



## BiogasCluster

### Descritivos gerais da ferramenta de Clusterização



ORGANIZAÇÃO DAS NAÇÕES UNIDAS  
PARA O DESENVOLVIMENTO INDUSTRIAL



GLOBAL ENVIRONMENT FACILITY  
INVESTING IN OUR PLANET

MINISTÉRIO DO  
DESENVOLVIMENTO REGIONAL

MINISTÉRIO DO  
MEIO AMBIENTE

MINISTÉRIO DE  
MINAS E ENERGIA

MINISTÉRIO DA  
AGRICULTURA, PECUÁRIA  
E ABASTECIMENTO

MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



## Parceiros do Projeto



## Parceiros nesta Atividade



## Comitê Diretor do Projeto



ORGANIZAÇÃO DAS NAÇÕES UNIDAS  
PARA O DESENVOLVIMENTO INDUSTRIAL



GLOBAL ENVIRONMENT FACILITY  
INVESTING IN OUR PLANET

MINISTÉRIO DO  
DESENVOLVIMENTO REGIONAL

MINISTÉRIO DO  
MEIO AMBIENTE

MINISTÉRIO DE  
MINAS E ENERGIA

MINISTÉRIO DA  
AGRICULTURA, PECUÁRIA  
E ABASTECIMENTO

MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



[www.gefbiogas.org.br](http://www.gefbiogas.org.br)

This project/program is funded by the Global Environment Facility

# Projeto “Aplicações do Biogás na Agroindústria Brasileira” (GEF Biogás Brasil)



Este documento está sob a licença Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License. Citações ao material deste documento devem ser da seguinte forma:

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES (Brasil); UNIDO, ORGANIZAÇÃO DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO INDUSTRIAL; UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ. Biogás Cluster: Descritivos gerais da ferramenta de Clusterização. MCTI. Brasília-DF. (Projeto Aplicações do Biogás na Agroindústria Brasileira: GEF Biogás Brasil).

## COMITÊ DIRETOR DO PROJETO

Fundo Global para o Meio Ambiente

Ministério da Ciência, Tecnologia e Inovações

Organização das Nações Unidas para o Desenvolvimento Industrial

Ministério da Agricultura, Pecuária e Abastecimento

Ministério de Minas e Energia

Ministério do Meio Ambiente

Ministério do Desenvolvimento Regional

Centro Internacional de Energias Renováveis

Itaipu Binacional

## PARCEIROS DO PROJETO

Serviço Brasileiro de Apoio às Micro e Pequenas Empresas

Associação Brasileira do Biogás

Empresa Brasileira de Pesquisa Agropecuária

## FICHA TÉCNICA

### Nome do produto:

Biogás Cluster: Descritivos gerais da ferramenta de Clusterização

### Atividade vinculada:

2.1.1 Validação dos modelos de negócios de biogás e biometano nas agroindústrias, inclusive esquemas associativos de produção de biogás.

### Publicado por:

Ministério da Ciência, Tecnologia e Inovações  
Organização das Nações Unidas para o Desenvolvimento Industrial

### Entidade(s) diretamente envolvida(s):

Universidade Tecnológica Federal do Paraná

### Autoria:

Antonio Carlos de Francisco  
Thalita Monteiro Obal  
Jovani Taveira De Souza  
Rômulo Henrique Gomes De Jesus  
Cassiano Moro Piekarski  
Daniel Poletto Tesser  
Fabio Neves Puglieri  
Julia Eunice Fernandes (UTFPR)  
Karoline Aquino Santos (UTFPR)  
Pâmela Brocardo Cruz (UTFPR)  
André Luiz Przybysz

### Coordenação:

Antonio Carlos de Francisco e Cassiano Moro Piekarski

### Editoração:

Nicole Mattiello

Data da publicação: Brasília, Agosto de 2021





## APRESENTAÇÃO

O Projeto “Aplicações do Biogás na Agroindústria Brasileira” (GEF Biogás Brasil) reúne o esforço coletivo de organismos internacionais, setor privado, entidades setoriais e do Governo Federal em prol da diversificação da matriz energética do país por meio do biogás.

O Projeto é liderado pelo Ministério da Ciência, Tecnologia e Inovações (MCTI), implementado pela Organização das Nações Unidas para o Desenvolvimento Industrial (UNIDO), financiado pelo Fundo Global para o Meio Ambiente (GEF), e conta com o Centro Internacional de Energias Renováveis (CIBiogás) como principal entidade executora.

O objetivo do Projeto é reduzir a emissão de gases de efeito estufa, fortalecendo as cadeias de valor e inovação tecnológica ligadas à produção de biogás. Por meio de ações concretas, o Projeto amplia a oferta de energia e combustível no Brasil a partir da geração de biogás e biometano, fortalecendo as cadeias nacionais de fornecimento de tecnologia no setor e facilitando investimentos.

O biogás é uma fonte renovável de energia elétrica, energia térmica e combustível. Seu processamento também resulta em biofertilizantes de alta qualidade para uso agrícola. A gestão sustentável dos resíduos orgânicos provenientes da agroindústria e de ambientes urbanos por meio da produção de biogás traz um diferencial competitivo para a economia brasileira. Desenvolver a cadeia de

valor do biogás significa investir em uma economia circular envolvendo inovação e novas oportunidades de negócios. Indústrias de equipamentos e serviços, concessionárias de energia e gás, produtores rurais e administrações municipais estão entre os beneficiários do Projeto, que conta com US \$7,828,000 em investimentos diretos.

Com abordagem inicial na Região Sul e no Distrito Federal, o Projeto gera impactos positivos para todo o país. As atividades do Projeto incluem a atuação direta junto a empresas, cooperativas e entidades da governança do biogás para implementar acordos de cooperação, fazer análises de mercado, desenvolver modelos de negócio inovadores e atrair investimentos nacionais e internacionais.

O Projeto também investe diretamente na otimização de plantas de biogás mais eficientes, seguras e com modelos replicáveis, entregando ao mercado exemplos práticos de sucesso operacional. Além disso, o Projeto desenvolve ferramentas digitais e atividades de capacitação que atualizam e dinamizam o setor, facilitando o desenvolvimento de projetos executivos de biogás. Em paralelo, especialistas do Projeto desenvolvem estudos técnicos com dados inéditos que apoiam o avanço de políticas públicas favoráveis ao biogás. Dessa forma, o Projeto entrega para o mercado brasileiro mais competitividade, fomentando o biogás como um grande catalizador de novas oportunidades.





# **BiogasCluster**

## **Descritivos gerais da ferramenta de Clusterização**

**Data da Publicação:**

**Julho/2020**



# Sumário

<b>1. Visão Geral .....</b>	<b>8</b>
<b>2. Termos utilizados .....</b>	<b>8</b>
<b>3. Manual da Ferramenta .....</b>	<b>9</b>
<b>3.1. IMPORTAÇÃO DE DADOS.....</b>	<b>12</b>
<b>3.2. OPÇÕES DE FILTRO .....</b>	<b>15</b>
<b>3.3. ANÁLISE DA SOLUÇÃO .....</b>	<b>18</b>
<b>3.4. SATISFAÇÃO DO USUÁRIO COM OS RESULTADOS.....</b>	<b>20</b>
<b>3.5. ROTA À UNIDADE CENTRALIZADORA.....</b>	<b>23</b>
<b>3.6. CONTATO DOS DESENVOLVEDORES.....</b>	<b>25</b>
<b>4. Etapas do Algoritmo .....</b>	<b>26</b>
<b>4.1. PRÉ-PROCESSAMENTO DOS DADOS .....</b>	<b>26</b>
<b>4.1.1. Etapa 1: Seleção de unidades baseada no potencial de produção de biogás.....</b>	<b>27</b>
<b>4.1.2. Etapa 2: Exclusão das unidades outliers.....</b>	<b>27</b>
<b>4.1.3. Etapa 3: Preferência para localização da unidade centralizadora de biogás.....</b>	<b>28</b>
<b>4.2. PROCESSAMENTO DOS DADOS.....</b>	<b>28</b>
<b>5. Limitações .....</b>	<b>30</b>
<b>6. Case de sucesso da Ferramenta .....</b>	<b>30</b>
<b>6.1. A EMPRESA.....</b>	<b>30</b>
<b>6.2. OS RESULTADOS .....</b>	<b>31</b>
<b>6.3. CONSIDERAÇÕES FINAIS DO CASE. ....</b>	<b>40</b>
<b>7. Considerações Finais .....</b>	<b>41</b>

<b>Apêndice 1</b> .....	42
<b>Apêndice 2</b> .....	48

# Sumário Executivo

## PORTUGUÊS

A ferramenta denominada “BiogasCluster” (acessível pelo seguinte link: <https://biogasclustering.shinyapps.io/Cluster/>) foi projetada e desenvolvida a partir de um algoritmo hierárquico para clusterização. O algoritmo e a ferramenta foram estruturados pela equipe de professores e pesquisadores da Universidade Tecnológica Federal do Paraná ([UTFPR](#)) – Campus Ponta Grossa, juntamente ao Programa de Pós-Graduação em Engenharia de Produção ([PPGEP](#)) e o Laboratório de Estudos em Sistemas Produtivos Sustentáveis ([LESP](#)). A BiogasCluster é inédita no mercado e foi pensada e projetada para que seja flexível, de fácil utilização e intuitiva. O principal objetivo da BiogasCluster é fornecer ao tomador de decisão informações confiáveis, abrangentes e flexíveis com vistas ao desenvolvimento territorial de uma determinada região. Ela forma e organiza os produtores da cadeia agroindustrial em clusters para a produção de biogás, levando em consideração: localização geográfica, potencial de produção de biogás, proximidades da rede de gasoduto e dos maiores consumidores de bioenergia, dentre outros fatores. Com isso, a ferramenta oferece alternativas para os produtores que não possuem potencial para produção individual de biogás se organizarem em clusters com capacidade produtiva, dando destino aos resíduos produzidos em suas propriedades, minimizando assim, os possíveis impactos ambientais decorrentes da atividade, gerando uma nova oportunidade de negócio e aumentando a competitividade da cadeia agroindustrial.

**Palavras-chave:** biogás; cluster; bioenergia; agroindústria; desenvolvimento territorial.

## ENGLISH

The tool called “BiogasCluster” (<https://biogasclustering.shinyapps.io/Cluster/>) was designed and developed using a hierarchical algorithm for clustering. The algorithm and the tool were structured by the team of professors and researchers from the Universidade Tecnológica Federal do Paraná ([UTFPR](#)) – Campus Ponta Grossa, together with the Postgraduate Program in Production Engineering ([PPGEP](#)) and the Laboratory for the Study of Sustainable Productive Systems ([LESP](#)). BiogasCluster is unprecedented in the market and was thought and designed to be flexible, easy to use and intuitive. The



## BIOGASCLUSTER

main objective of BiogasCluster is to provide decision makers with reliable, comprehensive and flexible information aimed at the territorial development of a given region. It trains and organizes producers in the cluster agro-industrial chain for the production of biogas, taking into consideration: geographic location, potential for biogas production, proximity to the gas pipeline network and the largest bioenergy consumers, among other factors. The tool offers alternatives for producers who do not have the potential for individual biogas production to organize themselves into clusters with productive capacity, giving destination to the waste produced on their properties, thus minimizing the possible environmental impacts resulting from the activity, generating a new business opportunity and increasing the competitiveness of the agro-industrial chain.

**Keywords:** biogas; cluster; bioenergy; agribusiness; territorial development.

### 1. Visão Geral

O presente relatório tem o objetivo de apresentar o descritivo da ferramenta BiogasCluster, incluindo principalmente o manual de uso, bem como as etapas do algoritmo. O aplicativo de apresentação dos resultados está temporariamente hospedado no site: <https://biogasclustering.shinyapps.io/Cluster/>, e poderá ser localizado em [www.gefbiogas](http://www.gefbiogas)

É importante ressaltar que durante todo o seu desenvolvimento a ferramenta teve o acompanhamento de parceiros, como: UNIDO, CIBiogás, SEBRAE, ITAI e Castrolanda, que por meio de reuniões presenciais e virtuais observaram os resultados da ferramenta e deram feedbacks que contribuíram para a implementação de melhorias ao código auxiliando a molda-lo para sua versão atual.

As seções a seguir apresentam os termos utilizados, o manual da ferramenta, o descritivo do algoritmo desenvolvido para obtenção dos resultados, as limitações da ferramenta, o case da aplicação da ferramenta na base de dados de uma empresa e as considerações finais.

### 2. Termos utilizados

Os termos descritos a seguir serão utilizados ao decorrer deste relatório, portanto é de suma importância que o leitor conheça seu contexto para o melhor entendimento da ferramenta apresentada.

- a) Unidade: Objeto desejado para formação de cluster, a unidade deve ser referenciada geograficamente e informada na importação de dados. A unidade pode ser, por exemplo, propriedade rural, unidade industrial, entre outras possibilidades.
- b) Unidade centralizadora de biogás: Biodigestor, usina de beneficiamento de biogás, hub de biomassa, central de tratamento de resíduos, entre outras possibilidades.
- c) Abrangência: Maior distância dentre as unidades de um cluster e sua respectiva unidade centralizadora de biogás (em metros).

- d) Ponto ideal (PI): Mínima amplitude de abrangência, máxima produção mínima de biogás.
- e) Filtros: são atributos do problema que podem passar por alterações a partir de escolhas feitas pelo usuário.

### 3. Manual da Ferramenta


Ao acessar a ferramenta aparece uma mensagem (Figura 1), com intuito de apresentar o objetivo principal da ferramenta, que é identificar agrupamentos ideais de unidades na agroindústria com potencial de criação de parceria estratégica para produção de biogás, e informar que os inputs da ferramenta são feitos pelo usuário. Para fechar a mensagem e iniciar a navegação nas funcionalidades da BiogasCluster o usuário deve selecionar o botão .



Figura 1: Mensagem de apresentação da ferramenta.

Na página inicial, o usuário poderá escolher diferentes opções de seleção, conforme ilustrado na Figura 2 (ver figura na página seguinte).

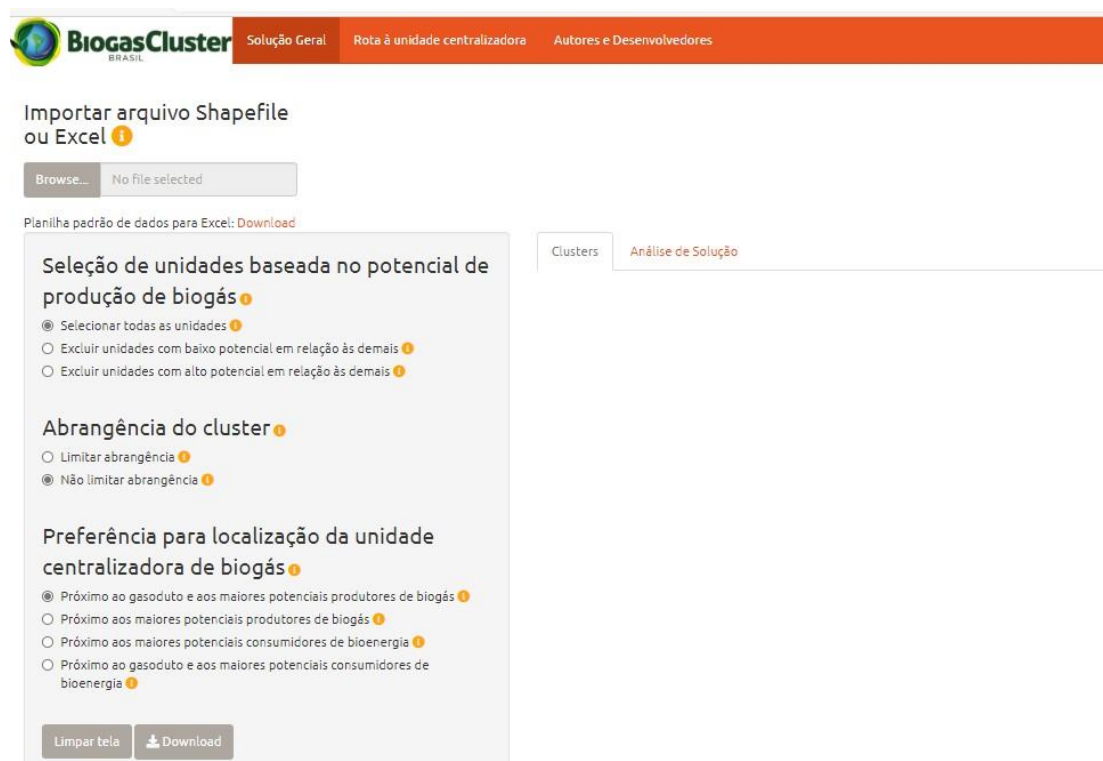


Figura 2: Interface inicial da ferramenta.

De modo geral a ferramenta BiogasCluster pode ser dividida em 6 etapas, sendo elas:

- a) Importação de dados;
- b) Opções de filtro;
- c) Resultado dos clusters;
- d) Análise de solução;
- e) Satisfação com a solução;
- f) Rota à unidade centralizadora.

Estas etapas possuem uma ordem, como demonstrado no fluxograma a seguir (ver na página seguinte).

