

**FERNANDO MALAVASI
LUCAS OLIVEIRA REIS**

**SIMULAÇÃO E OTIMIZAÇÃO DO ATENDIMENTO
EMERGENCIAL DE UMA DISTRIBUIDORA DE
GÁS NATURAL**

São Paulo
2022

**FERNANDO MALAVASI
LUCAS OLIVEIRA REIS**

**SIMULAÇÃO E OTIMIZAÇÃO DO ATENDIMENTO
EMERGENCIAL DE UMA DISTRIBUIDORA DE
GÁS NATURAL**

Trabalho apresentado à Escola Politécnica
da Universidade de São Paulo para ob-
tenção do Título de Engenheiro Mecatrônico.

São Paulo
2022

**FERNANDO MALAVASI
LUCAS OLIVEIRA REIS**

**SIMULAÇÃO E OTIMIZAÇÃO DO ATENDIMENTO
EMERGENCIAL DE UMA DISTRIBUIDORA DE
GÁS NATURAL**

Trabalho apresentado à Escola Politécnica
da Universidade de São Paulo para ob-
tenção do Título de Engenheiro Mecatrônico.

Orientador:
André Kubagawa Sato

São Paulo
2022

RESUMO

Segundo dados divulgados pela Agência Nacional de Petróleo e Gás (ANP), a produção de gás natural no Brasil atinge cerca de 134 milhões de metros cúbicos todos os dias. Parte desta demanda é destinada a população geral através de distribuidoras de gás natural encanado. No entanto, um dos maiores entraves para o setor é o vazamento de gás. Sendo assim, a Agência Reguladora de Serviços Públicos do Estado de São Paulo (ARSESP) criou uma norma para garantir a segurança da população, o desafio proposto pela agência é que as distribuidoras devem atender os vazamentos em até 60 minutos após a notificação deste vazamento, caso contrário, uma multa é aplicada à distribuidora. A aplicação de uma norma como esta permite desenvolver estratégias para atacar este problema da forma mais eficiente possível, que, apesar de serem promissoras soluções, representam um grande desafio. O maior desafio é o dimensionamento de uma equipe para atender os vazamentos. Isso pode ser realizado por meio de um algoritmo capaz de selecionar de maneira ótima a localização e a escala de trabalho da equipe responsável pelos atendimentos. Assim, este projeto buscou desenvolver algoritmos de otimização e clusterização que, em conjunto, poderão resultar em um dimensionamento ótimo da equipe. Além disso, este projeto avaliou a qualidade do dimensionamento através de uma simulação da operação. Este projeto é concluído com o dimensionamento da equipe realizado e validado pelo simulador.

Palavras-Chave – Alocação de recursos, Constraint Programming, Agglomerative Clustering, Roteirização

ABSTRACT

According to data released by the National Oil and Gas Agency (ANP), natural gas production in Brazil reaches around 134 million cubic meters every day. Part of this demand is destined to the general population through piped natural gas distributors. However, one of the biggest obstacles for the sector is gas leakage. Therefore, the Regulatory Agency for Public Services of the State of São Paulo (ARSESP) created a rule to guarantee the safety of the population, the challenge proposed by the agency is that the distributors must respond to leaks within 60 minutes after notification of this leak, otherwise, a fine is applied to the distributor. The application of a standard like this allows the development of strategies to tackle this problem in the most efficient way possible, which, despite being promising solutions, represent a great challenge. The biggest challenge is sizing a team to deal with leaks. This can be done using an algorithm capable of optimally selecting the location and work schedule of the team responsible for solving the problem. Thus, this project sought to develop optimization and clustering algorithms that, together, may result in optimal team sizing. In addition, this project evaluated the dimensioning quality through an operation simulation. This project is concluded with the sizing of the team done and validated by the simulator.

Keywords – Resource allocation, Constraint Programming, Agglomerative Clustering, Vehicle Routing

LISTA DE FIGURAS

1	RMSP recortada em zonas	20
2	Média diária de atendimentos por zona	22
3	Média de chamados por dia da semana	23
4	Distribuição de chamados ao longo do dia	24
5	Regiões selecionadas como de baixa densidade (zonas periféricas)	25
6	Distribuição geográfica dos chamados gerados	26
7	Média de chamados por dia por município	27
8	Comparativo da média de chamados por dia por município para 7 dias . .	28
9	Comparativo da média de chamados por dia por município para 70 dias .	28
10	Comparativo da média de chamados por dia por município para 700 dias .	29
11	Comparativo da média de chamados por dia por município para 1400 dias .	29
12	Comparativo da média de chamados por dia por zona para 7 dias	30
13	Comparativo da média de chamados por dia por zona para 70 dias	30
14	Comparativo da média de chamados por dia por zona para 700 dias	31
15	Comparativo da média de chamados por dia por zona para 1400 dias	31
16	Comparativo da distribuição horária dos chamados para 7 dias	32
17	Comparativo da distribuição horária dos chamados para 70 dias	32
18	Comparativo da distribuição horária dos chamados para 700 dias	33
19	Comparativo da distribuição horária dos chamados para 1400 dias	33
20	Tabela mostrando todas as possíveis escalas em um dia	34
21	Localização dos técnicos que compõem uma escala com 1 técnico. Na legenda, 'C0' representa o centroide dos pontos identificados com '0'	35

22	Localização dos técnicos que compõem uma escala com 3 técnicos. Na legenda, 'C0', 'C1' e 'C2' representam os centroides dos pontos identificados com '0', '1' e '2' respectivamente	35
23	Localização dos técnicos que compõem uma escala com 4 técnicos. Na legenda, 'C0', 'C1', 'C2' e 'C3' representam os centroides dos pontos identificados com '0', '1', '2' e '3' respectivamente	36
24	Localização dos técnicos que compõem uma escala com 6 técnicos. Na legenda, 'C0', 'C1', 'C2', 'C3', 'C4' e 'C5' representam os centroides dos pontos identificados com '0', '1', '2', '3', '4' e '5' respectivamente	36
25	Localização dos técnicos que compõem uma escala com 7 técnicos. Na legenda, 'C0', 'C1', 'C2', 'C3', 'C4', 'C5' e 'C6' representam os centroides dos pontos identificados com '0', '1', '2', '3', '4', '5' e '6' respectivamente	37

LISTA DE TABELAS

1	Primeiras linhas da base de propriedades das zonas	21
2	Primeiras linhas da base de chamados gerados	26
3	Primeiras linhas da base das escalas de cada técnico	38
4	Primeiras linhas da base de resultados da simulação	40
5	Tempo de atendimento médio para cada município	41
6	Primeiras linhas da base de resultados da simulação	43
7	Tempo de atendimento médio para cada município	44

SUMÁRIO

1	Introdução	9
1.1	Motivação	9
1.2	Objetivos gerais	10
1.3	Estrutura do texto	10
2	Revisão Bibliográfica	11
3	Metodologia	13
3.1	Gerador de chamados	13
3.1.1	Média de chamados por dia por município	13
3.1.2	Média de chamados por dia da semana	13
3.1.3	Média de chamados por dia por zona	14
3.1.4	Distribuição horária dos chamados	14
3.1.5	Tratamento das zonas de baixa densidade demográfica	14
3.1.6	Validação do gerador	15
3.2	Quantidade ideal de técnicos	15
3.2.1	Otimizador de escalas	16
3.3	Localização ideal dos técnicos	17
3.4	Simulador	17
4	Resultados	19
4.1	Dados utilizados	19
4.1.1	Zonas de atendimento	19
4.1.2	Propriedades das zonas	20
4.1.3	Distribuição horária dos chamados	23

4.1.4	Tratamento das zonas de baixa densidade demográfica	24
4.2	Gerador de chamados	25
4.2.1	Validação do gerador	26
4.3	Otimizador de escalas	34
4.4	Localização ideal dos técnicos	34
4.5	Fator de correção da distância entre chamado e técnico	38
4.6	Estruturação dos resultados	38
4.7	Resultado Preliminar	39
4.7.1	Discussão de Resultados	41
4.8	Resultado Corrigido	42
4.8.1	Discussão de Resultados	44
5	Discussões	46
6	Conclusão	48
Referências		49

1 INTRODUÇÃO

A produção de gás natural no Brasil vem crescendo todos os anos, a cada dia são produzidos cerca de 134 milhões de metros cúbicos, segundo dados divulgados pela Agência Nacional de Petróleo e Gás (ANP) [1]. Este aumento da produção se deve principalmente a crescente demanda por gás natural no Brasil, principalmente na RMSP (Região Metropolitana de São Paulo). Grande parte desta demanda é destinada à indústria, porém existe também uma alta demanda pela população geral.

Pensando nessa crescente demanda pela população, algumas distribuidoras fornecem gás natural encanado. A grande vantagem em optar pelo gás encanado é que o fornecimento é contínuo, dessa forma o cliente não precisa se preocupar com a recarga ou com a falta de gás repentina em sua residência [2]. Porém, um dos maiores problemas enfrentados pela distribuidora é o vazamento de gás, que pode ocorrer tanto na rede de distribuição quanto nos pontos finais, como no interior das residências ou nas cozinhas de restaurantes. Quando é confirmado um vazamento, o órgão regulador, ARSESP (Agência Reguladora de Serviços Públicos do Estado de São Paulo), determina que a distribuidora atenda a esta emergência em um tempo médio máximo 60 minutos para cada município [3, 4]. Portanto, a distribuidora precisa determinar quantos técnicos ela deve ter disponível e em quais localizações geográficas distribuí-los, para que seja possível atender estas emergências dentro do tempo regulamentado pelo órgão regulador. Pensando nisso, será desenvolvido um simulador para emular esta operação, e um otimizador para auxiliar no dimensionamento da equipe.

1.1 Motivação

A razão deste trabalho é possibilitar o atendimento emergencial dentro dos 60 minutos estabelecidos pelo órgão regulador para todos os municípios da RMSP. Esta medida é fundamental para garantir a segurança da população, uma vez que a operação lida com um produto altamente volátil como o gás natural. A ideia é atender esta meta por meio de uma

otimização de parâmetros como quantidade, localização e escalas dos técnicos em operação e uma eventual simulação da operação para confirmar que estes chamados sejam atendidas dentro do tempo estabelecido pelo órgão regulador. Tal dimensionamento otimizado garante maior eficiência à distribuidora, reduzindo assim os custos com funcionários e minimizando eventuais valores de multa por não cumprimento das normas do órgão regulador. Além de melhorar a qualidade do serviço prestado aos clientes, reduzindo o tempo de espera caso uma emergência ocorra.

1.2 Objetivos gerais

Criar um simulador capaz de emular um período típico de operação, levando em consideração parâmetros como a frequência de surgimento dos chamados, a localização e o horário de cada chamado, o número de técnicos disponíveis para atender estes chamados, localização geográfica de cada técnico, a escala de trabalho de cada técnico e o tempo que ele leva para resolver um atendimento e, por fim, calcular o tempo de deslocamento dos técnicos até os locais das ocorrências. Além disso, o simulador deve sinalizar se este tempo de deslocamento até a chegada do técnico está dentro dos 60 minutos estipulados pelo órgão regulador.

Depois, será realizado um processo de otimização sobre os parâmetros do simulador, como quantidade, localização e escala de cada técnico, com o objetivo de minimizar o custo operacional e melhorar a qualidade de atendimento, balanceando os custos com os técnicos e procurando atender a maior parte dos clientes dentro do tempo regulamentado, minimizando a chance de ocorrência de multa.

1.3 Estrutura do texto

A estrutura da monografia segue com uma apresentação sobre o Estado da Arte, a Metodologia, os Resultados, as Discussões e uma Conclusão. No Capítulo 2 são citadas alguns dos principais trabalhos relacionados ao tema e como eles se relacionam a este trabalho. Já no Capítulo 3, são apresentadas algumas técnicas relevantes existentes na atualidade e como elas podem ser utilizadas em conjunto neste trabalho. No Capítulo 4 podem ser observados todos os resultados obtidos ao longo do projeto. O Capítulo 5 traz alguns pontos de melhoria e refinamentos que podem ser implementados em projetos futuros. Por fim, o Capítulo 6 é a conclusão do trabalho.

2 REVISÃO BIBLIOGRÁFICA

A simulação deve ser capaz de criar chamados aleatoriamente no mapa e, dadas as posições dos técnicos livres, ou seja, aqueles que não estão atendendo um chamado ou que não estão em deslocamento até um chamado, selecionar aquele que consiga chegar mais rapidamente ao local da ocorrência.

Neste caso, o simulador se comporta como um sistema a eventos discretos (SED), o que envolve a criação de agentes que interagem com o ambiente através de um modelo de compartilhamento de recursos. Cada agente alocado tem seu tempo de indisponibilidade bem definido, como aplicado em [5], no caso de transporte de madeira com caminhões de uma transportadora. Para isso, deve-se levar em consideração a interação entre todos os agentes disponíveis para que o simulador performe de forma estável e robusta como um todo, como aplicado em [6]. Para tanto, a linguagem Python de programação fornece uma solução viável computacionalmente, pois possibilita simular a interação entre vários agentes simultaneamente por longos períodos de tempo sem que seja necessário um recurso computacional elevado [7]. A simulação tal qual foi descrita anteriormente é uma aplicação da Teoria das Filas, que consiste em estudar a probabilidade da formação de filas por meio de pesquisas exatas [8]. A ideia central da Teoria de Filas é que uma fila ocorre sempre que a demanda por um recurso é maior do que a disponibilidade desse recurso no sistema, o que pode ser facilmente traduzido ao problema que este trabalho se propõe a resolver. Além disso, é possível utilizar como inspiração uma biblioteca desenvolvida em R que implementa esta lógica de simulação, demonstrado por [9]. Nela são apresentados tanto casos com uma quantidade fixa de prestadores de serviço quanto uma quantidade variável no tempo. Isto torna a simulação mais condizente com o problema de simulação dos técnicos, dado que eles tem escalas que duram um período limitado durante o dia.

