UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Câmpus de Sorocaba

PROGRAMA DE PÓS-GRADUAÇÃO em
Ciências Ambientais

PEDRO ANTONIO GALVÃO JUNIOR

MODELOS HÍBRIDOS APLICADOS A CONSTRUÇÃO DE ÍNDICE DE CLASSIFICAÇÃO DE NÍVEIS DE RISCO DE QUEIMADAS NO BRASIL

Sorocaba

2021

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Câmpus de Sorocaba

PROGRAMA DE PÓS-GRADUAÇÃO em
Ciências Ambientais

PEDRO ANTONIO GALVÃO JUNIOR

MODELOS HÍBRIDOS APLICADOS A CONSTRUÇÃO DE ÍNDICE DE CLASSIFICAÇÃO DE NÍVEIS DE RISCO DE QUEIMADAS NO BRASIL

Dissertação apresentada como requisito para a obtenção do título de Mestre em Ciências Ambientais da Universidade Estadual Paulista "Júlio de Mesquita Filho" na Área de Geoprocessamento e Modelagem Matemática.

Orientadora: Profa. Dra. Sandra Regina Monteiro Masalskiene Roveda.

Coorientador: Prof. Dr. Henrique Ewbank de Miranda Vieira.

Sorocaba

2021

G182m

Galvão, Pedro Antonio Junior
Modelos híbridos aplicados a construção de índice de classificação de níveis de risco de queimadas no Brasil / Pedro Antonio Junior Galvão. -- Sorocaba, 2022
105 p. : il., tabs., fotos, mapas

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Instituto de Ciência e Tecnologia, Sorocaba
Orientadora: Sandra Regina Monteiro Masalskiene Roveda
Coorientador: Henrique Ewbank de Miranda Vieira

1. Ciências ambientais. 2. Clima e florestas. 3. Lógica simbólica e matemática. 4. Banco de dados relacionais. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Ciência e Tecnologia, Sorocaba. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Sorocaba

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: MODELOS HÍBRIDOS APLICADOS A CONSTRUÇÃO DE ÍNDICE DE CLASSIFICAÇÃO DE NÍVEIS DE RISCO DE QUEIMADAS NO BRASIL

AUTOR: PEDRO ANTONIO GALVÃO JUNIOR

ORIENTADORA: SANDRA REGINA MONTEIRO MASALSKIENE ROVEDA

COORIENTADOR: HENRIQUE EWBANK DE MIRANDA VIEIRA

Aprovado como parte das exigências para obtenção do Título de Mestre em CIÊNCIAS AMBIENTAIS, área: Diagnóstico, Tratamento e Recuperação Ambiental pela Comissão Examinadora:

Prof. Dr. HENRIQUE EWBANK DE MIRANDA VIEIRA (Participaçao Virtual)
Departamento de Engenharia de Producao / Centro Universitário Facens

Prof. Dr. DARLLAN COLLINS DA CUNHA E SILVA (Participaçao Virtual)
Departamento de Engenharia Ambiental / Unesp - Instituto de Ciência e Tecnologia de Sorocaba

Profª. Drª. ROSANA SUELI DA MOTTA JAFELICE (Participaçao Virtual)
Departamento de Matemática / Universidade Federal de Uberlândia (UFU)

Sorocaba, 20 de dezembro de 2021

Aos meus familiares, em especial meus pais,

minha esposa e filhos.

AGRADECIMENTOS

À Universidade Estadual Paulista “Júlio de Mesquita Filho”, na pessoa do diretor do Instituto de Ciência e Tecnologia de Sorocaba, Prof. Dr. Roberto Wagner Lourenço.

Ao Programa de Pós-graduação em Ciência Ambientais, na pessoa do coordenador Prof. Dr. Admilson Írio Ribeiro.

Aos docentes do Programa de Pós-Graduação em Ciência Ambientais.

À Prof. Dra. Sandra Regina Monteiro Masalskiene Roveda, pelo empenho, respeito, compreensão, confiança e amizade.

Ao Prof. Dr. Henrique Ewbank de Miranda Vieira, pelo empenho, respeito, compreensão, confiança e amizade.

Ao Prof. Dr. Darllan Collins da Cunha e Silva, por sua atenção, conselhos e amizade.

Ao Prof. Dr. José Arnaldo Frutuoso Roveda, por sua atenção e amizade.

Aos demais Professores que a vida me proporcionou.

“Se a sua vida for a melhor coisa que já te aconteceu, acredite,

você tem mais sorte do que pode imaginar.”