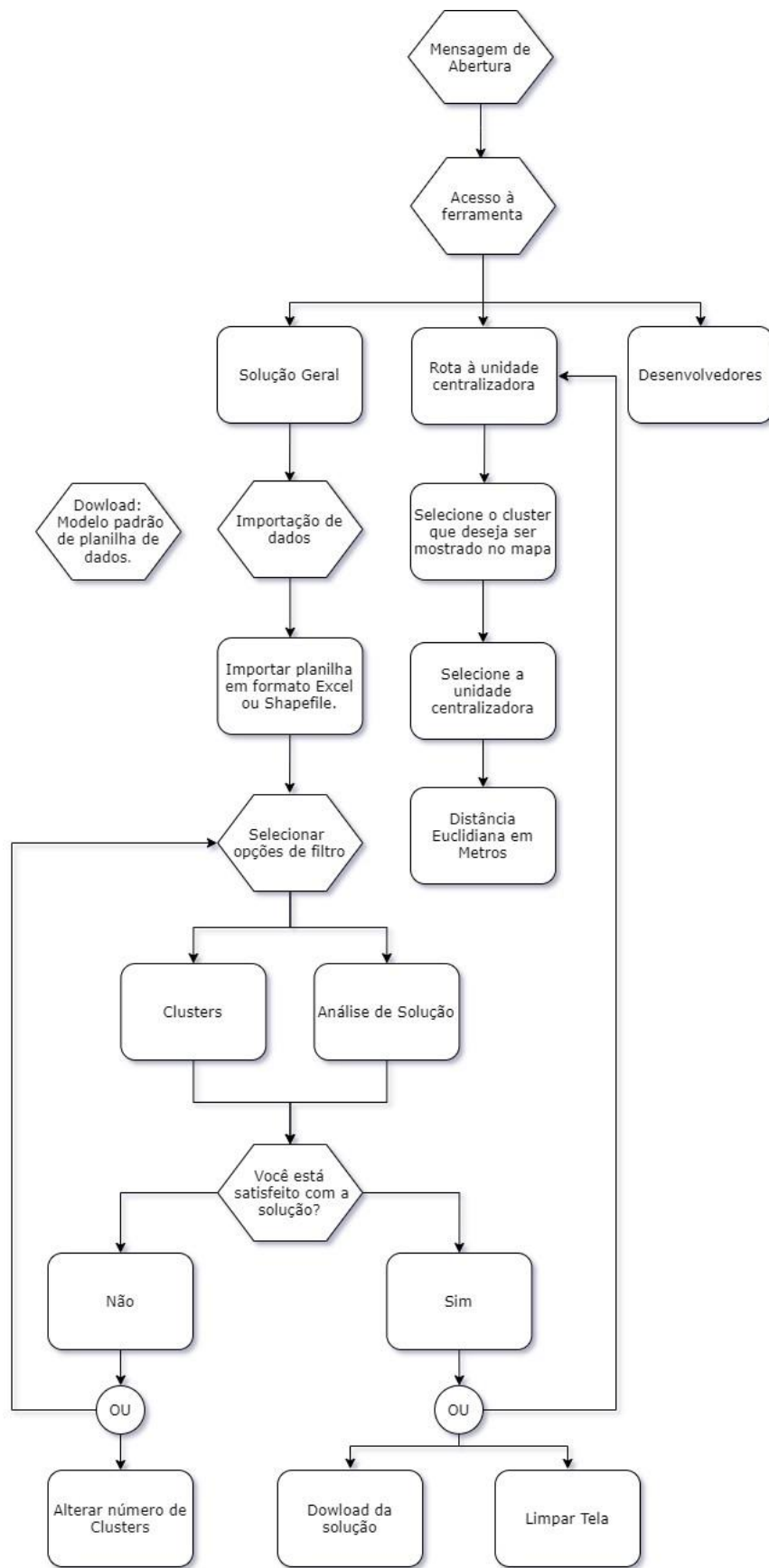




Figura 3: Fluxograma das etapas da ferramenta.

Nas seções 3.1, 3.2, 3.3, 3.4 e 3.5, as etapas anteriormente citadas terão seus detalhes descritos.

### 3.1. IMPORTAÇÃO DE DADOS

A importação de dados é essencial para o funcionamento da ferramenta, pois a partir dos dados importados pelo usuário o algoritmo poderá calcular a solução geral e a rota à unidade centralizadora.

Dessa forma, com a seção de solução geral selecionada (Figura 4), o próximo passo é clicar no botão **Download**, conforme destacado na Figura 5, que iniciará automaticamente o download do modelo de planilha de dados, que será salvo no dispositivo do usuário como “Base-de-dados-padrão.xlsx”.

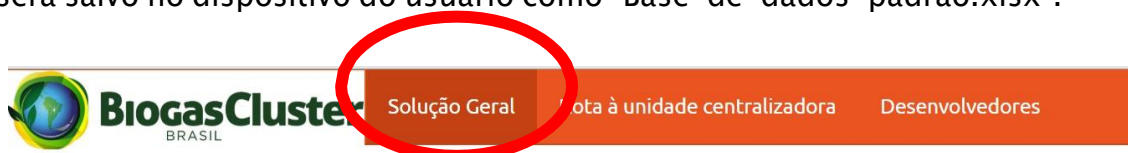


Figura 4: Solução geral.

#### Importar arquivo Shapefile ou Excel

Planilha padrão de dados para Excel: **Download**

Figura 5: Download do Modelo de Planilha.




Ao entrar em  **Base-de-dados-padrão** o usuário deverá incluir os dados no formato da Tabela 1 (na página seguinte), para que o algoritmo determine a solução de clusterização dos dados importados.

Tabela 1 – Modelo de tabela de importação dos dados de entrada.

Unidade	Latitude	Longitude	Distância	Biogás	Consumo

Os dados que compõem a tabela de importação são:

- a) Unidade: Objeto desejado para formação de cluster;
- b) Latitude: Coordenada geográfica latitude (geographic coordinate system, SIRGAS 2000);
- c) Longitude: Coordenada geográfica longitude (geographic coordinate system, SIRGAS 2000);
- d) Distância: Menor distância euclidiana entre unidade centralizadora de biogás e gasoduto (metros);
- e) Biogás: Potencial volume produção de biogás (m<sup>3</sup>/dia);
- f) Consumo: Potencial de consumo de bioenergia das unidades.

Com a base de dados salva, o usuário poderá efetivamente realizar a importação dos dados. Para isso basta clicar no botão  a esquerda do ícone de importação de dados (Figura 5) e selecionar o arquivo no dispositivo do usuário. Quando a barra  aparecer, como no exemplo abaixo (Figura 6, na página seguinte), significa que os dados já foram salvos na ferramenta e o algoritmo está pronto para gerar a clusterização.

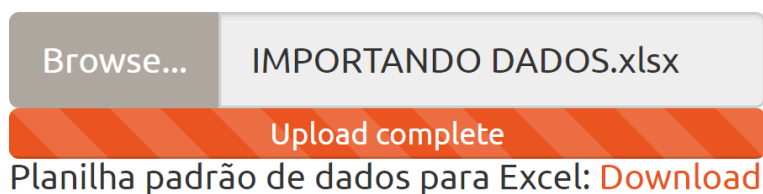


Figura 6: Upload completo de dados.

Caso a importação não seja concluída devido a algum problema com a base de dados, aparecerá uma mensagem de “Erro”. O usuário deverá fechá-la para realizar uma nova tentativa de importação dos dados.

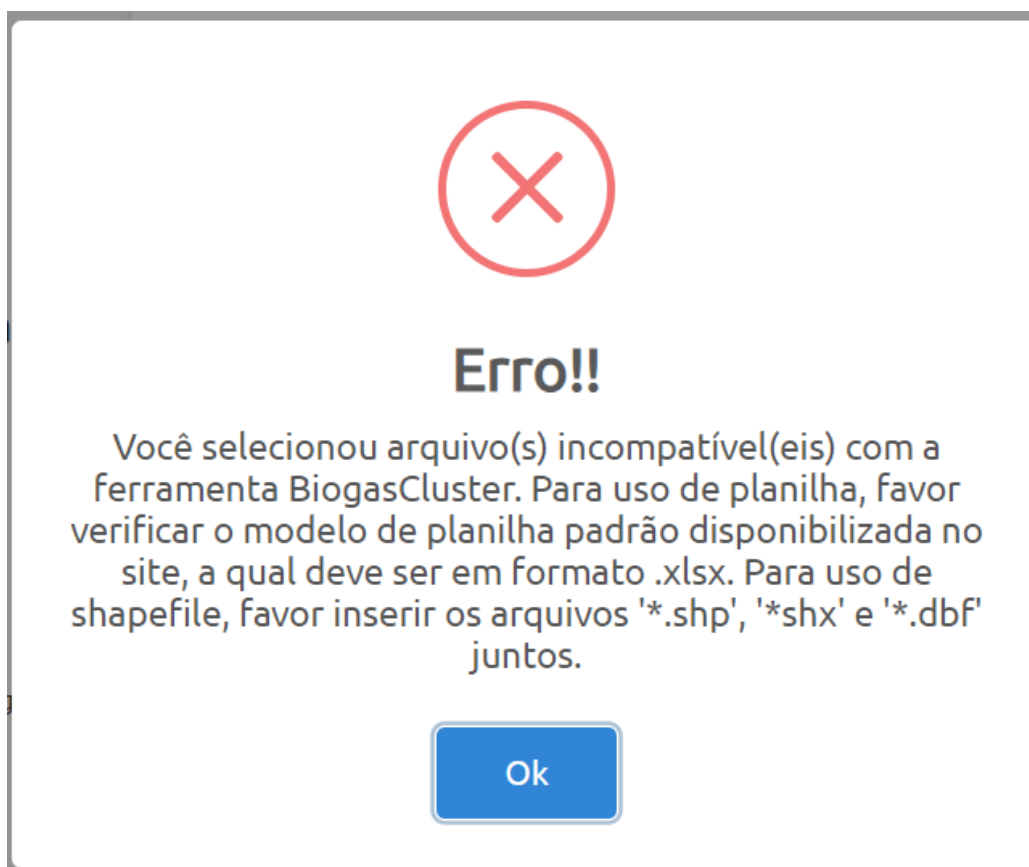


Figura 7: Erro de importação de dados.

Quando o arquivo importando na ferramenta for shapefile deve-se inserir os arquivos (.shp), (.shx) e (.dbf) do respectivo shapefile. Para que a ferramenta visualize os dados dos arquivos importados, este deve possuir uma tabela de atributo, conforme Tabela 2 (na página seguinte).


Tabela 2 – Modelo de tabela de importação dos dados de entrada shapefile.

FID	Shape	Unidade	Latitude	Longitude	Distância	Biogás	Consumo

Os dados preenchidos na tabela de atributo possuem as mesmas características dos dados usados para importação de arquivo Excel.

### 3.2. OPÇÕES DE FILTRO

Com os dados importados, o próximo passo é selecionar as opções de filtro que permitirão obter diversos resultados de clusters, criando diferentes cenários para auxiliar os tomadores de decisão.

Em cada uma das opções de filtro, há um ícone de informação . Ao colocar o cursor do mouse sobre esse ícone, a ferramenta fornecerá ao usuário uma explicação detalhada da opção selecionada, auxiliando o entendimento do usuário e aumentando a efetividade da ferramenta.

Há três tipos de filtros:

1. Filtro 1: possibilita ao usuário selecionar todas as unidades ou excluir unidades consideradas como outliers no conjunto de dados de entrada, conforme Figura 8.
  - a. Selecionar todas as unidades: todas as unidades importadas no BiogasCluster são selecionadas para formação de clusters.
  - b. Excluir unidades com baixo potencial em relação às demais: exclui unidades com baixo potencial de produção de biogás em relação às demais (outliers inferiores).
  - c. Excluir unidades com alto potencial em relação às demais: exclui unidades com alto potencial de produção de biogás em relação às demais (outliers superiores).

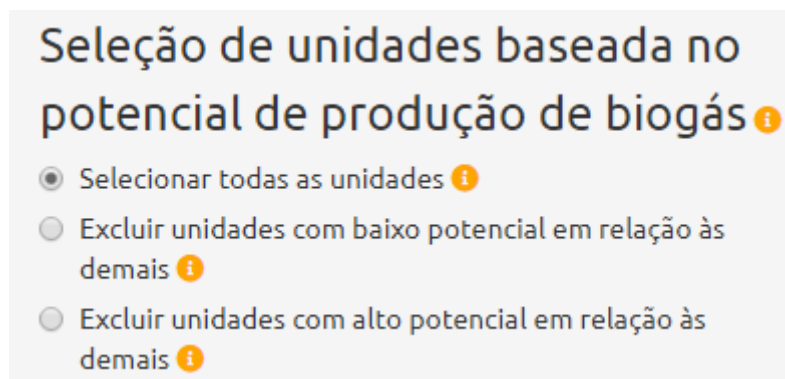


Figura 8: Seleção de unidades baseada no potencial de produção de biogás.

2. Filtro 2: a abrangência do cluster é definida pelo usuário conforme Figura 9. Neste filtro o usuário pode escolher entre limitar a abrangência e não limitar a abrangência. As opções de seleção de abrangência são:

- a. Limitar abrangência: Exclui formação de clusters que possuem unidades georreferenciadas dispersas, isto é, abrangência maior que o outlier superior das possíveis soluções de clusters.
- b. Não limitar abrangência: Não limita abrangência dos clusters. (Não define abrangência mínima e/ou máxima dos clusters)

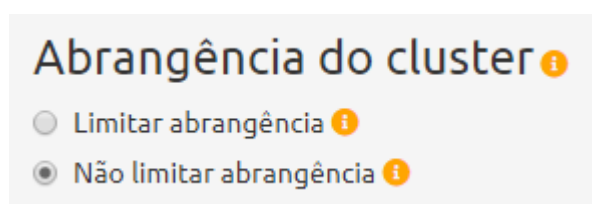


Figura 9: Definir abrangência do cluster.

3. Filtro 3: o algoritmo permite ao usuário expressar sua preferência pela localização da unidade centralizadora baseado na proximidade ao gasoduto, potenciais produtores de biogás e potenciais consumidores de biogás, conforme Figura 10 (na página seguinte).



## Preferência para localização da unidade centralizadora de biogás





- Próximo ao gasoduto e aos maiores potenciais produtores de biogás 
- Próximo aos maiores potenciais produtores de biogás 
- Próximo aos maiores potenciais consumidores de bioenergia 
- Próximo ao gasoduto e aos maiores potenciais consumidores de bioenergia 

Figura 10: Preferência para localização da unidade centralizadora de biogás.

As opções de seleção de preferência para localização da unidade centralizadora são:

- a. Próximo ao gasoduto e aos maiores potenciais produtores de biogás: Tende a localizar a unidade centralizadora próximo aos maiores potenciais produtores de biogás importados por arquivo no BiogasCluster e ao gasoduto.
- b. Próximo aos maiores potenciais produtores de biogás: Tende a localizar a unidade centralizadora próximo aos maiores potenciais produtores de biogás importados por arquivo no BiogasCluster.
- c. Próximo aos maiores potenciais consumidores de bioenergia: Tende a localizar a unidade centralizadora próximo aos potenciais consumidores de bioenergia importados por arquivo no BiogasCluster.
- d. Próximo ao gasoduto e aos maiores potenciais consumidores de bioenergia: Tende a localizar a unidade centralizadora próximo aos potenciais consumidores de bioenergia importados por arquivo no BiogasCluster e ao gasoduto.

O tomador de decisão deve selecionar as opções que melhor se enquadram à sua realidade.

### 3.3. ANÁLISE DA SOLUÇÃO

Após selecionados os filtros, o usuário possui três formas de analisar a solução:

1. Tabela: a tabela apresentada abaixo exibe um exemplo de resultado proposto após a implementação do algoritmo (Figura 11).

Cluster	Unidades	Abrangência (metros)	Potencial de Biogás (m <sup>3</sup> /dia)
1	1, 2, 3, 4, 11, 12, 13, 14, 15, 16, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30	4,814.17	3,984.19
2	5, 6, 7, 8, 9, 10, 17, 18, 19, 20	2,629.05	1,254.98

Showing 1 to 2 of 2 entries Previous  Next

Figura 11: Tabela de saída dos resultados gerados pela ferramenta.

A tabela exibe:

- a. Cluster: Identificação do cluster;
  - b. Unidades: Unidades que pertencem ao respectivo cluster;
  - c. Abrangência (metros): Maior distância dentre as unidades de um cluster e sua respectiva unidade centralizadora de biogás;
  - d. Potencial de biogás (m<sup>3</sup>/dia): Expressa a quantidade total potencial de biogás do respectivo cluster.
2. Mapa: Outra forma de apresentação do resultado é por meio da visualização de um mapa (Figura 12), nos quais são representados os agrupamentos estabelecidos através do processo de clusterização.

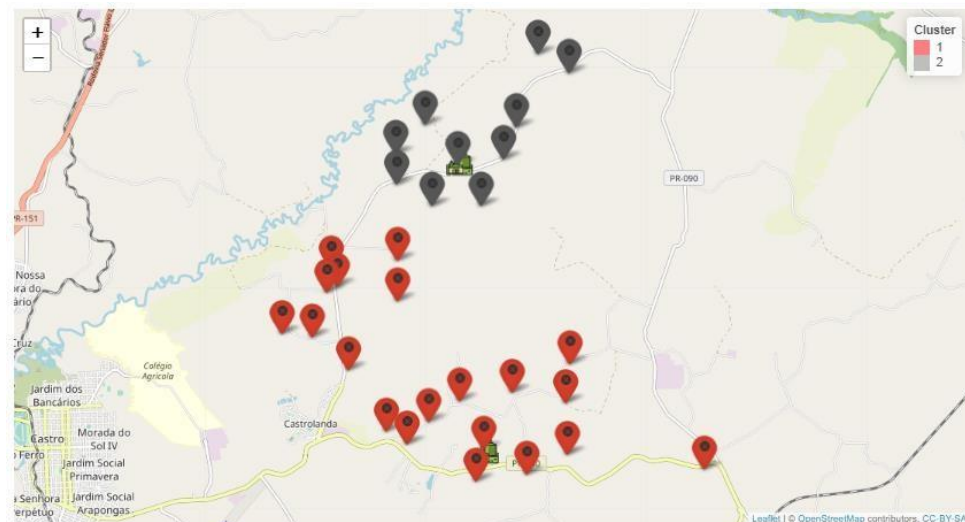


Figura 12: Mapa dos clusters gerados.

