



Wellington Luiz Antonio

Serviço Computacional para Interpolação Espacial de Dados Meteorológicos

Recife

2019

Wellington Luiz Antonio

Serviço Computacional para Interpolação Espacial de Dados Meteorológicos

Monografia apresentada ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

Universidade Federal Rural de Pernambuco – UFRPE

Departamento de Estatística e Informática

Curso de Bacharelado em Sistemas de Informação

Orientador: Glauco Estácio Gonçalves

Coorientador: Victor Wanderley Costa de Medeiros

Recife

2019

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

- A635s Antonio, Wellington Luiz
 Serviço Computacional para Interpolação Espacial de Dados Meteorológicos / Wellington Luiz Antonio. -
 2019.
 48 f. : il.
- Orientador: Glauco Estácio Gon Gonçalves.
 Coorientador: Victor Wanderley Costa de Medeiros.
 Inclui referências.
- Trabalho de Conclusão de Curso (Graduação) - Universidade Federal Rural de Pernambuco,
 Bacharelado em Sistemas da Informação, Recife, 2019.
1. Variáveis Ambientais. 2. Aprendizagem de Máquina. 3. Software como Serviço. 4. Interpolação
 Espacial. I. Gonçalves, Glauco Estácio Gon, orient. II. Medeiros, Victor Wanderley Costa de, coorient. III.
 Título

Wellington Luiz Antonio

Serviço Computacional para Interpolação Espacial de Dados Meteorológicos

Monografia apresentada ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Sistemas de Informação.

Trabalho aprovado. Recife, 11 de Dezembro de 2019:

BANCA EXAMINADORA

Prof. Dr. Glauco Estácio Gonçalves

Departamento de Estatística e Informática
(DEINFO)/UFRPE
Orientador

**Prof. Dr. Victor Wanderley Costa de
Medeiros**

Departamento de Estatística e Informática
(DEINFO)/UFRPE
Convidado 1

Prof. Dr. Thieres George Freire da Silva

Programa de Pós-Graduação em
Produção Vegetal (PGPV)/UFRPE
Convidado 2

Recife
2019

A Jesus meu grande amigo, à Aline Roberta minha amada e ajudadora que tanto contribuiu para realização deste trabalho. Dedico também a minha família, aos que fazem parte da igreja assembleia de Deus em Angélicas – Vicência – PE, aos integrantes do Laboratório de Pesquisa em Infraestrutura Computacional - Juá Labs – UFRPE e do Núcleo de Tecnologia da Informação – NTI – UFPE, aos que compõem o PET políticas públicas e do Ciranda da Ciência da UFRPE. Além, é claro, de todos os meus amigos e conhecidos.

Agradecimentos

Agradeço ao Deus trino, que sempre me protegeu e me indicou os melhores caminhos a serem seguidos para chegar até onde estou, eu, um aluno de escola pública que viveu toda sua infância, adolescência e parte da juventude criando boi para sustento da família, nunca poderia imaginar que um dia estaria concluindo um curso superior em tecnologia da informação, algo que desde criança me fascina. Foi esse Deus maravilhoso que me concedeu esse sonho.

Agradeço aos meus queridos e estimados orientadores por serem exemplos a serem seguidos tanto como professores quanto pesquisadores. Agradeço-vos pelas orientações que me foram dadas no Juá Labs, desde o início da pesquisa até a conclusão desse trabalho. Essas orientações e ensinamentos me tornaram um profissional melhor e foram fundamentais para que eu conseguisse concluir esse trabalho, sou eternamente grato por tudo que fizeram e estão fazendo.

À minha querida e amada noiva, Aline Roberta, que me ajuda em todo tempo, mesmo com todos os compromissos e disciplinas do seu curso superior, sempre me apoiou nos momentos difíceis do meu curso superior e principalmente no último período. Amo você e sou grato por fazer parte da minha vida.

À minha família que mesmo sem saber exatamente o que faço, me apoiam. Família essa que nunca imaginou que teria um filho que iria fazer um curso superior em uma universidade federal pública, eles à chamam de escola. Amo vocês, sou grato por terem me criado e por me ajudarem nas mais variadas situações dessa vida.

A todos da igreja Assembleia de Deus em Angélicas – Vicência – PE, principalmente ao diácono Jasiel Ribeiro e a toda mocidade que sempre me apoiaram e encorajaram a vim fazer este curso em Recife, sou grato por toda a paciência e orações para comigo.

Aos que fazem parte do Laboratório em Infraestrutura Computacional – Juá labs, especialmente aos meus amigos queridos, José Clodoalves e Diego Bezerra, que sempre estão dispostos a me ajudar, obrigado por revisarem todo o texto deste trabalho. Sou grato por vocês e pelo Juá labs, pois, esse laboratório é de grande relevância quando se trata de aprendizado e geração de conhecimento.

Aos integrantes do Núcleo de Tecnologia da Informação – NTI – UFPE, em especial aos que constituem a coordenação de Qualidade de Software – CQS e a Coordenação de Serviço ao Usuário – CSU, que sempre me apoiaram e me ensinaram na prática como funciona um ambiente de colaboração em prol de um objetivo, sou

agradecido por tudo que aprendi com vocês. Agradeço a Jacinto Filipe e Ana Paula pela grande ajuda na revisão do resumo em inglês, além disso, agradeço a Daniel Ramos pela revisão das figuras dispostas nesse trabalho.

Aos que compõem o PET (Programa de Educação Tutorial) política públicas e Ciranda da Ciência da UFRPE, que me ensinaram a trabalhar em um grupo diversificado com opiniões variadas, mas com em prol de um objetivo em comum. Sou grato a todos, sobretudo aos tutores e orientadores Michael Lee Sundheimer, Alexandro Tenorio e Giselle Nanes e aos meus queridos amigos e PETianos: Lidiane Maria, Rodrigo Caetano, Nomager Nunes e Fernanda Rodrigues, que são de uma importância ímpar para a minha vida.

Aos que constituem a Pró-Reitoria de Gestão Estudantil e Inclusão – PRO-GESTI, agradeço por ter me acolhido quando cheguei nessa universidade sem saber nada sobre residência universitária, sei que não é fácil manter uma residência estudantil com tantas pessoas. Sou grato por todo apoio e atenção que a me foram dados, acredito que sem esses recursos o meu sonho não teria se realizado, pois, sei o quão difícil é uma pessoa sair do interior do estado para vim estudar em Recife, obrigado por tudo mesmo!

As meninas da CEU-F – Casa da Estudante Universitária Feminina da UFPE, em especial a Maria Isabel, Emillie Bianca, Elizamar Mariano e Laissa Paloma, que me aturaram enquanto fazia, junto com Aline Roberta, a revisão deste trabalho. Agradeço por toda paciência e amizade de vocês.

A todos os meus amigos, especialmente a Bruno Santos, Lucas Vieira, Jobson Pereira, Daniel Ramos, Joel Silva e Joelma Maria, que sempre me ajudaram durante esse percurso da minha história na universidade.

Por fim, agradeço aos que participaram e me ajudaram direta ou indiretamente na minha vida acadêmica e pessoal durante todo o decorrer do meu percurso na graduação.

“Por isso mesmo, empenhem-se para acrescentar à sua fé a virtude; à virtude o conhecimento; ao conhecimento o domínio próprio; ao domínio próprio a perseverança; à perseverança a piedade; à piedade a fraternidade; e à fraternidade o amor.”
(2 Pedro 1:5-7)

Resumo

A interpolação espacial é uma técnica de suma importância para diversas atuações: meteorologia, hidrologia, zoneamento agrícola, caracterização de áreas de risco à saúde, sociodemográfica, entre outras. Por meio da interpolação pode-se modelar a superfície de uma variável espacialmente distribuída a partir de um conjunto finito de dados pontuais conhecidos, no caso de dados meteorológicos para agricultura, por exemplo, a interpolação permite observar como as variáveis meteorológicas se comportam em uma determinada propriedade rural, o que poderia servir como base para o manejo de irrigação nesta propriedade. Devido à demanda cada vez maior pelo uso da interpolação espacial, propôs-se desenvolver um serviço de interpolação espacial, escalável baseado em tecnologias e padrões do estado da arte em sistemas distribuídos para interpolação espacial de dados meteorológicos associados à agricultura. Para o alcance de tal objetivo, buscou-se desenvolver um serviço web que implementasse três algoritmos diferentes para a interpolação da evapotranspiração de referência, sendo eles: *Inverse distance weighted* (IDW), *Ordinary Kriging* (OK) e *Random Forest* (RF), os dois primeiros são algoritmos amplamente utilizados na espacialização da evapotranspiração de referência, sendo conhecidos por produzirem baixos erros na interpolação. Já o último algoritmo utilizado é originário da aprendizagem de máquina e tem sido empregado em estudos recentes como alternativa para a interpolação espacial de variáveis ambientais, obtendo também resultados promissores na estimativa da evapotranspiração. O serviço web de interpolação espacial proposto foi implementado e seu desempenho avaliado por meio de medição. Este serviço foi implantado em ambiente de produção por meio de um contêiner *Docker*, e um aplicativo móvel foi desenvolvido para integrar e demonstrar as principais funcionalidades do serviço web. O presente serviço pode ser aplicado em diversas áreas, contudo nesse trabalho atentou-se mais para o setor agrícola, por ser o setor para o qual esse estudo encontra-se direcionado. Os principais beneficiários do serviço web incluem pesquisadores e desenvolvedores de *software* que, por sua vez, a partir da aplicação do serviço, podem desenvolver estudos que beneficiarão o agricultor. Durante a realização desse trabalho, buscou-se também avaliar como o serviço desenvolvido poderia ser útil para a promoção do desempenho e da escalabilidade com relação ao cálculo da interpolação espacial e geração de modelos espaciais. Destacou-se também, a importância desse *software* como ferramenta de apoio para outras pesquisas ou mesmo para outros *software*, como por exemplo, o Aquaprev, que utiliza, além de outros parâmetros, a evapotranspiração e a interpolação espacial para estimar o tempo de irrigação de uma determinada cultura.

Palavras-chave: Variáveis Ambientais, Aprendizagem de Máquina, Software como Serviço, Interpolação Espacial.

Abstract

The spatial interpolation is an essential technique for several fields, such as meteorology, hydrology, agricultural zoning, characterization of health risk areas, sociodemographic, among others. Through interpolation, it is possible to model a surface of a spatially distributed variable from a finite set of known data points. In the case of weather data for agriculture, for instance, interpolation allows us to observe how weather variables behave on a given rural property, which could serve as input for irrigation management on this property. Due to the increasing demand for the use of spatial interpolation, this work proposed the development of a scalable service based on technologies and standards of state-of-the-art in distributed systems, for spatial interpolation of meteorological data associated with agriculture. In order to achieve this goal, we developed a web service based on three different reference evapotranspiration interpolation algorithms, namely: Inverse distance weighted (IDW), Ordinary Kriging (OK) and Random Forest (RF). The first two are widely used algorithms in the spatialization of reference evapotranspiration and they are known to produce low interpolation errors. The third algorithm is originated from Machine Learning. It has been used in recent studies as an alternative for spatial interpolation of environmental variables. This last algorithm has also been obtaining promising results in the estimation of evapotranspiration. The spatial interpolation web service proposed was developed and its performance was evaluated through measurement. This service was deployed on a production environment using Docker container and a mobile application was developed to integrate and show the main functionalities of the web service. The developed service can be applied in several areas. However, in this work more attention was paid to the agricultural sector, as this one is the sector to which this study is focused on. The main beneficiaries of the web service are researchers and developers, which, in turn, are able to develop studies that will benefit the farmer from the application of the service. During this work, we also sought to evaluate how the developed service could be useful for the promotion of the performance and the scalability with respect to spatial interpolation calculus and generation of spatial models. We also highlighted the importance of this software as a support tool for other researches or even for other software, such as Aquaprev, which uses, among other parameters, evapotranspiration and spatial interpolation to estimate the irrigation time of a given crop.

Keywords: Environmental Variables, Machine Learning, Software as a Service, Spatial Interpolation.

Lista de ilustrações

Figura 1 – Exemplo de arquitetura, onde o sistema web de interpolação espacial pode ser inserido.	25
Figura 2 – Fluxograma simplificado da execução do recurso <i>Model</i>	25
Figura 3 – Exemplo de JSON de dados enviados para geração do modelo por meio da URL <i>model</i>	26
Figura 4 – Exemplo de JSON de dados retornado após a geração do modelo. .	28
Figura 5 – Fluxograma simplificado da execução do recurso <i>Interpolation</i> . . .	28
Figura 6 – Exemplo de JSON de dados enviados por meio do recurso <i>interpolation</i>	29
Figura 7 – Exemplo de JSON de dados retornado após a interpolação das coordenadas.	29
Figura 8 – Retorno com a Interpolação Espacial no aplicativo móvel	32
Figura 9 – Tempo de Resposta por Usuários Virtuais	36
Figura 10 – Tempo de Resposta por Algoritmo	37
Figura 11 – Tempo de Resposta por Tipo de Requisição	38
Figura 12 – Tempo de Resposta por Pontos do Modelo	39
Figura 13 – Interação entre os Algoritmos e o tipo de requisição (Mix Modelo/Interpolação) no tempo	41
Figura 14 – Interação entre a quantidade de Pontos do Modelo e a quantidade de Usuários Virtuais no Tempo de Resposta	42
Figura 15 – Interação entre a quantidade de Usuários Virtuais e os Algoritmos no tempo	43

Lista de tabelas

Tabela 1 – Configuração do computador onde foi implantado o serviço web de interpolação espacial de dados meteorológicos	30
Tabela 2 – Fatores utilizados no experimento	33
Tabela 3 – Parâmetros para o plano de teste do Serviço	34
Tabela 4 – Resultado da análise de variância (ANOVA)	40

Lista de abreviaturas e siglas

API	Application Programming Interface
HTTP	Hypertext Transfer Protocol
IDW	Inverse Distance Weighted
JSON	JavaScript Object Notation
OK	Ordinary Kriging
POWER	Prediction of Worldwide Energy Resource
RF	Random Forest
RT	Regression Tree
REST	Representational State Transfer
SOA	Service-Oriented Architecture
URL	Uniform Resource Locator
XML	eXtensible Markup Language

Sumário

	Lista de ilustrações	10
1	INTRODUÇÃO	14
1.1	Motivação	15
1.2	Objetivos	16
1.3	Organização do trabalho	16
2	CONCEITOS E TECNOLOGIAS	17
2.1	Arquitetura Orientada à Serviços - SOA	17
2.2	Virtualização	18
2.3	Interpolação Espacial	19
2.4	Projeto de Experimentos	22
3	SERVIÇO DE INTERPOLAÇÃO ESPACIAL	24
3.1	Implementação do Serviço de Interpolação Espacial	24
3.2	Implantação do Serviço de Interpolação Espacial	30
3.3	Aplicativo Móvel	31
4	EXPERIMENTOS	33
4.1	Método para Avaliação de Desempenho do Serviço de Interpolação Espacial	33
4.2	Resultados da Avaliação de Desempenho do Serviço de Interpolação Espacial	35
4.2.1	Impacto Individual de cada fator	35
4.2.2	Interações entre os fatores	40
5	CONCLUSÃO	44
5.1	Contribuições	45
5.2	Trabalhos Futuros	45
	REFERÊNCIAS	46

1 Introdução

O uso de técnicas de interpolação espacial é de fundamental relevância, pois, permite a construção de modelos estatístico-computacionais para a estimativa do valor de um atributo espacialmente distribuído. A interpolação espacial é, em essência, uma técnica a qual pressupõe que os valores de um atributo tendem a assemelhar-se em locais mais próximos e costumam ser mais distintos quando estão em locais mais distantes entre si (JAKOB; YOUNG, 2016). Tais técnicas são consideradas clássicas para fornecer dados à tomada de decisão espacialmente dependente, como é o caso dos dados meteorológicos que podem estimar, por exemplo, medidas de evapotranspiração obtidas, *in loco*, de uma dada localidade (SILVA, 2006).