Daniel Gobri

RESUMO

O fogo sempre exerceu uma grande atração sobre os seres humanos. As queimadas de maneira geral proporcionam impactos sociais e ambientais nos locais onde ocorrem. Diversas localidades brasileiras, especialmente nos meses mais secos do ano, estão mais suscetíveis a este fenômeno. Para este trabalho foi desenvolvido um índice capaz de classificar os níveis de risco de foco em áreas geograficamente localizadas em nosso país, por meio dos sistemas neuro-fuzzy. O índice proposto poderá ser utilizado como ferramenta de apoio e auxílio a diversos órgãos ou institutos de pesquisa, que necessitam identificar a possibilidade de ocorrência de queimadas. Com o propósito de ajudar nas medidas para redução de emissores atmosféricos, satisfazendo o objetivo 15 da Agenda 30 definido pela ONU em 2015, o qual, visa estimular ações de conservação, recuperação e uso sustentável de ecossistemas em especial. Os resultados obtidos por este índice se mostraram promissores, apresentando valores referentes a média de erros absolutos abaixo de 3% aplicados na predição do risco de queimada para o período máximo de até 3 dias.

Palavras-chaves: Modelagem Fuzzy. Modelos Híbridos. Queimadas. Índices. Dados.

ABSTRACT

Fire has always exerted a great attraction on humans. Fires generally provide social and environmental impacts in the places where they occur. Several Brazilian localities, especially in the driest months of the year, are more susceptible to this phenomenon. For this work, an index was developed capable of classifying the levels of risk of focus in geographically located areas in our country, through the neuro-fuzzy systems. The proposed index can be used as a tool to support and assist various research agencies or institutes, which need to identify the possibility of burning. With the purpose of assisting in measures to reduce atmospheric emitters, meeting objective 15 of Agenda 30 defined by the ONU in 2015, which aims to stimulate conservation actions, recovery and sustainable use of ecosystems in particular. The results obtained by this index were promising, presenting values referring to the average of absolute errors below 3% applied in the prediction of the risk of burning for the maximum period of up to 3 days.

Keywords: Fuzzy modeling. Hybrid models. Burned. Indexes. Data.

LISTA DE FIGURAS

Figura 1 - Arquitetura padrão de um sistema ANFIS. 36
Figura 2 - Localização do município de Sorocaba no estado de São Paulo. 39
Figura 3 – Os 10 municípios com as maiores quantidades de queimadas..... 40
Figura 4 - Distribuição de queimadas por ano ocorridas no Município de Sorocaba. 41
Figura 5 - Fluxograma – Construção do ICRQB. 42
Figura 6 - Etapas do processo de preparação das bases de dados. 46
Figura 7 - Validação Cruzada 5-folds. 49
Figura 8 – Nomenclatura – TrainingGroup. 49
Figura 9 - Estrutura criada para os Modelos 1 e 4. 52
Figura 10 - Menu Tasks – Import Data. 55
Figura 11 - Origem de dados. 55
Figura 12 - Tratamento de Dados..... 56
Figura 13 - Ferramenta Import Wizard. 61
Figura 14 - Relação de variáveis criadas na Toolbox Workspace. 61
Figura 15 - AnfisEdit carregado dentro do ambiente MatLab. 62
Figura 16 - Visualização dos dados de treinamento carregados. 64
Figura 17 - Visualização da estrutura criada para o FIS. 65
Figura 18 - Épocas de treinamento de dados realizada. 65
Figura 19 - Retorno da execução do Test Now..... 67
Figura 20 - Visualização dos dados de checagem carregados. 68
Figura 21 - Época de treinamento de dados realizada. 68
Figura 22 - Test FIS realizado. 69
Figura 23 - Ferramenta AnfisEdit. 71
Figura 24 - Funções de pertinência existentes no Modelo 3. 73
Figura 25 - Gráfico de Superfície - Dia sem chuva e Precipitação. 74