O mapa exibe o referenciamento geográfico das unidades e a identificação dos clusters por cores.

3. Análise de Solução: É possível realizar a análise da solução dos possíveis números de clusters por meio das cartas de controle de variabilidade da abrangência e variabilidade de produção de biogás dos clusters. Estas permitem a análise da sensibilidade da abrangência e produção de biogás com relação a variação do número de clusters.
  - a. A carta de controle “Variabilidade da abrangência dos clusters” apresenta a dispersão das abrangências em relação variação do número de clusters, conforme Figura 13.

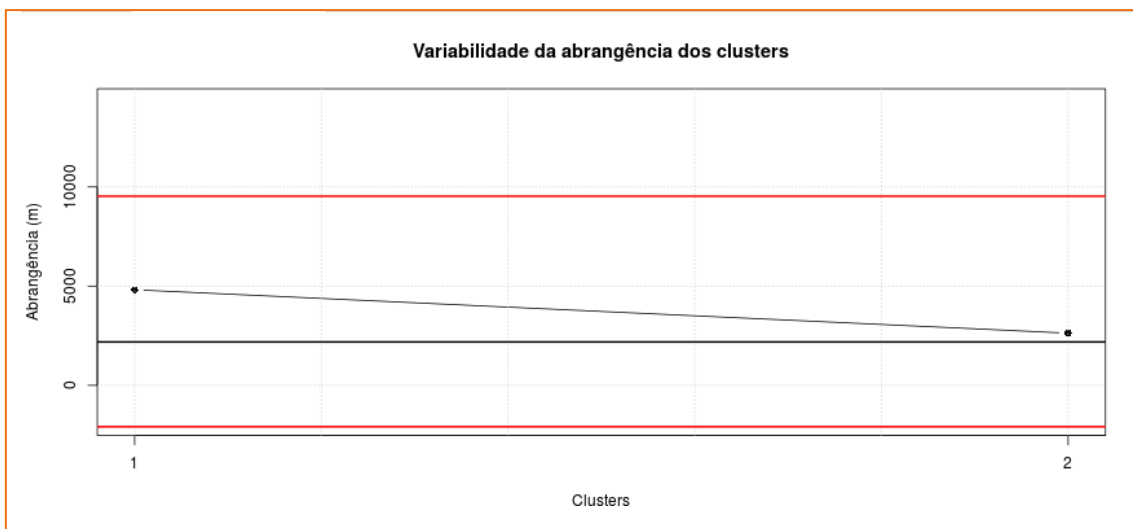


Figura 13: Variabilidade da abrangência dos clusters.

- b. A carta de controle “Variabilidade da produção de biogás dos clusters” representa a dispersão da produção de biogás em relação a variação dos números de clusters, conforme Figura 14.

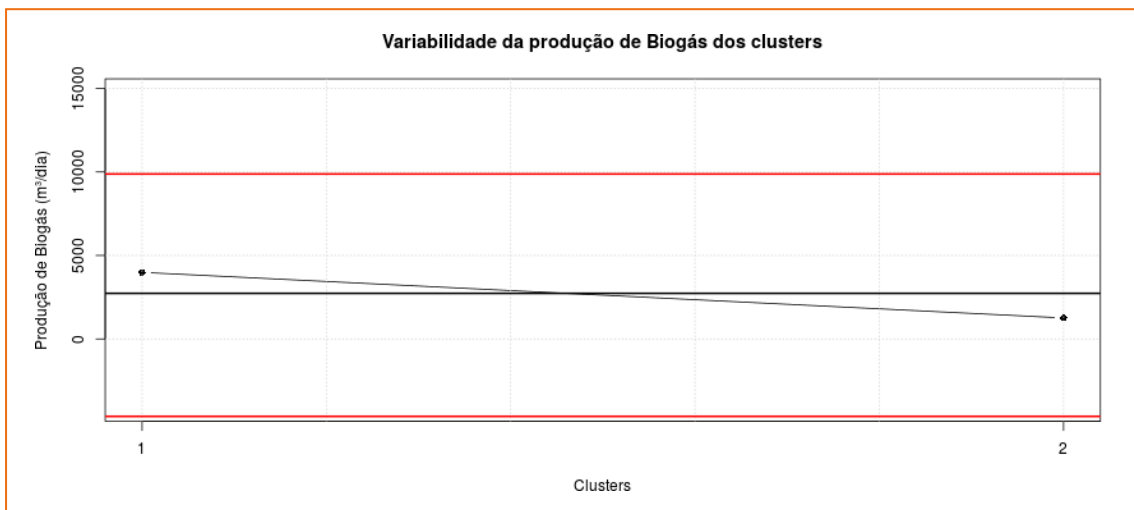


Figura 14: Variabilidade da produção de biogás dos clusters.

Para ter acesso às duas primeiras análises, o usuário deve clicar no botão  e para visualizar a terceira e a quarta análise de resultados o usuário deve selecionar a opção .

### 3.4 SATISFAÇÃO DO USUÁRIO COM OS RESULTADOS.

A ferramenta BiogasCluster permite ao usuário, após verificar a análise dos resultados, decidir se o resultado encontrado é satisfatório ou não para a realidade do tomador de decisão.

Dessa forma, após a seleção das opções aparecerá a pergunta “Você está satisfeito com a solução?”, como descrito na Figura 15 (na página seguinte).

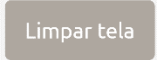
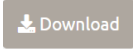
Você está satisfeito com a solução?

Sim

Não

Limpar tela    Download

Figura 15: Usuário satisfeito com os resultados.

Se o usuário optar por “sim” (Figura 15) poderá então selecionar os botões  , que apagará todos os resultados, e  que realizará automaticamente um download dos resultados encontrados para aquela solução.

Caso o usuário não esteja satisfeito com os resultados encontrados ele selecionará a opção “Não” (Figura 16) e então aparecerá um campo para mais duas opções de modificação do resultado (Figura 17, na página seguinte), sendo alterar as opções de filtro anteriormente selecionadas ou alterar o número de clusters da solução.

Você está satisfeito com a solução?

Sim

Não

Selecione uma das opções abaixo:

Nenhuma opção selecionada ▼

Figura 16: Não satisfação com os resultados.



Não

Selecione uma das opções abaixo:

Nenhuma opção selecionada ▲

Altere as opções de filtro

Altere o número de clusters

Figura 17: Alternativas para a não satisfação com os resultados.

Se o usuário optar por alterar as opções de filtro (Figura 18), devem-se modificar as opções anteriormente selecionadas para que a ferramenta gere um novo resultado.

Não

Selecione uma das opções abaixo:

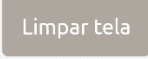

Altere as opções de filtro ▼

Figura 18: Alterar as opções de filtro.

Caso o usuário não deseje mudar as opções de filtro, ele terá a alternativa de alterar o número de clusters da solução (Figura 19, na página seguinte). De acordo com o número de cluster escolhido, o resultado será recalculado e novas análises aparecerão para o usuário.



Figura 19: Alterar o número de clusters.

Por fim, o usuário poderá optar pelos botões  e , para excluir o cenário ou salvá-lo.

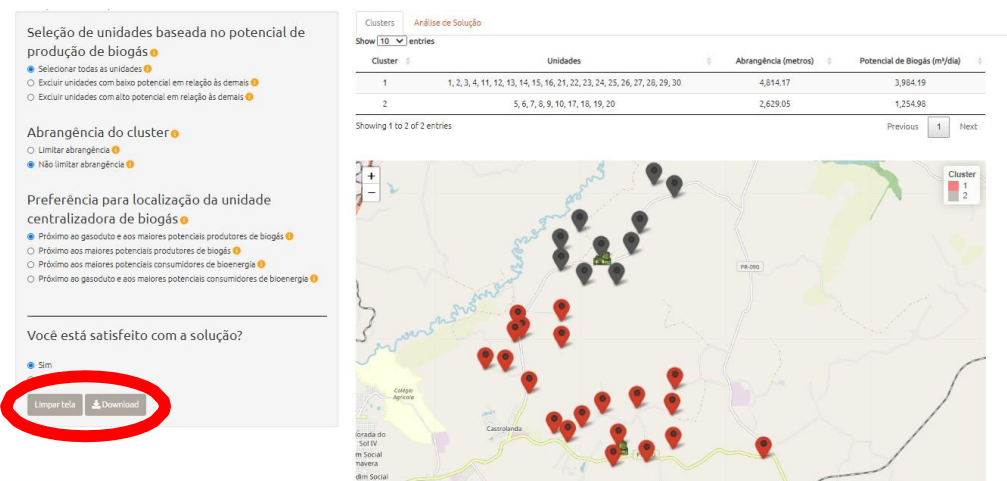


Figura 20: Interface após cálculo do cluster.

### 3.5. ROTA À UNIDADE CENTRALIZADORA

Outro benefício desta ferramenta é a possibilidade de identificar a rota de uma determinada unidade até sua respectiva unidade centralizadora, bem como calcular a distância entre elas. Basta clicar em “Rota à unidade centralizadora”, como demonstrado na Figura 21 (página seguinte).



Figura 21: Rota à unidade centralizadora.

Para o cálculo, a ferramenta primeiro requer que o usuário selecione o cluster em que a unidade georreferenciada desejada está inserida (Figura 22). Esta opção refere-se à seleção de um dos clusters gerados após a escolha da solução ideal por parte do usuário.

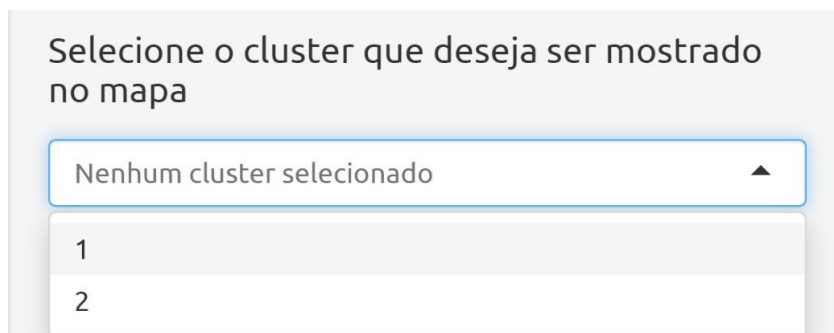


Figura 22: Seleção do cluster.

A ferramenta oferece a opção de escolher a unidade para cálculo da distância até a unidade centralizadora (Figura 23), onde apresenta todas as unidades do cluster selecionado na opção anterior.

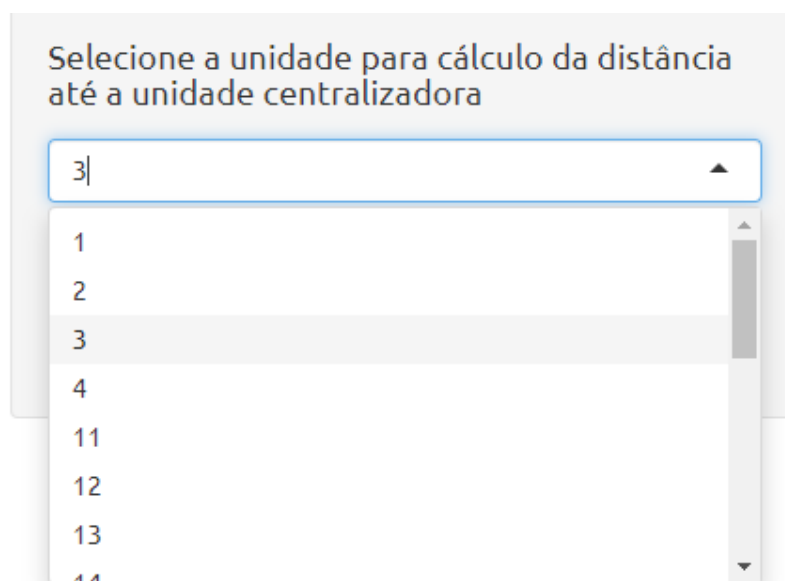


Figura 23: Seleção de unidade georreferenciada.

A Figura 24 abaixo apresenta a distância euclidiana (A), em metros, e o mapa (B) entre a unidade escolhida e a respectiva unidade centralizadora.

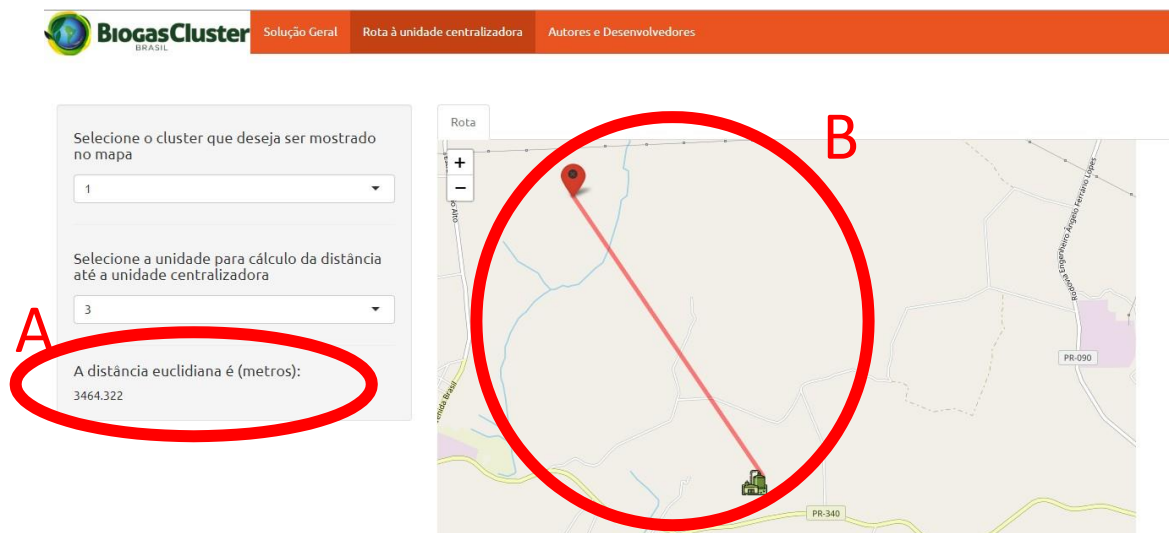


Figura 24: Interface após cálculo de distância à unidade centralizadora.

### 3.6. CONTATO DOS DESENVOLVEDORES

A ferramenta também mostra o nome dos desenvolvedores, como ilustrado na Figura 25 (na página seguinte). Ao clicar nos nomes dos autores, o usuário será direcionado ao currículo lattes do respectivo autor. Se o usuário clicar no título Laboratório de Estudos sobre Sistemas Produtivos Sustentáveis, ele será encaminhado diretamente ao site do laboratório.



Figura 25: Interface de Autores e Desenvolvedores da ferramenta.

## 4. Etapas do Algoritmo

Os pseudocódigos referentes ao algoritmo do método de clusterização presente na ferramenta BiogasCluster são descritos nesta seção, sendo o código implementado em R apresentado no Apêndice 1.

### 4.1. PRÉ-PROCESSAMENTO DOS DADOS

Inicialmente, os dados recebidos pelo input do usuário no BiogasCluster passam por pré-processamento, que consiste na seleção das unidades que farão parte dos clusters baseada nas etapas 1 a 3 apresentadas a seguir.



#### 4.1.1. Etapa 1: Seleção de unidades baseada no potencial de produção de biogás

A seleção das unidades que serão consideradas para o agrupamento é realizada conforme input do usuário, descrito em 3.2, e está detalhada no Algoritmo 1.

Algoritmo 1: Seleção das unidades a serem agrupadas (DadosUsuario, Filtro1)

---

```

1. Dados <- input(DadosUsuario)
2. Se Filtro 1="selecionar todas as unidades", faça:
    2.1 Dados<-Dados
3. Se Filtro 1="Excluir unidades com baixo potencial em
    relação às demais", faça:
    3.1 Determine o valor de  $L$  usando (1)
        
$$L = Q1 - 1,5 * (Q3 - Q1), \quad (1)$$

        onde:  $Q1 = \text{quartil } 1(\text{Dados}(\text{"Biogás"}))$  e  $Q3 = \text{quartil } 3(\text{Dados}(\text{"Biogás"}))$ 
    3.2 Dados<-subconjunto(Dados, "Biogás">L)
4. Se Filtro 1="Excluir unidades com alto potencial em
    relação às demais:", faça
    4.1 Determine o valor de  $U$  usando (2)
        
$$U = Q3 + 1,5 * (Q3 - Q1), \quad (2)$$

        onde,  $Q1 = \text{quartil } 1$  e  $Q3 = \text{quartil } 3$ 
    4.2 Dados<-subconjunto(Dados, "Biogás"<U)
    
```

---

#### 4.1.2. Etapa 2: Exclusão das unidades outliers

A definição de unidade outlier e tratamento destas está descrito no Algoritmo 2.