Exemplos de aplicabilidade da interpolação espacial na resolução de situações-problema podem ser observados na estimativa espacial de variáveis ambientais (LI; HEAP, 2014; LOPES; MELO; LEAL, 2017), na previsão de precipitação para determinada localidade, em análises sociodemográficas (JAKOB; YOUNG, 2016) e na elaboração de mapas hidrogeológicos paramétricos (VARGAS et al., 2019). No caso da agrometeorologia, por exemplo, a interpolação espacial pode auxiliar no manejo de irrigação fazendo estimativas da evapotranspiração em uma propriedade rural, a partir de dados meteorológicos coletados em regiões próximas. Desta forma, é possível reduzir os custos com o manejo de irrigação que, de outra forma, exigiria uso de sensores específicos, *software* adequado e pessoal técnico especializado para a operação e manutenção de equipamentos. Em síntese, a aplicabilidade da interpolação espacial pode abranger as mais diversas aplicações, desde a meteorologia, hidrologia, sociologia, zoneamento agrícola até a caracterização de áreas de risco à saúde, por exemplo.

Propõe-se com este estudo o desenvolvimento e a implantação de um serviço web de interpolação espacial utilizando-se de três algoritmos para a interpolação da (ET_o), são eles: *Inverse Distance Weighted* (IDW), *Ordinary Kriging* (OK) e *Random Forest* (RF). Sendo os dois primeiros algoritmos amplamente utilizados na interpolação de diversas variáveis ambientais (JAKOB; YOUNG, 2016; LOPES; MELO; LEAL, 2017; VARGAS et al., 2019), conhecidos por produzirem baixos erros de interpolação. Já o último algoritmo, por outro lado, tem mostrado desempenho superior aos métodos clássicos (LI et al., 2011; APPELHANS et al., 2015; JÚNIOR et al., 2019).

Uma das principais vantagens do serviço proposto é a possibilidade de interpolação de uma ampla gama de variáveis tais como temperatura ambiente, umidade relativa do ar, velocidade do vento, etc. Pretendeu-se, com isso, tornar o serviço útil para o desenvolvimento de aplicações em áreas diversas, mas, com foco no setor da

agricultura e meio ambiente.

Esse *software* pode ser utilizado como ferramenta de apoio para outras pesquisas ou mesmo para outros *software*. Desta forma, os principais beneficiários do presente serviço incluem pesquisadores e desenvolvedores de *software* que, por sua vez, a partir da aplicação do presente serviço, podem desenvolver estudos que beneficiarão o agricultor na ponta da cadeia produtiva e até mesmo a agroindústria, proporcionando a estes redução de custos correlacionados ao serviço prestado pelo *software* desenvolvido neste trabalho, a exemplo disso tem-se o Aquaprev, que fará uso da interpolação espacial, dada como retorno pelo *software* desenvolvido, para estimar a quantidade de água necessária para a irrigação de uma determinada cultura.

Como resultados, além da especificação, do desenvolvimento, da avaliação de desempenho e da implantação em ambiente de produção realizados no serviço de interpolação espacial, implementou-se também, como suporte para o uso do serviço web, um aplicativo móvel que possibilitou demonstrar as suas principais funcionalidades. Este aplicativo, integrado ao serviço voltado para o manejo de irrigação, pode ser útil para se determinar o tempo de irrigação em uma determinada plantação, a partir dos dados interpolados de evapotranspiração de referência da área analisada.

1.1 Motivação

Na era da informação, a rapidez, a precisão, a segurança e a estabilidade de serviços computacionais são requisitos essenciais e cada vez mais procurados por organizações em aplicações diversas. Desse modo, a necessidade de se obter em tempo real dados para a solução de problemas é cada vez mais constante.

Estações meteorológicas, por exemplo, captam dados do ambiente de interesse por meio de sensores a fim de possibilitar a posterior aplicação desses em determinados serviços solicitantes. Contudo, tais estações possuem limitação quanto ao seu raio de alcance, o que pode prejudicar a coleta de dados e fazer com que sua atuação fique comprometida. Nesse contexto, a interpolação espacial se apresenta como alternativa para permitir inferir o valor de variáveis ambientais para além do alcance das estações.

Nos últimos anos, há uma forte tendência do uso de sensores para a coleta de dados, no entanto, a aquisição e implantação dos equipamentos necessários é custosa, o que leva a uma dificuldade da instalação massiva destes sensores. Mas, por outro lado, a técnica de interpolação espacial é vantajosa porque permite estimar pontos desejados, a partir das medidas em pontos conhecidos. Entretanto, cada vez mais a variedade e o volume de solicitações a um determinado sistema cresce requerendo assim, que continue a operar de forma eficiente, sem que a estabilidade no atendimento à essas solicitações seja comprometida. Logo, a importância da utilização de

um serviço web escalável se torna necessária.

Dessa forma, por meio dos objetivos da [seção 1.2](#), propõe-se mitigar os problemas referentes a realização da interpolação espacial enfrentados em diversas aplicações por meio do desenvolvimento e implantação de um serviço web de interpolação espacial.

1.2 Objetivos

Objetivou-se, com este trabalho, desenvolver um serviço de interpolação espacial escalável, baseado em tecnologias e padrões do estado da arte em sistemas distribuídos para interpolação espacial de dados meteorológicos associados à agricultura. Logo, para atingir o objetivo principal, buscou-se realizar os seguintes objetivos específicos, que foram:

- Avaliar métodos estatístico-computacionais para interpolação espacial de dados meteorológicos como um serviço;
- Implantar o serviço web de interpolação espacial em um ambiente de produção.
- Integrar o serviço a uma aplicação móvel para interpolação, em tempo real, da evapotranspiração de referência para uma localidade.

1.3 Organização do trabalho

O trabalho está organizado em cinco capítulos, além do capítulo de introdução. O [Capítulo 2](#), expõe as tecnologias e conceitos tomados como alicerce para a implementação do serviço web de interpolação espacial. O [Capítulo 3](#), mostra a proposta de especificação, implementação, implantação do serviço web, bem como, descreve o aplicativo móvel desenvolvido para demonstrar as principais funcionalidades do serviço. O [Capítulo 4](#), detalha os experimentos para avaliação de desempenho do serviço web e discute os resultados obtidos. E, por fim, o [Capítulo 5](#), diz respeito as conclusões obtidas após a análise dos resultados. além de apresentar as contribuições e os trabalhos futuros estimados para este trabalho.

2 Conceitos e Tecnologias

Este capítulo descreve as tecnologias e conceitos tomados como alicerce para a implementação do serviço web de interpolação espacial desenvolvido neste trabalho. Na [seção 2.1](#), apresenta-se os conceitos e tecnologias usados na Arquitetura Orientada à Serviços (SOA), usufruídos como base para a elaboração do serviço web de interpolação proposto neste trabalho. A [seção 2.2](#), mostra as ferramentas e definições acerca da virtualização, a qual serviu como infraestrutura tecnológica para implantar o serviço em produção. Na [seção 2.3](#), são abordados os conceitos da interpolação espacial, os algoritmos convencionais e modernos utilizados para o cálculo da interpolação espacial. Por fim, a [seção 2.4](#), descreve os conceitos e formas de planejamento de experimentos comumente utilizados.

2.1 Arquitetura Orientada à Serviços - SOA

A integração entre sistemas permite que dados possam ser processados entre vários dispositivos com tecnologias distintas. Entretanto, tal integração pode ocasionar problemas se não for bem implementada ([SERMAN, 2010](#)).

Os *Web Services* (Serviços Web) tem como objetivo atender as necessidades encontradas na integração de sistemas, viabilizando a implementação das regras definidas pela SOA (*Service-Oriented Architecture* - Arquitetura Orientada a Serviço). Pois, a SOA implementa o conceito de serviço e é uma arquitetura de *software* que permite desenvolver sistemas que criam e consomem serviços entre si, a partir de interações que tem como princípio fundamental a interoperabilidade dos elementos de *software* por meio de interfaces padronizadas ([KOU BAA, 2019](#)).

Levando em consideração a tendência atual do uso de serviços para integração de sistemas, a utilização da SOA, torna-se essencial para o desenvolvimento de *software* que necessite comunicar-se com outros. Além disso, a SOA também estimula o reuso das funcionalidades, beneficiando-se de serviços previamente criados.

Os *Web Services* seguem a arquitetura orientada à serviços utilizando protocolos e padrões já estabelecidos no funcionamento da *World Wide Web* ([KOU BAA, 2019](#)), como é o caso do XML (*eXtensible Markup Language* - Extensível de Marcação Genérica), do JSON (*JavaScript Object Notation* - Notação de Objetos JavaScript) e do próprio HTTP (*Hypertext Transfer Protocol* - Protocolo de Transferência de Hipertexto). O uso do protocolo HTTP se deve principalmente por ser agnóstico em relação ao dado carregado, permitindo o tráfego de dados codificados em diferentes formatos

e por aderir aos princípios REST (*Representational State Transfer* - Transferência de Estado Representacional), que definem um estilo de arquitetura de *software* para Web Services simples, portáteis, escaláveis e de bom desempenho (MALESHKOVA et al., 2019).

Note-se que o REST não pode ser entendido como um protocolo, mas sim como uma arquitetura de *software* com regras para o desenvolvimento de *Web Services* (ARCURI, 2019). Tal arquitetura vem conquistando espaço e popularidade no desenvolvimento de *software* orientado a serviços (ZHANG et al., 2019; ARCURI, 2019), sendo muito utilizada por grandes empresas de tecnologia, como Google¹, Twitter² e Amazon³.

No desenvolvimento de uma interface REST, os objetos são o foco e cada um é unicamente representado na forma de um recurso e endereçado por meio de uma URL (*Uniform Resource Locator*). Cada recurso pode ser manipulado por meio dos métodos estabelecidos pelo padrão HTTP (POST, GET, PUT etc), os quais são ressignificados para a execução de operações sobre o objeto endereçado. Assim, o envio de uma mensagem GET sobre uma determinada URL específica, por exemplo, faz com que seja retornada a informação acerca do objeto associado à URL.

2.2 Virtualização

A virtualização é uma técnica muito utilizada como infraestrutura tecnológica atualmente, porque permite o compartilhamento de recursos de hardware por diversos serviços – e, até mesmo, sistemas operacionais completos com bom isolamento lógico entre eles. Desta forma, a virtualização cria uma camada de abstração de *software* para o serviço, que opera com controle restrito sobre os recursos de hardware do servidor físico (SILVA, 2017).

De forma geral, a virtualização pode ser classificada em: convencional (tradicional) e leve (TRINDADE; COSTA, 2018). A virtualização convencional é realizada pelo *software* denominado *hypervisor* (hipervisor), que atua como um intermediário responsável por gerenciar os recursos físicos do servidor para os serviços nele hospedados. Nesse caso, os serviços são hospedados em máquinas virtuais (*virtual machines*). Essa abstração permite a execução de múltiplos sistemas operacionais não modificados sobre o servidor físico, isto é, os sistemas operacionais destas máquinas virtuais acessam o hardware virtual, criado pelo *hypervisor*, como se estivessem acessando diretamente o hardware físico. Por sua vez, o *hypervisor* administra o acesso deste hardware virtual aos recursos físicos propriamente.

¹ <https://developers.google.com/docs/api/>

² <https://developer.twitter.com/en/docs>

³ <https://docs.aws.amazon.com/AmazonS3/latest/API/Welcome.html>

A virtualização leve, diferentemente da anterior, permite o isolamento de serviços diretamente sobre um mesmo sistema operacional, cada serviço é hospedado no sistema operacional sob a forma de um contêiner, que é implementado como um processo isolado no sistema operacional. Esse isolamento permite que os diferentes serviços operem sobre um mesmo sistema operacional, atuando diretamente sobre seu *Kernel* (TRINDADE; COSTA, 2018; SILVA, 2017). Esta é uma alternativa de virtualização considerada leve, porque permite o isolamento dos serviços ao mesmo tempo em que entrega maior eficiência, já que não obriga o emprego de um *hypervisor* e hardware virtualizado, tal como na virtualização convencional.

Uma plataforma de virtualização leve largamente utilizada atualmente é o *Docker*. Essa é *Open Source*, foi escrita na linguagem GO⁴, surgiu em 2013 e permite o gerenciamento do ciclo de vida completo dos contêineres, desde sua criação e execução até sua eventual parada e deleção (TRINDADE; COSTA, 2018; SILVA, 2017; CUNHA et al., 2019).