Outra aplicação de teoria das filas foi em [10], que simula a situação de passageiros e taxistas em um aeroporto, situação parecida a de chamados e técnicos da companhia de gás.

Para escolher o técnico que atende a um chamado no menor tempo possível, é necessário

solucionar um problema de roteirização [11]. Pode-se utilizar um Sistema de Informação Geográfica (SIG), que é um conjunto de programas, metodologias e dados integrados, capaz de promover a análise de dados georreferenciados [12, 13], como aplicado em [14], trabalho que expõe a dificuldade na criação de um simulador de rotas para uma operação logística que tem por objetivo transportar cargas de um local até outro. Para solucionar este problema, o autor avaliou duas abordagens na criação dos modelos, a primeira delas considera apenas os locais de interesse conectados de forma direta, com a distância entre um local e outro calculada utilizando dados históricos da região analisada. A outra abordagem considera também as principais vias que poderiam ser utilizadas para se deslocar entre os locais de interesse. Para validar o modelo foram utilizados os dados de localização obtidos com um GPS (*Global Positioning System*) de um veículo em um dia típico, depois, o autor comparou esses dados de distância e tempo de rota com os dados obtidos no simulador.

3 METODOLOGIA

A seguir serão apresentadas as metodologias e fundamentações teóricas envolvendo o gerador de chamados, como foram obtidas a quantidade e a localização ideal dos técnicos e o simulador. O objetivo deste capítulo é esclarecer como os dados foram utilizados e quais foram as soluções encontradas em cada tópico.

3.1 Gerador de chamados

O primeiro passo para solucionar o problema proposto neste trabalho é, dadas as bases de dados históricos da companhia, que serão explicadas em detalhe na Seção 4, gerar uma base de chamados fictícios que atendam às propriedades dessas bases históricas.

3.1.1 Média de chamados por dia por município

A primeira propriedade que deve ser mantida é a média de chamados por dia por município. Para que os chamados fictícios gerados sejam fidedignos aos chamados reais da companhia de gás, é necessário garantir que a quantidade média de chamados por dia para cada município seja equivalente em ambas as bases.

3.1.2 Média de chamados por dia da semana

A fim de garantir ainda mais que estes chamados fictícios sejam equivalentes aos chamados reais, deve-se garantir que a volumetria de chamados por dia da semana também seja preservada. Isso é importante pois existe uma diferença significativa na volumetria para cada dia da semana, o que impacta diretamente na alocação dos técnicos ao longo do dia.

3.1.3 Média de chamados por dia por zona

Além de garantir que os chamados gerados sejam equivalentes aos chamados reais em relação aos municípios, é necessário garantir que sejam equivalentes em relação às zonas também. Como o município de São Paulo, que é responsável pela maioria dos chamados, é subdividido em várias zonas, deve-se garantir que cada zona mantenha esta propriedade.

3.1.4 Distribuição horária dos chamados

Cada chamado tem uma localização e um horário, portanto, além de garantir as propriedades de localização, deve-se garantir também a distribuição horária dos chamados. Na Seção 4.1.3 será apresentada uma função de densidade de probabilidade para determinar como serão determinados os horários de um chamado. Além disso, será utilizada uma distribuição uniforme para determinar a minutagem dentro de cada intervalo de uma hora.

3.1.5 Tratamento das zonas de baixa densidade demográfica

Além de garantir as propriedades citadas nas Seções 3.1.1-3.1.4, ainda é necessário tratar as zonas que apresentam baixa volumetria. Zonas com grande parte de sua região com pouca densidade demográfica apresentam um problema para o simulador, pois acabam sendo gerados chamados em regiões distantes dos grandes centros, onde ninguém mora, forçando os técnicos a fazerem longas jornadas e desperdiçar tempo em uma situação que não aconteceria na vida real. Para solucionar esse problema, com a ajuda da plataforma ReSolution [15], os dados de densidade demográfica da RMSP foram levados em consideração. O surgimento de chamados nas regiões delimitadas como problemáticas segue uma distribuição normal centrada no centro populacional da região e a variância seguindo a dispersão destes centro demográficos. O surgimento de chamados nas demais zonas foram definidos a partir de uma distribuição uniforme.

As zonas de baixa densidade demográfica escolhidas são as que englobam os municípios de São Bernardo do Campo, Guarulhos, Mogi das Cruzes, Santo André, Diadema, Arujá, Suzano, Santana de Parnaíba, Cotia, Embu das Artes, Itapecerica da Serra, Itapevi, Ribeirão Pires e Rio Grande da Serra. Essas regiões podem ser consultadas na Seção 4.1.4.

3.1.6 Validação do gerador

Uma vez gerados os chamados, é necessário comparar se as propriedades de fato se mantiveram. Para a etapa de validação, será utilizado o erro percentual entre os valores real e gerado.

Uma das perguntas centrais nesta etapa de validação é determinar qual deve ser a quantidade de chamados gerados para se obter um resultado estatisticamente significativo. Esta preocupação surge do fato de que alguns municípios apresentam uma volumetria muito baixa, como discutido na Seção 3.1.5.

A diferença da volumetria de São Paulo quando comparada aos municípios de menor quantidade de chamados chega a ser ordens de grandeza maior. Portanto, é necessário gerar uma quantidade suficientemente grande de chamados para que os dados sejam significativos para os municípios de menor volumetria.

Feito isso, será feita uma comparação entre cenários obtidos a partir de diferentes quantidades de dias simulados, então será escolhida uma quantidade de dias para trabalhar ao longo do trabalho.

3.2 Quantidade ideal de técnicos

Até agora foram explicados os métodos através dos quais a base de chamados será obtida. O próximo passo é determinar a menor quantidade possível de técnicos que consigam atender a todos os chamados dentro das normas estabelecidas pelo órgão regulador. Vale relembrar que a norma diz: "o tempo de atendimento médio dos chamados deve ser de no máximo 60 minutos para cada município que compõe a RMSP".

Sabe-se ainda que a quantidade total de técnicos disponíveis em um dia depende das escalas desses técnicos, ou seja, depende de suas jornadas de trabalho. Dessa forma, o objetivo agora é determinar qual deve ser a quantidade mínima de técnicos disponíveis em cada intervalo de hora. A partir deste resultado, será possível obter a quantidade total de técnicos por meio de um algoritmo de otimização que será explicado na Seção 3.2.1.

Para auxiliar na resolução deste problema, uma série de premissas foram estabelecidas:

- é necessário determinar quantos técnicos devem estar disponíveis em cada intervalo de hora;
- sabe-se que se a quantidade de chamados em uma zona ultrapassar a quantidade de

técnicos disponíveis nessa zona, haverá atraso no atendimento;

- será adotada uma probabilidade aceitável de ter mais chamados numa zona do que técnicos disponíveis, isto é, uma probabilidade aceitável de ter atraso; e
- dada uma quantidade esperada de chamados em um intervalo de hora, deseja-se saber X tal que a probabilidade de pelo menos X desses chamados caírem na mesma zona seja menor ou igual a probabilidade aceitável.

A partir destas premissas foi determinada a quantidade mínima necessária de técnicos para cada intervalo de hora, o que servirá de entrada para o otimizador de escalas, que será explicado na Seção 3.2.1.

Como a quantidade esperada de chamados em um intervalo de hora é conhecida, a quantidade máxima de chamados que poderiam cair em cada zona foi determinada de forma que a probabilidade de cair um chamado a mais fosse menor ou igual a probabilidade aceitável. Assim, a partir da segunda premissa, foi possível encontrar a quantidade de técnicos que deveriam estar presentes em cada zona em cada intervalo de hora. Portanto, a quantidade total de técnicos que deveriam estar disponíveis em cada intervalo de hora, que é um parâmetro fundamental para a Seção 3.2.1, pode ser obtida somando a quantidade de técnicos em todas as zonas.

3.2.1 Otimizador de escalas

A partir da quantidade mínima de técnicos que devem estar disponíveis em cada intervalo de hora, o desafio é determinar quantos técnicos devem ser alocados em cada uma das possíveis escalas de forma que, para cada intervalo de hora, a quantidade mínima de técnicos nesse intervalo seja satisfeita. Para tanto, foi utilizado um algoritmo de *Constraint Programming*, que é o nome dado à identificação de soluções viáveis a partir de um conjunto muito grande de candidatos, sendo que o problema pode ser modelado em termos de restrições arbitrárias [16].

O que o otimizador faz é determinar quantos técnicos devem estar em cada escala com o objetivo de minimizar a quantidade total de técnicos, mas atendendo ainda à quantidade mínima de técnicos por hora.

3.3 Localização ideal dos técnicos

Uma vez encontrada a quantidade ideal e a escala de cada um dos técnicos, o último passo antes da simulação é determinar onde estes técnicos devem estar localizados. Para isso, foi utilizado um algoritmo de clusterização denominado *Agglomerative Clustering*, cujo funcionamento se dá mesclando recursivamente pares de clusters de dados da amostra [17].

Para determinar a localização dos técnicos, os chamados foram clusterizados em quantidades de clusters equivalentes à quantidade de técnicos em cada escala. Dessa forma, é possível garantir que os técnicos de uma determinada escala estão distribuídos da melhor forma possível. Depois que os clusters foram obtidos, o centroide, isto é, a localização do técnico, foi calculado utilizando a mediana geométrica, que é uma métrica capaz de fornecer um centroide com a propriedade de que o somatório das distâncias de cada ponto do cluster até o centroide seja minimizado. Depois, é criada uma planilha de alocação contendo as localizações e escalas dos técnicos.

3.4 Simulador

O objetivo da simulação é validar tudo o que foi feito até agora. O simulador utiliza como entradas os chamados gerados e a planilha com a alocação dos técnicos. O resultado da simulação é uma planilha contendo todos os chamados simulados, os técnicos responsáveis pelo atendimento de cada chamado, o tempo que o técnico levou para chegar até o local da ocorrência e o tempo que a ocorrência levou para ser concluída. Com essas informações, é possível calcular o tempo médio de atendimento dos chamados para cada município e checar se o dimensionamento de equipe atendeu à norma do órgão regulador.

Dada uma quantidade de dias que se deseja simular, os chamados que ocorreram nesses dias e a alocação dos técnicos, o simulador é capaz de disparar chamados e atribuir técnicos para o atendimento desses chamados.

A simulação é feita através de uma iteração sobre os dias e os minutos do dia. Para cada minuto o simulador filtra apenas os chamados que ocorreram naquele minuto e checa se existe algum técnico disponível para realizar o atendimento. Normalmente a quantidade de técnicos disponíveis para atender um chamado é alta, portanto, o simulador escolhe, dentre todos os técnicos disponíveis, aquele que chega mais rápido até o local do chamado. Assim que um técnico é atribuído para um atendimento, ele fica bloqueado por um tempo equivalente ao tempo de deslocamento até o chamado mais o tempo que o chamado demora

para ser resolvido. Uma vez que o técnico finaliza o chamado, sua localização é atualizada com a localização do chamado que ele acabou de atender, assim, a simulação continua até o último minuto do último dia.

Algumas premissas foram implementadas para que a operação pudesse ser mais eficiente e condizente com a realidade da companhia. Por exemplo, ao procurar pelo técnico que chega mais rápido até o local da ocorrência, o simulador checa também se algum técnico que está prestes a acabar o serviço consegue chegar mais rápido do que algum técnico que esteja livre. Com isso, se o chamado ocorrer ao lado de um técnico ocupado, o simulador é inteligente o suficiente para atribuir o chamado a este técnico ao invés de chamar um técnico que esteja mais distante do chamado e que demoraria mais para chegar.

Além disso, a distância entre o chamado e o técnico é calculada a partir da distância em linha reta entre eles multiplicada por um fator de correção que será apresentado na Seção 4.5. Feito isso, o tempo de deslocamento do técnico até o local da ocorrência é determinado a partir de uma velocidade média de deslocamento na RMSP de 20km/h [18].

Ao longo do trabalho foram utilizados dois parâmetros fixos, a velocidade média de deslocamento na RMSP, discutida anteriormente, e o tempo de atendimento de um chamado. O tempo de atendimento nada mais é do que, a partir do momento que o técnico chega no local da ocorrência, o tempo que ele demora para resolver o problema.

4 RESULTADOS

Neste capítulo serão apresentados os dados utilizados, como os resultados são obtidos e algumas comparações entre eles. Depois, no Capítulo 5 serão colocadas algumas características em comum entre os resultados e alguns pontos de melhoria para próximos trabalhos neste tema.

4.1 Dados utilizados

Para o desenvolvimento do projeto foram utilizadas bases de dados contendo informações históricas e estatísticas sobre o tema proposto. Esta Seção tem por objetivo apresentar estas bases e expor alguns entendimentos sobre os dados.

4.1.1 Zonas de atendimento

A inspiração para a divisão das zonas na Figura 1 veio da divisão original da companhia de gás. O município de São Paulo foi quebrado em bairros, onde cada bairro é considerado como uma zona específica. Já os municípios periféricos da RMSP não foram quebrados, o município todo é classificado como uma zona única.