Algoritmo 2: Exclusão de unidades outliers (Dados)

---

```

1.  $(x,y)$ <-Dados("Longitude", "Latitude")
2. Determine  $d = [d_{ij}]$ , usando (3)
        
$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

3. Calcular  $m$  e  $sd$ , usando (4) e (5)
        
$$m = \text{média}(d) \quad (4)$$

        
$$sd = \text{desvio padrão}(d) \quad (5)$$

    
```

4. Calcular o limite:

$$\text{limite} = m + 3 * sd \quad (6)$$

5. `UnidadeOutlier<-subconjunto (Dados[i,],  $d_{ij} > \text{limite}$ )`

6. `Dados<-subconjunto (Dados, <>UnidadeOutlier)`

---

Após a definição do número ideal de clusters, as unidades outliers voltam ao conjunto de dados e são incluídas nos clusters que estiverem mais próximo à respectiva unidade centralizadora.

### 4.1.3. Etapa 3: Preferência para localização da unidade centralizadora de biogás

De acordo com a preferência de localização para a unidade centralizadora indicada em 3.2., o peso  $P_i$  para a determinação das coordenadas das respectivas unidades centralizadoras dos clusters é determinado como:

a) Próximo ao gasoduto e aos maiores potenciais produtores de biogás

$$P_i = (\text{potencial de produção de biogás}) / (\text{distância ao gasoduto}) \quad (7)$$

b) Próximo aos maiores potenciais produtores de biogás

$$P_i = \text{potencial de produção de biogás} \quad (8)$$

c) Próximo aos maiores potenciais consumidores de bioenergia

$$P_i = \text{potencial consumo de bioenergia} \quad (9)$$

d) Próximo ao gasoduto e aos maiores potenciais consumidores de bioenergia

$$P_i = (\text{potencial consumo de energia}) / (\text{distância ao gasoduto}) \quad (10)$$

Sendo que as medidas: “ potencial de produção de biogás ”, “ distância ao gasoduto ” e “ potencial consumo de bioenergia ” são normalizadas quando incluídas aos pares como peso.

## 4.2. PROCESSAMENTO DOS DADOS

Após a etapa de pré-processamento, os dados são processados para formação dos clusters. O método de formação de cluster BiogasCluster é apresentado no Algoritmo 3.

---

**Algoritmo 3: BiogasCluster**


---

1. Determine o número ótimo de cluster, usando o Algoritmo 4

$k < \text{--número ótimo de clusters}$

2. Agrupe as propriedades em  $k$  clusters, usando o Algoritmo 5

3. Usuário satisfeito com a solução?

3.1 Sim: FIM

3.2 Não

3.2.1 Defina outro input de filtros.

3.2.2 Retorne ao passo 2, utilizando o número de cluster a escolha do usuário.

---



---

**Algoritmo 4: Determinação do número ótimo de clusters (Dados, Filtro 2)**


---

1. Clusterização hierárquica aglomerativa, usando o Algoritmo 5.

2. Para cada cluster de cada solução  $S_k$ , determine as coordenadas  $(x_c, y_c)$  da unidade centralizadora, conforme (11) e (12)

$$x_c = \frac{\sum_{i=1}^n P_i * x_i}{\sum_{i=1}^n P_i} \quad (11)$$

$$y_c = \frac{\sum_{i=1}^n P_i * y_i}{\sum_{i=1}^n P_i} \quad (12)$$

3. Se Filtro 2= "Limitar abrangência", faça

3.1 Calcule:

$$u = q_3 + 1,5 * (q_3 - q_1) \quad (13)$$

onde  $q_1$  e  $q_3$  são o quartil 1 e 3, respectivamente, do conjunto  $S_k\{\text{abrangência máxima}\}$

3.2 Selecione as soluções  $S_k$  para as quais se verifique:

$$\text{abrangência máxima} < u \quad (14)$$

4. Determinação do Ponto Ideal

*Ponto Ideal*

= (mínima amplitude de abrangência, máxima produção mínima de biogás)

5. Determinação do ponto solução mais próximo ao Ponto Ideal

**Algoritmo 5: Clusterização hierárquica aglomerativa**

1. Defina cada propriedade como um cluster
  2. Calcule a distância  $d_{ij}$  entre os clusters  $d_{ij}$   
= *média das distâncias entre as propriedades do cluster  $i$  e as propriedades do cluster  $j$*
  3. Agrupe os clusters  $i$  e  $j$  que apresentem a distância mínima  $d_{ij}$ , formando  $k$  clusters.
  4. Uma solução  $S_k$  é definida como:  
 $S_k$   
= {*abrangência máxima, amplitude de abrangência, produção mínima de biogás*}
  5. Repita os passos 2 a 4 até que todos os clusters estejam agrupados num único cluster.
- 

## 5. Limitações

Uma das limitações da ferramenta está relacionada com o tempo de processamento, que depende do tamanho da base de dados e do dispositivo utilizado para o processamento dos mesmos, visto que o algoritmo foi implementado computacionalmente em R, tendo assim um tempo e um custo computacional que é gasto fora do servidor atual (Shiny).

Além disso, o Shiny, onde a ferramenta está hospedada, possui algumas restrições por ser uma conta gratuita, o que pode impedir algumas funcionalidades, dependendo do tamanho da base de dados.

## 6. Case de sucesso da Ferramenta

Para testar os cálculos da ferramenta, utilizou-se os dados de uma empresa parceira dos desenvolvedores, como detalhado nas seções 6.1., 6.2. e 6.3.

### 6.1. A EMPRESA

A Cooperativa é uma das mais conhecidas do Brasil e atua nos setores de agricultura, suinocultura, pecuária de leite e ovinocultura. Suas atuações

no mercado abrangem 30 municípios e sua sede está localizada em Castro, no Paraná.

Nos últimos anos, a cooperativa viu a oportunidade de investir na produção de biogás. Para isso, começou a trabalhar em conjunto com seus cooperados, capacitando-os e buscando ferramentas inovadoras que atendam às necessidades do modelo de negócio. Uma das medidas adotadas foi a parceria com o Laboratório de Estudos em Sistemas Produtivos Sustentáveis (LESP), grupo de pesquisa da Universidade Tecnológica Federal do Paraná (UTFPR), campus Ponta Grossa.

Como uma de suas propostas, o LESP apresentou a ferramenta BiogasCluster, com o objetivo de aumentar a vantagem competitiva das unidades produtoras, bem como a viabilidade da aplicação desse modelo por meio do agrupamento de fazendas para a produção conjunta de biogás. Os executivos, engenheiros, técnicos e cooperados da empresa em questão, acompanharam a construção da ferramenta desde a sua idealização até a sua conclusão, contribuindo com opiniões e considerações técnicas para a melhoria do algoritmo, possibilitando que os cálculos da ferramenta fossem mais consistentes com a realidade por considerarem a experiência de profissionais que trabalham com fazendas geradoras de biogás.

Para testar a efetividade do algoritmo, a equipe de desenvolvedores utilizou os dados disponibilizados pela cooperativa, por serem confiáveis e atuais. Na próxima seção, serão apresentados os resultados encontrados pela ferramenta a partir da base de dados enviada.

## 6.2. OS RESULTADOS

A empresa disponibilizou informações referentes a Latitude, Longitude e produção de biogás de 30 fazendas cooperadas na base de dados, que para o estudo são denominadas de unidades ou unidades georreferenciadas. Além disso, um dos desenvolvedores do algoritmo buscou informações sobre as linhas de gasoduto presentes na região, possibilitando o cálculo de distâncias entre as linhas e as unidades. Os dados foram organizados em uma planilha de Excel e importados para a ferramenta BiogasCluster.

Como não foram disponibilizados dados relacionados ao consumo de energia, o número de combinações dos filtros possíveis para clusterização foi limitado a 12, como mostrado na Tabela 3.

Tabela 3: Tabela de cenários.

<b>Cenários</b>	<b>Opções de filtro selecionadas para o cálculo do cluster.</b>
1	Selecionar todas as unidades, Limitar a abrangência, Próximo ao gasoduto e aos maiores potenciais produtores de biogás.
2	Excluir unidades com baixo potencial em relação às demais, Limitar a abrangência, Próximo ao gasoduto e aos maiores potenciais produtores de biogás.
3	Excluir unidades com alto potencial em relação às demais, Limitar a abrangência, Próximo ao gasoduto e aos maiores potenciais produtores de biogás.
4	Selecionar todas as unidades, Não limitar abrangência, Próximo ao gasoduto e aos maiores potenciais produtores de biogás.
5	Excluir unidades com baixo potencial em relação às demais, Não limitar abrangência, Próximo ao gasoduto e aos maiores potenciais produtores de biogás.
6	Excluir unidades com alto potencial em relação às demais, Não limitar abrangência, Próximo ao gasoduto e aos maiores potenciais produtores de biogás.
7	Selecionar todas as unidades, Limitar a abrangência, Próximo aos maiores potenciais produtores de biogás.
8	Excluir unidades com baixo potencial em relação às demais, Limitar a abrangência, Próximo aos maiores potenciais produtores de biogás.
9	Excluir unidades com alto potencial em relação às demais, Limitar a abrangência, Próximo aos maiores potenciais produtores de biogás.
10	Selecionar todas as unidades, Não limitar abrangência, Próximo aos maiores potenciais produtores de biogás.
11	Excluir unidades com baixo potencial em relação às demais, Não limitar abrangência, Próximo aos maiores potenciais produtores de biogás.
12	Excluir unidades com alto potencial em relação às demais, Não limitar abrangência, Próximo aos maiores potenciais produtores de biogás.

Em reunião, os tomadores de decisão mediante conhecimento da realidade das unidades utilizadas, neste case, selecionaram estrategicamente as opções: todas as unidades e localizar a unidade centralizadora próxima ao gasoduto e aos maiores produtores de biogás, limitando assim o cluster ideal a análise dos resultados do cenário 1 e cenário 4, que são apresentados abaixo (Figuras 26 a 35).

Cluster	Unidades	Abrangência (metros)	Potencial de Biogás (m³/dia)
1	1, 2, 3, 4, 15, 21, 22, 23	2,201.34	1,216.07
2	5, 6, 7, 8, 9, 10, 17, 18, 19, 20	2,629.05	1,254.98
3	11, 12, 13, 14, 16, 24, 25, 26, 27, 28, 29, 30	3,978.04	2,768.12

Figura 26: Tabela de resultados do cenário 1.

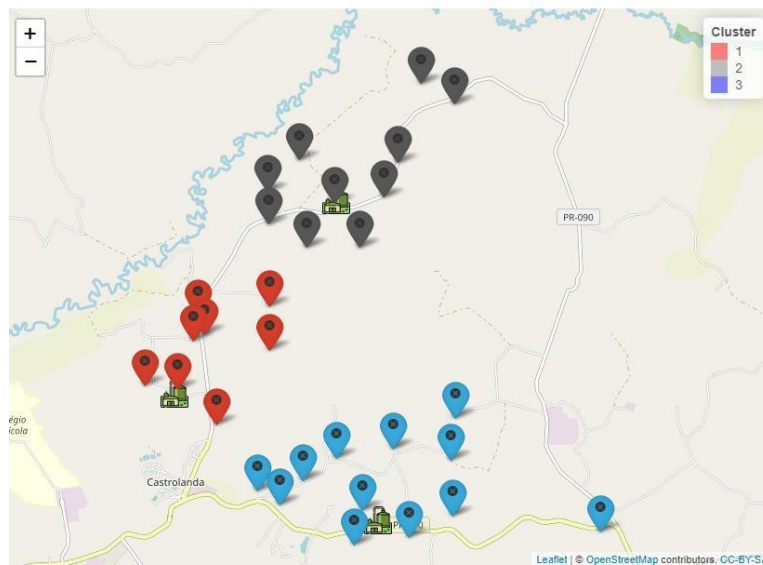


Figura 27: Mapa de clusters do cenário 1.





Figura 28: Variabilidade de abrangência dos clusters do cenário 1.

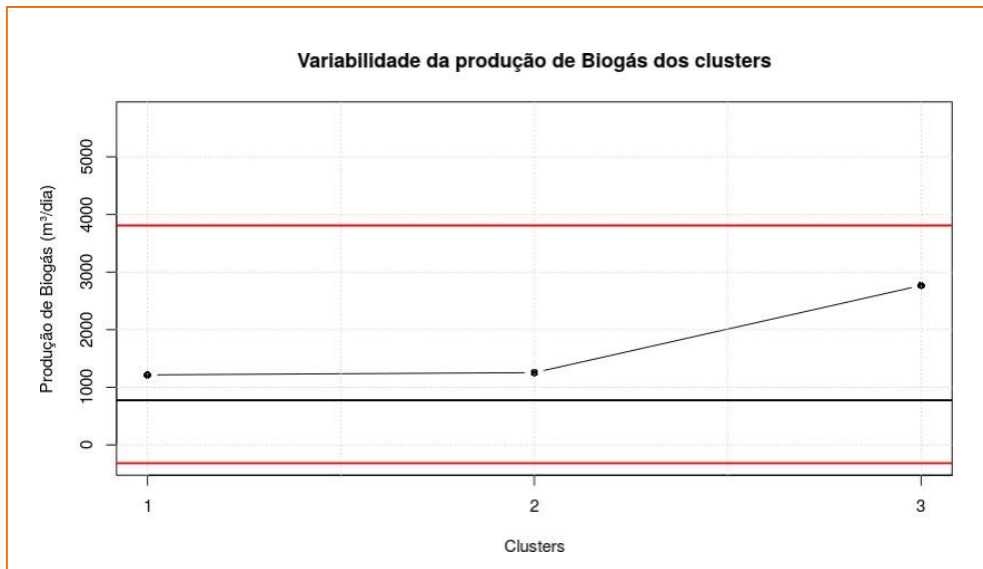


Figura 29: Variabilidade da produção de biogás dos clusters do cenário 1.

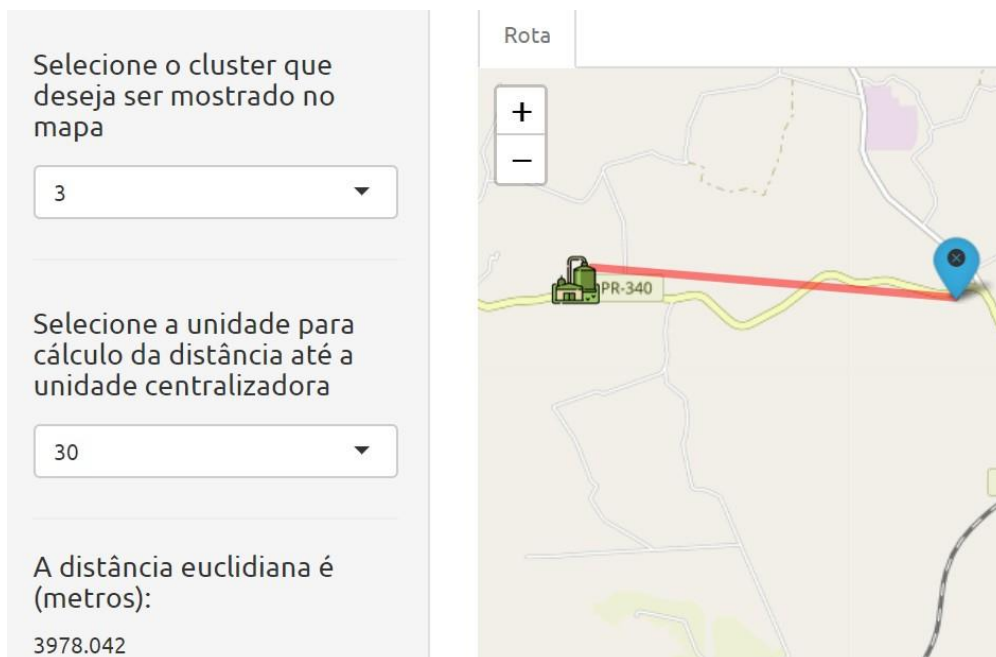


Figura 30: Maior distância euclidiana entre uma unidade e sua unidade centralizadora para o cenário 1.

Cluster	Unidades	Abrangência (metros)	Potencial de Biogás (m³/dia)
1	1, 2, 3, 4, 11, 12, 13, 14, 15, 16, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30	4,814.17	3,984.19
2	5, 6, 7, 8, 9, 10, 17, 18, 19, 20	2,629.05	1,254.98

Figura 31: Tabela de resultados do cenário 4.