O *Docker* estabelece o conceito de imagem, que é modelo-base de um contêiner contendo um serviço e suas dependências, os contêineres, por sua vez, são instâncias de uma imagem. Note-se que a partir do momento em que é instanciado, um contêiner inicia como uma cópia da imagem que lhe serve de base, mas passa a ter um ciclo de vida próprio e divergirá da imagem conforme seja usado.

As imagens do *Docker* são criadas por meio de um arquivo – denominado *Dockerfile*, que descreve as dependências necessárias para o funcionamento da aplicação, e podem ser armazenadas localmente ou em um repositório em nuvem privada ou pública, um exemplo é o *Docker Hub*⁵, hospedando sua imagem neste repositório, o desenvolvedor de um serviço pode disponibilizá-lo para o público em geral, tendo a portabilidade garantida, já que o *Dockerfile* será utilizado para organizar as dependências do serviço (SILVA, 2017).

2.3 Interpolação Espacial

A computação oferece grande potencialidade na redução de custos em propriedades agrícolas. Diversos serviços já existentes procuram oferecer informação e funcionalidades que auxiliem a tomada de decisão diária, tais serviços estão sendo usados, por exemplo, na previsão de chuva ou mesmo para prever a evapotranspiração de determinada região, de modo que se possa compensar, a partir da irrigação, as perdas de água de uma cultura.

⁴ <http://www.golangbr.org/>

⁵ <https://hub.docker.com>

Um exemplo desse tipo de serviço é o Dark Sky⁶, um serviço que entrega medidas meteorológicas e previsão do tempo para todo o planeta. Serviços como o Dark Sky, fundamentam-se fortemente na interpolação espacial, que é uma técnica que permite a elaboração de estimativas em pontos desconhecidos de um espaço a partir de medidas realizadas em pontos conhecidos. Tal técnica pressupõe que os valores de um atributo tendem a se assemelhar em locais próximos e, da mesma forma, costumam ser mais distintos quando estão em locais mais distantes entre si (JAKOB; YOUNG, 2016).

Outro exemplo é o aplicativo EVAPO, que tem o objetivo de estimar de forma precisa e em tempo real a evapotranspiração potencial. Ele utiliza o serviço NASA-POWER⁷ com a finalidade de obter os dados meteorológicos necessários para o cálculo da evapotranspiração potencial (MALDONADO; VALERIANO; ROLIM, 2019). O projeto NASA-POWER ou simplesmente POWER (*Prediction of Worldwide Energy Resource*) é um conjunto de dados meteorológicos obtidos por meio de sistemas de satélites que buscam auxiliar no desenvolvimento de trabalhos na área de energias renováveis, construção de eficiência energética e necessidades agrícolas.

Outro aplicativo tomado como exemplo é o Aquaprev⁸, que é um *software* que visa estimar a quantidade de água necessária para irrigar uma determinada cultura, utilizando-se a evapotranspiração de referência diária e a interpolação espacial, ele encontra-se atualmente desenvolvido em sua versão beta disponível na loja virtual *Play Store*⁹ para download.

Um ponto importante a ressaltar é que, embora o EVAPO e Aquaprev sejam aplicativos desenvolvidos com propósitos semelhantes, estes se diferenciam pelo fato do Aquaprev proporcionar ao seu usuário informação que vai além da estimativa de evapotranspiração. A partir de dados sobre a cultura e área plantada, fornecidos pelo usuário, o Aquaprev pode fornecer o tempo que se deve permanecer irrigando uma determinada plantação.

Tais serviços utilizam diferentes algoritmos para interpolação espacial, dentre os quais, pode-se citar: *Inverse Distance Weighted* (IDW), *Ordinary Kriging* (OK) e *Random Forest* (RF), enquanto os dois primeiros algoritmos são reconhecidos por sua eficiência e, por isso, são convencionalmente utilizados nos cálculos da interpolação espacial de variáveis ambientais em trabalhos acadêmicos, o terceiro (RF) tem apresentado desempenho superior com relação as mesmas variáveis ambientais, em estudos recentes (JÚNIOR et al., 2019; LI et al., 2011; APPELHANS et al., 2015).

⁶ <https://darksky.net/about>

⁷ <https://power.larc.nasa.gov/>

⁸ <https://play.google.com/store/apps/details?id=com.jualabs.aquaprev.app>

⁹ <https://play.google.com/store/apps>

O algoritmo IDW caracteriza-se como um dos mais utilizados para interpolação espacial de variáveis ambientais (JÚNIOR et al., 2019). O pressuposto básico desta técnica é que a estimativa de um ponto a ser medido sofre influência local de pontos próximos diminuindo quando estimado em relação a pontos mais distantes (LI; HEAP, 2008). O IDW usa a média ponderada dos valores próximos ao ponto onde se busca estimar, os pesos da média tentam modelar a dependência espacial dos valores com a respectiva distância formada entre os pontos originais e o ponto estimado.

Tem-se também a técnica de *Kriging*, onde a média é constante, desconhecida e possui um estimador linear, ela ajusta uma função matemática a um número específico de pontos, ou a todos os pontos, dentro de um raio estabelecido. A partir desta função, é possível determinar o valor interpolado para cada ponto desejado. Vale ressaltar que, o estimador de *Kriging* varia conforme o modelo adotado para a função aleatória utilizada, essa é usualmente decomposta em um componente residual, que é modelado como uma função aleatória estacionária com média igual zero e covariância que muda em função de um componente residual e outro de tendência, onde a medida é dada a partir do valor esperado da variável aleatória no local (CARVALHO; VIEIRA, 2001).

Existem diversos estimadores de *Kriging* que se diferenciam de acordo com o modelo considerado para o componente de tendência, alguns desses modelos são: *Simple Kriging*, *Ordinary Kriging*, *Universal Kriging*, *Indicator Kriging*. Dentre esses, um que se destaca é o *Ordinary Kriging* (OK), pois, ele consegue acompanhar as flutuações de dados considerando as variações locais da média.

Outra técnica utilizada para estimar é a *Random Forest*, que trata-se de uma combinação de preditores do tipo árvore de decisão, treinados aleatoriamente, e cujo resultado final, é baseado no voto da maioria (JUSTA, 2019). Pode-se dizer ainda, que o RF é um algoritmo de aprendizagem de máquina flexível que retorna, geralmente, bons resultados mesmo sem o ajuste de hiper-parâmetros (DONGES, 2018).

A *Random Forest* (RF) pode ser entendida como sendo uma generalização do algoritmo *Regression Tree* (RT). Sendo assim, pode-se dizer que RF baseia-se em múltiplas RTs para geração de valores interpolados, esses, por fim, são submetidos à análise de frequência, que retorna o valor mais frequente que será adotado como previsão final (JÚNIOR et al., 2019). Além disso, a RF pode ser utilizada para dois tipos de tarefas: classificação e regressão, onde os algoritmos de classificação são usados quando a variável dependente é categórica e os de regressão quando a variável dependente é contínua, como é o caso do presente estudo.

2.4 Projeto de Experimentos

Estudos experimentais acerca de fenômenos naturais ou do comportamento de sistemas complexos podem ser afetados por um grande número de fatores, a investigação da influência desses fatores exige o rigoroso seguimento de protocolos para evitar que erros experimentais possam levar a conclusões incorretas, bem como cuidado para reduzir os custos relacionados a execução de tais experimentos (NETO; SCARMINIO; BRUNS, 2010).

Para essas situações, o uso de projetos experimentais respaldados em princípios estatísticos é a forma mais eficaz e eficiente de se obter o máximo de informações úteis da população de interesse, sem incorrer em um número elevado de experimentos. Assim, consegue-se obter uma economia considerável de esforços de trabalho despendidos em coletas de dados excessivas (JAIN, 1990).

A importância de um bom planejamento e de uma análise adequada reflete diretamente na melhoria da performance dos experimentos, pois, auxilia na categorização das consequências dos fatores envolvidos, deixando clara sua influência sobre os dados. Sendo assim, o planejamento de experimentos pode ser considerado a atividade estatística mais importante da experimentação, visto que, a sua não execução de forma apropriada impedirá retirar conclusões sobre o objeto de estudo (NETO; SCARMINIO; BRUNS, 2010).

Existem 3 tipos de planejamento de experimentos comumente utilizados, que são: o simples, o fatorial fracionário e o fatorial completo. No planejamento simples tem-se a variação de um fator, por vez, a fim de avaliar seu efeito individual sobre o fenômeno ou sistema em estudo. O experimento fatorial fracionário é usado quando não é possível utilizar um experimento fatorial completo, por conta dos custos financeiros ou de tempo, os quais são ocasionados pelo elevado número de níveis e fatores a serem combinados (NETO; SCARMINIO; BRUNS, 2010). Já em experimentos fatoriais completos, busca-se fazer uso de todas as combinações possíveis, em todos os níveis, de todos os fatores e, assim, avaliar quantitativamente a influência de fatores sobre a resposta de interesse e também como as interações entre esses fatores influenciam-se mutuamente (JAIN, 1990).

Esses tipos de planejamentos de experimentos proporcionam um embasamento para a execução de um experimento estatístico. No entanto, almeja-se também, obter uma melhor análise dos fatores que influenciam a população analisada, a partir de um número mínimo de experimentações possíveis. Para isso, utiliza-se da técnica estatística ANOVA, cujo o foco está no bom planejamento estatístico, para a estimação de médias das populações analisadas no experimento (CASELLA; BERGER, 2010).

A ANOVA constitui-se de uma técnica estatística de análise da variância, onde

divide-se as variáveis em partes e busca-se avaliar a influência dessas na variável de interesse de forma prática e concisa, dando aos tomadores de decisão uma ferramenta eficaz na análise de dados ([JAIN, 1990](#)).

3 Serviço de Interpolação Espacial

Este capítulo apresenta a proposta de especificação, implementação, implantação e demonstração do serviço web de interpolação espacial. A [seção 3.1](#) detalha sobre a especificação e implementação. Já a [seção 3.2](#) descreve a implantação do serviço em um ambiente de produção. Por fim, a [seção 3.3](#) apresenta o desenvolvimento de um aplicativo móvel, que teve como objetivo a demonstração das principais funcionalidades do serviço web desenvolvido neste trabalho.

3.1 Implementação do Serviço de Interpolação Espacial

O serviço de interpolação espacial, proposto no presente estudo, baseou-se nos princípios da SOA, descritos na [seção 2.1](#). Esta arquitetura tenta demonstrar as regras e protocolos de uso gratuito utilizados para a implementação do serviço web (*Web Services*), buscando ao mesmo tempo fazer com que o serviço desenvolvido tenha a maior independência possível de outros serviços que o utilizam como fonte. Além disso, a partir da SOA buscou-se estimular o reuso de funções que foram desenvolvidas previamente, tornando o código menos propenso a erros e diminuindo o tempo de construção do serviço.

Para a implementação do serviço de interpolação espacial, optou-se pela utilização da linguagem R ([R Core Team, 2018](#)), por se tratar de uma linguagem aberta e gratuita com a finalidade de resolução de problemas gráficos e de computação estatística, que possui pacotes já validados e aceitos pela comunidade acadêmica.

Além de dispor de funções especiais para exibir, armazenar e organizar dados espaciais, a utilização da linguagem R ([R Core Team, 2018](#)) se justifica pelo grau de estabilidade e confiabilidade das bibliotecas para interpolação espacial que o trabalho demanda. Outra razão pela qual optou-se pela utilização da referida linguagem, foi devido a sua característica de desenvolvimento aberto (*open source*), que garante um contínuo aprimoramento de bibliotecas e rotinas, como é o caso de rotinas para interpolação espacial.

A [Figura 1](#) apresenta um exemplo de arquitetura do sistema, onde o serviço desenvolvido neste trabalho pode ser inserido. Pois, ela ilustra como o serviço de interpolação espacial (*Spatial Interpolation Service*) interage com outros componentes da arquitetura, tais como, micros serviços (*microservices*) para previsões meteorológicas e aplicativos para usuários finais.

O serviço de interpolação espacial contém dois recursos apresentados no flu-

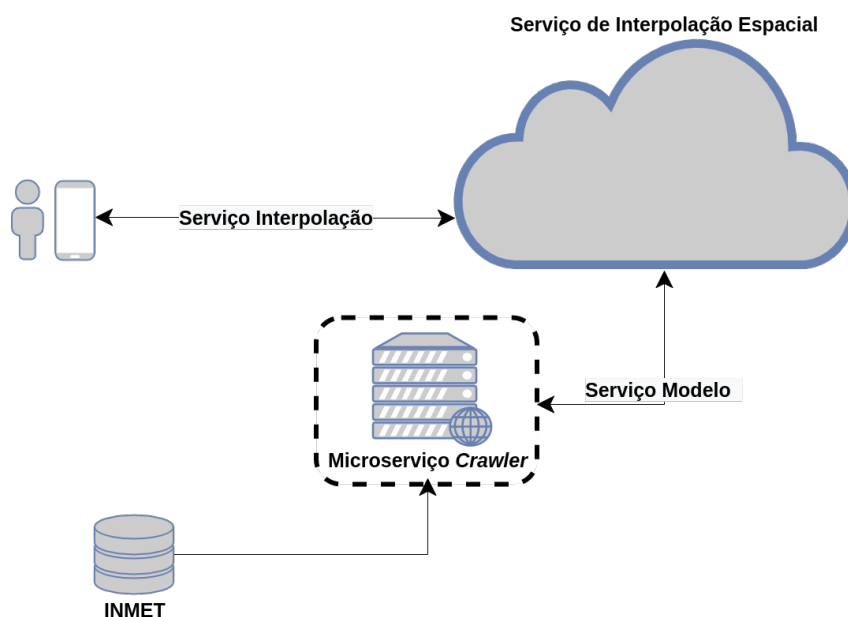


Figura 1 – Exemplo de arquitetura, onde o sistema web de interpolação espacial pode ser inserido.