LISTA DE TABELAS

Tabela 1 - Restrições na somatória da Fórmula de Monte Alegre 26
Tabela 2 - Graus de perigo de incêndio classificados pelo INPE 27
Tabela 3 - Tabela 3 – Descrição das variáveis selecionadas 44
Tabela 4 - Descrição das variáveis Banco de Dados de Queimadas do INPE 45
Tabela 5 - Descrição das variáveis Banco de Dados de Precipitações..... 46
Tabela 6 - Grupos de Treinamento..... 50
Tabela 7 – Descrição das variáveis de entrada e saída selecionadas no Modelo 1 51
Tabela 8 - Descrição das variáveis de entrada e saída selecionadas no Modelo 2 51
Tabela 9 - Descrição das variáveis de entrada e saída selecionadas no Modelo 3 51
Tabela 10 - Descrição das variáveis de entrada e saída selecionadas no Modelo 4..... 51
Tabela 11 - Descrição das variáveis de entrada e saída selecionadas no Modelo 5..... 52
Tabela 12 - Descrição das variáveis de entrada e saída selecionadas no Modelo 6..... 52
Tabela 13 - Resumo dos Resultados – Modelos período de 1 dia 71
Tabela 14 - Resumo dos Resultados – Modelos período de 3 dias..... 72
Tabela 15 - Variável risco de fogo observado combinada com a saída calculada..... 75
Tabela 16 - Níveis de Classificação de Risco de Fogo 76
Tabela 17 - Matriz de Confusão..... 77
Tabela 18 - Resultados classes estatísticas Sensibilidade e Especificidade 77

LISTA DE QUADROS

Quadro 1 - Estudos sobre indicadores de risco de queimadas 30

LISTA DE EQUAÇÕES

Equação 1 – Função de Pertinência de um subconjunto fuzzy 32
Equação 2 - ANFIS – Camada 1 – Entrada 36
Equação 3 - ANFIS - Camada 2 - Operador lógico fuzzy 37
Equação 4 - ANFIS – Camada 3 – Grau de ativação normalizado 37
Equação 5 - ANFIS - Camada 4 – Multiplicação para saída 37
Equação 6 - ANFIS - Camada 5 - Soma 37

LISTA DE ABREVIAÇÕES E SIGLAS

ANFIS	= Sistema de Inferência Neuro-Fuzzy Adaptável.
CPTEC	= Centro de Previsão do Tempo e Estudos Climáticos.
CSV	= Arquivo separado por vírgulas.
DAT	= Arquivo de dados.
ETL	= Extração, Transformação e Carga.
FMA	= Fórmula de Monte Alegre.
FMAhoje	= Valor calculado do índice Fórmula de Monte Alegre para hoje.
FMAOntem	= Valor calculado do índice Fórmula de Monte Alegre do dia anterior.
FIS	= *Fuzzy Inference System* (Sistema de inferência Fuzzy).
IBGE	= Instituto Brasileiro de Geografia e Estatística.
INPE	= Instituto Nacional de Pesquisas Espaciais.
INMET	= Instituto Nacional de Meteorologia.
IPEF	= Instituto de Pesquisas e Estudos Florestais (IPEF).
ICRQB	= Índice de Classificação de Níveis de Risco de Queimadas no Brasil.
MCTIC	= Ministério da Ciência Tecnologia e Inovações.
PPA	= Plano Plurianual do Governo Federal.
ONU	= Organização das Nações Unidas.
RF	= Risco de Fogo Previsto.
RFO	= Risco de Fogo Observado.
RFF	= Risco de Fogo Futuro.
RMS	= Região Metropolitana de Sorocaba.
SGBD	= Sistema Gerenciador de Banco de Dados.
TrainingGroup	= Grupo de Treinamento.
TXT	= Arquivo no formato de texto.
UFPR	= Universidade Federal do Paraná.
UTF	= Formato de Transformação Unicode.

SUMÁRIO

1. INTRODUÇÃO 17
1.1 Objetivo Geral 20
1.2 Objetivos específicos 20
2. REVISÃO BIBLIOGRÁFICA 21
2.1 Proteção Florestal 21
2.2 Risco de Fogo 22
2.3 Ocorrência de Queimadas 23
3. ÍNDICES DE CLASSIFICAÇÃO DE NÍVEIS DE RISCO DE FOGO 25
3.1 Índice Fórmula De Monte Alegre 25
3.2 Índice Risco de Fogo Observado (RFO) 27
3.3 Estudos Relacionados aos Indicadores de Risco de Fogo 29
4. TEORIA DOS CONJUNTOS FUZZY 32
4.1 Sistemas de Inferência Fuzzy 33
4.2 Redes Neurais 34
4.3 Modelos Híbridos 35
5. METODOLOGIA E TRATAMENTO DOS DADOS 39
5.1 Área de estudo 39
5.2 Desenvolvimento do Modelo 40
5.3 Construção do Índice de Classificação de Níveis de Riscos de Queimada no Brasil 41
5.4 Seleção das variáveis 43
5.5 Base de dados 45
5.6 Preparação da Base de Dados 46
5.7 Validação Cruzada 47
5.8 Modelos para Predição do Risco de Fogo 50
5.9 Dados Brutos – Extração, Transformação e Carga 53
5.10 Tratamento dos Dados 54
5.11 Limpeza e Integração 54
5.12 Redução e Transformação 56
5.13 Carregamento de Dados (Observações) 57
5.14 Ambiente de Desenvolvimento para Computação Técnica MatLab 57
5.15 Implementando o Modelo Híbrido ANFIS para Classificação do Risco de Fogo 58
5.16 Carregando dados no MatLab 59
5.17 Utilizando a ferramenta AnfisEdit para materializar o ANFIS 62
5.18 Realizando o Treinamento dos Dados (Training) 63