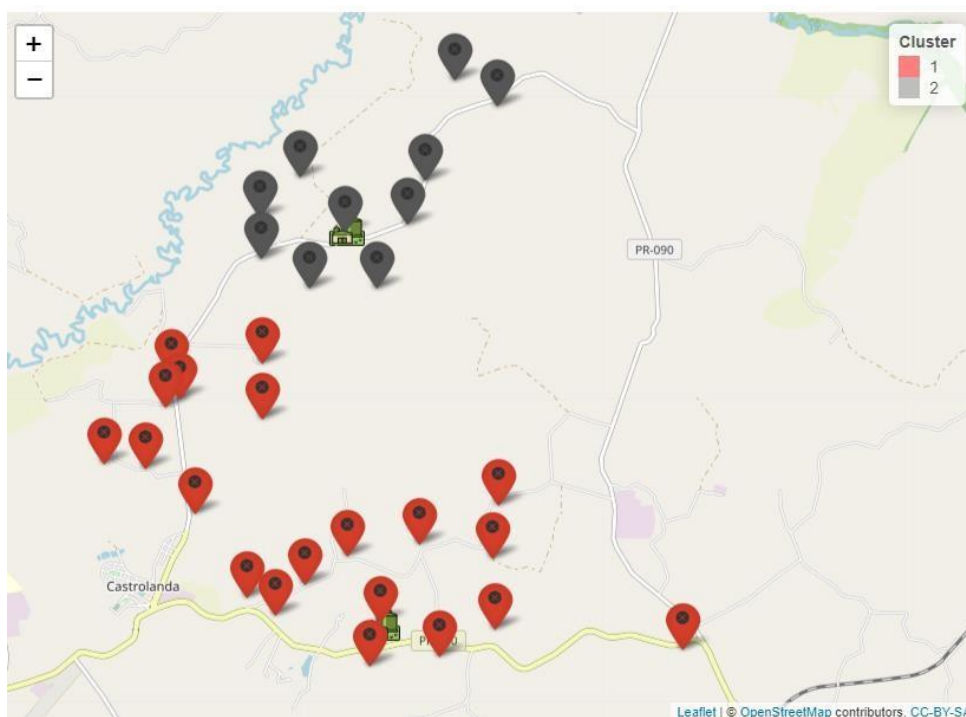


Figura 32: Mapa de clusters do cenário 4.

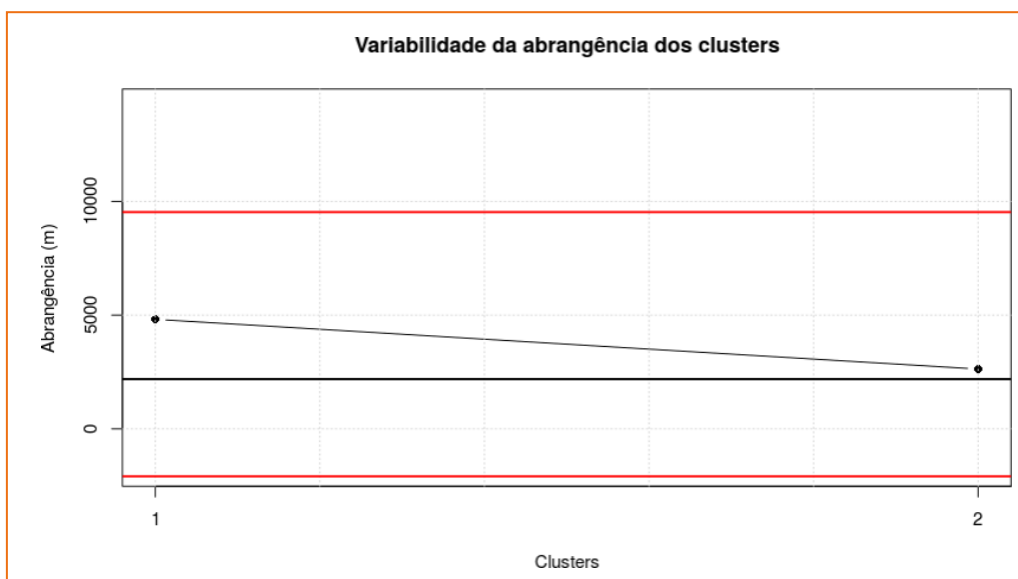


Figura 33: Variabilidade de abrangência dos clusters do cenário 4.

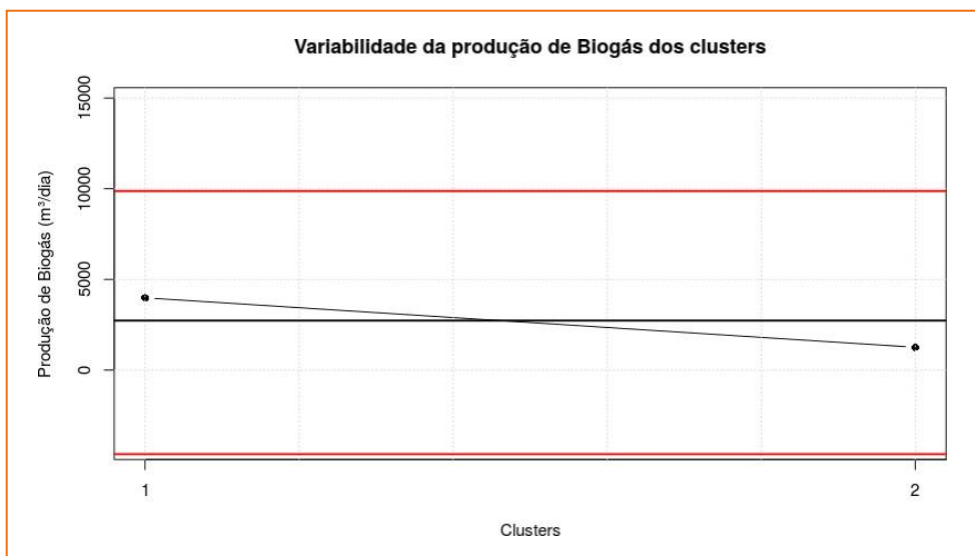


Figura 34: Variabilidade da produção de biogás dos clusters do cenário 4.

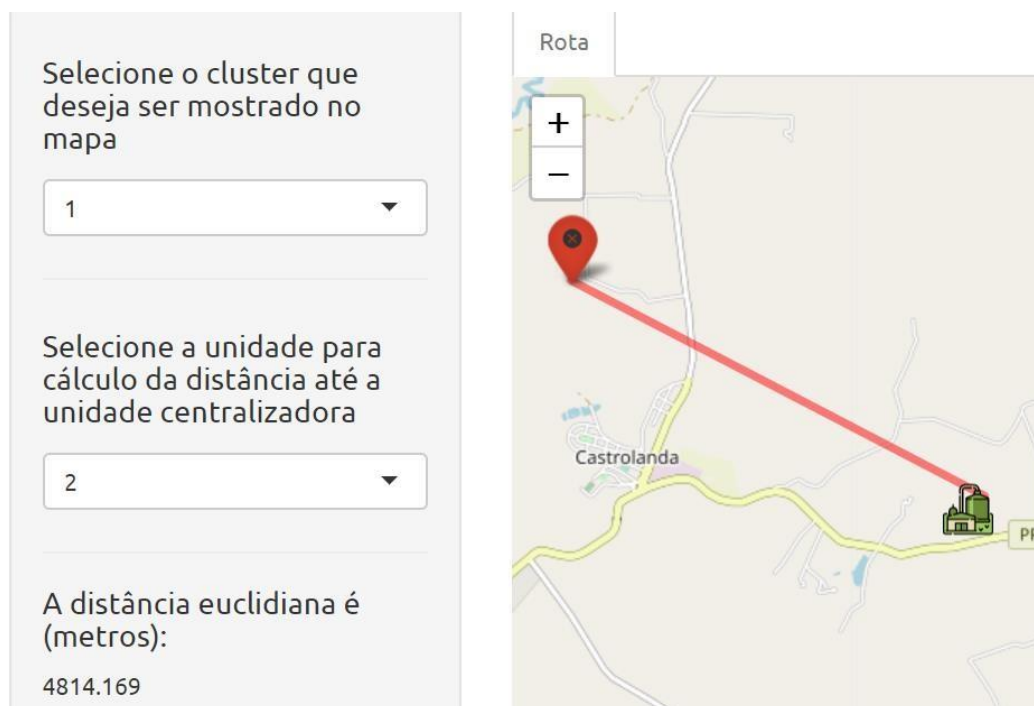


Figura 35: Maior distância euclidiana entre uma unidade e sua unidade centralizadora para o cenário 4.

As cartas de controle “Variabilidade da abrangência dos clusters” (Figuras 28 e 33) e “Variabilidade da produção de biogás dos clusters” (Figuras 29 e 34), mostram que os clusters gerado pelo cenário 4 são mais condizentes, pois seus resultados, quando comparados, estão mais equilibrados do que no cenário 1.

Para testar a opção de satisfação da ferramenta BiogasCluster, os pesquisadores também geraram um novo cenário. Após selecionada a opção “Não”, optou-se por alterar o número de clusters para 5, utilizando os filtros do cenário 4: “Selecionar todas as unidades”, “Não limitar abrangência”, “Próximo ao gasoduto e aos maiores potenciais produtores de biogás”. E os resultados encontrados foram:

Cluster	Unidades	Abrangência (metros)	Potencial de Biogás (m³/dia)
1	1, 2, 3, 4, 15, 21, 22, 23	2,201.34	1,216.07
2	5, 6, 7, 8, 9, 18, 19, 20	1,782.27	939.86
3	10, 17	403.69	315.11
4	11, 12, 14, 16, 26, 27, 28, 29, 30	3,424.50	2,472.78
5	13, 24, 25	630.97	295.34

Figura 36: Tabela de resultados do novo cenário.

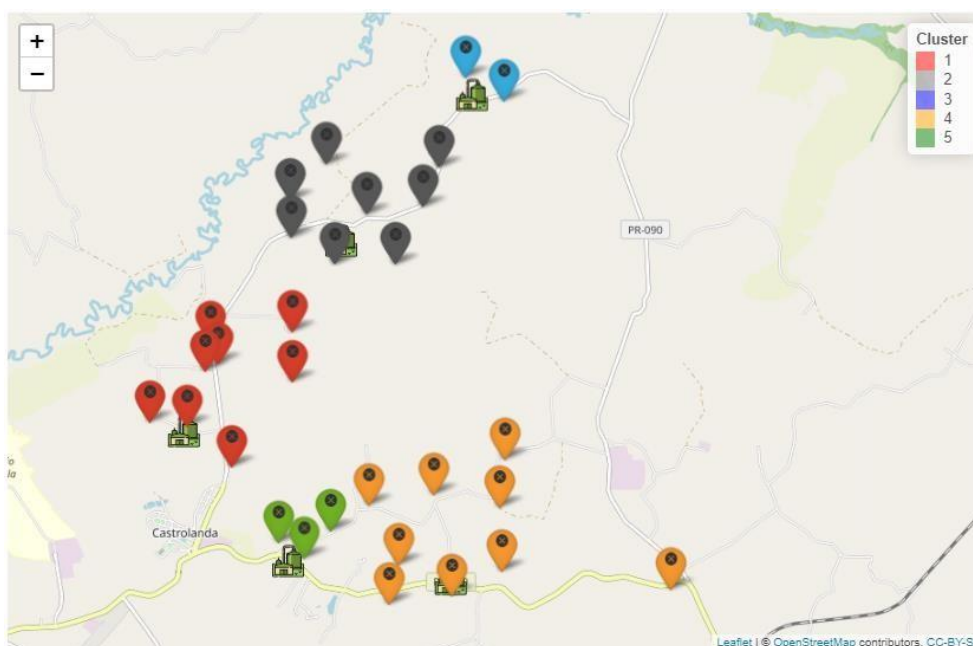


Figura 37: Mapa de clusters do novo cenário.

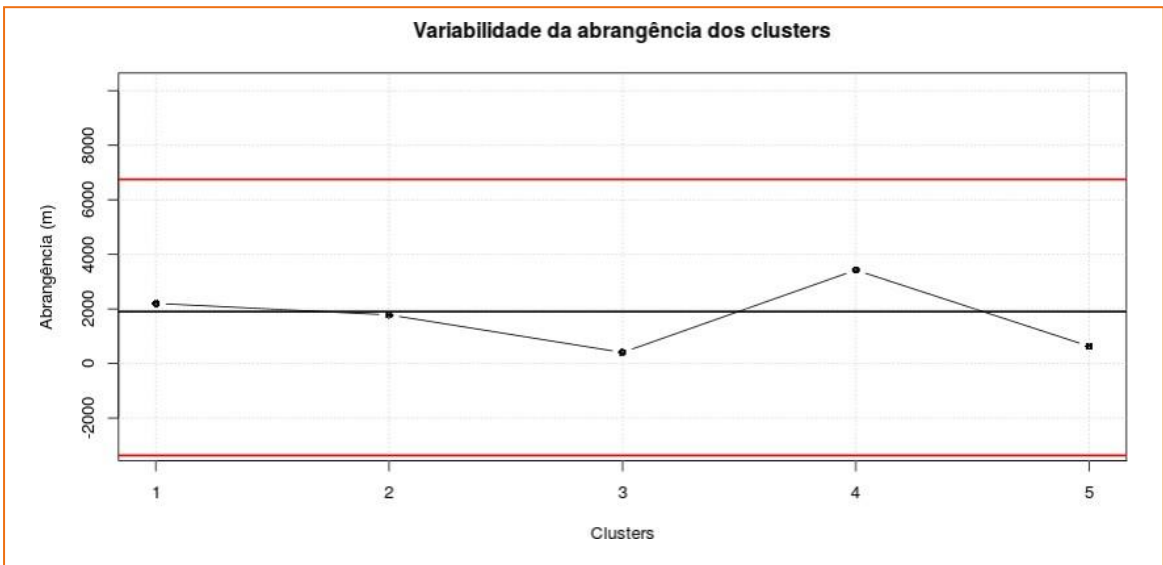


Figura 38: Variabilidade de abrangência dos clusters do novo cenário.

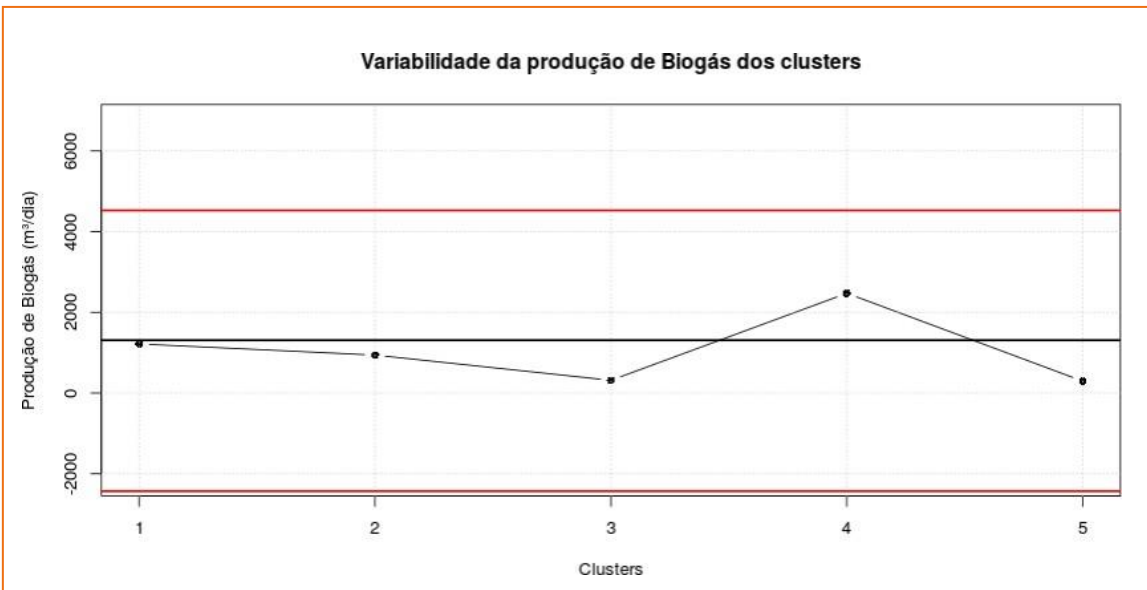


Figura 39: Variabilidade da produção de biogás dos clusters do novo cenário.

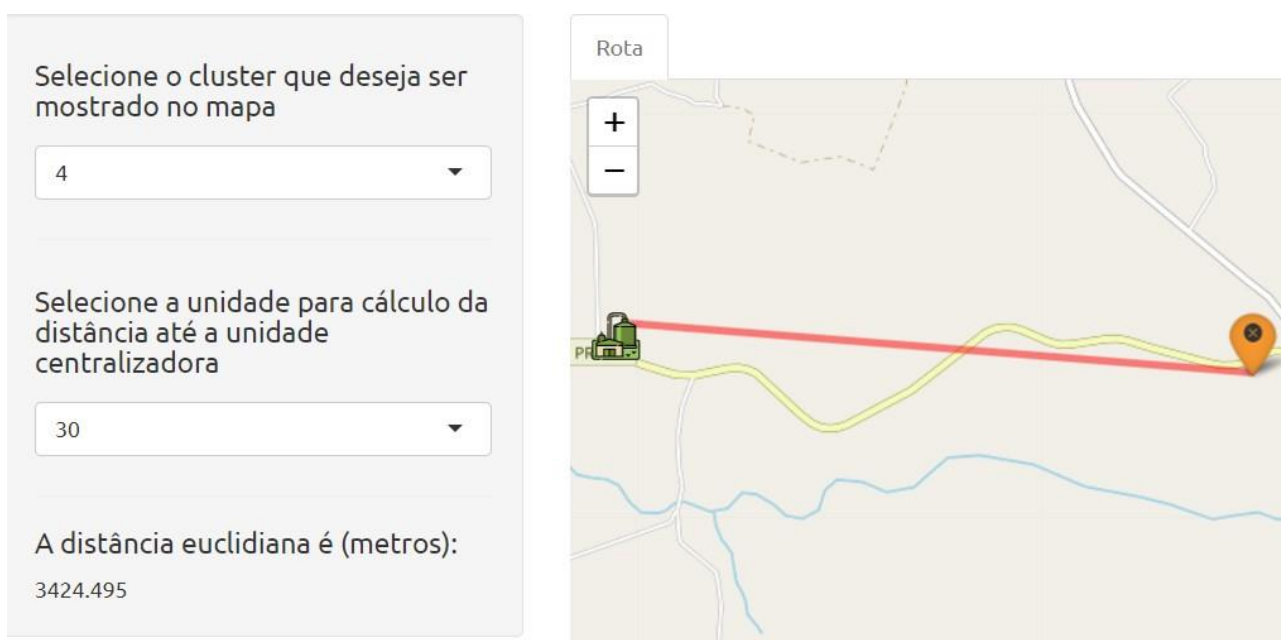


Figura 40: Maior distância euclidiana entre uma unidade e sua unidade centralizadora para o novo cenário.