Fonte: Próprio Autor

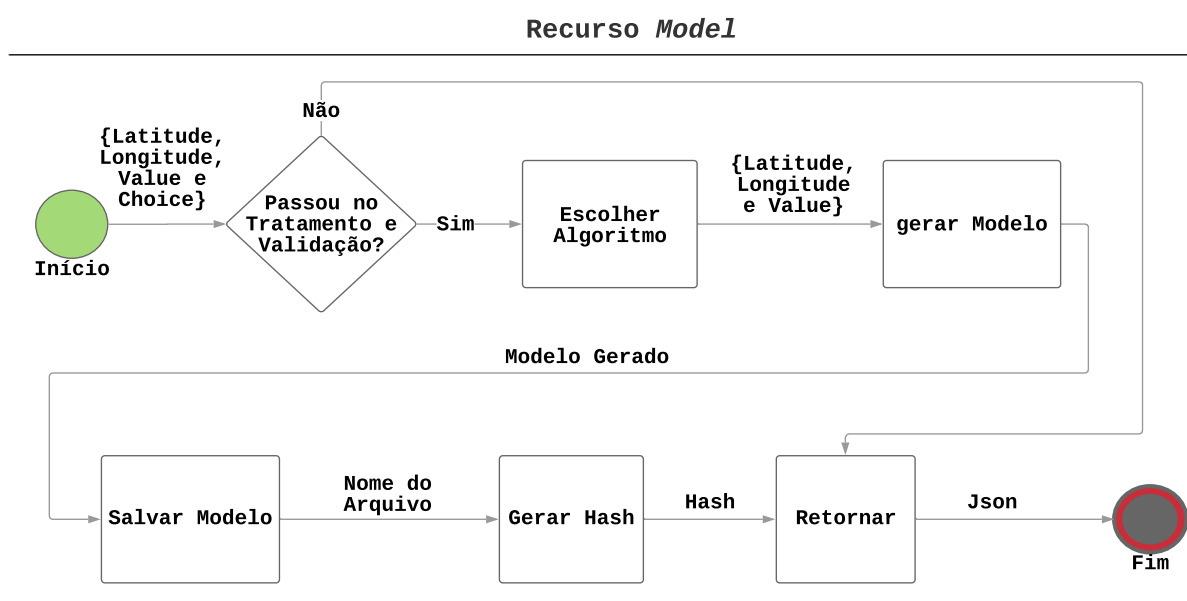


Figura 2 – Fluxograma simplificado da execução do recurso *Model*

Fonte: Próprio Autor

xograma da Figura 2 e Figura 5, que são: o de geração de um modelo de interpolação espacial (*Model*), Figura 2, que cria um modelo de interpolação a partir de um conjunto de observações espaciais e o da interpolação espacial (*Interpolation*), que realiza a interpolação propriamente dita para um conjunto de coordenadas fornecidas.

Para o uso do recurso *Model*, Figura 2, faz-se necessário que o cliente do ser-

```
1  {"meteorologicalData":  
2    [  
3      {  
4        "Latitude": -3.1211,  
5        "Longitude": -40.0873,  
6        "value": 5.4241  
7      },  
8      {  
9        "Latitude": -3.1212,  
10       "Longitude": -40.0875,  
11       "value": 5.4243  
12     },  
13     {  
14       "Latitude": -3.1216,  
15       "Longitude": -40.0878,  
16       "value": 5.4245  
17     }  
18   ],  
19   "choice": "idw"  
20 }
```

Figura 3 – Exemplo de JSON de dados enviados para geração do modelo por meio da URL *model*.

Fonte: Próprio Autor

viço, usuário ou outro serviço, envie uma requisição utilizando o método POST ao serviço, acompanhada de um conjunto de dados meteorológicos georeferenciados (*latitude*, *longitude* e *value*) e codificados no formato JSON, os quais baseiam-se no parâmetro *value*, que é a variável que se deseja interpolar posteriormente, e no tipo de algoritmo (*choice*) que se deseja utilizar para a geração do modelo.

A [Figura 3](#) exemplifica o formato JSON (*JavaScript Object Notation* - Notação de Objetos JavaScript) aceito pelo serviço web, ilustrado na [Figura 2](#), sendo enviado para criação de um modelo baseado no algoritmo IDW. Neste exemplo, *meteorologicalData* é o parâmetro que representa o conjunto de dados meteorológicos georeferenciados e, além disso, na [Figura 3](#) é possível ver um outro parâmetro que também é passado na requisição, o parâmetro *choice* o qual permite ao cliente do serviço, usuário ou outro serviço, escolher qual algoritmo, dos descritos na [seção 2.3](#), ele quer usar para gerar o modelo naquele momento, ainda na [Figura 3](#), foram utilizados apenas 3 valores passados como conjunto de dados georeferenciados por simplicidade, dado que as requisições devem conter um número maior do que este para geração do modelo com um baixo erro.

A biblioteca *Plumber*¹ foi escolhida, pois, oferece boa escalabilidade ao serviço web, além de oferecer uma rápida implantação do serviço, o que não ocorre em soluções alternativas, como é o caso da *OpenCPU*² e *DeployR*³.

Já a biblioteca *randomforest* é um pacote de geração e manipulação de modelos Random Forest. A escolha por esta biblioteca justifica-se por se tratar de uma implementação que tem sido utilizada em diversos trabalhos anteriores (Júnior et al., 2019; ABREU, 2016; CHAGAS et al., 2015).

Para o armazenamento de valores específicos do ambiente do sistema, utilizou-se o pacote *config*⁴, o qual facilita no gerenciamento de valores de configuração específicos do ambiente, como é o caso do uso de valores distintos para ambientes de desenvolvimento, teste e produção. O serviço buscou tratar e validar os dados recebidos pela interface do serviço, organizando-os de forma a ordenar as posições dos valores recebidos em uma ordem pré-definida pelo sistema, para assegurar que os dados recebidos sejam realmente os de interesse definidos pelo software.

Os dados enviados ao recurso *Model*, Figura 2, pelo cliente do serviço, passaram por três validações: detecção de campos vazios no conjunto de dados, verificação da quantidade de dados recebidos, que serviu para verificar se o número de dados é adequado para o uso do método escolhido e se é tratável computacionalmente pelo serviço, sendo esses valores máximos e mínimos são parâmetros configurados antes da inicialização do serviço, e validação do algoritmo de interpolação escolhido.

A geração do modelo em si, mostrado na Figura 2, utilizou-se da opção *choice* do JSON, Figura 3, para escolher o algoritmo de interpolação espacial para a geração do modelo. Como resultado da chamada a este recurso, é enviado ao cliente um *hash* que serve como um identificador do modelo, que fica armazenado no servidor onde o serviço está sendo executado em formato próprio do R (*.Rds*).

O *hash* fornecido, conforme Figura 2, é formado pelo *timestamp* no qual foi gerado o modelo, esse utilizando data e hora no padrão POSIX time, concatenado ao tipo de algoritmo escolhido pelo usuário para gerar o modelo. Esse *hash* é retornado para o usuário em formato JSON, como exemplificado na Figura 4. Desta forma, os clientes do serviço (usuário ou sistemas) que desejem fazer uma chamada para a interpolação (*interpolation*) fazendo uso do modelo gerado anteriormente, terão que passar como parâmetro o *hash* na requisição.

O segundo recurso implementado foi a operação de interpolação propriamente dita, alcançada por meio da URL (*interpolation*), ilustrada na Figura 5, que também

¹ <https://www.rplumber.io/>

² <https://docs.microsoft.com/pt-br/machine-learning-server/deployr/deployr-about>

³ <https://www.opencpu.org/>

⁴ <https://cran.r-project.org/web/packages/config/vignettes/introduction.html>

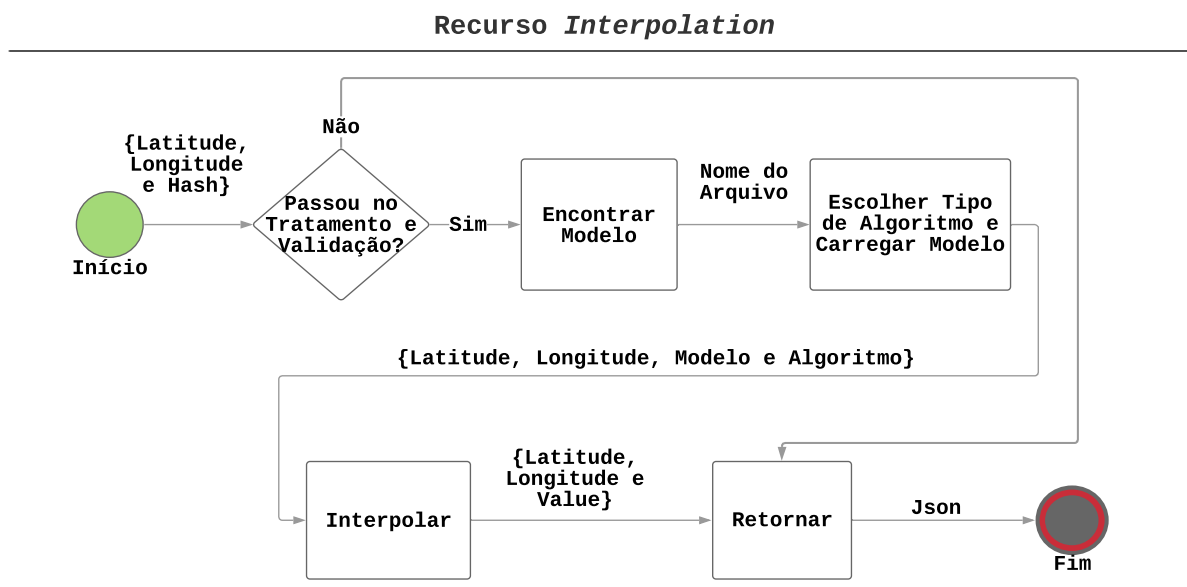
```

1  [
2      "b2sgMjAxOS0wMy0yNCAxNjowNjo1NS42Mjc2MTc="
3  ]

```

Figura 4 – Exemplo de JSON de dados retornado após a geração do modelo.

Fonte: Próprio Autor

Figura 5 – Fluxograma simplificado da execução do recurso *Interpolation*

Fonte: Próprio Autor

utiliza-se do método POST para o envio do JSON com as coordenadas (*coordinates*) geográficas. Um exemplo do JSON utilizado no recurso *interpolation*, demonstrado na Figura 5, pode ser visto na Figura 6, que mostra um exemplo de conjunto de dados meteorológicos georeferenciados com *latitude* e *longitude*, que representam o ponto ou conjunto de pontos que se deseja interpolar. Além disso, na Figura 6, pode-se observar o segundo parâmetro que é o *hash* gerado pela requisição *model*, Figura 2, utilizada para gerar o modelo baseado em um dos algoritmos estudados neste trabalho. Esse *hash*, Figura 6, é utilizado para recuperar o arquivo (*.Rds*) que armazena um modelo gerado, anteriormente, no disco rígido (HD) do servidor (computador), onde o serviço está sendo executado.

Após a chamada do recurso *interpolation*, Figura 5, o modelo gerado a partir do fluxo, ilustrado na Figura 2, correspondente ao passado pelo *hash*, é carregado em memória e a interpolação é realizada por meio da função *predict*, presente no pacote *stats*. A escolha por esta biblioteca se deu por se tratar de uma biblioteca genérica para predições relacionadas a qualquer tipo de modelo.

```
1  {"coordinates":  
2    [  
3      {  
4        "latitude": -3.8965,  
5        "longitude": -40.1320  
6      }  
7    ],  
8    "hash": "b2sgMjAxOS0wMy0yNCAxNjowNjo1NS42Mjc2MTc=",  
9  }
```

Figura 6 – Exemplo de JSON de dados enviados por meio do recurso *interpolation*.

Fonte: Próprio Autor

Com a conclusão da interpolação, [Figura 5](#), após a execução da função *predict* que retorna o valor da estimativa para a variável a qual se deseja interpolar, o serviço retorna para o cliente um conjunto de dados meteorológicos georeferenciados em formato JSON. Na [Figura 7](#), observa-se que o conjunto de dados retornado ao cliente, contém as coordenadas enviadas com os valores interpolados correspondentes a cada par de coordenadas repassadas inicialmente.

```
1  {  
2    "latitude": -3.8965,  
3    "longitude": -40.1320,  
4    "value": 24.4023  
5  }
```

Figura 7 – Exemplo de JSON de dados retornado após a interpolação das coordenadas.

Fonte: Próprio Autor

Vale ressaltar que a escolha por separar a geração do modelo de sua efetiva utilização para interpolação, foi feita com o intuito de dar maior generalidade ao serviço, possibilitando a seus usuários a reutilização de um modelo gerado repetidas vezes para diferentes conjuntos de pontos de interpolação. Por isso, não é objetivo do serviço a consulta aos *hashes* gerados, ficando a cargo destes clientes (usuários ou sistemas) o gerenciamento e armazenamento dos *hashes* retornados pelo serviço de interpolação espacial.

3.2 Implantação do Serviço de Interpolação Espacial

Após a implementação e avaliação do serviço, descrito na [seção 3.1](#) e na [seção 4.1](#) respectivamente, buscou-se realizar o terceiro objetivo específico, que é a implantação do serviço desenvolvido, a fim de disponibilizá-lo como forma de contribuição para uso por parte da sociedade.

Para tanto, na busca de fazer o uso do servidor (máquina que hospeda o serviço web de interpolação espacial) de forma eficiente e eficaz, fez-se uso de contêiner de *software*, o qual proporciona ao serviço web a capacidade de ser escalável, tendo uma ou mais instâncias do serviço, como já mencionado na [seção 2.2](#). Além disso, essa abordagem exige pouca implementação e recursos para tornar o serviço web disponível para uso (CUNHA et al., 2019).

A utilização do *Docker*, no presente estudo, foi feita de modo a construir uma imagem com todas as dependências em conjunto com o serviço desenvolvido, a fim de torná-lo portátil e viabilizar a possibilidade de escalar o serviço de forma vertical e horizontal, possibilitando seu uso em um ambiente de produção de forma eficiente mesmo que sejam feitas solicitações e cargas de trabalho que exijam mais recursos.

Para a implantação dessa imagem portátil com todas as dependências instaladas e prontas para serem utilizadas, fez-se a aquisição de uma máquina junto ao Núcleo de Tecnologia da Informação (NTI) da Universidade Federal Rural de Pernambuco (UFRPE). As configurações dessa máquina estão descritas na [Tabela 1](#).