5.19 Realizando a Checagem dos Dados (Checking) 67
6. RESULTADOS E DISCUSSÃO 70
7. CONSIDERAÇÕES FINAIS 79
8. REFERÊNCIAS 81
9. APÊNDICE 84
9.1 Figura 1 - Fonte de Dados – Amostra – Dados do INPE. 84
9.2 Figura 2 - Fonte de Dados – Amostra – Dados do INMET. 84
9.3 Figura 3 - Banco de Dados – Modelo Híbrido – Tabelas. 85
9.4 Figura 4 - Modelos - Arquivos – TrainingGroup, Testing e CheckingGroup. 85
9.5 Figura 5 – Gráfico – RMS - Queimadas por Município. 86
10. ANEXOS 87
10.1 Tabela 1 - Base dados – Amostra aleatória de observações selecionadas para o estudo. 87
10.2 Script 1 – Microsoft SQL Server - Criação – Modelo 3. 88
10.3 Script 2 – MatLab – Carregamentos dos Modelos. 103
10.4 Script 3 – MatLab – Avaliação dos Modelos. 105

1. INTRODUÇÃO

Com o passar dos anos, o fogo exerceu uma grande atração sobre os seres humanos. Para muitos, sendo considerado uma forma que permitiu a evolução da espécie humana, a qual se encontra diretamente ligada a ela até hoje. Este fenômeno tem influenciado padrões e processos ambientais no mundo todo, alterando a distribuição e estrutura da vegetação, criando novos ciclos de emissão do carbono e afetando diretamente o clima em todo planeta (WHITLOCK et al., 2003; BOND et al., 2005; MALMSTRÖM et al., 2010).

Estabelecer uma forma de aprender, criando novos conhecimentos através da ocorrência de queimadas, nas mais diversas áreas geográficas do Brasil, vem a cada dia se tornando uma importante ferramenta no combate a destruição da fauna e flora. A possibilidade de criação de um sistema capaz de ajudar a prever ou até mesmo detectar a presença de uma possível queimada ou risco de fogo chama a atenção dos mais diversos estudiosos e especialistas nesta área de estudos.

Os estudos e formas de classificação de índices como indicadores de risco de fogo iniciou-se em meados dos anos de 1930 no Canadá e nos Estados Unidos. Por meio destes estudos iniciais, foi possível dividir os índices em cumulativos e não cumulativos.

Os índices classificados como não cumulativos baseiam-se somente nas condições do tempo vigentes no dia, entretanto, os índices cumulativos levam em consideração as condições climáticas de uma sucessão de dias, para um período máximo de até 30 dias correntes.

Soares, testou vários elementos meteorológicos, dentre eles a umidade relativa do ar, em diferentes equações de perigo de incêndio, queimadas ou risco de fogo. Dentre estas equações se destaca a Fórmula de Monte Alegre, a qual foi desenvolvida e validada para a região de Telêmaco Borba pertencente ao Estado do Paraná.

Desde o início da década de 1980, o Instituto Nacional de Pesquisas Espaciais (INPE), realiza diariamente atividades relacionadas ao monitoramento das queimadas em toda América do Sul e principalmente no Brasil. Em 1998, foi criada a Divisão de Sistemas e Satélites Ambientais do INPE, com objetivo do desenvolvimento, aprimoramento e criação de métodos para conseguir em tempo hábil, estimar o risco de fogo (SETZER, 2019).

Para estabelecer uma forma de realizar esta estimativa, torna-se necessário a utilização dos chamados fatores climáticos, tais como a precipitação, umidade relativa do ar, temperatura do ar, entre outros.

A utilização de sistemas para obtenção de dados espaciais através do uso de Satélites como o Noaa12 e Noaa16, surgem como ferramentas importantes no auxílio para utilização dos dados ofertados por estes sistemas.