Observando as cartas de controle “Variabilidade da abrangência dos clusters” (Figura 38) e “Variabilidade da produção de biogás dos clusters” (Figura 39) para o teste, conclui-se que o cenário 4 calculado pela ferramenta ainda possui resultados mais próximos à linha média, ou seja, segue sendo o mais indicado.

Esta indicação de solução ideal para a aplicação foi chancelada pelos gestores e cooperados da empresa em reunião, que relataram que o número de clusters e o agrupamento indicado pela ferramenta no cenário 4 atende ao que esperavam, considerando suas observações e conhecimentos da região.

### 6.3. CONSIDERAÇÕES FINAIS DO CASE

Esta parceria com a cooperativa possibilitou que os desenvolvedores aumentassem a credibilidade da ferramenta, pois a partir das contribuições feitas pela empresa, pontos de melhorias foram levantados e inseridos ao algoritmo.

O feedback recebido dos gestores da empresa foi positivo, no sentido de que os resultados encontrados condizem fielmente com a realidade e



pelo conhecimento que possuem da região a viabilidade de efetivação desses clusters é alta, o que é crucial para a garantia da confiabilidade da ferramenta.

## 7. Considerações Finais

Nos experimentos computacionais realizados, a ferramenta de clusterização “BiogasCluster” se mostrou eficiente na geração de grupos para posterior implantação de uma unidade centralizadora de biogás. Além de permitir a flexibilização de parâmetros escolhidos pelo usuário, denominados filtros, é uma ferramenta que abrange múltiplos objetivos, quando da determinação do número ótimo de clusters tendo em vista a minimização da amplitude da abrangência, que possibilita maior homogeneidade dos clusters, e a maximização da produção mínima de biogás dos clusters. Para apresentação dos resultados da ferramenta, o aplicativo desenvolvido é de fácil compreensão ao usuário e contempla, dentre outras funcionalidades apresentadas neste relatório, a análise dos resultados obtidos por meio de tabelas, gráficos e mapas.

## Apêndice 1

O algoritmo apresentado a seguir está armazenado em um repositório privado que pode ser acessado no link: [https://github.com/GefBrasilBiogas/biogasclustering/blob/master/APP\\_FINAL](https://github.com/GefBrasilBiogas/biogasclustering/blob/master/APP_FINAL).

Tabela 3: Algoritmo de clusterização implementado em R.

```
#Bibliotecas
library(rdist)
library(geosphere)
library(kableExtra)
library(readxl)
library(tidyverse)
library(dplyr)
library(shinythemes)
library(DT)
library(leaflet)
library(rgdal)

#usando dados input como shapefile
shape<-readOGR(dsn="C:/", layer="")
DadosEntrada<-shape@data

#usando dados input como xlsx
#DadosEntrada<-read_excel('C:/./xlsx')

DadosCompleto<-DadosEntrada

#####Pre-processamento da matriz de
entrada#####

#####Exclusao de propriedades muito distante das
demais (outliers) e que afetam muito a
clusterizacao#####

coordenadas<-
subset(DadosCompleto,select=c("Latitude","Longitude"))
#matriz contendo latitude e longitude
d<-distm(coordenadas[,2:1]) #calculo da distancia de
coordenadas geodesicas
diag(d)<-1000000 #distancia infinita da propriedade
para ela mesma
min_distancia<-as.matrix(apply(d,MARGIN=2,FUN=min))
limite<-mean(min_distancia)+3*sd(min_distancia)
#calcular o limite aceitavel de distancia minima entre as
propriedades
```

```
#procurar na matriz d propriedade que tenham distancia
minima maior que o limite
search_vec <- function(mat, vec, dim = 1, tol = 1e-7,
fun = all)
  which(apply(mat, dim, function(x) fun((x - vec) >
tol)))
  indice_excluir<-
search_vec(min_distancia,limite,fun=any)
  if(is_empty(indice_excluir)==FALSE){
  coordenadas<-coordenadas[-c(indice_excluir),]
  DadosCompleto<-DadosCompleto[-c(indice_excluir),]
  d<-d[-c(indice_excluir),-c(indice_excluir)]}

#####Filtrol#####
####
#filtrol vale 1, se forem utilizadas todas as
propriedades,
#vale 2 para excluir propriedades que tem baixo
potencial de producao
#vale 3 para excluir propriedades que tem alto
potencial de producao

Q1<-matrix(quantile(DadosCompleto$Biogas, probs =
0.25))
Q3<-matrix(quantile(DadosCompleto$Biogas, probs =
0.75))
L<-Q1-1.5*(Q3-Q1)
S<-Q3+1.5*(Q3-Q1)
if (Filtrol==2){
  d<-
subset(d,DadosCompleto$Biogas>L[1],DadosCompleto$Biogas>L
[1])
  DadosCompleto<-subset(DadosCompleto,Biogas>L[1])
}
if (Filtrol==3){
  d<-
subset(d,DadosCompleto$Biogas<S[1],DadosCompleto$Biogas<S
[1])
  DadosCompleto<-subset(DadosCompleto,Biogas<S[1])
}
coordenadas<-
subset(DadosCompleto,select=c("Latitude","Longitude"))

##### Filtro 3
#####
#definicao do peso para localizacao do centro de massa
#(filtro3=1, proximo aos produtores e gasoduto, =2,
proximo aos produtores)
```

```

    if (Filtro3==1) {DadosCompleto$peso<-
(DadosCompleto$Biogas/max (DadosCompleto$Biogas)) * (max (Dad
osCompleto$`Distancia`)/DadosCompleto$`Distancia`)
    }else{if (Filtro3==2) {DadosCompleto$peso<-
DadosCompleto$Biogas}
    if (Filtro3==3) {DadosCompleto$peso<-
DadosCompleto$Consumo}
    if (Filtro3==4) {DadosCompleto$peso<-
(DadosCompleto$Consumo/max (DadosCompleto$Consumo)) * (max (D
adosCompleto$`Distancia`)/DadosCompleto$`Distancia`)}}

#####
#####

d<-as.dist(d)
fit.average<-hclust(d,method="average") #clusterizacao
hierarquica com media da distancia
npropriedades<-dim(DadosCompleto)[1]

#considerando soluções que apresentem cluster com no
mínimo 2 unidades
p<-1
clusters<-cutree(fit.average, p)
nclusters<-matrix(table(clusters))
while (min(nclusters)>1) {
  p<-p+1
  clusters<-cutree(fit.average, p)
  nclusters<-matrix(table(clusters))}
p<-p-1

##### buscando valor otimo de k
#####
  media<-matrix(nrow=p-1,ncol=3)
  colnames(media)<-c("Abrangencia","Amplitude de
abrangencia","Biogas")
  l<-1
  for(k in 2:p){

#corte pela definicao do numero de clusters

  clusters<-cutree(fit.average, k)
  nclusters<-matrix(table(clusters)) #para indicar
quantas propriedades ha em cada cluster
  DadosCompleto$cluster <- clusters #inserindo coluna
com determinacao dos clusters

#centro de massa = unidade cetralizadora

  centro_massa<-matrix(nrow=k,ncol=2)
  for(i in 1:k){

```

## BIOGASCLUSTER

```
        centro_massa[i,]<-
c(weighted.mean(subset(DadosCompleto,cluster==i)$Latitude
,subset(DadosCompleto,cluster==i)$peso),

weighted.mean(subset(DadosCompleto,cluster==i)$Longitude,
subset(DadosCompleto,cluster==i)$peso))}
    coordenadas$cluster<-clusters #incluindo indice dos
clusters
    centro_massa<-
cbind(centro_massa,matrix(c(1:k),ncol=1)) #incluindo o
indice dos clusters

#####calculo da abrangencia dos clusters
#calculo das distancias = abrangencia dos cluster,
considerada como a maior distancia entre as propriedades
e o centro de massa

    abrangencia<-matrix(nrow=k,ncol=1)
    for(i in 1:k){
        abrangencia[i,]<-
max(data.matrix(dist(rbind(subset(coordenadas,cluster==i)
[,2:1],centro_massa[i,2:1]))) [nclusters[i]+1,])}

    #soma de biogas dos clusters
    soma_biogas<-matrix(nrow=k,ncol=1)
    for(i in 1:k){
        soma_biogas[i,]<-
sum(subset(DadosCompleto,cluster==i) ["Biogas"])
    }

    #amplitude de abrangencia e producao minima de
biogas
    media[1,]<-c(max(abrangencia),max(abrangencia)-
min(abrangencia),min(soma_biogas))
    l<-l+1
    }

    q1<-matrix(quantile(media[,1], probs = 0.25))
    q3<-matrix(quantile(media[,1], probs = 0.75))
    s<-q3+1.5*(q3-q1)
    if(Filtro2==1 & s[1]>0){#limitando abrangencia dos
clusters
        Media<-subset(media, media[,1] <=s[1])
    } else{Media<-media}

    Media<-Media[,2:3]

#definicao do ponto ideal, considerando o minima
amplitude de abrangencia e maxima producao minima de
biogas
```

```
PontoIdeal<-
as.matrix(t(c(min(Media[,1]),max(Media[,2]))))
  aux_solucao<-as.matrix(dist(rbind(Media,PontoIdeal)))

#calculando a distancia das solucoes dos clusters ao
ponto ideal
  distancia_solucao<-
min(aux_solucao[as.matrix(dim(Media))[1,1]+1,1:as.matrix(
dim(Media))[1,1]])
  a<-
which(aux_solucao[dim(Media)[1]+1,]==distancia_solucao)
  SolucaoFinal<-Media[a[1],]

  search_vec3 <- function(mat, vec, dim = 1, tol = 1e-
7, fun = all)
  which(apply(mat, dim, function(x) fun(abs(x - vec)
< tol)))

  b<-
as.matrix(search_vec3(media[,2:3],SolucaoFinal,fun=any))#
valor otimo de numero de clusters
  k<-b[1]+1 #numero otimo de clusters

#se o usuario nao estiver satisfeito com a solucao e
quiser escolher o numero de clusters
  if(decisao==2 & NaoSatisfacao==2){
  k<-K}

#precisa repetir o codigo para obter as matrizes
abrangencia e soma_biogas referentes ao k otimo ou
inserido pelo usuario
#clusterizacao

#Corte pela definicao do numero de clusters
clusters<-cutree(fit.average, k)
nclusters<-matrix(table(clusters)) #para indicar
quantas propriedades ha em cada cluster
  DadosCompleto$cluster <- clusters #inserindo coluna com
determinacao dos clusters

centro_massa<-matrix(nrow=k,ncol=2)
for(i in 1:k){
  centro_massa[i,]<-
c(weighted.mean(subset(DadosCompleto,cluster==i)$Latitude
,subset(DadosCompleto,cluster==i)$peso),
weighted.mean(subset(DadosCompleto,cluster==i)$Longitude,
subset(DadosCompleto,cluster==i)$peso)) }
```

```

centro_massa<-cbind(centro_massa,matrix(c(1:k),ncol=1))
#incluindo o indice dos clusters

#incluir propriedades que foram excluidas no pre-
processamento por serem muito distantes das demais
  if (is_empty(indice_excluir)==FALSE){
    for(i in 1:dim(as.array(indice_excluir))){
      distanciaPropriedade<-
t(distm(DadosEntrada[indice_excluir[i],c("Longitude","Lat
itude")],centro_massa[,2:1]))
      aa<-min(distanciaPropriedade)
      search_vec2<-function(mat, vec, dim = 1, tol = 1e-7,
fun = all)
        which(apply(mat, dim, function(x) fun((x - vec)
==0)))
      cc<-search_vec2(distanciaPropriedade,aa,fun=any)
      DadosCompleto<-
rbind(DadosCompleto,cbind(DadosEntrada[indice_excluir[i],
],peso=0,cluster=cc[1]))
      nclusters[cc[1]]<-nclusters[cc[1]]+1
    }
}

#####
#####
  coordenadas<-
DadosCompleto[c("Latitude","Longitude","cluster")]

#Abrangencia
abrangencia<-matrix(nrow=k,ncol=1)
for(i in 1:k){
  abrangencia[i,]<-
max(distm(subset(coordenadas,cluster==i)[,2:1],centro_mas
sa[i,2:1]))}

abrangencia<-cbind(abrangencia,matrix(c(1:k),ncol=1))
colnames(abrangencia)<-c("Abrangencia_m","cluster")

#Soma de biogas dos clusters
soma_biogas<-matrix(nrow=k,ncol=1)
for(i in 1:k){
  soma_biogas[i,]<-
sum(subset(DadosCompleto,cluster==i)["Biogas"])
}
soma_biogas<-cbind(soma_biogas,matrix(c(1:k),ncol=1))
colnames(soma_biogas)<-
c("Potencial_Biogas_m3dia","cluster")

```



## Apêndice 2

O algoritmo apresentado a seguir está armazenado em um repositório privado que pode ser acessado no link: [https://github.com/GefBrasilBiogas/biogasclustering/blob/master/APP\\_FINAL\\_SHINY](https://github.com/GefBrasilBiogas/biogasclustering/blob/master/APP_FINAL_SHINY).

Tabela 4: Algoritmo de clusterização implementado em um aplicativo Web usando o Shiny do software R.

```
#Bibliotecas
library(rdist)
library(geosphere)
library(kableExtra)
library(readxl)
library(tidyverse)
library(shiny)
library(dplyr)
library(shinythemes)
library(DT)
library(leaflet)
library(leaflet.sync)
library(sf)
library(osrm)
library(cartography)
library(shinyBS)
library(shinyWidgets)
library(rgdal)
library(openxlsx)
library(writexl)

function.clusterizacao<-
function(DadosEntrada,K,Filtro1,Filtro2,decisao,Filtro3,F
iltro4,Filtro5,NaoSatisfacao){

  DadosCompleto<-DadosEntrada

  #####Pre-processamento da matriz de
  entrada#####

  #####Exclusao de propriedades muito distante das
  demais (outliers) e que afetam muito a
  clusterizacao#####
```

```
coordenadas<-
subset(DadosCompleto,select=c("Latitude","Longitude"))
#matriz contendo latitude e longitude
  d<-dism(coordenadas[,2:1]) #calculo da distancia de
coordenadas geodesicas
  diag(d)<-1000000 #distancia infinita da propriedade
para ela mesma
  min_distancia<-as.matrix(apply(d,MARGIN=2,FUN=min))
  limite<-mean(min_distancia)+3*sd(min_distancia)
#calcular o limite aceitavel de distancia minima entre as
propriedades

  #procurar na matriz d propriedade que tenham distancia
minima maior que o limite
  search_vec <- function(mat, vec, dim = 1, tol = 1e-7,
fun = all)
  which(apply(mat, dim, function(x) fun((x - vec) >
tol)))
  indice_excluir<-
search_vec(min_distancia,limite,fun=any)
  if(is_empty(indice_excluir)==FALSE){
    coordenadas<-coordenadas[-c(indice_excluir),]
    DadosCompleto<-DadosCompleto[-c(indice_excluir),]
    d<-d[-c(indice_excluir),-c(indice_excluir)]}

#####Filtrol#####
####
  #filtrol vale 1, se forem utilizadas todas as
propriedades,
  #vale 2 para excluir propriedades que tem baixo
potencial de producao
  #vale 3 para excluir propriedades que tem alto
potencial de producao

  Q1<-matrix(quantile(DadosCompleto$Biogas, probs =
0.25))
  Q3<-matrix(quantile(DadosCompleto$Biogas, probs =
0.75))
  L<-Q1-1.5*(Q3-Q1)
  S<-Q3+1.5*(Q3-Q1)
  if (Filtrol==2){
    d<-
subset(d,DadosCompleto$Biogas>L[1],DadosCompleto$Biogas>L
[1])
    DadosCompleto<-subset(DadosCompleto,Biogas>L[1])
  }
  if (Filtrol==3){
```

```

d<-
subset(d,DadosCompleto$Biogas<S[1],DadosCompleto$Biogas<S
[1])
  DadosCompleto<-subset(DadosCompleto,Biogas<S[1])
}
coordenadas<-
subset(DadosCompleto,select=c("Latitude","Longitude"))

##### Filtro 3
#####
#definicao do peso para localizacao do centro de massa
#(filtro3=1, proximo aos produtores e gasoduto, =2,
proximo aos produtores)

  if(Filtro3==1){DadosCompleto$peso<-
(DadosCompleto$Biogas/max(DadosCompleto$Biogas))* (max(Dad
osCompleto$`Distancia`)/DadosCompleto$`Distancia`)
  }else{if(Filtro3==2){DadosCompleto$peso<-
DadosCompleto$Biogas}
  if(Filtro3==3){DadosCompleto$peso<-
DadosCompleto$Consumo}
  if(Filtro3==4){DadosCompleto$peso<-
(DadosCompleto$Consumo/max(DadosCompleto$Consumo))* (max(D
adosCompleto$`Distancia`)/DadosCompleto$`Distancia`)}

#####
#####

d<-as.dist(d)
fit.average<-hclust(d,method="average") #clusterizacao
hierarquica com media da distancia
npropriedades<-dim(DadosCompleto)[1]

#considerando soluções que apresentem cluster com no
mínimo 2 unidades
p<-1
clusters<-cutree(fit.average, p)
nclusters<-matrix(table(clusters))
while (min(nclusters)>1) {
  p<-p+1
  clusters<-cutree(fit.average, p)
  nclusters<-matrix(table(clusters))}
p<-p-1

##### buscando valor otimo de k
#####
  media<-matrix(nrow=p-1,ncol=3)
  colnames(media)<-c("Abrangencia","Amplitude de
abrangencia","Biogas")
  l<-1