Descrição	Configuração
Computador	Intel(R) Xeon(R) Gold 5120 CPU @ 2.20GHz 14 núcleos, RAM de 8 GB e HD de 280 GB

Tabela 1 – Configuração do computador onde foi implantado o serviço web de interpolação espacial de dados meteorológicos

Fonte: Próprio Autor

A instalação do serviço, via imagem Docker, foi realizada remotamente na máquina do NTI com as configurações mencionadas na [Tabela 1](#), onde notou-se algumas das vantagens do uso das funcionalidades do Docker, pois, em menos de três minutos todo o serviço estava implantado e pronto para uso. Após essa instalação, pode-se disponibilizar o serviço por meio de endereço eletrônico válido, tornando o serviço acessível para que qualquer pessoa, com conhecimento prévio sobre os padrões de funcionamento do serviço, possa utilizá-lo.

3.3 Aplicativo Móvel

Buscou-se, ainda, por meio do presente trabalho, integrar o serviço desenvolvido a uma aplicação móvel que pudesse demonstrar a funcionalidade do serviço proposto. Essa integração foi feita em um aplicativo móvel criado para a plataforma Android. Já, para o desenvolvimento do aplicativo utilizou-se a ferramenta Android Studio⁵, que é um ambiente de desenvolvimento integrado de construção de aplicativos para a plataforma Android.

É importante destacar que o aplicativo foi desenvolvido exclusivamente para testar as funcionalidades do serviço criado. O aplicativo demonstra o uso do recurso *interpolation* disposto pelo serviço de interpolação espacial e, portanto, não possui uma interface planejada para proporcionar uma melhor experiência ao usuário do aplicativo.

Para o desenvolvimento do aplicativo utilizou-se a linguagem de programação Kotlin⁶, dado que essa é uma linguagem multiplataforma. Associado à linguagem Kotlin, fez-se uso de duas APIs (*Application Programming Interface*), a API para o serviço Google Maps⁷, que permite a manipulação de mapas, e a API Google Places⁸, que possibilita obter as coordenadas geográficas de locais específicos informados na consulta. Desta forma, é possível que o usuário escolha as coordenadas sobre a qual deseja interpolar, a partir da indicação com um toque mais demorado na tela.

A Figura 8a mostra a tela principal do aplicativo móvel com as características supracitadas. Após a escolha do local, o aplicativo utiliza a coordenada escolhida para o preenchimento dos dois primeiros campos (latitude e longitude) do formulário presente na próxima tela (ilustrada na Figura 8b). Já o terceiro campo é preenchido por um *hash* identificador do modelo salvo no serviço. Este modelo será utilizado para a realização da interpolação para o respectivo ponto escolhido. Sendo o modelo de interpolação gerado a partir de 500 pontos de evapotranspiração conhecidos, o qual será útil para estimar o valor da evapotranspiração de referência diária para o ponto escolhido pelo usuário.

Na Figura 8c tem-se o retorno dado pelo serviço para o aplicativo com o valor da interpolação espacial, obtida para a coordenada que foi dada como parâmetro de entrada pelo usuário, sendo esta, uma medida para a evapotranspiração de referência diária nessa coordenada que antes era desconhecida.

A partir do valor obtido para a interpolação espacial, o usuário pode calcular o tempo de irrigação diário necessário para uma determinada cultura que está localizada neste ponto, atendendo, assim, o segundo objetivo específico proposto por este

⁵ <https://developer.android.com/studio>

⁶ <https://kotlinlang.org/>

⁷ <https://developers.google.com/maps/documentation/?hl=pt-br>

⁸ <https://developers.google.com/places/web-service/intro>

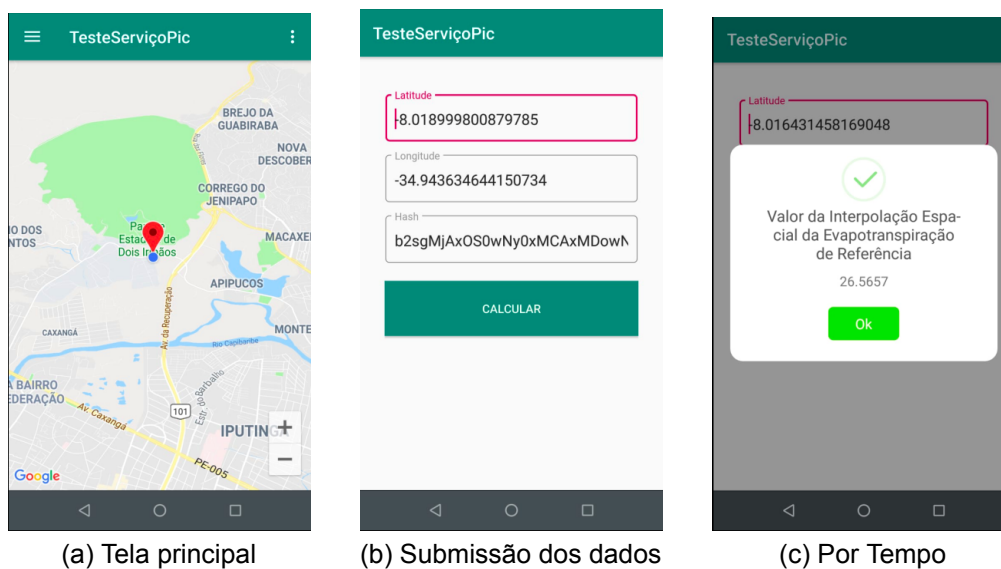


Figura 8 – Retorno com a Interpolação Espacial no aplicativo móvel
Fonte: Próprio Autor

estudo.

4 Experimentos

Neste capítulo serão abordados, na [seção 4.1](#) e [seção 4.2](#), o método e os resultados da avaliação de desempenho do serviço web de interpolação espacial proposto neste trabalho. Vale salientar que este trabalho não buscou validar os resultados gerados pelos algoritmos, visto que este tipo de validação já fora objeto de avaliações realizadas por outros autores ([JÚNIOR et al., 2019](#); [LI; HEAP, 2014](#); [XAVIER; KING; SCANLON, 2016](#)).

4.1 Método para Avaliação de Desempenho do Serviço de Interpolação Espacial

O plano de testes apresentado, nesta seção, buscou avaliar o desempenho do serviço web desenvolvido. Basicamente, realizou-se uma observação do comportamento em termos de tempo de resposta do serviço em diferentes condições, para, assim, sondar e identificar gargalos no serviço.

Para a realização dos testes utilizou-se a ferramenta *Apache JMeter*¹, que é um *software open source* baseado na linguagem de programação *Java*², definido para realização de testes de estresse e desempenho em sistemas do tipo cliente/servidor, tais como o sistema proposto neste trabalho. Com este *software* consegue-se fazer diversas simulações com muita robustez, por exemplo: simulações de carga, servidores, aplicações e até simulações de objetos ([SANTOS; NETO, 2008](#)).

O plano de teste envolveu o uso de requisições para geração do modelo e requisições para interpolação. Na [Tabela 2](#) são apresentados os fatores do experimento com seus respectivos níveis.

Fatores	Níveis
Pontos do Modelo	100, 200, 300, 400 e 500
Algoritmos	IDW, OK e RF
Usuários Virtuais	1, 5, 25 e 50
Mix Modelo/Interpolação	0%/100%, 25%/75%, 50%/50%, 75%/25% e 100%/0%

Tabela 2 – Fatores utilizados no experimento

Fonte: Próprio Autor

¹ <https://jmeter.apache.org/>

² https://www.java.com/pt_BR/download/faq/develop.xml

O fator Pontos do Modelo refere-se a quantidade de pontos para a geração do modelo, ou seja, o número de amostras espaciais da variável ambiental sob estudo. Esse primeiro fator, variou de 100 a 500 dados em intervalos de 100. Desta forma, conseguiu-se estabelecer como a quantidade de dados submetidos afeta o desempenho da geração do modelo.

Já o segundo fator (Algoritmos) representa o algoritmo utilizado nas requisições do experimento, sendo escolhido um dos 3 algoritmos implementados neste trabalho: IDW, OK e RF. Neste contexto, observou-se como a escolha do algoritmo impacta no tempo de geração do modelo e na interpolação espacial.

O fator Usuários Virtuais corresponde a quantidade de requisições simultâneas realizadas ao serviço. Esse fator variou entre 1, 5, 25 e 50 requisições simultâneas. Pretendeu-se, com isso, compreender a quantidade limite de usuários que o serviço poderia suportar, e estipular valores máximos a serem utilizados nas etapas de validação do serviço.

Por fim, o fator Mix Modelo/Interpolação está relacionado ao percentual ao qual cada tipo de requisição é solicitado dentro do experimento, nesse caso, cada nível indica a porcentagem de requisições feitas ao recurso *model* e a porcentagem de requisições ao recurso *interpolation* respectivamente. Assim, nos experimentos realizados para o nível 0%/100% do fator em questão, por exemplo, contou-se apenas com a execução de requisições do tipo *interpolation*.

Os principais parâmetros do experimento são apresentados na [Tabela 3](#).

Parâmetros	Valor
Requisições totais	500
Tempo de inicialização	0
Pontos de Interpolação	1
Computador	Processador Intel Xeon 24 núcleos, RAM de 16 GB e HD de 1 TB

Tabela 3 – Parâmetros para o plano de teste do Serviço
Fonte: Próprio Autor

A quantidade de requisições totais representa o número de requisições que foram feitas ao serviço de interpolação espacial. Observou-se que, a obtenção da quantidade de amostra ideal para os parâmetros na situação analisada, é indispensável. Pois, deve-se ter um número mínimo na composição da amostra que represente a população total de requisições do experimento, para, portanto, conseguir determinar a confiabilidade das estatísticas obtidas. Este número foi conseguido por meio de dois experimentos prévios com a finalidade de obtenção da amostra mínima necessária, de modo a minimizar os erros de amostragem utilizando 95% de confiança (JAIN, 1990).

O Tempo de Inicialização determinou o intervalo de tempo entre requisições simultâneas. Nestes tipos de experimentos este tempo é tido como zero, a fim de aumentar a intensidade da demanda por recursos nos casos simultâneos. Já os Pontos de Interpolação determinam a quantidade de pontos utilizados em cada requisição do tipo *interpolation*, que neste experimento fora fixado em um ponto. Tais pontos, supracitados, são coordenadas (latitude e longitude) do território brasileiro, geradas automaticamente pelo site *Random Point Generator*³.

O serviço de interpolação espacial e a ferramenta de medição JMeter foram hospedados em um mesmo computador, para eliminar impactos causados pela rede no desempenho do serviço. O computador possui processador Intel Xeon com 24 núcleos de processamento, 16 GB de memória RAM, 1 TB de armazenamento e sistema operacional Linux.

O experimento proposto nesta seção foi executado como um experimento completamente fatorial, onde cada requisição foi tratada como uma execução independente do experimento. A medida de desempenho avaliada no presente estudo é nomeada como tempo de resposta e inclui o tempo de disparo de uma requisição pelo cliente, sendo, o processamento dessa requisição pelo serviço e o retorno dado por esse ao cliente. Utilizou-se, nesta análise, o tempo médio de resposta das requisições referente a cada tratamento do experimento.

A partir da avaliação de desempenho, buscou-se entender qual a relação e impacto estabelecidos entre os possíveis fatores que podem influenciar na performance do serviço de interpolação espacial no referido tempo de resposta. A análise estatística para determinar a presença de interações significativas em relação ao tempo de resposta, baseou-se no teste ANOVA (JAIN, 1990), descrito na [seção 2.4](#).

4.2 Resultados da Avaliação de Desempenho do Serviço de Interpolação Espacial

Os resultados são apresentados a seguir e estão divididos em duas seções. Na Seção [4.2.1](#), tem-se apresentado e discutido o impacto individual de cada fator no tempo de resposta. Já na Seção [4.2.2](#), discute-se as relações entre estes fatores e seu impacto no tempo de resposta.

4.2.1 Impacto Individual de cada fator

Para a análise inicial dos fatores em estudo, inicialmente fez-se uso do *boxplot* (Diagrama de caixa), que é um tipo de gráfico usado para avaliar a distribuição empí-

³ <http://www.geomidpoint.com/random/>

rica dos dados. A [Figura 9](#), mostra como a variável tempo de resposta se comporta frente a variações no número de usuários virtuais. Claramente, é possível perceber que o aumento do número de usuários virtuais tem pequeno impacto sobre a mediana do tempo de resposta, indicando que para a maior parte das requisições, o tempo de resposta é baixo. Contudo, observa-se que há um forte impacto na variabilidade do tempo de resposta, obtendo-se, para mais de 25 usuários, um tempo de resposta consideravelmente alto, mais de 10 segundos. Contudo, deve-se observar que este aumento tem forte relação com o fator Algoritmos, conforme será discutido na [Seção 4.2.2](#).