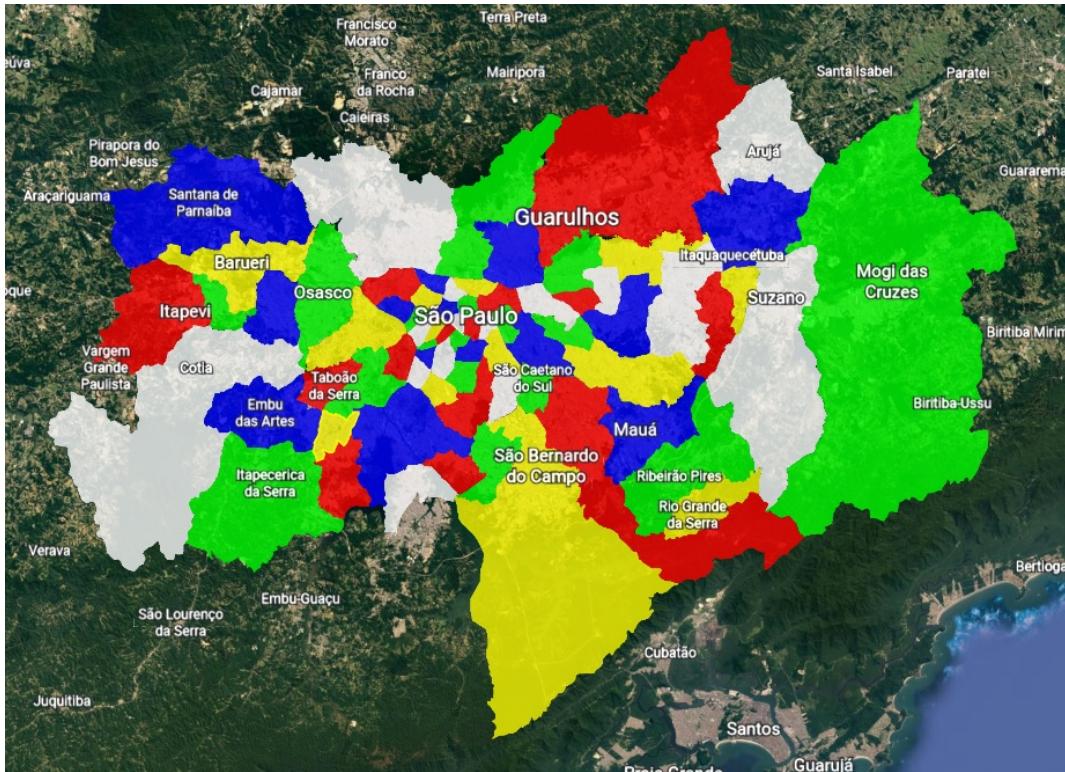


Figura 1: RMSP recortada em zonas

Utilizar um município inteiro como uma zona única, sem nenhuma subdivisão, pode gerar problemas com o fato de grande parte de sua área ser pouco povoada e não gerar chamados. Apesar disso, esta divisão de zonas será mantida, e as soluções encontradas para esse problema serão discutidas na Seção 3.1.5.

4.1.2 Propriedades das zonas

Em complemento às zonas apresentadas anteriormente, será utilizada uma base de dados contendo todas as propriedades conhecidas destas zonas. Para cada uma delas, sabe-se a quantidade média de chamados para cada dia da semana e o município ao qual a zona pertence. A Tabela 1 mostra as primeiras linhas desta base, cujo significado das colunas são:

- **zona:** número do identificador correspondente à cada zona;
- **dia da semana:** dia da semana correspondente;
- **notas esperadas:** quantidade de chamados esperados para uma determinada zona em um determinado dia da semana; e
- **municipio:** município ao qual a zona pertence.

Esta base será utilizada em praticamente todas as análises no decorrer do projeto. Ela fornece a volumetria histórica de chamados que será posteriormente utilizada no gerador de chamados.

Tabela 1: Primeiras linhas da base de propriedades das zonas

zona	dia da semana	notas esperadas	municipio
0	Seg	15,18	SAO PAULO
0	Ter	11,03	SAO PAULO
0	Qua	12,32	SAO PAULO
0	Qui	10,43	SAO PAULO
0	Sex	12,87	SAO PAULO
0	Sab	8,58	SAO PAULO
0	Dom	4,66	SAO PAULO
1	Seg	3,36	SAO PAULO
1	Ter	2,35	SAO PAULO
1	Qua	2,49	SAO PAULO
1	Qui	2,35	SAO PAULO
1	Sex	1,61	SAO PAULO
1	Sab	1,29	SAO PAULO
1	Dom	0,64	SAO PAULO
2	Seg	4,10	COTIA
2	Ter	3,69	COTIA
2	Qua	3,50	COTIA

Com esses valores é possível montar a Figura 2, relacionando cada uma das zonas de atendimento da RMSP com a sua média diária de chamados. É possível notar alta volumetria de chamados nas grandes zonas que contém municípios inteiros, pelo seu baixo detalhamento, assim como nas zonas do centro da cidade de São Paulo, pela sua alta demanda.

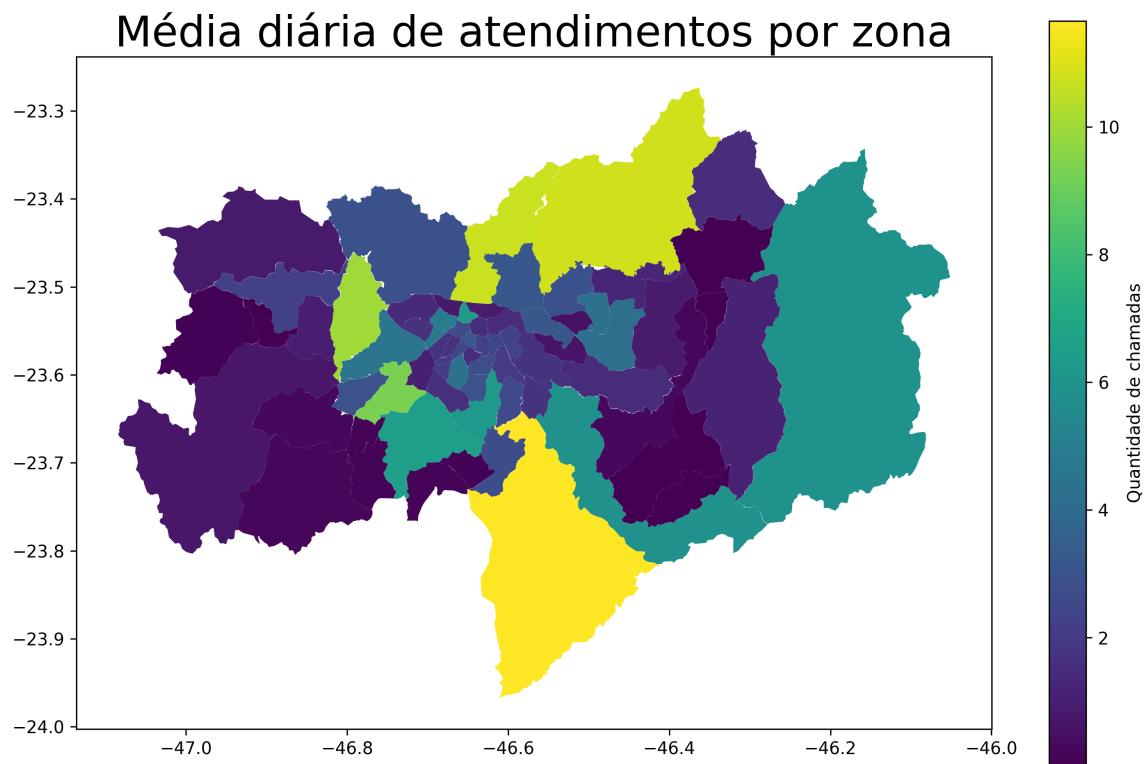


Figura 2: Média diária de atendimentos por zona

Também é possível montar a média de chamados da RMSP para cada dia da semana, Figura 3. É possível observar que o dia da semana com a maior quantidade de chamados é segunda-feira, com a quantidade diminuindo ao longo da semana, até chegar no seu mínimo no domingo.

Média de chamados por dia da semana

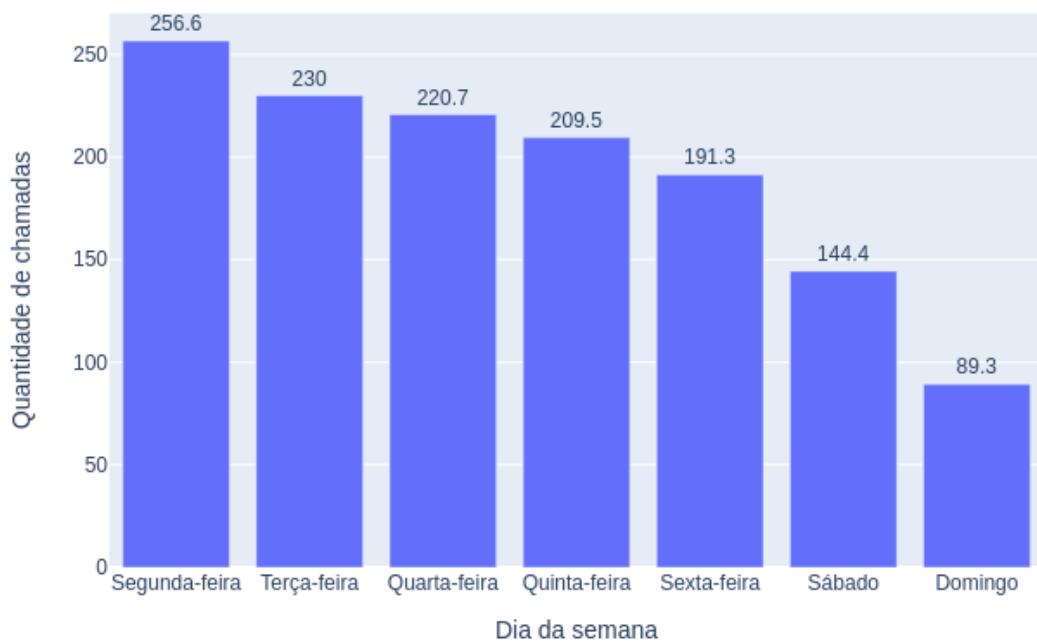


Figura 3: Média de chamados por dia da semana

4.1.3 Distribuição horária dos chamados

Será utilizada também uma base contendo a distribuição média de chamados da RMSP ao longo do dia, discretizada em períodos de uma hora. A Figura 4 mostra esta distribuição.

Distribuição de chamados ao longo do dia

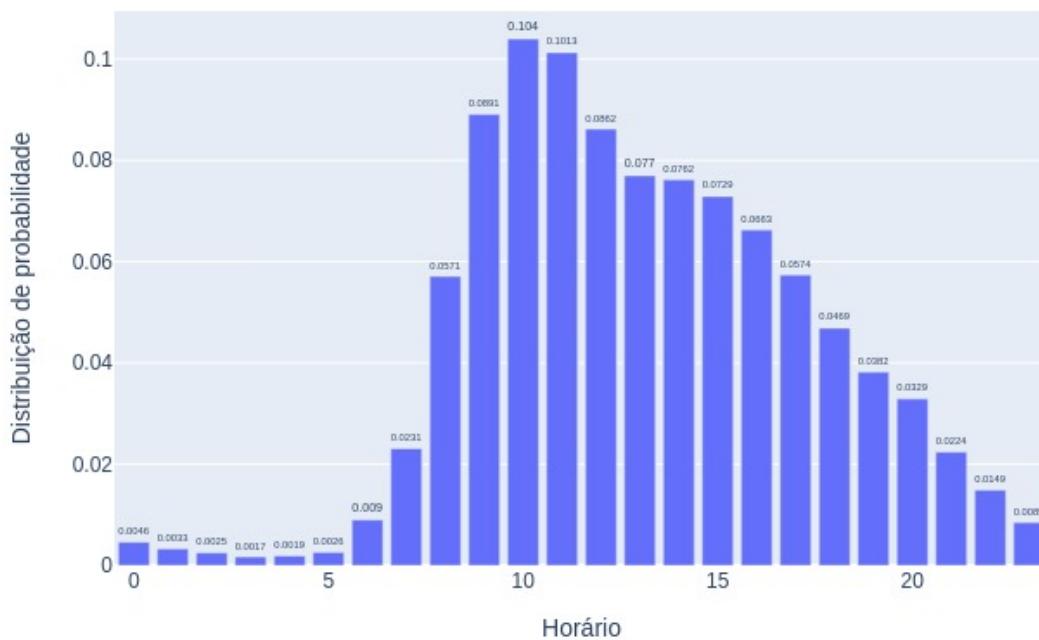


Figura 4: Distribuição de chamados ao longo do dia

É possível observar que existe um horário de pico de chamados que ocorre por volta das 10h da manhã. Além disso, a quantidade de chamados é bastante reduzida de madrugada.

4.1.4 Tratamento das zonas de baixa densidade demográfica

As zonas de baixa densidade demográfica podem ser vistas na Figura 5. Assim como explicado na Seção 3.1.5, essas zonas tem uma área não habitada muito grande, o que torna necessário selecionar apenas o centro populacional dessas regiões.

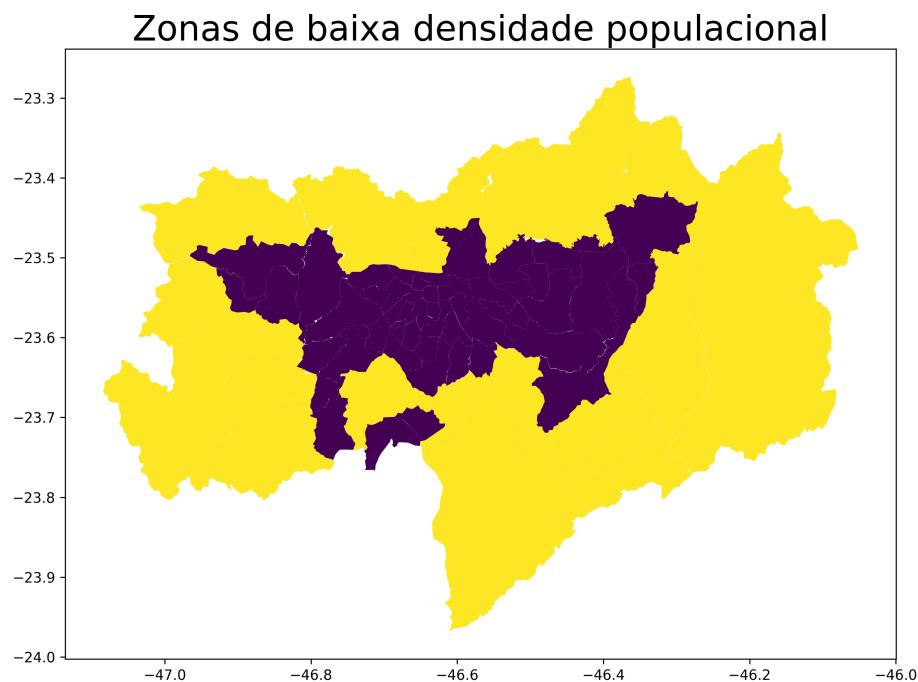


Figura 5: Regiões selecionadas como de baixa densidade (zonas periféricas)

4.2 Gerador de chamados

Primeiramente é selecionada a quantidade de dias que se deseja gerar. Os dados são iniciados com o primeiro dia sendo uma segunda-feira. Para cada dia se itera pelas zonas de atendimento, gerando a quantidade necessária de chamados por zona, sua localização e horário. Depois são gerados os chamados do próximo dia e assim por diante até que a quantidade de dias requisitados seja atingida.