```

## BIOGASCLUSTER

```
for(k in 2:p){

#corte pela definicao do numero de clusters

    clusters<-cutree(fit.average, k)
    nclusters<-matrix(table(clusters)) #para indicar
quantas propriedades ha em cada cluster
    DadosCompleto$cluster <- clusters #inserindo coluna
com determinacao dos clusters

#centro de massa = unidade cetralizadora

    centro_massa<-matrix(nrow=k,ncol=2)
    for(i in 1:k){
        centro_massa[i,]<-
c(weighted.mean(subset(DadosCompleto,cluster==i)$Latitude
,subset(DadosCompleto,cluster==i)$peso),

weighted.mean(subset(DadosCompleto,cluster==i)$Longitude,
subset(DadosCompleto,cluster==i)$peso))}
        coordenadas$cluster<-clusters #incluindo indice dos
clusters
        centro_massa<-
cbind(centro_massa,matrix(c(1:k),ncol=1)) #incluindo o
indice dos clusters

#####calculo da abrangencia dos clusters
#calculo das distancias = abrangencia dos cluster,
considerada como a maior distancia entre as propriedades
e o centro de massa

        abrangencia<-matrix(nrow=k,ncol=1)
        for(i in 1:k){
            abrangencia[i,]<-
max(data.matrix(dist(rbind(subset(coordenadas,cluster==i)
[,2:1],centro_massa[i,2:1]))) [nclusters[i]+1,])}

        #soma de biogas dos clusters
        soma_biogas<-matrix(nrow=k,ncol=1)
        for(i in 1:k){
            soma_biogas[i,]<-
sum(subset(DadosCompleto,cluster==i) ["Biogas"])
        }

        #amplitude de abrangencia e producao minima de
biogas
        media[1,]<-c(max(abrangencia),max(abrangencia)-
min(abrangencia),min(soma_biogas))
        l<-l+1
    }
}
```

```
q1<-matrix(quantile(media[,1], probs = 0.25))
q3<-matrix(quantile(media[,1], probs = 0.75))
s<-q3+1.5*(q3-q1)
if(Filtro2==1 & s[1]>0){#limitando abrangencia dos
clusters
  Media<-subset(media, media[,1] <=s[1])
} else{Media<-media}

Media<-Media[,2:3]

#definicao do ponto ideal, considerando o minima
amplitude de abrangencia e maxima producao minima de
biogas

PontoIdeal<-
as.matrix(t(c(min(Media[,1]),max(Media[,2]))))
aux_solucao<-as.matrix(dist(rbind(Media,PontoIdeal)))

#calculando a distancia das solucoes dos clusters ao
ponto ideal
distancia_solucão<-
min(aux_solução[as.matrix(dim(Media))[1,1]+1,1:as.matrix(
dim(Media))[1,1]])
a<-
which(aux_solução[dim(Media)[1]+1,]==distancia_solucão)
SoluçãoFinal<-Media[a[1],]

search_vec3 <- function(mat, vec, dim = 1, tol = 1e-
7, fun = all)
  which(apply(mat, dim, function(x) fun(abs(x - vec)
< tol)))

b<-
as.matrix(search_vec3(media[,2:3],SoluçãoFinal,fun=any))#
valor otimo de número de clusters
k<-b[1]+1 #numero otimo de clusters

#se o usuario nao estiver satisfeito com a solucao e
quiser escolher o numero de clusters
if(decisao==2 & NaoSatisfacao==2){
k<-K}

#precisa repetir o codigo para obter as matrizes
abrangencia e soma_biogas referentes ao k otimo ou
inserido pelo usuario
#clusterizacao

#Corte pela definicao do numero de clusters
clusters<-cutree(fit.average, k)
```

## BIOGASCLUSTER

```
nclusters<-matrix(table(clusters)) #para indicar
quantas propriedades ha em cada cluster
DadosCompleto$cluster <- clusters #inserindo coluna com
determinacao dos clusters

centro_massa<-matrix(nrow=k,ncol=2)
for(i in 1:k){
  centro_massa[i,]<-
c(weighted.mean(subset(DadosCompleto,cluster==i)$Latitude
,subset(DadosCompleto,cluster==i)$peso),

weighted.mean(subset(DadosCompleto,cluster==i)$Longitude,
subset(DadosCompleto,cluster==i)$peso))}
  centro_massa<-cbind(centro_massa,matrix(c(1:k),ncol=1))
#incluindo o indice dos clusters

#incluir propriedades que foram excluidas no pre-
processamento por serem muito distantes das demais
  if (is_empty(indice_excluir)==FALSE){
    for(i in 1:dim(as.array(indice_excluir))){
      distanciaPropriedade<-
t(distm(DadosEntrada[indice_excluir[i],c("Longitude","Lat
itude")],centro_massa[,2:1]))
      aa<-min(distanciaPropriedade)
      search_vec2<-function(mat, vec, dim = 1, tol = 1e-7,
fun = all)
        which(apply(mat, dim, function(x) fun((x - vec)
==0)))
      cc<-search_vec2(distanciaPropriedade,aa,fun=any)
      DadosCompleto<-
rbind(DadosCompleto,cbind(DadosEntrada[indice_excluir[i],
],peso=0,cluster=cc[1]))
      nclusters[cc[1]]<-nclusters[cc[1]]+1
    }}

#####
#####
  coordenadas<-
DadosCompleto[c("Latitude","Longitude","cluster")]

#Abrangencia
abrangencia<-matrix(nrow=k,ncol=1)
for(i in 1:k){
  abrangencia[i,]<-
max(distm(subset(coordenadas,cluster==i)[,2:1],centro_mas
sa[i,2:1]))}

abrangencia<-cbind(abrangencia,matrix(c(1:k),ncol=1))
colnames(abrangencia)<-c("Abrangencia_m","cluster")
```