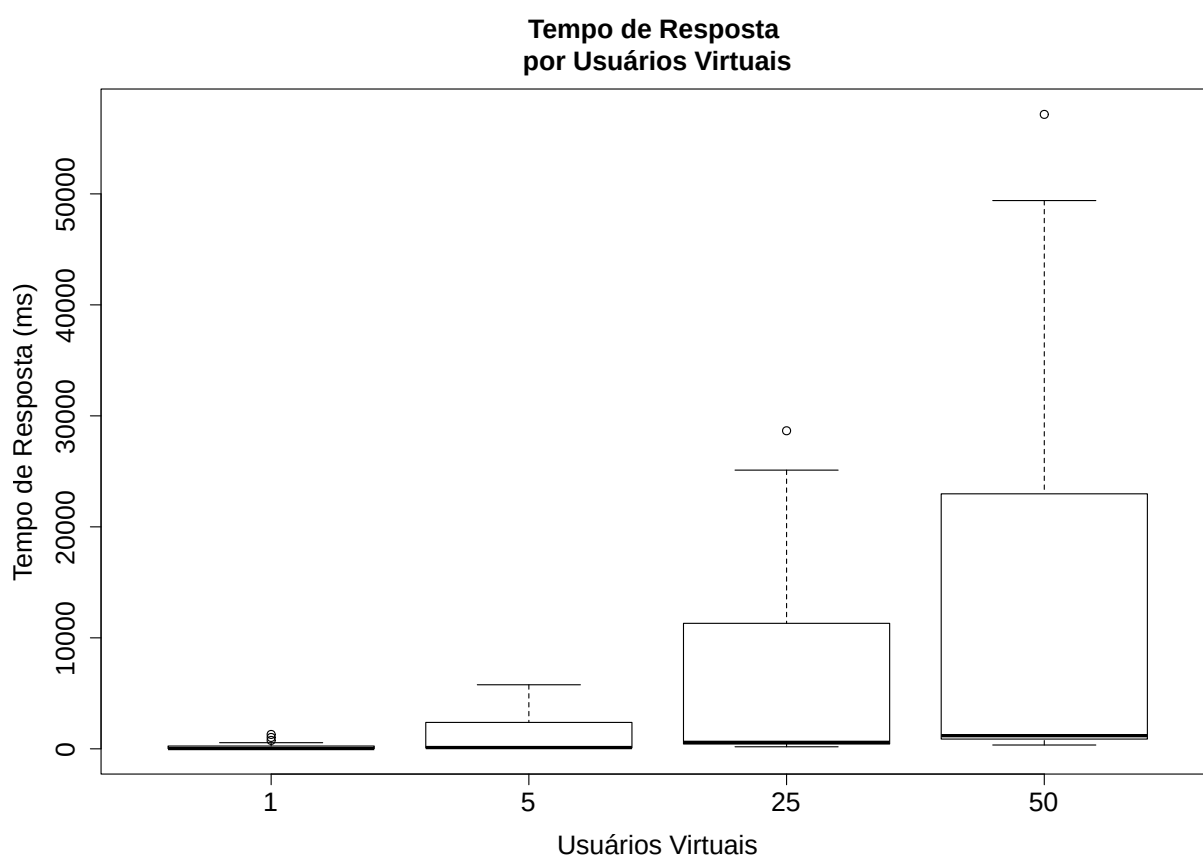


Figura 9 – Tempo de Resposta por Usuários Virtuais
Fonte: Próprio Autor

Outro fator analisado, foi como o tempo de resposta se comporta para cada um dos algoritmos avaliados, o qual encontra-se ilustrado na [Figura 10](#), onde verifica-se que o tempo de resposta é fortemente impactado pela escolha do algoritmo de interpolação. Pois, enquanto os algoritmos IDW e OK apresentaram tempos de resposta baixos (IDW com mediana de 379.8392 ms e OK com mediana de 540.8264 ms), o algoritmo RF apresentou mediana superior a 10 segundos (RF com mediana de 11695.7 ms), alcançando, inclusive, valores maiores do que 50 segundos (RF com tempo máximo de 57162.27 ms). Este resultado corrobora com o observado em ([JúNIOR et al.](#),

2019), o qual mostrou um tempo médio de resposta do RF, aproximadamente 8 vezes maior do que o apresentado pelo algoritmo OK.

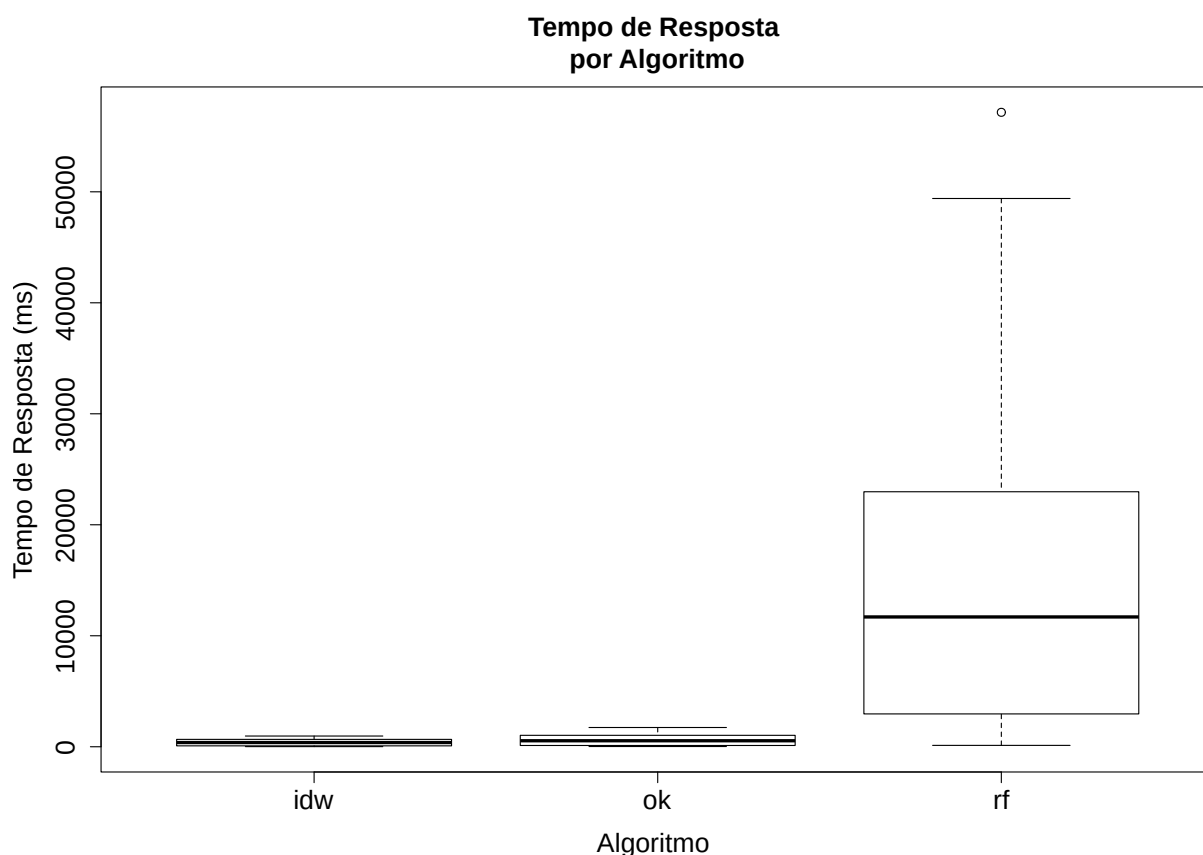


Figura 10 – Tempo de Resposta por Algoritmo

Fonte: Próprio Autor

Também se observa grande variabilidade no tempo de resposta do algoritmo RF. Tal variabilidade pode estar associada a forma como este algoritmo trabalha, já que esse é um classificador de aprendizagem de máquina supervisionado, baseado na construção de um conjunto de árvores de decisão, que pode demandar um grande esforço computacional do sistema, principalmente, nos casos em que há um maior número de usuários requisitando a geração de novos modelos.

Na Figura 11 encontra-se representado no eixo das abcissas o Mix modelo/interpolação, onde o primeiro par ordenado, da esquerda para direita, indica que não há requisições para criação de modelos (0%) e que, no caso, todas as requisições (100%) são de interpolação. Já segundo item aponta que 25% das requisições são para criação do modelo e 75% são para interpolação. O terceiro mostra que há um equilíbrio entre as requisições de cada tipo, 50% para cada, e os demais seguem esta ideia, até que se tenha o oposto da proporção inicial com 100% das requisições para criação do modelo e nenhuma (0%) para interpolação.

O resultado mostra um aumento da variabilidade do tempo de resposta, com

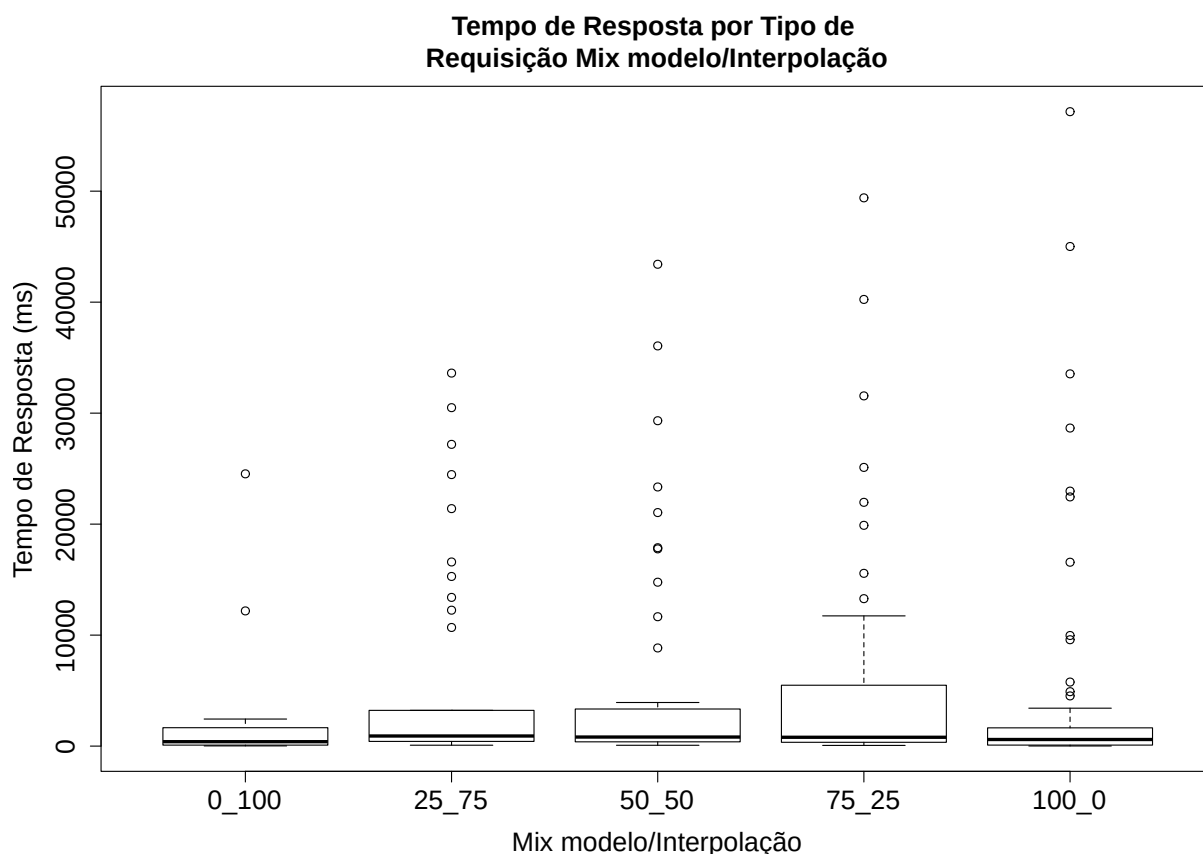


Figura 11 – Tempo de Resposta por Tipo de Requisição
Fonte: Próprio Autor

consequente aumento da média de cada caso, do nível 0_100 (média de 3481.95 ms) ao nível 25_75 (média de 5544.315 ms), decorrente do fato de que os algoritmos para geração dos modelos são, comumente, de maior complexidade computacional, e assim, um maior número de requisições para geração de modelos, exigirá maior tempo de processamento do servidor no qual o serviço está hospedado. Por outro lado, observa-se uma diminuição do tempo de resposta quando se passa para o nível 100_0 (média de 4799.085 ms), cujo resultado se aproxima do encontrado no nível 0_100 (média de 3481.95 ms), mas, sendo maior que este. Apesar de todas as requisições serem para geração do modelo, neste nível deve-se considerar que a mudança na proporção de requisições de cada tipo, impacta, principalmente, nos tipos de operações que são feitas em disco.

Neste experimento, todos os modelos são armazenados/recuperados diretamente em disco, como explicado na [seção 3.1](#), portanto, a mudança no mix altera a quantidade de operações de escrita/leitura em disco. Os casos que fazem apenas o mesmo tipo de operação, somente leitura (nível 0_100) e somente escrita (100_0), promovem um melhor desempenho, devido a maior liberdade que o sistema operacional e a controladora de disco têm em otimizar as operações, sendo a operação de

escrita ligeiramente mais lenta do que a de leitura.

Nos casos intermediários, o aumento da quantidade de operações de escrita, em comparação com o caso 0_100, leva a um aumento gradativo do tempo de resposta atingindo, de modo geral, valores médios maiores, até mesmo do que o nível 100_0. Isso é decorrente da menor liberdade para otimização que o sistema operacional e a controladora de disco têm para as operações de leitura e escrita.

A [Figura 12](#) ilustra o resultado do tempo de resposta para cada um dos níveis do fator Pontos do Modelo. O nível com 0 pontos, indica os casos onde só há requisições para interpolação sobre um modelo já criado, equivalente ao Mix Modelo/Interpolação 0_100, já discutido. Pode-se observar deste resultado, que o tempo de resposta aumenta ligeiramente ao variar-se o número de pontos do modelo, saindo de uma mediana de 530.6118 ms com 100 pontos e atingindo uma mediana 884.7898 ms com 500 pontos. O que é esperado, dado que o aumento dos níveis indica a criação de modelos cada vez maiores.