A base de dados gerada ter a seguinte estrutura:

- **surgimento:** horário de surgimento do chamado em minutos;
- **hora:** parcela de hora do surgimento do chamado;
- **lat:** latitude da localização do chamado;
- **lon:** longitude da localização do chamado;
- **zona:** identificador da zona onde o chamado está contido;
- **municipio:** município da zona da em que o chamado está contido;

- **dia da semana:** dia da semana do surgimento do chamado; e
- **dia:** dia em que o chamado surge, começando em 0.

Tabela 2: Primeiras linhas da base de chamados gerados

surgimento	hora	lat	lon	zona	municipio	dia da semana	dia
758	12	-23.56	-46.66	0	SAO PAULO	0	0
675	11	-23.56	-46.65	0	SAO PAULO	0	0
1003	16	-23.57	-46.64	0	SAO PAULO	0	0
1039	17	-23.57	-46.64	0	SAO PAULO	0	0
409	6	-23.54	-46.66	1	SAO PAULO	0	0

Parte da base de chamados gerados pode ser observada na Tabela 2 e a localização deste chamado na RMSP pode ser observada na Figura 6.

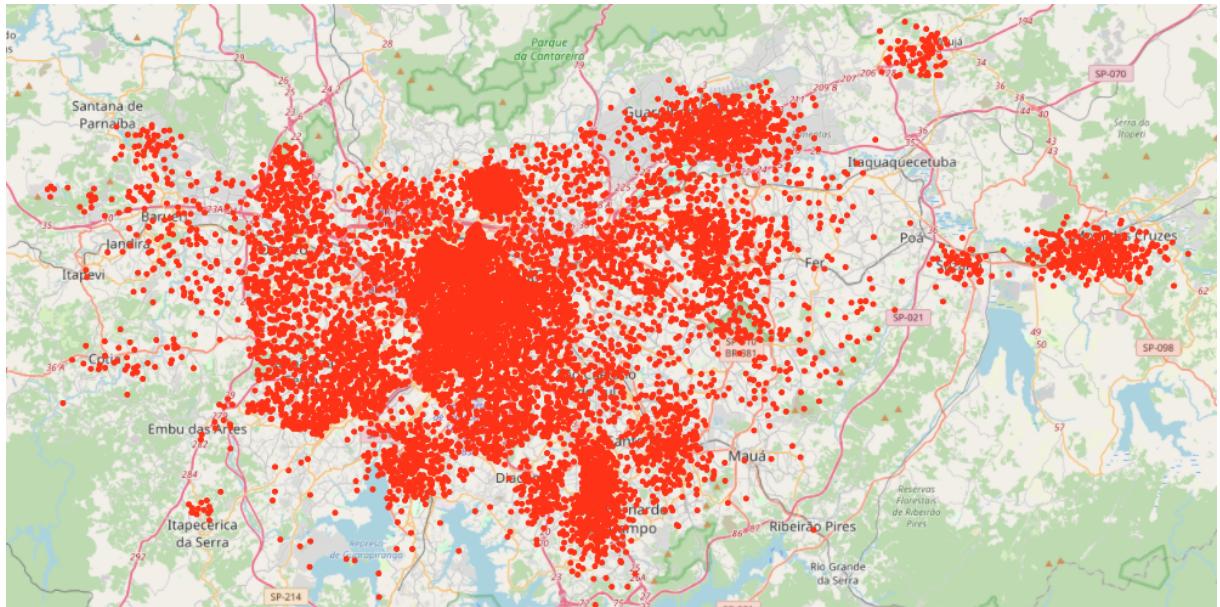


Figura 6: Distribuição geográfica dos chamados gerados

Pode-se observar a diferença entre a distribuição uniforme de chamados nas zonas do centro e a distribuição normal nas zonas periféricas menos populosas.

4.2.1 Validação do gerador

Assim como explicado na Seção 3.1.6, existe uma diferença grande entre a volumetria do município de São Paulo e dos municípios periféricos, assim como se pode observar na Figura 7.

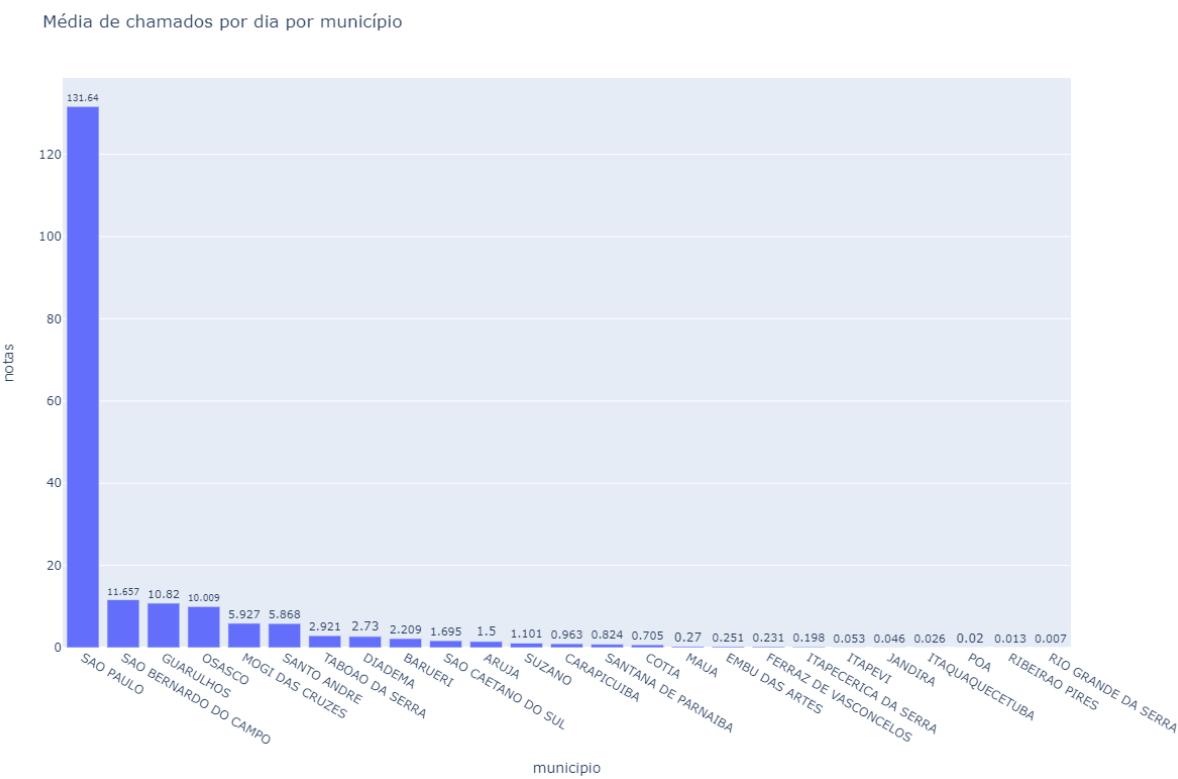


Figura 7: Média de chamados por dia por município

A seguir será apresentada uma comparação entre os chamados esperados e os chamados gerados para as métricas descritas nas Seções 3.1.1-3.1.4. Os gráficos comparativos ilustram o impacto da quantidade de dias simulados nas métricas de comparação, quanto mais dias simulados, maior a chance de chamados nas regiões de baixa volumetria serem gerados e menor é o erro percentual dessas regiões.

