the Eta CPTEC/HadCM3 climate change projections: climatology and regional analyses for the amazon, São Francisco and the Paraná river basins. *Climate Dynamics*, p. 1829–1848, 2012.
- [65] IPCC. *Intergovernmental Panel on Climate Change*. 2013. Disponível em <http://www.ipcc.ch/>. Acesso em: 10 Fev. 2017. Disponível em: [<http://www.ipcc.ch/>](http://www.ipcc.ch/).
- [66] CHOU, S. C. et al. Assessment of climate change over south america under RCP 4.5 and 8.5 downscaling scenarios. *American Journal of Climate Change*, p. 512–527, 2014.
- [67] Andy Chung. *RequireJS*. [S.l.], 2011. Disponível em: <http://requirejs.org>.
- [68] Olivier Combe. *ocLazyLoad*. [S.l.], 2013. Disponível em: <https://oclazyload.readme.io>.
- [69] Sequelize. *SequelizeJS*. [S.l.], 2010. Disponível em: <https://sequelize.readthedocs.io/en/v3/>.
- [70] NETO, W. *Construindo APIs testáveis com Node.js*. 2017. Disponível em <https://leanpub.com/construindo-apis-testaveis-com-nodejs/>. Acesso em: 02 Out. 2017. Disponível em: <https://leanpub.com/construindo-apis-testaveis-com-nodejs>.
- [71] GROUP, N. N. *Evidence-Based User Experience Research, Training, and Consulting*. 1998. Disponível em <https://www.nngroup.com/>. Acesso em: 30 Set. 2017. Disponível em: <https://www.nngroup.com/>.
- [72] EAESP, F. *Centro de estudos em sustentabilidade*. 2014. Disponível em <http://www.gvces.com.br/adaptaclima-objetivos>. Acesso em: 25 Out. 2017. Disponível em: <http://www.gvces.com.br/adaptaclima-objetivos>.