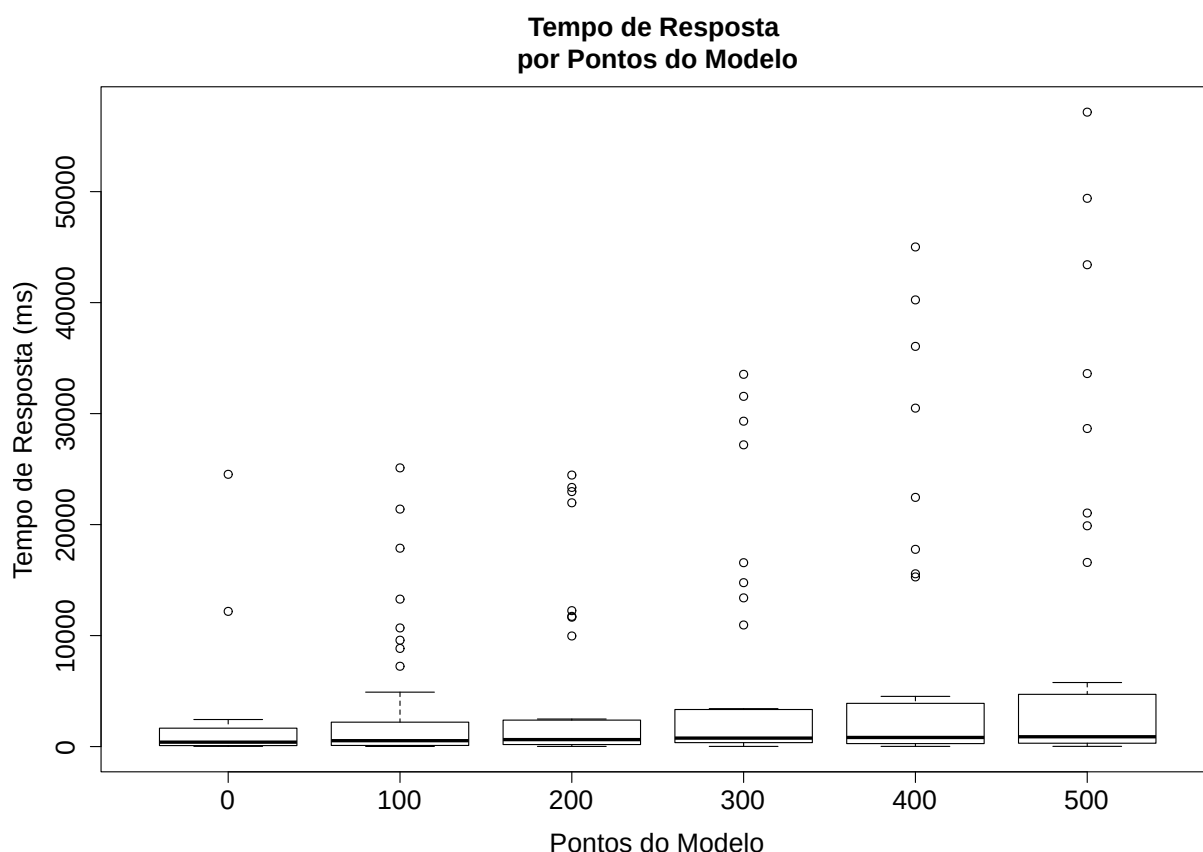


Figura 12 – Tempo de Resposta por Pontos do Modelo

Fonte: Próprio Autor

Para uma melhor análise do problema proposto por este estudo recorreu-se também ao teste de análise de variância (ANOVA), descrita na [seção 2.4](#), para avaliar a influência estatística de cada fator sobre o tempo de resposta frente aos fatores:

Usuários Virtuais, Algoritmos, Mix Modelo/Interpolação e Pontos do Modelo. A [Tabela 4](#) apresenta os resultados do teste. Onde para cada fator são demonstrados a variação explicada pelo fator (coluna SQ) e sua respectiva contribuição percentual, bem como o valor da estatística de teste F (coluna Valor F) e o p-valor, que indica a probabilidade da evidência contra a hipótese nula, que no caso do presente estudo, é expressa como sendo o não impacto de um fator sobre o tempo de resposta do serviço.

Fator	SQ	Contribuição %	Valor F	p-valor
Mix Modelo/Interpolação	91395282	0,41	0,503	0,733
Algoritmos	9366991365	42,34	103,133	0,0000000 000000002
Pontos do Modelo	555080773	2,51	3,056	0,018
Usuários Virtuais	3435880804	15,53	25,22	0,0000000 000000863
Resíduos	8673712530	-	-	-

Tabela 4 – Resultado da análise de variância (ANOVA)

Fonte: Próprio Autor

Sendo o p-valor menor que nível de significância de 0,05 para os fatores Algoritmos, Usuários Virtuais e Pontos do Modelo. Conclui-se que há uma associação estatisticamente significativa, a 95% de nível de confiança, entre o tempo de resposta e estes fatores. Contudo, nota-se que o fator Pontos do Modelo traz menor impacto ao tempo de resposta, com uma contribuição de 2,51%, reafirmando o que já fora observado através dos diagramas de caixa. Já no caso do fator Mix Modelo/Interpolação, o p-valor não nos permite rejeitar a hipótese nula. Desse modo, há indícios de que o fator Mix Modelo/Interpolação exerce uma fraca influência sobre o tempo de resposta, sendo sua contribuição percentual de apenas 0,41%.

4.2.2 Interações entre os fatores

Neste trabalho, investigou-se ainda, como a interação entre os fatores propostos impacta no tempo de resposta do serviço. A [Figura 13](#) mostra como o tempo médio de resposta de cada um dos algoritmos se comporta com a mudança no Mix Modelo/Interpolação. Apesar dos resultados do algoritmo RF estarem em uma ordem de grandeza diferente daquela dos demais algoritmos, o comportamento geral dos algoritmos é similar: aumento gradativo do tempo de resposta do nível 0_100 ao nível 75_25 e diminuição do tempo para o caso 100_0. Sendo este comportamento motivado, principalmente, pelo custo computacional dos algoritmos para geração dos modelos e pela quantidade de operações de leitura e escrita em disco, conforme já destacado na [subseção 4.2.1](#).

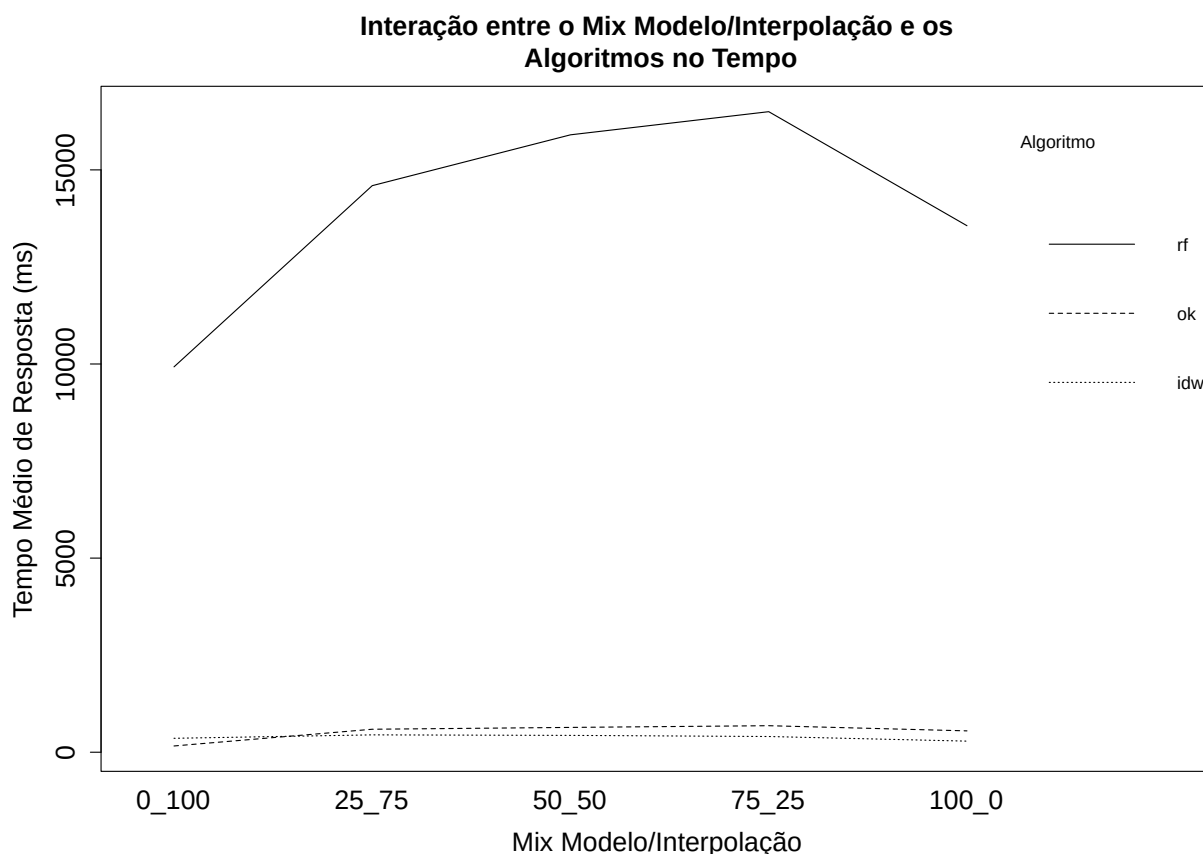


Figura 13 – Interação entre os **Algoritmos** e o tipo de requisição (**Mix Modelo/Interpolação**) no tempo

Fonte: Próprio Autor

A relação entre os fatores Pontos do modelo e Usuários Virtuais sobre o tempo médio de resposta, é mostrada na [Figura 14](#). Expondo que enquanto que o fator Pontos do Modelo promove, de forma geral, um aumento gradual do tempo de resposta, o fator Usuários Virtuais impacta fortemente o tempo médio de resposta, fazendo com que esse mude em ordens de grandeza. Mostrando em um caso particular com 50 usuários virtuais, que o tempo médio de resposta cresce sobremaneira com o aumento dos Pontos do Modelo.

O caso com 0 Pontos do Modelo é a condição em que não há requisições para geração do modelo (i.e., o caso com Mix Modelo/Interpolação ao nível 0_100). Desta forma, pode-se inferir que o tempo médio de resposta mais alto do algoritmo RF (ver [Figura 13](#)) tem também relação com quantidades de 25 e 50 usuários, em outras palavras, quando a quantidade de usuários é mais baixa (1 e 5), o tempo médio de resposta é baixo para todos os algoritmos avaliados, enquanto que, os resultados dos casos de 25 e 50 usuários são influenciados, principalmente, pelo algoritmo RF.

Esta observação é confirmada pela [Figura 15](#), que mostra o impacto da quantidade de Usuários Virtuais e dos Algoritmos sobre o tempo de resposta, no qual o

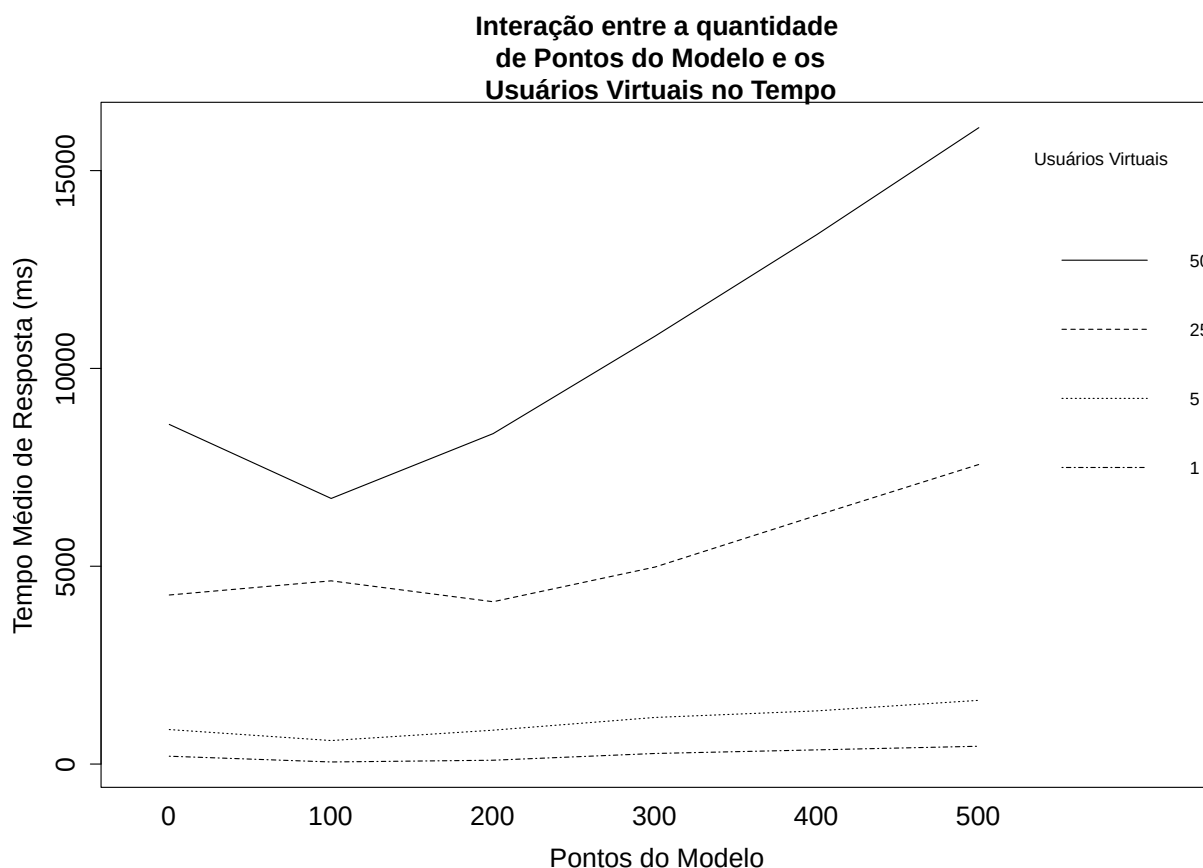


Figura 14 – Interação entre a quantidade de **Pontos do Modelo** e a quantidade de **Usuários Virtuais** no Tempo de Resposta

Fonte: Próprio Autor

algoritmo RF, diferentemente dos demais, foi expressivamente impactado pela variação do número de Usuários Virtuais.

Sendo assim, a partir da observação dos gráficos supracitados nessa seção percebeu-se que tanto os algoritmos quanto o número de usuários virtuais são os fatores que mais impactam no tempo de resposta do serviço, de forma que quanto maior o número de usuários, maior é o tempo de resposta do serviço de interpolação espacial, no qual o algoritmo RF apresentou desempenho inferior ao dos demais algoritmos.

Além disso, é possível observar que o aumento do número de usuários virtuais é muito mais impactante na métrica, do que o aumento da quantidade de pontos no modelo, corroborando com os resultados obtidos pelo teste ANOVA.