Média de chamados por dia por município



Figura 8: Comparativo da média de chamados por dia por município para 7 dias

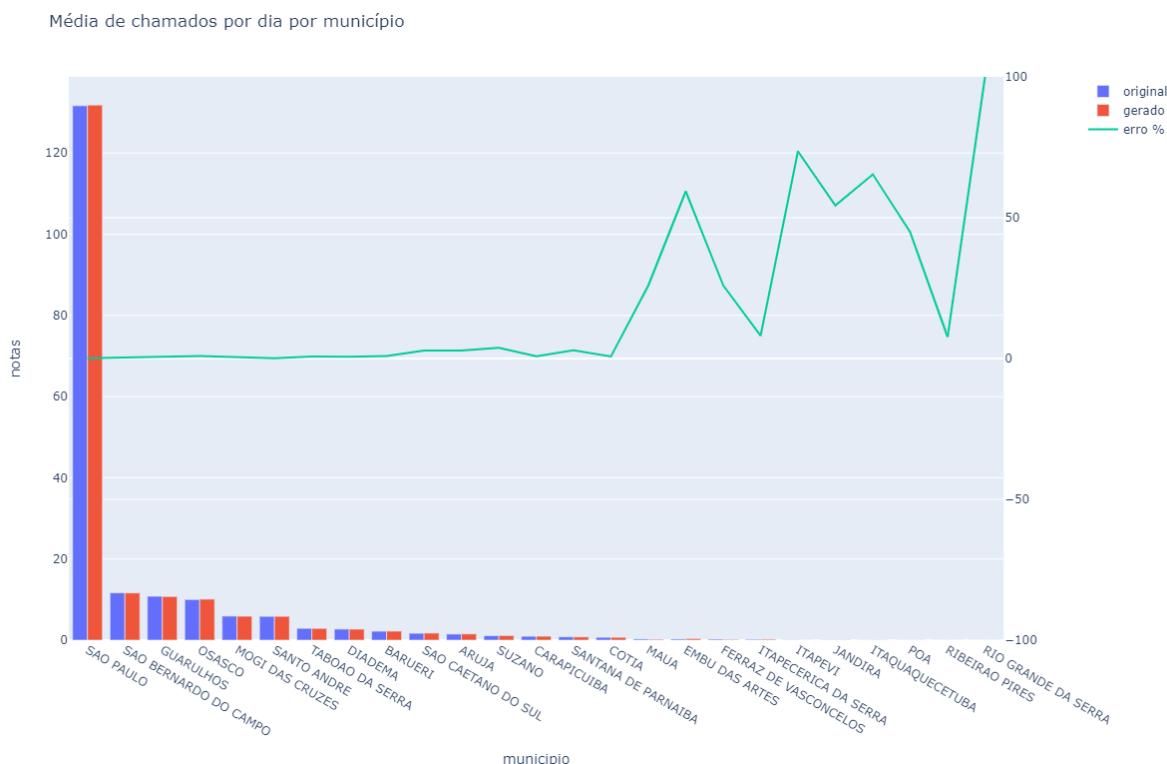


Figura 9: Comparativo da média de chamados por dia por município para 70 dias

Média de chamados por dia por município

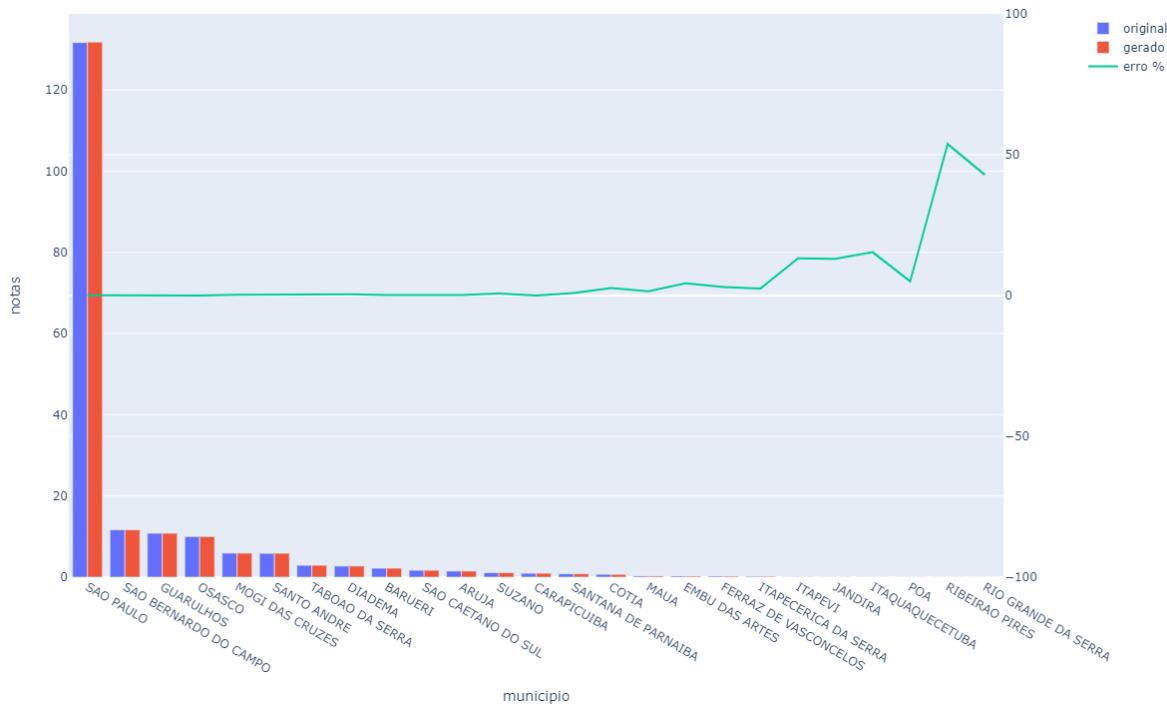


Figura 10: Comparativo da média de chamados por dia por município para 700 dias

Média de chamados por dia por município

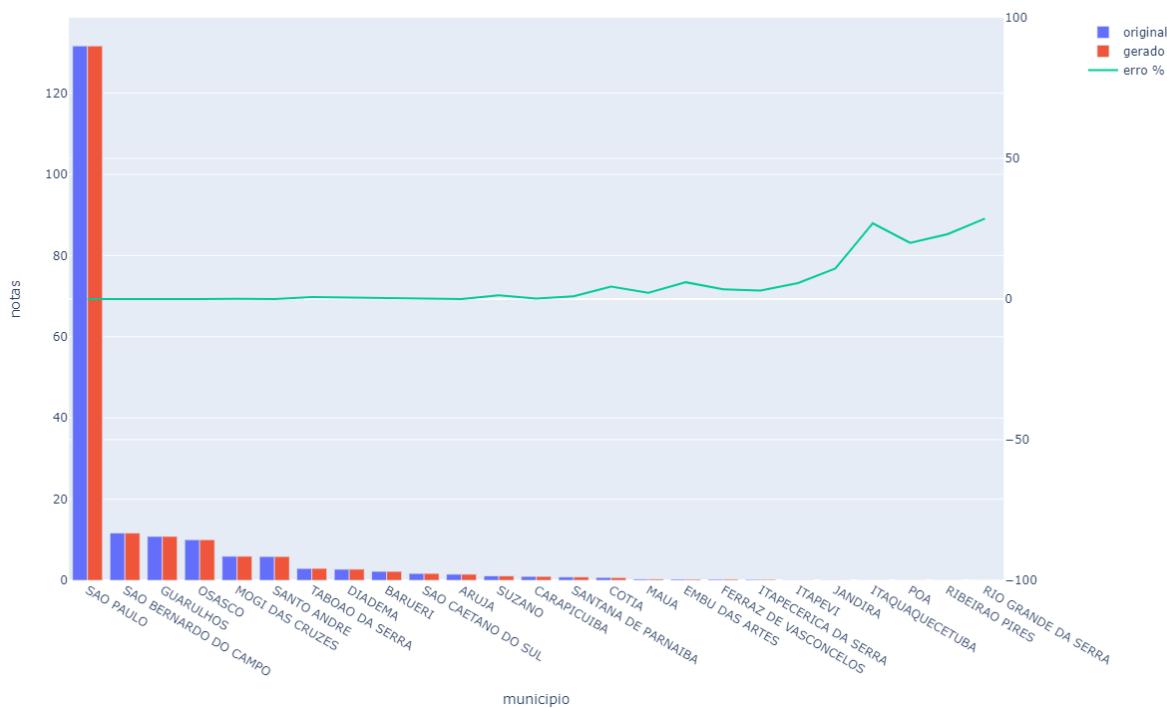


Figura 11: Comparativo da média de chamados por dia por município para 1400 dias

Média de chamados por dia por zona

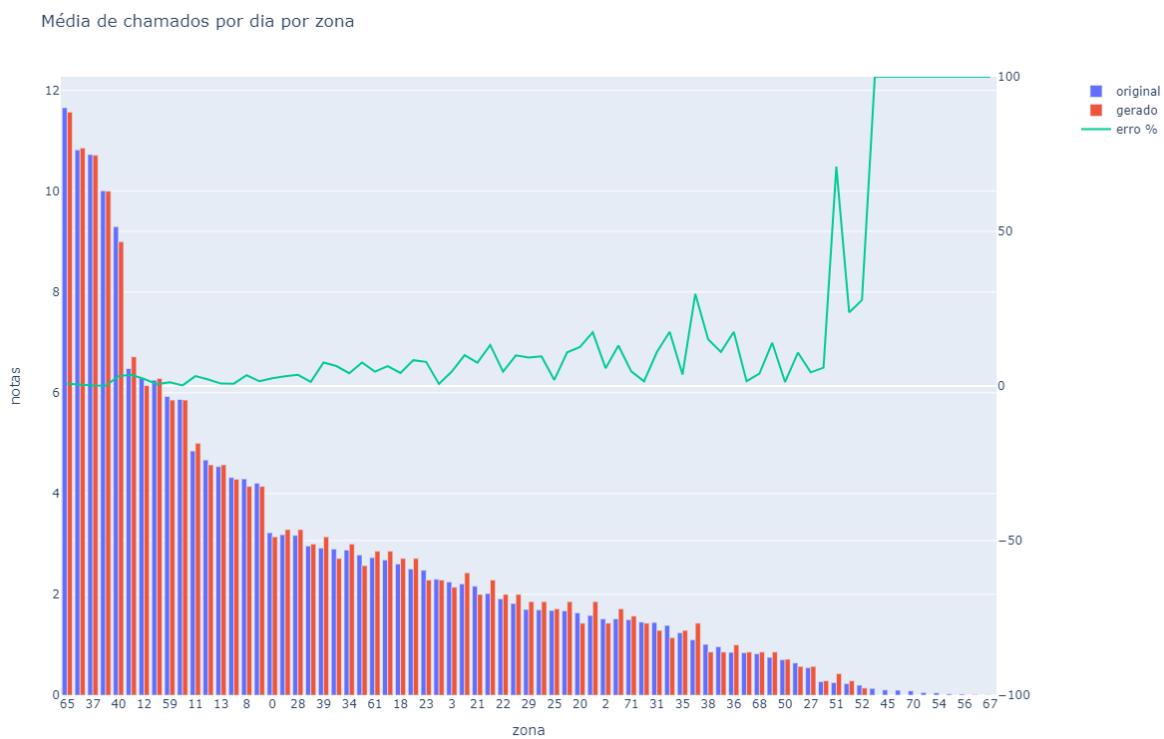


Figura 12: Comparativo da média de chamados por dia por zona para 7 dias

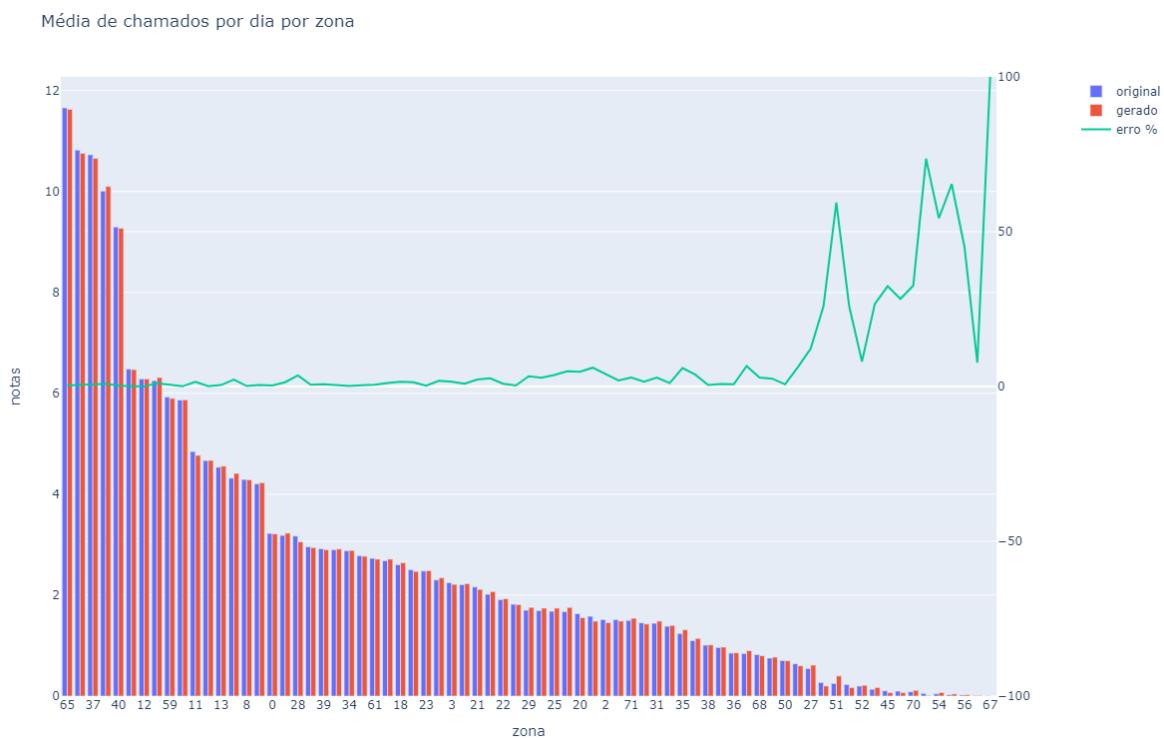
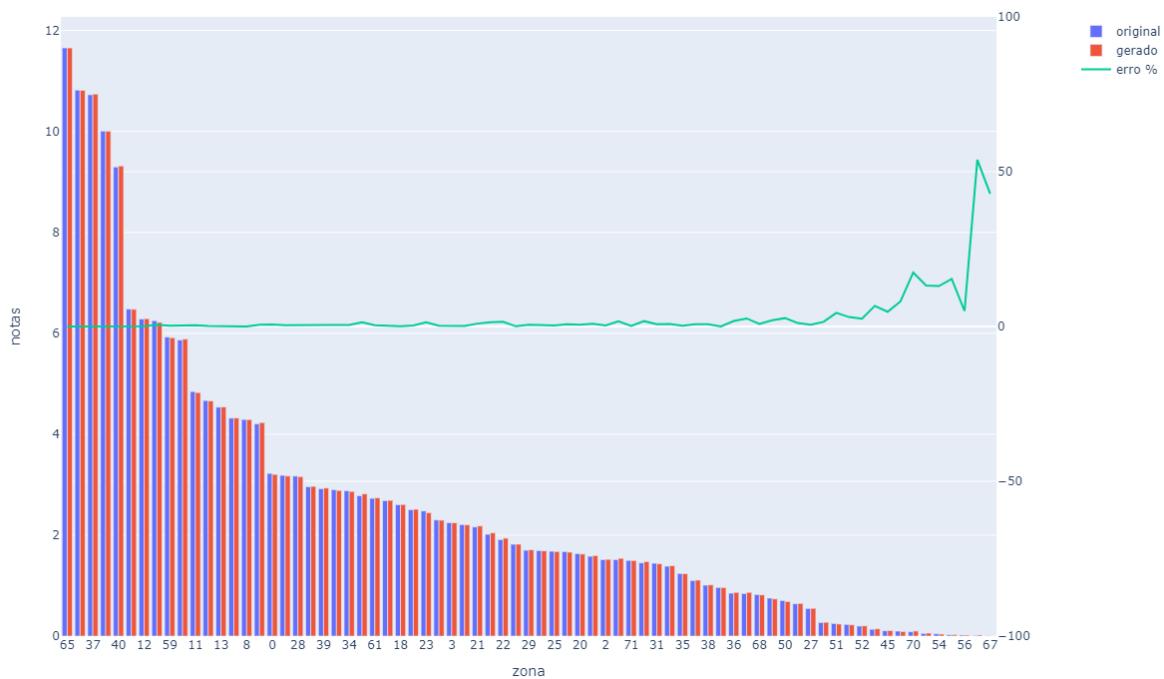
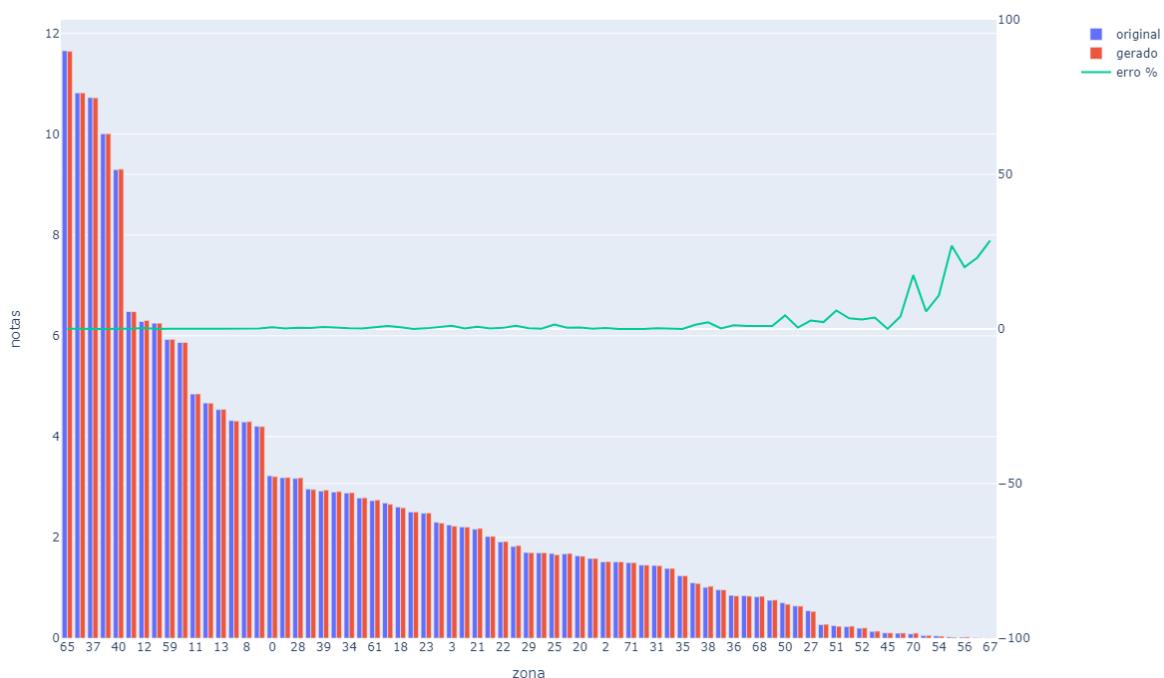