## BIOGASCLUSTER

```
#Soma de biogas dos clusters
soma_biogas<-matrix(nrow=k,ncol=1)
for(i in 1:k){
  soma_biogas[i,]<-
sum(subset(DadosCompleto,cluster==i)["Biogas"])
}
soma_biogas<-cbind(soma_biogas,matrix(c(1:k),ncol=1))
colnames(soma_biogas)<-
c("Potencial_Biogas_m3dia","cluster")

#Tabela de Saida
dados_tabela <- Reduce(merge, list(DadosCompleto,
abrangencia, soma_biogas))
dados_tabela <-
dados_tabela[order(dados_tabela$cluster,
as.numeric(dados_tabela$Unidade)),]
dados_tabela_1 <- aggregate(. ~ cluster + Abrangencia_m
+ Potencial_Biogas_m3_dia,
dados_tabela[,c("cluster","Unidade","Abrangencia_m","Pote
ncial_Biogas_m3_dia")], toString)
colnames(dados_tabela_1)<-c("Cluster", "Abrangência
(metros)", "Potencial de Biogás (m³/dia)", "Unidades")
df1<-as.data.frame(centro_massa)
colnames(df1) <-c("Latitude", "Longitude", "cluster")
DadosCompleto<-
DadosCompleto[,c("Unidade","Latitude","Longitude","Distan
cia","Biogas","Consumo","cluster")]

#Especificar cluster e propriedade para gerar os mapas
df_spec_clust <- df1[df1$cluster == Filtro4,]
DadosCompleto_spec_prop<-
DadosCompleto[DadosCompleto$Unidade == Filtro5,]

#Mapa usando leaflet
ai_colors <-
c("red","gray","blue","orange","green","beige","darkgreen
","lightgreen", "lightred", "darkblue","lightblue",
"purple","darkpurple","pink",
"cadetblue","white","darkred", "lightgray","black")
clust_colors <- ai_colors[DadosCompleto$cluster]
icons <- awesomeIcons(
  icon = 'ios-close',
  iconColor = 'black',
  library = 'ion',
  markerColor = clust_colors)

# Criar Ícone do Biodigestor:
leafIcons <- icons(
  iconUrl = ifelse(df1$cluster,
```



```
"https://image.flaticon.com/icons/svg/542/542461.svg"
),
  iconWidth = 30, iconHeight = 40,
  iconAnchorX = 25, iconAnchorY = 12)

html_legend <- "<img
src='https://image.flaticon.com/icons/svg/542/542461.svg'
>"

# primeiro mapa
m1<-leaflet(df1) %>% addTiles() %>% #criar primeiro
mapa
  addMarkers(~Longitude, ~Latitude, icon = leafIcons)
%>%
  addAwesomeMarkers(lat=~DadosCompleto$Latitude, lng =
~DadosCompleto$Longitude, icon=icons,
label=~as.character(DadosCompleto$cluster)) %>%
  addLegend( position = "topright", title="Cluster",
colors = ai_colors[1:max(DadosCompleto$cluster)], labels =
unique(DadosCompleto$cluster))
plotL1<-m1

#segundo mapa
if(nrow(df_spec_clust)>0){
  clust_colors <- ai_colors[df_spec_clust$cluster]
  icons <- awesomeIcons(
    icon = 'ios-close',
    iconColor = 'black',
    library = 'ion',
    markerColor = clust_colors)
  m2<-leaflet(df_spec_clust) %>% addTiles() %>% #cria
segundo mapa
  addMarkers(~Longitude, ~Latitude, icon = leafIcons)
%>%
  addAwesomeMarkers(leaflet(DadosCompleto_spec_prop)
%>% addTiles(), lat=~DadosCompleto_spec_prop$Latitude,
lng = ~DadosCompleto_spec_prop$Longitude, icon=
icons,label=~cluster)#%>%

  for(i in 1:nrow(df_spec_clust)){
    df_line <-
rbind(DadosCompleto_spec_prop[,c("Latitude", "Longitude")]
,
df_spec_clust[i,c("Latitude", "Longitude")])
m2 <- m2 %>%
  addPolylines(data = df_line,
lat=~Latitude,
```

```
        lng = ~Longitude,
        color="red")
    }
    plotL2<-m2} else plotL2 <- NULL

#### Calcula distância euclidiana entre uma unidade e
sua respectiva unidade centralizadora
    DISTANCIA<- merge(DadosCompleto,df1,by = c("cluster"),
    suffixes = c("_DadosCompleto","_df1"))

    (DISTANCIA$distancia <- purrr::pmap_dbl(.l =
list(DISTANCIA$Longitude_DadosCompleto,
DISTANCIA$Latitude_DadosCompleto,
DISTANCIA$Longitude_df1,
DISTANCIA$Latitude_df1),
.f = ~distm(c(..1,..2),c(..3,..4)))

    # Utilizado para download do resultado da solucao
escolhida pelo usuario
    DadosCompleto<-
rbind(DadosCompleto,cbind(Unidade="UnidadeCentralizadora"
,Latitude=centro_massa[,1],Longitude=centro_massa[,2],Dis
tancia=0,Biogas=0,Consumo=0,cluster=centro_massa[,3]))

    return(list(
    "Data" = dados_tabela_1,
    "CartaAbrangencia" = abrangencia,
    "CartaBiogas" = soma_biogas,
    "Plot1" = plotL1,
    "Plot2" =plotL2,
    "Data1" = dados_tabela,
    "DIST" = DISTANCIA,
    "Resultado" = DadosCompleto

    ))
}

#### Funções para gerar as Cartas de Controle
function.CartaAbrangencia <- function(abrangencia) {
    LIC <- mean(abrangencia[, 1]) - 3 *
(mean(abs(diff(abrangencia[, 1]))) / 1.1284)
    LSC <- mean(abrangencia[, 1]) + 3 *
(mean(abs(diff(abrangencia[, 1]))) / 1.1284)
    plot(abrangencia[, 1],type = "b",pch = 16, ylim = c(LIC
- 0.1 * LIC, LSC + 0.5 * LSC), axes = FALSE,
        main = "Variabilidade da abrangência dos
clusters",xlab = "Clusters",ylab = "Abrangência (m)")
    axis(1, at = 1:35)
    axis(2)
    box()
}
```

```
grid()
abline(h = mean(abs(diff(abrangencia[, 1]))),lwd = 2)
abline(h = LSC, lwd = 2, col = "red")
abline(h = LIC, lwd = 2, col = "red")

function.CartaBiogas <- function(soma_biogas) {
  LIC <- mean(soma_biogas[, 1]) - 3 *
(mean(abs(diff(soma_biogas[, 1]))) / 1.1284)
  LSC <- mean(soma_biogas[, 1]) + 3 *
(mean(abs(diff(soma_biogas[, 1]))) / 1.1284)
  plot(soma_biogas[, 1],type = "b",pch = 16,ylim = c(LIC
- 0.1 * LIC, LSC + 0.5 * LSC),axes = FALSE,
      main = "Variabilidade da produção de Biogás dos
clusters",xlab = "Clusters",ylab = "Produção de Biogás
(m³/dia)")
  axis(1, at = 1:35)
  axis(2)
  box()
  grid()
  abline(h = mean(abs(diff(soma_biogas[, 1]))),lwd = 2)
  abline(h = LSC, lwd = 2, col = "red")
  abline(h = LIC, lwd = 2, col = "red")
}

#Descrições dos filtros e opções dos filtros

Des_Formato<-paste("Para o formato shapefile, é
necessário selecionar os três arquivos juntos (.shp,
.shx, .dbf). Já para o excel, é aceito apenas o formato
.xlsx", sep = "<br>")

DES_Filtro1<-paste("Os filtros se referem a seleção das
unidades georeferenciadas que foram importadas por
arquivo no BiogasCluster. Estas unidades podem ser, por
exemplo, propriedades rurais ou unidades industriais.",
sep = "<br>")
DES_Filtro1_opcao1<-paste("Todas as unidades importadas
no BiogasCluster são selecionadas para formação de
clusters.", sep = "<br>")
DES_Filtro1_opcao2<-paste("Exclui unidades com baixo
potencial de produção de biogás em relação às demais
(outliers inferiores).", sep = "<br>")
DES_Filtro1_opcao3<-paste("Exclui unidades com alto
potencial de produção de biogás em relação às demais
(outliers superiores).", sep = "<br>")

DES_Filtro2<-paste("Os filtros se referem à abrangência
dos clusters, definida como a distância máxima dentre as
```

```
distâncias entre as unidades e a respectiva unidade
centralizadora do cluster.", sep = "<br>")
DES_Filtro2_opcao1<-paste("Exclui formação de clusters
que possuem unidades georeferenciadas dispersas, isto é,
abrangência maior que o outlier superior das possíveis
soluções de clusters. Além disso, não permite clusters de
abrangência nula, ou seja, com apenas uma unidade.", sep
= "<br>")
DES_Filtro2_opcao2<-paste("Não limita abrangência dos
clusters.", sep = "<br>")

DES_Filtro3<-paste(" Os filtros se referem à localização
da unidade centralizadora. Esta unidade centralizadora
pode ser um biodigestor, hub de biomassa, central de
tratamento de resíduo entre outras possibilidades.", sep
= "<br>")
DES_Filtro3_opcao1<-paste("Tende a localizar a unidade
centralizadora próximo aos maiores potenciais produtores
de biogás importados por arquivo no BiogasCluster e ao
gasoduto.", sep = "<br>")
DES_Filtro3_opcao2<-paste("Tende a localizar a unidade
centralizadora próximo aos maiores potenciais produtores
de biogás importados por arquivo no BiogasCluster.", sep
= "<br>")
DES_Filtro3_opcao3<-paste("Tende a localizar a unidade
centralizadora próximo aos potenciais consumidores de
bioenergia importados por arquivo no BiogasCluster.", sep
= "<br>")
DES_Filtro3_opcao4<-paste("Tende a localizar a unidade
centralizadora próximo aos potenciais consumidores de
bioenergia importados por arquivo no BiogasCluster e ao
gasoduto.", sep = "<br>")

popoverTempate <- '<div class="popover popover-lg"
role="tooltip"><div class="arrow"></div><h3
class="popover-title"></h3><div class="popover-
content"></div></div>'

ui <- shiny::navbarPage(theme = shinytheme("united"),
collapsible = TRUE,title = div(img(src="https://hub-
s.tech/wp-
content/uploads/2020/06/biogascluster1.jpg",style="margin
-top: -14.52px;margin-left:-15px;margin-right:-
14px;margin-bottom:0px; padding-right:-200px;padding-
bottom:30px,padding-left:0px", height = 53)),

tabPanel("Solução Geral",
tags$head(tags$style('#tab {margin-top:-23px;}')),
fileInput("data",h3(tags$span("Importar arquivo Shapefile
ou Excel"),
```

```
tags$span(icon("info-circle"), id = "icon1", style =
"color: orange")),multiple = T,accept =
c(".xlsx",".shp",".shx",".dbf")),uiOutput("tab"),
      bsPopover("icon1",
"Formato adequado para importação de arquivos",
Des_Formato, placement = "right",options = list(template
= popoverTempate)),

tags$style(HTML(".popover.popover-lg {width: 500px; max-
width: 500px;}")),
sidebarLayout(
sidebarPanel(

radioButtons("filtrol",
tagList(
tags$span("Seleção de unidades baseada no potencial de
produção de biogás",style = "font-size: 24px; font-
weight: normal;"),
tags$span(icon("info-circle"), id = "icon2", style =
"color: orange;")),
choiceValues = c(1,2,3),
choiceNames=list(
tagList(
tags$span("Selecionar todas as unidades"),
tags$span(icon("info-circle"), id = "icon3", style =
"color: orange;")),
tagList(
tags$span("Excluir unidades com baixo potencial em
relação às demais"),
tags$span(icon("info-circle"), id = "icon4", style =
"color: orange;")),
tagList(
tags$span("Excluir unidades com alto potencial em relação
às demais"),
tags$span(icon("info-circle"), id = "icon5", style =
"color: orange;")))),
selected=1),

bsPopover("icon2", "Seleção de unidades baseada no
potencial de produção de biogás", DES_Filtrol, placement
= "right",options = list(template = popoverTempate)),
bsPopover("icon3", "Selecionar todas as unidades",
DES_Filtrol_opcao1, placement = "right",options =
list(template = popoverTempate)),
bsPopover("icon4", "Excluir unidades com baixo potencial
em relação às demais", DES_Filtrol_opcao2, placement =
"right",options = list(template = popoverTempate)),
bsPopover("icon5", "Excluir unidades com alto potencial
em relação às demais", DES_Filtrol_opcao3, placement =
"right",options = list(template = popoverTempate)),br(),
```

```
radioButtons("filtro2",
tagList(
tags$span("Abrangência do cluster",style = "font-size:
24px; font-weight: normal;"),
tags$span(icon("info-circle"), id = "icon6", style =
"color: orange;")),
choiceValues = c(1,2),
choiceNames=list(
tagList(
tags$span("Limitar abrangência"),
tags$span(icon("info-circle"), id = "icon7", style =
"color: orange;")),
tagList(
tags$span("Não limitar abrangência"),
tags$span(icon("info-circle"), id = "icon8", style =
"color: orange;"))),selected = 2),

bsPopover("icon6", "Abrangência do cluster", DES_Filtro2,
placement = "right",options = list(template =
popoverTempate)),
                                bsPopover("icon7",
"Limitar abrangência", DES_Filtro2_opcao1, placement =
"right",options = list(template = popoverTempate)),
                                bsPopover("icon8",
"Não limitar abrangência", DES_Filtro2_opcao2, placement
= "below",options = list(template = popoverTempate)),

br(),

radioButtons("Filtro3",

tagList(

tags$span("Preferência para localização da unidade
centralizadora de biogás",style = "font-size: 24px; font-
weight: normal;"),

tags$span(icon("info-circle"), id = "icon9", style =
"color: orange;")),
choiceValues = c(1,2,3,4),

choiceNames=list(
tagList(

tags$span("Próximo ao gasoduto e aos maiores potenciais
produtores de biogás"),

tags$span(icon("info-circle"), id = "icon10", style =
"color: orange;")),
```

```
tagList(  
  
tags$span("Próximo aos maiores potenciais produtores de  
biogás"),  
  
tags$span(icon("info-circle"), id = "icon11", style =  
"color: orange;"),  
tagList(  
  
tags$span("Próximo aos maiores potenciais consumidores de  
bioenergia"),  
  
tags$span(icon("info-circle"), id = "icon12", style =  
"color: orange;"),  
tagList(  
  
tags$span("Próximo ao gasoduto e aos maiores potenciais  
consumidores de bioenergia"),  
  
tags$span(icon("info-circle"), id = "icon13", style =  
"color: orange;")),  
selected=1),  
  
bsPopover("icon9", "Preferência para localização da  
unidade centralizadora de biogás", DES_Filtro3, placement  
= "right",options = list(template = popoverTempate)),  
bsPopover("icon10", "Próximo ao gasoduto e aos maiores  
potenciais produtores de biogás", DES_Filtro3_opcao1,  
placement = "right",options = list(template =  
popoverTempate)),  
bsPopover("icon11", "Próximo aos maiores potenciais  
produtores de biogás", DES_Filtro3_opcao2, placement =  
"right",options = list(template = popoverTempate)),  
bsPopover("icon12", "Próximo aos maiores potenciais  
consumidores de bioenergia", DES_Filtro3_opcao3,  
placement = "right",options = list(template =  
popoverTempate)),  
bsPopover("icon13", "Próximo ao gasoduto e aos maiores  
potenciais consumidores de bioenergia",  
DES_Filtro3_opcao4, placement = "right",options =  
list(template = popoverTempate)),  
br(),  
  
conditionalPanel(  
"output.fileUploaded == true",  
tags$hr(style="border-color: black;"),  
tags$p(h3("Você está satisfeito com a solução?")),  
radioButtons("satisfacao","", choices = list("Sim" =  
1,"Não " = 2),selected = 1)),
```

```
conditionalPanel(
  "input.satisfacao == '2'",
  selectInput("naosatisf", h4("Selecione uma das opções
  abaixo:"),
  choices = c("Nenhuma opção selecionada" = "", "Altere as
  opções de filtro" = 1, "Altere o número de clusters" =
  2), selected = ""),

conditionalPanel(
  "input.naosatisf == '2'",
  sliderInput("Slider", h5(""),
  min = 2, max = 31, value = ""))),

actionButton("reset", "Limpar tela"),
downloadButton("Download", "Download"),),

mainPanel(
  tabsetPanel(
    tabPanel("Clusters", (div(DTOutput("tabela"),
    style="font-size: 100%; width: 95%")),br(),br(),
    (leafletOutput("Leaf1",width = "95%", height = "600"))),
    tabPanel("Análise
  de
  Solução",plotOutput("CAAbrangencia"),plotOutput("CABiogas
  ")))

  ))),

  tabPanel("Rota à unidade centralizadora",
  br(),
  br(),

  sidebarLayout(
  sidebarPanel(
  selectInput("Filtro4", label = h4("Selecione o cluster
  que deseja ser mostrado no mapa"),""),
  tags$hr(),
  selectInput("Filtro5", label = h4("Selecione a unidade
  para cálculo da distância até a unidade
  centralizadora"),""),

  tags$hr(),
  h4("A distância euclidiana é (metros):"),
  textOutput("dist")),

  mainPanel(
  tabsetPanel(
  tabPanel("Rota", uiOutput("Mapa2"))))))),

  tabPanel("Autores e Desenvolvedores",
```



```
br(),

p(a("Antonio Carlos de Francisco",
href="http://lattes.cnpq.br/6457056051910603",
target="_blank"),style = "font-size:25px"),
p("e-mail: acfrancisco@utfpr.edu.br",style = "font-size:20px"),
p(a("Thalita Monteiro Obal",
href="http://lattes.cnpq.br/7574246674522809",
target="_blank"),style = "font-size:25px"),
p("e-mail: thalitaobal@utfpr.edu.br",style = "font-size:20px"),
p(a("Jovani Taveira de Souza",
href="http://lattes.cnpq.br/8929110034758487",
target="_blank"),style = "font-size:25px"),
p("e-mail: jovanisouza5@gmail.com",style = "font-size:20px"),
p(a("Rômulo Henrique Gomes de Jesus ", href="
http://lattes.cnpq.br/4244750019550590",
target="_blank"),style = "font-size:25px"),
p("e-mail: romulohenriquegomes@hotmail.com",style =
"font-size:20px"),
br(),
p(a("Laboratório de Estudos em Sistemas Produtivos
Sustentáveis ", href="https://hub-s.tech",
target="_blank"),style = "font-size:25px"),
p("LESP/UTFPR",style = "font-size:20px")),

tags$head(tags$style("#shiny-modal img { max-width: 100%;
}"))

)

server <- function(input, output,session) {

  #Mensagem introdutoria da ferramenta
  observeEvent("", {
    showModal(modalDialog(
      includeHTML("mensageminicial.html"),
      easyClose = TRUE,
      footer = tagList(
        actionButton(inputId = "ok", label = "OK")
      )
    )
  })

  observeEvent(input$ok, {
    removeModal()
  })
}
```

```
v <- reactiveValues(DadosEntrada = NULL, clear=FALSE)

### Ler os dados excel e shapefile da ferramenta
observeEvent(input$data, {
  if(any(grepl(".xlsx",input$data$name))){
    v$DadosEntrada <- read.xlsx(input$data$datapath)
  }else if(any(grepl(".shp",input$data$name))){
    shpDadosEntrada <- input$data
    erro <- F
    if(!any(grepl(".shx",input$data$name))){
      erro<-T
    }
    if(!any(grepl(".dbf",input$data$name))){
      erro<-T
    }
    if(erro){
      sendSweetAlert(
        session = session,
        title = "Erro!!",
        text = "Você selecionou arquivo(s)
incompatível(eis) com a ferramenta BiogasCluster. Para
uso de planilha, favor verificar o modelo de planilha
padrão disponibilizada no site, a qual deve ser em
formato .xlsx. Para uso de shapefile, favor inserir os
arquivos '*.shp', '*.shx' e '*.dbf' juntos.",
        type = "error"
      )
      return(NULL)
    }else{
      prevWD <- getwd()
      uploadDirectory <-
dirname(shpDadosEntrada$datapath[1])
      setwd(uploadDirectory)
      for (i in 1:nrow(shpDadosEntrada)){
        file.rename(shpDadosEntrada$datapath[i],
shpDadosEntrada$name[i])
      }
      shpName <-
shpDadosEntrada$name[grep(x=shpDadosEntrada$name,
pattern="*.shp")]
      shpName<-substr(shpName,1,nchar(shpName)-4)
      setwd(prevWD)
      shpFile<-
readOGR(dsn=uploadDirectory,layer=shpName)
      v$DadosEntrada<-shpFile@data
    }}else{
      sendSweetAlert(
        session = session,
        title = "Erro!!",
```

```
        text = "Você selecionou arquivo(s)
incompatível(eis) com a ferramenta BiogasCluster. Para
uso de planilha, favor verificar o modelo de planilha
padrão disponibilizada no site, a qual deve ser em
formato .xlsx. Para uso de shapefile, favor inserir os
arquivos '*.shp', '*.shx' e '*.dbf' juntos.",
        type = "error"
    )
    return(NULL)
}
v$clear <- TRUE
}))

ModeloClusterizacao<-reactive({
  req(v$DadosEntrada)
  out <- NULL
  tryCatch({
    out<-
function.clusterizacao(v$DadosEntrada,input$Slider,input$
filtro1,input$filtro2,input$satisfacao,input$Filtro3,inqu
t$Filtro4,input$Filtro5,input$naosatisf)
  },error = function(e){
    sendSweetAlert(
      session,
      title = "Erro!!",
      text = "Você selecionou arquivo(s)
incompatível(eis) com a ferramenta BiogasCluster. Para
uso de planilha, favor verificar o modelo de planilha
padrão disponibilizada no site, a qual deve ser em
formato .xlsx. Para uso de shapefile, favor inserir os
arquivos '*.shp', '*.shx' e '*.dbf' juntos.",
      type = "error"
    )
  })
  out
})

output$fileUploaded <- reactive({
  v$clear && !is.null(ModeloClusterizacao())
})
outputOptions(output, 'fileUploaded',
suspendWhenHidden=FALSE)

observeEvent(input$data, {
  updateSliderInput(session, "Slider", max =
nrow(v$DadosEntrada)-1)
})
```

## BIOGASCLUSTER

```
### Funcao para o botao limpar tela
observeEvent(input$reset, {
  v$DadosEntrada <- NULL
  v$clear <- FALSE
  updateRadioButtons(session,"filtro1", selected = 1)
  updateRadioButtons(session,"filtro2", selected = 2)
  updateRadioButtons(session,"satisfacao", selected =
1)
  updateRadioButtons(session,"Filtro3", selected = 1)
  updateRadioButtons(session,"Slider", selected = 5)
  updateRadioButtons(session, "Filtro4", selected =
" Nenhum cluster selecionado")
  updateRadioButtons(session,"Filtro5", selected = "")
  updateRadioButtons(session, "naosatisf", selected =
" Nenhuma opção selecionada")
})

#### Gera a tabela de saida
output$tabela <- renderDataTable({
  dados_tabela_1 <- req(ModeloClusterizacao())[[1]]
  x <-
datatable(dados_tabela_1[order(dados_tabela_1$Cluster),c(
1,4,2,3)],
           options = list(columnDefs =
list(list(className = 'dt-center', targets = 0:3)),
      paging =TRUE,searching
= FALSE,
      pageLength =
10,lenghtMenu=c(5,10,15,20),scrollx=T
      ), rownames = FALSE)%>%
formatRound(c(3:4), 2,mark = ",")%>%
  formatStyle(columns = c(3:4), 'text-align' =
'center')
  return(x)
})

#### Gera a Carta de Variabilidade da abrangencia dos
clusters
output$CAAbrangencia <- renderPlot({
function.CartaAbrangencia(req(ModeloClusterizacao())[[2]]
)
})

#### Gera a Carta de Variabilidade da produção de
Biogas dos clusters
output$CABiogas <- renderPlot({
  function.CartaBiogas(req(ModeloClusterizacao())[[3]])
})
```

```
### Gera primeiro mapa
output$Leaf1 <- renderLeaflet({
  req(ModeloClusterizacao())[[4]]
})

output$Leaf2 <- renderLeaflet({
  req(ModeloClusterizacao())[[5]]
})

### Gera segundo mapa
output$Mapa2 <- renderUI({
  if(input$Filtro4!="")
    leafletOutput("Leaf2",width = "95%", height =
"600") })

### Opcoes referentes ao Filtro 4 (Selecionar o cluster
que deseja ser mostrado no mapa)
observeEvent(c(v$DadosEntrada,
input$filtro1,input$filtro2,input$Filtro3,input$satisfaca
o,input$Slider,input$naosatisf),{
  abc <- req(ModeloClusterizacao())$Data)
  updateSelectInput(session,'Filtro4',
                    choices=c("Nenhum cluster
selecionado" = "", sort(unique(abc$Cluster))))
})

### Opções referentes ao Filtro 5 (Selecione a unidade
para calculo da distancia até a unidade centralizadora)
observeEvent(c(v$DadosEntrada,
input$filtro1,input$filtro2,input$Filtro3,input$satisfaca
o,input$Slider, input$Filtro4,input$naosatisf),{
  abc <- req(ModeloClusterizacao())$Data1) %>%
filter(cluster == as.numeric(input$Filtro4))
  updateSelectInput(session,'Filtro5',
                    choices = unique(abc$Unidade))
})

### Mostrar a distancia euclidiana no app entre a
unidade selecionada e a respectiva unidade centralizadora
output$dist <- renderText({
  DIST <- data.frame(ModeloClusterizacao())[[7]])
  DIST$distancia[DIST$cluster == input$Filtro4 &
DIST$Unidade == input$Filtro5]
})

### Download da planilha padrao de dados para Excel
```

## BIOGASCLUSTER

```
url <- a("Download", href="https://hub-s.tech/wp-
content/uploads/2020/06/Base-de-dados-padrao.xlsx")
output$tab <- renderUI({
  tagList("Planilha padrão de dados para Excel:", url)
})

### Download dos resultados obtidos pela ferramenta a
partir da escolha do usuário
output$Download <- downloadHandler(
  filename = function() { "basededados.xlsx"},
  content = function(file)
{write_xlsx(req(ModeloClusterizacao())$Resultado), path =
file})
)

}

shinyApp(ui = ui, server = server)
```

### **EQUIPE TÉCNICA**

Coordenador geral do projeto “*Development of innovative tools for locating and evaluating business opportunities related to biogas utilization for energetic purposes*”:

**Antonio Carlos de Francisco (UTFPR)**

Coordenador do desenvolvimento da ferramenta BiogasCluster:

**Antonio Carlos de Francisco (UTFPR)**

Desenvolvedores e idealizadores da ferramenta BiogasCluster:

**Antonio Carlos de Francisco**

**Thalita Monteiro Obal**

**Jovani Taveira De Souza**

**Rômulo Henrique Gomes De Jesus**

**Cassiano Moro Piekarski**

**Daniel Poletto Tesser**

**Fabio Neves Puglieri**

**Julia Eunice Fernandes (UTFPR)**

**Karoline Aquino Santos (UTFPR)**

**Pâmela Brocardo Cruz (UTFPR)**

**André Luiz Przybysz**

Equipe de suporte técnico e apoio no desenvolvimento:

**Gerson Ishikawa (UTFPR)**

**Eduardo Bittencourt Sydney (UTFPR)**

**Alessandra Cristine Novak Sydney (UTFPR)**

**Percio Luiz Karam De Miranda (UTFPR)**

**Carla Cristiane Sokulski (UTFPR)**

**Andre Luiz Przybysz (UTFPR)**

**Rodrigo Salvador (UTFPR)**

**Murilo Vetroni Barros (UTFPR)**

**Jovani Taveira De Souza (UTFPR)**

**Romulo Henrique Gomes De Jesus (UTFPR)**

**Karoline Aquino Santos (UTFPR)**

**Pâmela Brocardo Cruz (UTFPR)**





**ABiogás**  
Associação Brasileira do Biogás



**CIBIOGAS**  
ENERGIAS RENOVÁVEIS



MINISTÉRIO DO  
DESENVOLVIMENTO REGIONAL

MINISTÉRIO DO  
MEIO AMBIENTE

MINISTÉRIO DE  
MINAS E ENERGIA

MINISTÉRIO DA  
AGRICULTURA, PECUÁRIA  
E ABASTECIMENTO

MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



**Biogás**  
BRASIL