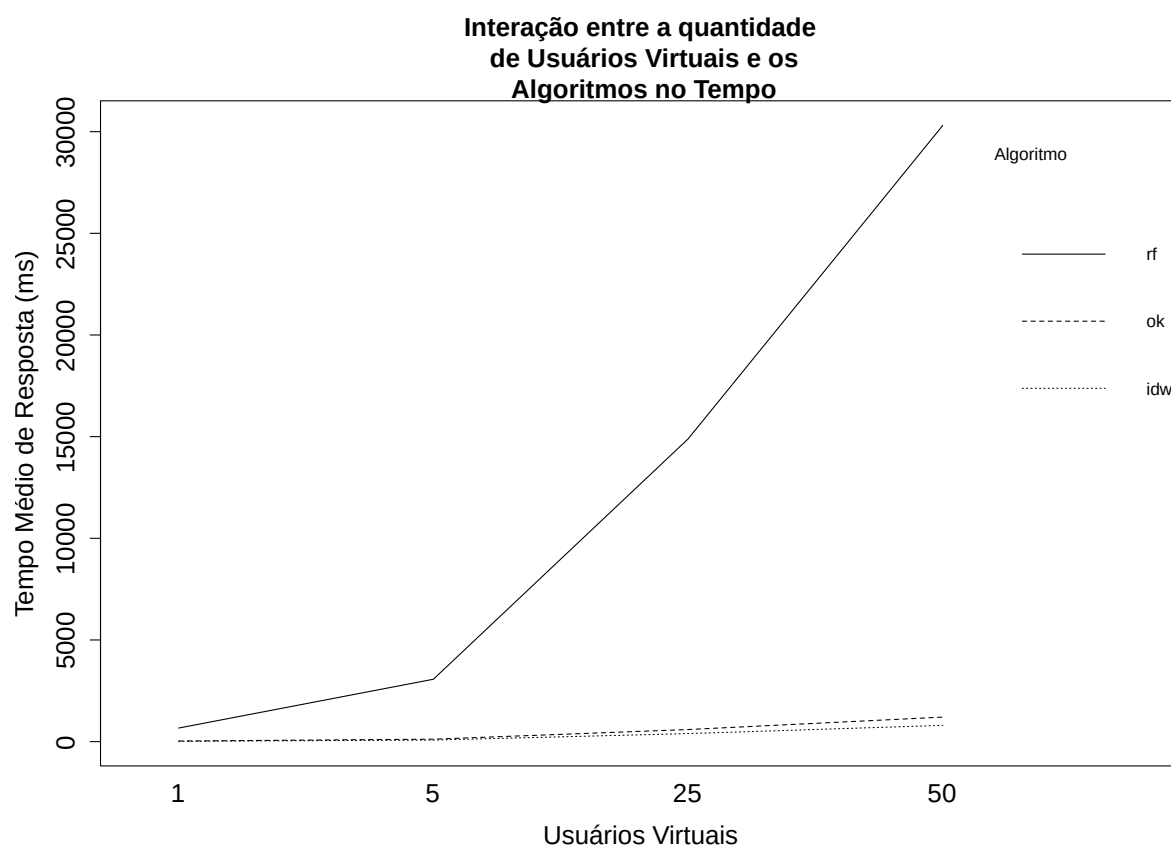


Figura 15 – Interação entre a quantidade de **Usuários Virtuais** e os **Algoritmos** no tempo

Fonte: Próprio Autor

5 Conclusão

A importância do presente estudo está no desenvolvimento de um serviço escalável, que pode ser aplicado na interpolação espacial que se fundamenta na computação móvel e em nuvem. Nesse sentido, desenvolveu-se um serviço Web, através da linguagem R ([R Core Team, 2018](#)), que se utiliza do estilo arquitetônico REST para disponibilizar recursos que geram modelos e interpolações espaciais. Tal serviço foi concebido para ser genérico, já que implementa três diferentes algoritmos de interpolação espacial (Krigagem, IDW e RF), que podem ser escolhidos pelo usuário do serviço no momento da geração do modelo.

Neste trabalho pôde-se implementar e avaliar o desempenho do serviço de interpolação espacial a partir de um experimento completamente fatorial, levando em conta quatro fatores: quantidade de pontos para composição do modelo espacial; os algoritmos de interpolação espacial; o número de usuários requisitando o serviço; e a proporção de requisições para geração do modelo e para interpolação. Outro ponto a se destacar é que este serviço foi implantado em ambiente de produção, além de ter sido integrado à um aplicação móvel.

De modo geral, os resultados da avaliação mostraram que os algoritmos clássicos – Krigagem e IDW – apresentaram uma melhor performance com relação ao tempo de resposta se comparado com o algoritmo RF, frente a variação e combinação dos outros três fatores avaliados. Este resultado indica a necessidade de um compromisso entre desempenho e acurácia na escolha do algoritmo para interpolação, pois enquanto o algoritmo RF tem obtido melhor acurácia do que os métodos convencionais ([JÚNIOR et al., 2019](#)), seu uso tem alto impacto em termos de tempo de resposta ([JÚNIOR et al., 2019](#); [LI; HEAP, 2014](#)). Tal resultado, indica que a busca por uma maior precisão deve vir acompanhada de uma cuidadosa implantação do serviço de interpolação - utilizando técnicas como replicação e balanceamento de carga - para que o usuário do serviço não venha a sofrer com o desempenho do algoritmo RF.

O serviço de interpolação desenvolvido contribui diretamente para a melhoria do aplicativo Aquaprev. Já que promove uma solução escalável e reutilizável para a interpolação espacial, que culmina num melhor desempenho do cálculo do tempo de irrigação, fornecido pelo Aquaprev. Desta forma, o serviço poderá tornar o aplicativo mais robusto com relação ao atendimento de usuários em larga escala.

Outro aspecto relevante, é que o serviço é genérico o suficiente para permitir não só a determinação de evapotranspiração de referência, mas também o cálculo de outras variáveis ambientais como temperatura ambiente, umidade relativa do ar,

velocidade do vento etc. Desta forma, abre-se um leque de possibilidades no desenvolvimento de soluções digitais para a agricultura, já que estas novas soluções poderão utilizar a interpolação como serviço, diminuindo seu tempo de desenvolvimento.

Como limitação do serviço, pode-se destacar que este é baseado exclusivamente na latitude e longitude. Esta limitação pode tornar o serviço inapropriado para interpolar variáveis que sejam fortemente impactadas por outras variáveis, tal como a precipitação cuja interpolação pode ser melhorada ao se considerar aspectos como altitude, relevo e cobertura do solo, por exemplo ([GOOVAERTS, 2000](#)). Contudo, esta limitação pode ser sobrepujada no futuro estendendo o serviço pelo emprego de métodos de interpolação que permitam o uso de variáveis auxiliares ([LI; HEAP, 2008](#)).

5.1 Contribuições

Este trabalho, torna-se pioneiro no melhor do nosso conhecimento, pelo fato de desenvolver um serviço de interpolação espacial que propicia uso em escala, por parte dos usuários/serviços que venham a se conectar para obter a interpolação espacial dos dados meteorológicos.

O código produzido, neste trabalho, foi enviado para registro junto ao Instituto Nacional de Propriedade Intelectual.

5.2 Trabalhos Futuros

Como trabalho futuro, pretende-se realizar ajustes no serviço, deixando as parametrizações dos algoritmos a disposição do requisitante. Além disso, almeja-se adicionar outros algoritmos de interpolação espacial para aumentar a versatilidade do serviço. Outro ponto que se espera fazer, é realizar novos experimentos de desempenho, desta vez em ambiente Docker, de modo a observar se o uso de um contêiner tem impacto sobre o desempenho do serviço. Pretende-se ainda, investigar a fundo cada etapa de execução do código do serviço, para delimitar, se houver, o ponto de gargalo, proporcionando uma melhor visão do problema e sua possível solução.

Referências

- ABREU, J. M. F. *Modelos de Data Mining para suporte a avaliações sensoriais do vinho*. Tese (Doutorado) — Instituto Superior de Engenharia do Porto, 2016. Citado na página 27.
- APPELHANS, T. et al. Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania. *Spatial Statistics*, v. 14, p. 91–113, nov. 2015. ISSN 22116753. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S2211675315000482>>. Citado 2 vezes nas páginas 14 e 20.
- ARCURI, A. Restful api automated test case generation with evomaster. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, ACM, v. 28, n. 1, p. 3, 2019. Citado na página 18.
- CARVALHO, J. de; VIEIRA, S. Avaliação e comparação de estimadores de krigagem para variáveis agrônômicas-uma proposta. *Embrapa Informática Agropecuária-Documents (INFOTECA-E)*, Campinas: Embrapa Informática Agropecuária, 2001., 2001. Citado na página 21.
- CASELLA, G.; BERGER, R. L. Inferência estatística. *São Paulo: Cengage Learning*, 2010. Citado na página 22.
- CHAGAS, C. d. S. et al. Mapeamento digital da ctc em solos do semiárido brasileiro. In: IN: CONGRESSO BRASILEIRO DE CIÊNCIA DO SOLO, 35., 2015, NATAL. O SOLO E SUAS *Embrapa Solos-Artigo em anais de congresso (ALICE)*. [S.I.], 2015. Citado na página 27.
- CUNHA, J. et al. A high-throughput shared service to estimate evapotranspiration using landsat imagery. *Computers & Geosciences*, 10 2019. Citado 2 vezes nas páginas 19 e 30.
- DONGES, N. *The Random Forest Algorithm*. 2018. Disponível em: <<https://towardsdatascience.com/the-random-forest-algorithm-d457d499ffcd>>. Citado na página 21.
- GOOVAERTS, P. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of hydrology*, Elsevier, v. 228, n. 1-2, p. 113–129, 2000. Citado na página 45.
- JAIN, R. *The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling*. [S.I.]: John Wiley & Sons, 1990. Citado 4 vezes nas páginas 22, 23, 34 e 35.
- JAKOB, A. A. E.; YOUNG, A. F. O uso de métodos de interpolação espacial de dados nas análises sociodemográficas. *Anais*, p. 1–22, 2016. Citado 2 vezes nas páginas 14 e 20.
- JUSTA, T. S. d. *O consumo de filmes em cinemas no Brasil: uma análise de florestas aleatórias*. Tese (Doutorado) — Universidade de São Paulo, 2019. Citado na página 21.

JÚNIOR, J. C. da S. et al. Random forest techniques for spatial interpolation of evapotranspiration data from brazilian's northeast. *Computers and Electronics in Agriculture*, v. 166, p. 105017, 2019. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169919302315>>. Citado 7 vezes nas páginas 14, 20, 21, 27, 33, 37 e 44.

KOUBAA, A. Service-oriented software architecture for cloud robotics. *arXiv preprint arXiv:1901.08173*, 2019. Citado na página 17.

LI, J.; HEAP, A. D. *A review of spatial interpolation methods for environmental scientists*. Canberra: Geoscience Australia, 2008. OCLC: 312275544. ISBN 978-1-921498-30-5 978-1-921498-28-2. Citado 2 vezes nas páginas 21 e 45.

LI, J.; HEAP, A. D. Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software*, v. 53, p. 173–189, mar. 2014. ISSN 13648152. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1364815213003113>>. Citado 3 vezes nas páginas 14, 33 e 44.

LI, J. et al. Application of machine learning methods to spatial interpolation of environmental variables. *Environmental Modelling & Software*, v. 26, n. 12, p. 1647–1659, dez. 2011. ISSN 13648152. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1364815211001654>>. Citado 2 vezes nas páginas 14 e 20.

LOPES, I.; MELO, J. M. M. D.; LEAL, B. G. ESPACIALIZAÇÃO DA TEMPERATURA DO AR PARA A REGIÃO DO SUBMÉDIO SÃO FRANCISCO. *IRRIGA*, v. 22, n. 1, p. 177, maio 2017. ISSN 1808-8546, 1413-7895. Disponível em: <<http://irriga.fca.unesp.br/index.php/irriga/article/view/2237>>. Citado na página 14.

MALDONADO, W.; VALERIANO, T. T. B.; ROLIM, G. de S. Evapo: A smartphone application to estimate potential evapotranspiration using cloud gridded meteorological data from nasa-power system. *Computers and Electronics in Agriculture*, v. 156, p. 187 – 192, 2019. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S016816991830680X>>. Citado na página 20.

MALESHKOVA, M. et al. Smart web services (smartws)–the future of services on the web. *arXiv preprint arXiv:1902.00910*, 2019. Citado na página 18.

NETO, B. B.; SCARMINIO, I. S.; BRUNS, R. E. *Como Fazer Experimentos-: Pesquisa e Desenvolvimento na Ciência e na Indústria*. [S.l.]: Bookman Editora, 2010. Citado na página 22.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2018. Disponível em: <<https://www.R-project.org/>>. Citado 2 vezes nas páginas 24 e 44.

SANTOS, I. de S.; NETO, P. d. A. dos S. Automação de testes de desempenho e estresse com o jmeter. 2008. Citado na página 33.

SERMAN, D. V. Orientação a projetos: uma proposta de desenvolvimento de uma arquitetura orientada a serviços. *JISTEM-Journal of Information Systems and Technology Management (Online)*, v. 7, n. 3, p. 619–638, 2010. Citado na página 17.

SILVA, A. C. d. Virtualização: um estudo sobre conceitos e técnicas. Universidade Federal Fluminense, 2017. Citado 2 vezes nas páginas 18 e 19.

SILVA, M. A. V. *Meteorologia e climatologia*. [S.l.: s.n.], 2006. Citado na página 14.

TRINDADE, L. V. P.; COSTA, L. H. M. Análise do desempenho da virtualização leve para ambientes com edge computing baseada em nfv. In: SBC. *Anais do XXXVI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*. [S.l.], 2018. Citado 2 vezes nas páginas 18 e 19.

VARGAS, T. D. et al. Aplicação do interpolador idw para elaboração de mapas hidrogeológicos paramétricos na região da serra gaúcha. *Scientia cum Industria*, v. 6, n. 3, p. 38–43, 2019. Citado na página 14.

XAVIER, A. C.; KING, C. W.; SCANLON, B. R. Daily gridded meteorological variables in Brazil (1980-2013): DAILY GRIDDED METEOROLOGICAL VARIABLES IN BRAZIL (1980-2013). *International Journal of Climatology*, v. 36, n. 6, p. 2644–2659, maio 2016. ISSN 08998418. Disponível em: <<http://doi.wiley.com/10.1002/joc.4518>>. Citado na página 33.

ZHANG, Y. et al. Covering-based web service quality prediction via neighborhood-aware matrix factorization. *IEEE Transactions on Services Computing*, IEEE, 2019. Citado na página 18.