Figura 13: Comparativo da média de chamados por dia por zona para 70 dias

Média de chamados por dia por zona

**Figura 14:** Comparativo da média de chamados por dia por zona para 700 dias

Média de chamados por dia por zona

**Figura 15:** Comparativo da média de chamados por dia por zona para 1400 dias

Distribuição horária dos chamados

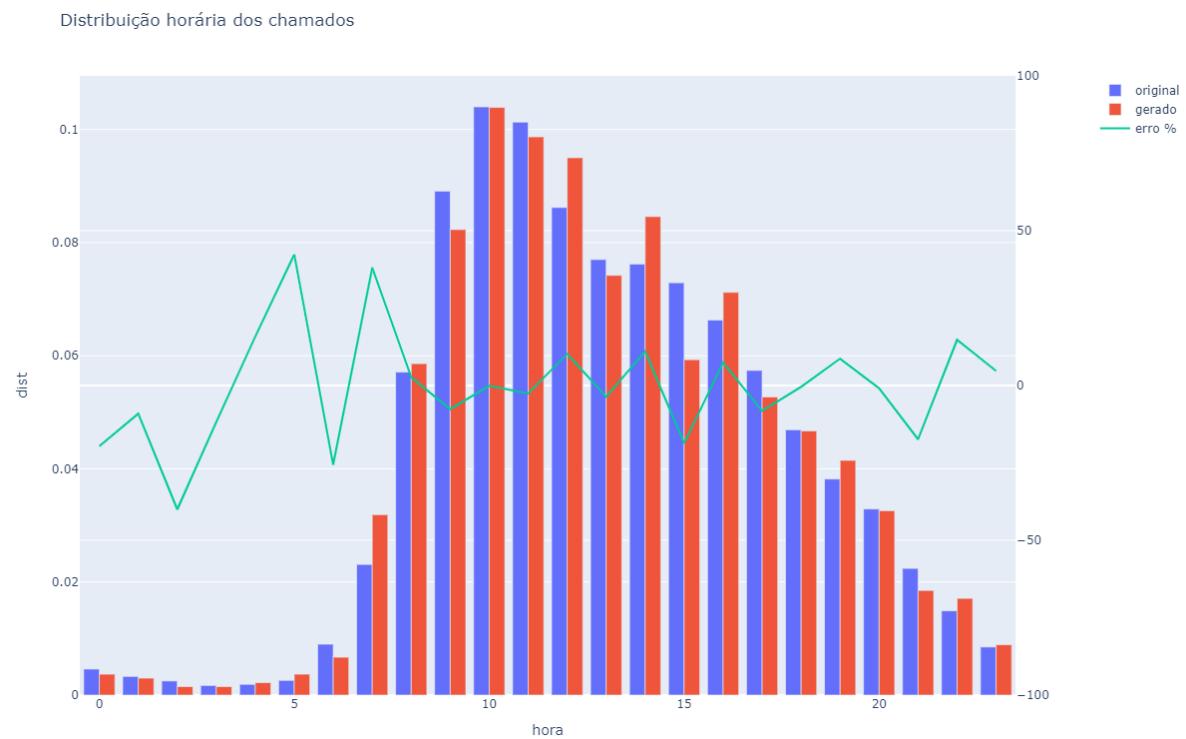


Figura 16: Comparativo da distribuição horária dos chamados para 7 dias

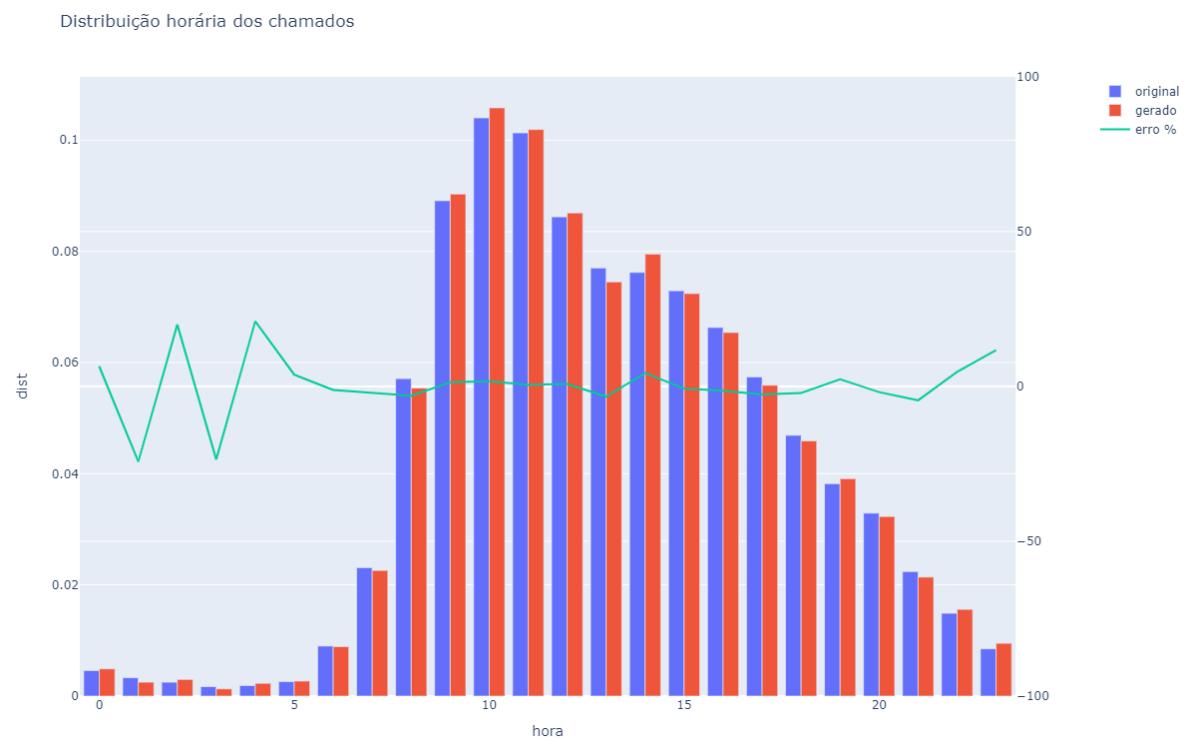
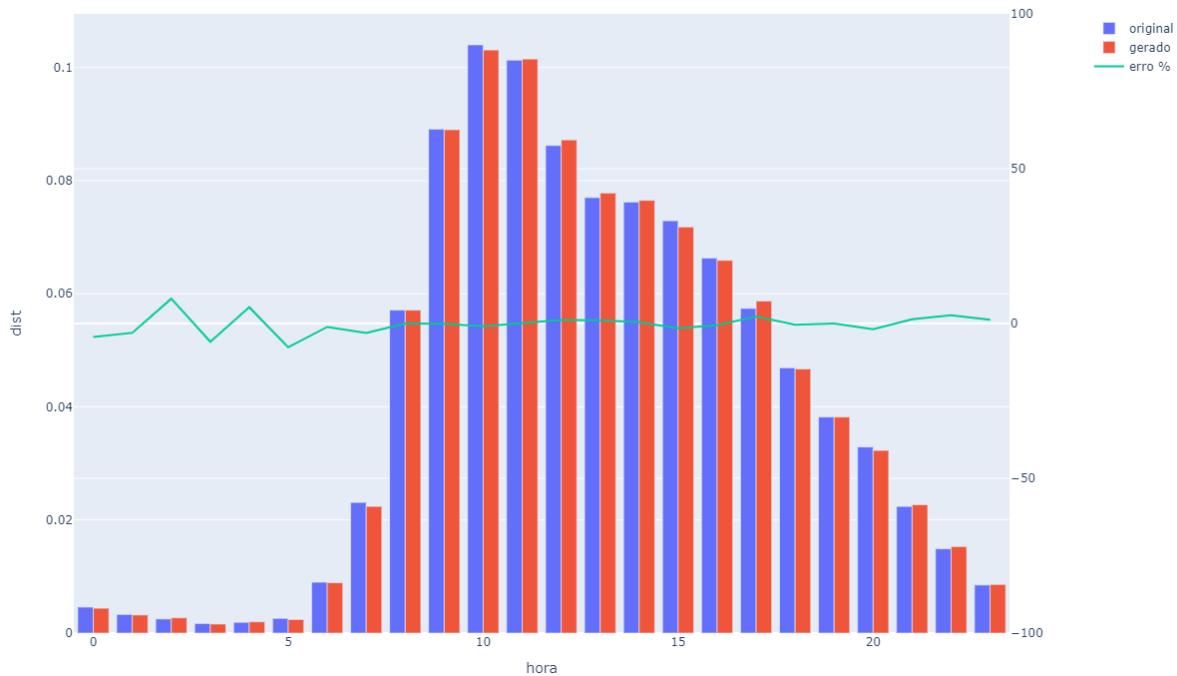
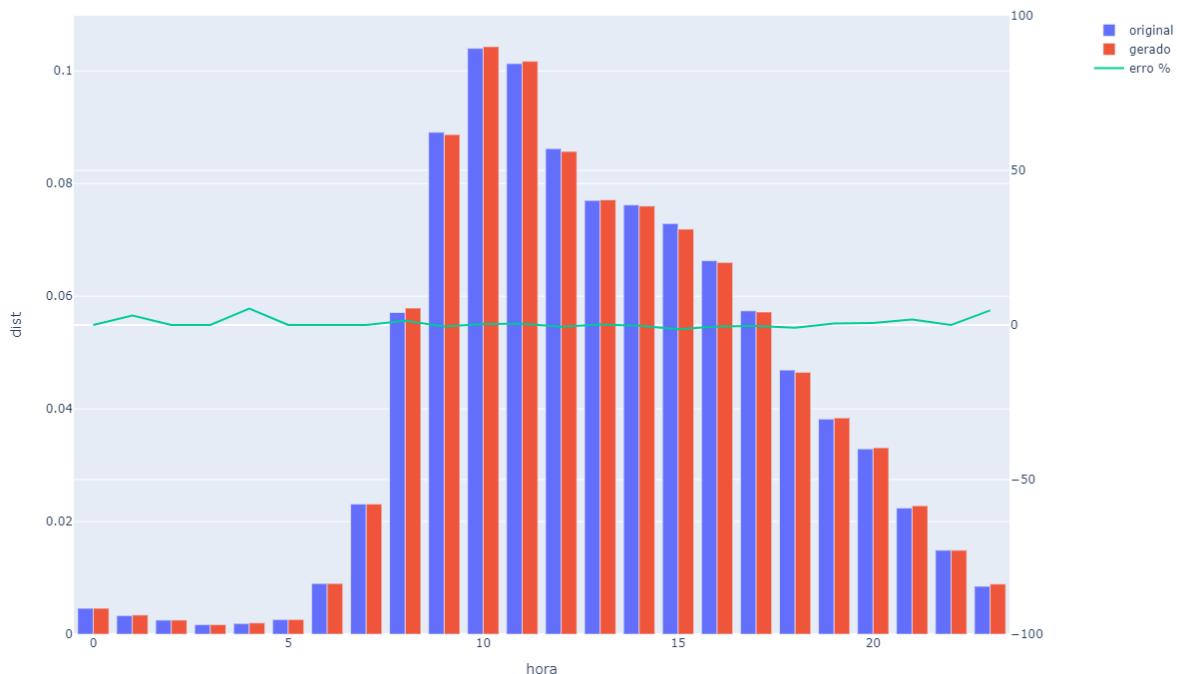


Figura 17: Comparativo da distribuição horária dos chamados para 70 dias

Distribuição horária dos chamados

**Figura 18:** Comparativo da distribuição horária dos chamados para 700 dias

Distribuição horária dos chamados

**Figura 19:** Comparativo da distribuição horária dos chamados para 1400 dias

Em todos os gráficos das Figuras 8-19 é possível observar uma tendência de queda no erro percentual entre os valores esperado e gerado. A quantidade de 1400 dias foi escolhida para a simulação pelo baixo erro percentual nas zonas de baixa volumetria, mantendo o tempo de simulação dentro de um valor razoável. Assim, a base utilizada no trabalho tem um total de 1400 dias e mais de 270 mil chamados.

4.3 Otimizador de escalas

De acordo com o que foi dito na Seção 3.2.1, na Figura 20 os números na segunda linha representam a quantidade mínima de técnicos que devem estar disponíveis em cada intervalo de hora (estes valores são apenas ilustrativos). Os vários **X** ilustram todas as possíveis escalas, de A até X, em um dia, dado que a jornada de trabalho dos técnicos é de 7 horas. O que o otimizador faz é determinar quantos técnicos devem ser alocados em cada escala com o objetivo de minimizar a quantidade total de técnicos, ainda atendendo à quantidade mínima de técnicos por hora.

	hora	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
	qtd	2	1	1	0	0	1	4	10	26	42	49	47	40	35	36	34	31	26	22	18	15	10	6	4
A		X																			X	X	X	X	X
B		X	X																		X	X	X	X	X
C		X	X	X																	X	X	X	X	X
D		X	X	X	X																X	X	X	X	X
E		X	X	X	X	X															X	X	X	X	X
F		X	X	X	X	X	X														X	X	X	X	X
G		X	X	X	X	X	X	X															X	X	X
H		X	X	X	X	X	X	X	X															X	X
I			X	X	X	X	X	X	X	X															
J				X	X	X	X	X	X	X	X														
K					X	X	X	X	X	X	X	X													
L						X	X	X	X	X	X	X	X												
M							X	X	X	X	X	X	X	X											
N								X	X	X	X	X	X	X	X										
O									X	X	X	X	X	X	X	X									
P										X	X	X	X	X	X	X	X								
Q											X	X	X	X	X	X	X	X							
R												X	X	X	X	X	X	X	X						
S													X	X	X	X	X	X	X	X					
T														X	X	X	X	X	X	X	X				
U															X	X	X	X	X	X	X	X			
V																X	X	X	X	X	X	X	X		
W																	X	X	X	X	X	X	X	X	
X																		X	X	X	X	X	X	X	X

Figura 20: Tabela mostrando todas as possíveis escalas em um dia

4.4 Localização ideal dos técnicos

Para determinar os agrupamentos, os chamados foram clusterizados em quantidades de clusters equivalentes à quantidade de técnicos em cada escala. Sendo assim, as localizações dos técnicos por escala podem ser vistas nas Figuras 21-25.

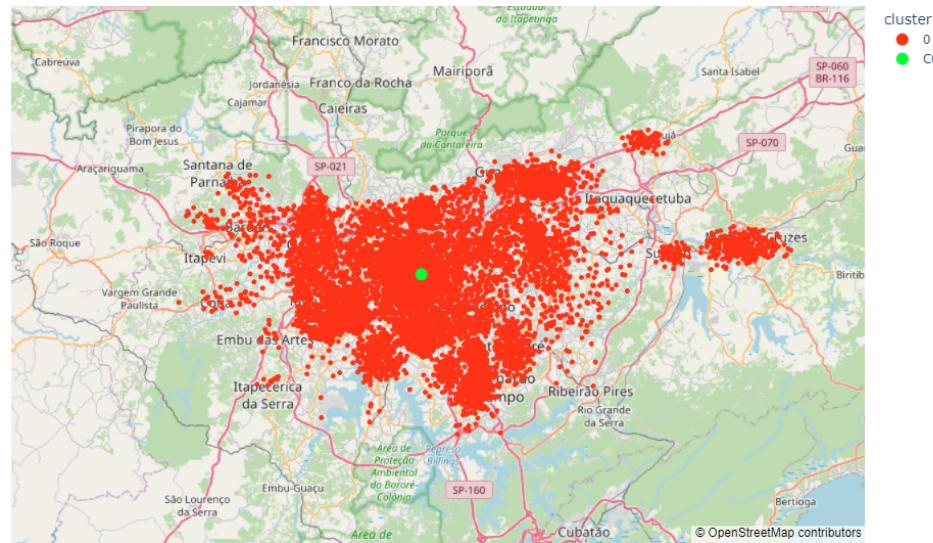


Figura 21: Localização dos técnicos que compõem uma escala com 1 técnico. Na legenda, 'C0' representa o centroide dos pontos identificados com '0'

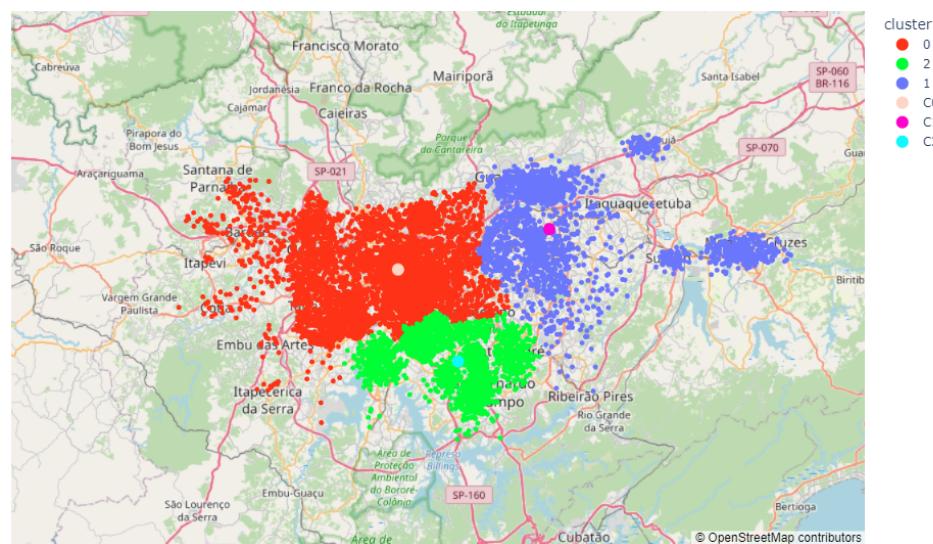


Figura 22: Localização dos técnicos que compõem uma escala com 3 técnicos. Na legenda, 'C0', 'C1' e 'C2' representam os centroides dos pontos identificados com '0', '1' e '2' respectivamente

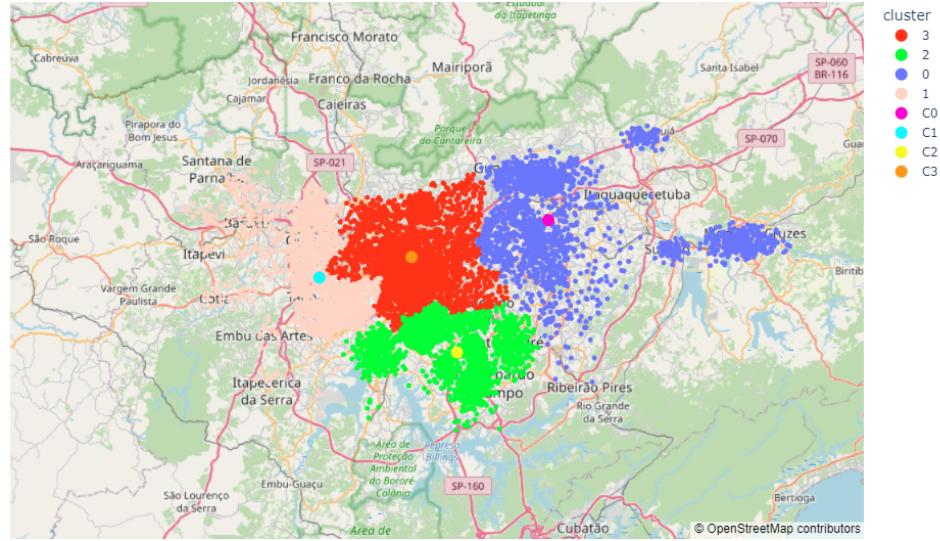


Figura 23: Localização dos técnicos que compõem uma escala com 4 técnicos. Na legenda, 'C0', 'C1', 'C2' e 'C3' representam os centroides dos pontos identificados com '0', '1', '2' e '3' respectivamente

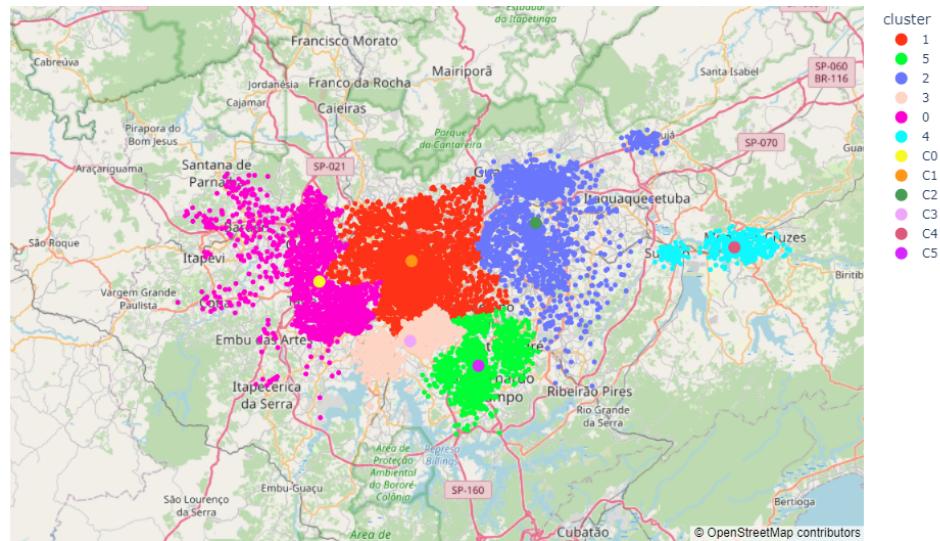


Figura 24: Localização dos técnicos que compõem uma escala com 6 técnicos. Na legenda, 'C0', 'C1', 'C2', 'C3', 'C4' e 'C5' representam os centroides dos pontos identificados com '0', '1', '2', '3', '4' e '5' respectivamente

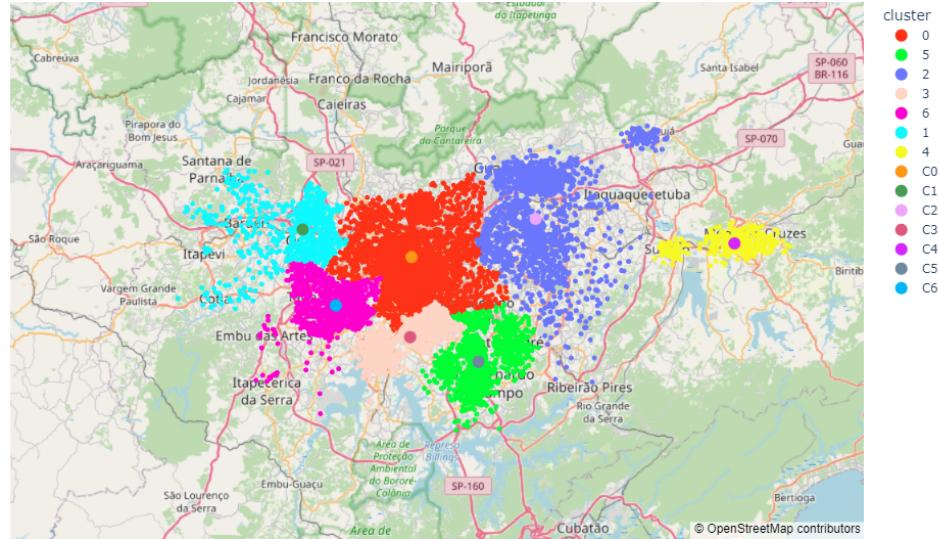


Figura 25: Localização dos técnicos que compõem uma escala com 7 técnicos. Na legenda, 'C0', 'C1', 'C2', 'C3', 'C4', 'C5' e 'C6' representam os centroides dos pontos identificados com '0', '1', '2', '3', '4', '5' e '6' respectivamente

Depois, é criada uma planilha de alocação contendo as localizações e escalas dos técnicos, a planilha tem as seguintes colunas:

- **id:** identificador do técnico;
- **hora inicial:** hora inicial da escala do técnico;
- **hora final:** hora final da escala do técnico;
- **lat:** latitude da localização do técnico; e
- **lon:** longitude da localização do técnico.

Na Tabela 3 é possível ver as primeiras linhas da base de escalas dos técnicos.

Tabela 3: Primeiras linhas da base das escalas de cada técnico

id	hora inicial	hora final	lat	lon
0	6	12	-23.58	-46.77
1	6	12	-23.55	-46.61
2	6	12	-23.67	-46.54
3	6	12	-23.53	-46.49
4	6	12	-23.52	-46.19
5	6	12	-23.65	-46.65
6	6	12	-23.50	-46.54
7	7	13	-23.55	-46.61
8	7	13	-23.67	-46.45
9	7	13	-23.61	-46.66

4.5 Fator de correção da distância entre chamado e técnico

Como a menor distância entre dois pontos é uma reta, considerar a distância entre o chamado e o técnico como tal é um subdimensionamento. Pensando nisso, foi criada uma amostra de 100 elementos contendo a distância entre dois pontos aleatórios na RMSP, tanto a distância em linha reta quanto a distância real. Depois, foram feitos testes de hipótese para avaliar o valor da média da relação entre distância real e distância em linha reta. A conclusão foi que, com 95% de significância, é possível dizer que esta relação é aproximadamente 1,6.

4.6 Estruturação dos resultados

Para cada execução do simulador, é gerada uma tabela de saída contendo as seguintes colunas:

- **id**: identificador do chamado;
- **surgimento**: minuto, dentro da simulação, no qual o chamado ocorreu. Esse valor pode variar entre 0 e 1439, que são todos os minutos contidos em um intervalo de 24 horas;
- **hora**: equivalente ao minuto de surgimento convertido em horas;

- **lat:**latitude do chamado;
- **lon:** longitude do chamado;
- **zona:** zona na qual o chamado ocorreu;
- **municipio:** municipio onde ocorreu o chamado;
- **dia da semana:** dia da semana no qual o chamado foi disparado;
- **dia:** em qual dia da simulação o chamado aconteceu;
- **reservado:** essa coluna é uma flag que indica se o chamado foi atribuído a um técnico que estava ocupado no momento em que o chamado ocorreu, ela é utilizada no código do simulador para ajudar na identificação de técnicos disponíveis em cada minuto;
- **tecnico responsável:** identificador do técnico atribuído ao chamado;
- **chegada:** minuto da simulação no qual o técnico chegou ao local da ocorrência. Essa coluna é uma das mais importantes do arquivo de saída, pois é a partir dela que o TAE será calculado;
- **termino:** minuto da simulação no qual o técnico finalizou um chamado e ficou disponível novamente para atender mais chamados; e
- **TAE:** Tempo de Atendimento médio Esperado. Essa coluna nada mais é do que a diferença entre o minuto de chegada e minuto de surgimento, que mede exatamente o tempo que o técnico responsável pelo atendimento levou para chegar até o local da ocorrência.

4.7 Resultado Preliminar

O primeiro resultado foi obtido com os seguintes parâmetros:

- foram simulados 200 dias de operação;
- foram distribuídos 80 técnicos pela RMSP, cada um atendendo na sua respectiva escala;
- velocidade média na RMSP de 20 km/h; e

- tempo de atendimento de um chamado de 40 minutos.

A Tabela 4 mostra as primeiras linhas do arquivo de saída da simulação. Algumas colunas foram omitidas para facilitar a visualização.

Tabela 4: Primeiras linhas da base de resultados da simulação

surgimento	municipio	chegada	termino	TAE
2	SAO PAULO	23	63	21
7	DIADEMA	59	99	52
18	SAO PAULO	58	98	40
29	SAO PAULO	64	104	35
75	GUARULHOS	137	177	62
360	SAO PAULO	365	405	5
371	SAO PAULO	402	442	31
382	SAO PAULO	401	441	19
386	SAO PAULO	443	483	57
393	COTIA	444	484	51
394	SAO PAULO	466	506	72
430	SAO PAULO	442	482	12

Vale ressaltar que cada linha da tabela 4 é um chamado. O que interessa de fato é saber se a média do tempo de atendimento para cada município ao longo do período simulado foi ou não menor do que 60 minutos, assim como dita o órgão regulador. Para tanto, o simulador também fornece como saída outra tabela contendo um agrupamento dos dados da Tabela 4 por município com a média da coluna **TAE**, que é sigla utilizada para se referir ao Tempo de Atendimento médio Esperado. Na tabela 5 é possível observar estes resultados:

Tabela 5: Tempo de atendimento médio para cada município

Município	TAE
ARUJA	48.08
BARUERI	31.89
CARAPICUIBA	33.95
COTIA	46.58
DIADEMA	34.33
EMBU DAS ARTES	50.41
FERRAZ DE VASCONCELOS	46.57
GUARULHOS	47.23
ITAPEVI	42.01
ITAQUAQUECETUBA	32.97
JANDIRA	29.47
MAUA	39.25
OSASCO	26.41
POA	53.70
RIBEIRAO PIRES	43.76
RIO GRANDE DA SERRA	-
SANTANA DE PARNAIBA	58.56
SANTO ANDRE	42.83
SAO BERNARDO DO CAMPO	49.61
SAO CAETANO DO SUL	42.46
SUZANO	63.80
TABOAO DA SERRA	39.04
MOGI DAS CRUZES	43.76
SAO PAULO	42.67

4.7.1 Discussão de Resultados

Analizando os resultados da tabela 5 é possível notar que em geral o tempo de atendimento médio foi bem abaixo dos 60 minutos, o que indica um forte potencial de redução do número de técnicos, isso significa que alguns parâmetros do otimizador podem ter sido muito conservadores, como a probabilidade aceitável de ter mais chamados do que técnicos em uma zona. Até porque, como visto anteriormente, o simulador tem capacidade para atribuir um técnico momentaneamente ocupado a um chamado, e analisando os

resultados é possível observar que na maioria das vezes em que isso ocorre, o técnico ocupado está quase terminando seu serviço atual, o que não gera um atraso tão grande no atendimento do próximo chamado.

Além disso, Suzano foi o único município que apresentou um TAE maior que 60 minutos, portanto, seria o único município onde o órgão regulador aplicaria uma multa. Uma possível solução para esse problema seria manualmente alterar a localização de alguns técnicos localizados nas proximidades de Suzano. Este processo poderia ser feito analisando os municípios em torno de Suzano e deslocando alguns técnicos de municípios que tiveram TAE mais baixo para que eles fiquem mais próximos do município afetado. Já Rio Grande da Serra não teve registro de chamados. Isso ocorreu porque a volumetria de chamados neste município é muito baixa, de fato é a menor volumetria entre todos os municípios, tanto que ao longo dos 200 dias de simulação não houve um chamado sequer neste município. Isso é preocupante, uma vez que, se a volumetria de dados for muito baixa em um município, não será possível tirar conclusões estatisticamente significativas para este município, portanto, uma possível solução para este problema seria executar o simulador para mais dias.

4.8 Resultado Corrigido

O segundo resultado foi obtido com os seguintes parâmetros:

- foram simulados 700 dias de operação;
- foram distribuídos 78 técnicos pela RMSP, cada um atendendo na sua respectiva escala;
- velocidade média na RMSP de 20 km/h;
- tempo de atendimento de um chamado de 40 minutos;
- a probabilidade aceitável de ter mais chamados do que técnicos foi alterada de 10% para 20%; e
- a localização de alguns técnicos foi manualmente alterada em direção aos municípios que tiveram os TAEs mais altos.

Assim como no cenário anterior, a Tabela 6 mostra as primeiras linhas do arquivo de saída da simulação. Algumas colunas foram omitidas para facilitar a visualização.

Tabela 6: Primeiras linhas da base de resultados da simulação

surgimento	municipio	chegada	termino	TAE
9	SAO PAULO	31	71	22
16	SAO PAULO	54	94	38
52	SAO PAULO	60	100	8
68	GUARULHOS	123	163	55
105	SAO PAULO	137	177	32
198	SAO PAULO	231	271	33
234	BARUERI	281	321	47
316	SAO PAULO	356	396	40
393	SAO PAULO	404	444	11
432	SAO PAULO	454	494	22
502	SAO PAULO	543	583	41
506	OSASCO	526	566	20

Na tabela 7 é possível observar os resultados:

Tabela 7: Tempo de atendimento médio para cada município

Município	TAE
ARUJA	53.43
BARUERI	44.07
CARAPICUIBA	30.61
COTIA	56.44
DIADEMA	17.14
EMBU DAS ARTES	45.65
FERRAZ DE VASCONCELOS	54.66
GUARULHOS	26.83
ITAPEVI	56.13
ITAQUAQUECETUBA	52.40
JANDIRA	56.24
MAUA	50.85
OSASCO	21.54
POA	52.92
RIBEIRAO PIRES	57.43
RIO GRANDE DA SERRA	50.50
SANTANA DE PARNAIBA	55.91
SANTO ANDRE	25.74
SAO BERNARDO DO CAMPO	14.83
SAO CAETANO DO SUL	25.05
SUZANO	49.22
TABOAO DA SERRA	21.20
MOGI DAS CRUZES	31.23
SAO PAULO	22.54

4.8.1 Discussão de Resultados

Como o número de técnicos foi reduzido, o tempo de atendimento médio aumentou. Apesar do tempo médio de atendimento ter aumentado como um todo, nenhum município teve TAE acima de 60 minutos desta vez, o que significa que a companhia conseguiria zerar o valor da multa.

No entanto, apesar do TAE estar abaixo dos 60 minutos para todos os municípios, alguns municípios tiveram um TAE bem menor do que outros. Uma possível solução para

este problema seria aplicar um refinamento na localização dos técnicos. Este refinamento poderia ser feito deslocando técnicos de municípios com TAEs baixos para municípios com TAEs mais altos. Detalhes dessa discussão podem ser vistas no Capítulo 5.

Quanto à Rio Grande da Serra, desta vez o município teve alguns chamados no período simulado, como era esperado.

5 DISCUSSÕES

Conforme foi possível observar no capítulo anterior, existem ainda alguns pontos de melhoria. Apesar do objetivo ter sido atendido, existe ainda a possibilidade de melhorar ou alterar algumas análises e métodos para que não haja mais a necessidade de refinamento dos resultados.

Alguns tópicos poderiam ser melhorados em projetos futuros, por exemplo, o método escolhido para o cálculo da distância entre um chamado e o técnico responsável pelo atendimento foi muito simplificado. Ao utilizar uma velocidade média para toda a RMSP, toda a informação do trânsito local é perdida, o que é uma informação importante no cálculo do tempo de deslocamento. Portanto, uma boa abordagem seria utilizar algum software capaz de calcular o tempo de uma viagem entre dois pontos levando em consideração o trânsito. Tanto que no gerador de chamados a abordagem escolhida para geração dos chamados foi míope às questões de arruamento e endereços. Se for desejado utilizar um software que utilize informações de trânsito para o cálculo do tempo de deslocamento, no futuro pode ser uma boa ideia checar se o chamado foi gerado numa área habitada. Apesar de ter usado dados geográficos com a localização dos centros das cidades periféricas da RMSP, alguns chamados ainda caíram em regiões não habitadas, o que pode não ser viável para o software.

Já ao determinar a localização dos técnicos, apesar do *Agglomerative Clustering* fornecer bons resultados, existem outras opções que poderiam ser testadas mais exaustivamente. O ideal seria utilizar algum algoritmo de clusterização capaz de ponderar a densidade de pontos na criação de um cluster, como o DBSCAN por exemplo. Um forte indício de que revisar o algoritmo de clusterização traria bons resultados é que, após o primeiro resultado, foi necessário ajustar manualmente algumas localizações. Uma alternativa a esse problema seria ajustar a localização dos técnicos de maneira automatizada. Ao invés de fazer ajustes manuais, seria possível criar vetores imaginários ligando todos os pares de municípios e deslocar os técnicos na direção desses vetores ponderando pelo TAE médio obtido no resultado anterior. Este seria um processo iterativo que talvez pudesse convergir

para uma solução ótima, valeria muito a pena investigar mais a fundo se este processo de fato convergiria.

Quanto ao simulador, existe a oportunidade de alterar a lógica de escolha de um técnico para um chamado. Talvez a escolha pelo técnico que chega mais rápido não seja a melhor opção, isso aconteceria por exemplo num cenário onde o técnico que chega mais rápido na verdade está saindo de uma zona de alta volumetria para atender um chamado numa zona de baixa volumetria, sendo que neste caso seria melhor alocar um técnico de uma zona de menor volumetria para este chamado, mesmo que demore um pouco mais, considerando que a qualquer momento pode surgir um chamado na zona mais movimentada.

Além disso, assim como a velocidade média, o tempo de atendimento de um chamado, que é o tempo que o técnico demora para resolver o problema, foi fixado, neste caso, em 40 minutos. Existe a oportunidade de variar este valor seguindo uma distribuição normal centrada em 40 minutos por exemplo, a fim de tornar a simulação mais fidedigna à operação real.

6 CONCLUSÃO

O objetivo deste trabalho foi dimensionar uma equipe para o atendimento emergencial de uma distribuidora de gás natural. Nesse sentido, foi possível atender a esse objetivo através do desenvolvimento de um simulador da operação e um otimizador da alocação dos técnicos.

Utilizando os dados históricos da companhia de gás, foi possível gerar chamados fictícios que serviram como insumo para o simulador. Essa abordagem trouxe resultados satisfatórios, pois possibilitou dimensionar uma equipe de forma a zerar os custos da companhia com multas provenientes do não cumprimento das normas do órgão regulador.

No entanto, foi possível verificar diversos pontos de melhoria que poderiam enriquecer ainda mais as análises realizadas neste trabalho. A possibilidade de refinar as análises abre margem a uma série de estudos cujos potenciais podem ir além do tema central deste trabalho, podendo até ser aplicados em outros problemas envolvendo alocação de recursos.

Sendo assim, pode-se inferir que todas as etapas intermediárias deste projeto, como o desenvolvimento do gerador de chamados, do otimizador de escalas, do algoritmo de localização ideal dos técnicos, até o produto final, como o simulador, não só atenderam às expectativas do trabalho, como criaram oportunidades para próximos estudos no futuro.

REFERÊNCIAS

- [1] CNN Brasil. *Produção de gás natural no Brasil em 2021 atinge recorde após alta de 5%*. [S.l.], 2022. Disponível em: <<https://www.cnnbrasil.com.br/business/producao-de-gas-natural-no-brasil-em-2021-atinge-recorde-apos-alta-de-5/>>. Acesso em: 2022-06-01.
- [2] Comgás. *Gás natural encanado faz toda diferença*. [S.l.], 2022. Disponível em: <<https://www.comgas.com.br/gas-natural-encanado-faz-toda-diferenca/>>. Acesso em: 2022-06-01.
- [3] SÃO PAULO (Estado). *Deliberação ARSESP Nº 1.257*. [S.l.], 2022. Disponível em: <<http://www.arsesp.sp.gov.br/LegislacaoArquivos/lld12572021.pdf>>. Acesso em: 2022-06-01.
- [4] Comgás. *Segurança no Fornecimento*. [S.l.], 2022. Disponível em: <<https://www.comgas.com.br/a-comgas/onde-atuamos/indicadores-arsesp/seguranca-no-fornecimento/>>. Acesso em: 2022-06-01.
- [5] KOGLER, C.; RAUCH, P. Contingency plans for the wood supply chain based on bottleneck and queuing time analyses of a discrete event simulation. *MDPI*, 2020. Disponível em: <<https://www.mdpi.com/1999-4907/11/4/396>>.
- [6] PINHO, T. et al. Routing and schedule simulation of a biomass energy supply chain through Simpy simulation package. *Applied Computing and Informatics*, 2018. Disponível em: <<https://www.emerald.com/insight/content/doi/10.1016/j.aci.2018.06.004/full/html>>.
- [7] . *Discrete event simulation for Python*. [S.l.], 2020. Disponível em: <<https://simpy.readthedocs.io/en/latest/>>. Acesso em: 2022-07-04.
- [8] Fernando Soares. *Teoria das filas: entenda como funciona no cotidiano*. [S.l.], 2018. Disponível em: <<https://blog.cmtecnologia.com.br/teoria-das-filas/>>. Acesso em: 2022-06-30.
- [9] EBERT, A. et al. Computationally efficient simulation of queues: The R package queuecomputer. *Journal of Statistical Software*, 2020. Disponível em: <<https://www.jstatsoft.org/article/view/v095i05>>.
- [10] JIA, W. et al. Modeling taxi drivers' decisions at airport based on queueing theory. *Elsevier*, 2022. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0739885921000652>>.
- [11] UTAMA, D. M. et al. The vehicle routing problem for perishable goods: A systematic review. *Cogent Engineering*, 2020. Disponível em: <<https://doi.org/10.1080/23311916.2020.1816148>>.

- [12] Ivan Leonardi. *Tudo que você precisa saber sobre o Sistema de Informação Geográfica.* [S.l.], 2020. Disponível em: <<http://geoeduc.com/2020/07/21/tudo-que-voce-precisa-saber-sobre-o-sistema-de-informacao-geografica/>>. Acesso em: 2022-07-04.
- [13] ZHANG, M. The use and value of Geographic Information Systems in transportation modeling. *Elsevier*, 2021. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780081026717103641>>.
- [14] LYU, Z. et al. Freight operations modelling for urban delivery and pickup with flexible routing: Cluster transport modelling incorporating discrete-event simulation and GIS. *MDPI*, 2021. Disponível em: <<https://www.mdpi.com/2412-3811/6/12/180>>.
- [15] ReSolution. *Resilient Systems for Land Use Transportation.* [S.l.], 2022. Disponível em: <<https://resolution.centrodametropole.fflch.usp.br/resolution/index.html>>. Acesso em: 2022-10-16.
- [16] Google. *Constraint Optimization.* [S.l.], 2022. Disponível em: <<https://developers.google.com/optimization/cp>>. Acesso em: 2022-12-01.
- [17] Scikit-Learn developers. *AgglomerativeClustering.* [S.l.], 2022. Disponível em: <<https://scikit-learn.org/stable/modules/generated/sklearn.cluster.AgglomerativeClustering.html>>. Acesso em: 2022-12-01.
- [18] Secretaria Municipal de Urbanismo e Licenciamento. *Velocidade Média no Trânsito RMSP.* [S.l.], 2022. Disponível em: <<https://www.prefeitura.sp.gov.br/cidade/secretarias/upload/urbanismo/infocidade/graficos/12,etc/2022-11-25>>.