Alguns dos principais índices relacionados a risco de fogo existentes atualmente foram estudados, dentre eles: Fórmula de Monte Alegre, Índices de *Haines*, Fator de Risco de *Ansgtröm*, Índice Hidrotermométrico e Risco Fogo Observado, conforme os estudos realizados por LEMOS et al. (2012).

Estes serviram como fonte de conhecimento no auxílio para fundamentação na escolha das variáveis tanto de entrada como também de saída. Ao compararmos como estes índices estão estruturados, bem como, a forma que sua aplicabilidade, divididos em índices não cumulativos e cumulativos.

O índice de Haines define como elemento de saída para seu resultado a variável denominada Risco de Propagação do Fogo, leva em consideração como fator de análise os dados relacionados a estabilidade atmosférica e precipitação (umidade).

O Fator de Risco de Angström, índice também não cumulativo, considera como fatores de análise de dados a temperatura do ar e a umidade, de acordo com a localidade que está sendo estudado em relação ao um determinado horário do dia. Seu resultado denominado fator de risco, pode ser obtido levando em consideração dados relacionados à diversos períodos de dias, meses ou até mesmo anos.

Tanto a Fórmula de Monte Alegre, como também o Índice Hidrotermométrico, fazem parte do grupo de índices cumulativos, que utilizam como elementos de análise de dados a umidade relativa e temperatura média, mas com algumas particularidades.

A Fórmula de Monte Alegre considera a utilização de dados do dia anterior em relação a área de estudo, para um determinado horário em específico, mas especificamente a umidade relativa do dia anterior (NUNES et al., 2012).

Por outro lado, o Índice de Haines, baseia-se na obtenção dos dados relacionados a um determinado horário em específico do local de estudo, relacionados a temperatura média e umidade relativa média. Dados que podem representar valores para um ou até 30 dias.

Ao pensarmos na maneira de predizer uma possível ocorrência da presença de queimada ou risco de fogo, fazemos referência aos princípios da estimativa meteorológica do chamado risco fogo para a queima da vegetação. A mesma queima é fruto da falta de chuvas ou precipitações nos mais variados tipos de fatores climáticos que promovem a diminuição da umidade no solo (SOARES; BATISTA, 2007).

Este trabalho tem por finalidade estabelecer uma forma de classificação de riscos de queimadas para o território brasileiro, através de um possível indicador denominado: Índice de Classificação de Níveis de Riscos de Queimadas no Brasil (ICRQB).

O qual consiste no uso de um Sistema de Inferência Adaptativa Neuro-Fuzzy (ANFIS), retroalimento pelos dados coletados no banco de dados de queimadas do INPE em conjunto com os dados do banco de dados pluviométricos do INMET, ambos definidos para o período dos anos de 2018, 2019 e 2020.

A partir do uso algoritmo de redes neurais artificiais, o qual, processa os dados dentro de ciclos de treinamento, checagem e testes, temos a capacidade de estabelecer um padrão de reconhecimento, denominado como máquina de aprendizado. Esta máquina de aprendizado, pode se tornar uma nova fonte de conhecimento, mediante aos resultados que ela apresenta.

Neste estudo, será criada uma máquina de aprendizado que terá a função preestabelecida de prover como forma de resultado a identificação da presença de risco de fogo ou queimada, para o período entre 1 ou 3 dias correntes.

Por fim, a estrutura deste estudo é composta de sete seções: Revisão Bibliográfica, Índices de Classificação de Níveis de Risco de Queimadas no Brasil, Lógica Fuzzy e o Desenvolvimento de Índices Ambientais, Metodologia e Tratamento de Dados, Resultados Preliminares e Considerações Finais.

A primeira seção, apresenta uma revisão bibliográfica de aspectos gerais sobre proteção florestal, risco de fogo e ocorrência de queimadas.

A segunda seção, destaca os índices de classificação de níveis de risco de fogo no brasil, destacando os índices Fórmula de Monte Alegre e Risco de Fogo Observado (RFO), os quais foram selecionados com fonte de aprendizado para posterior utilização em conjunto o modelo híbrido aqui proposto.

A terceira seção, enaltece os conceitos sobre Sistemas de Inferência Fuzzy, Redes Neurais e Modelos Híbridos.

A quarta seção, traz a caracterização da construção do índice para classificação de níveis de risco de fogo aplicado a cidade de Sorocaba, destacando as variáveis selecionadas, preparação, tratamento e manipulação das bases de dados do INPE e INMET. Em adicional, a descrição das ferramentas utilizadas para obtenção dos resultados deste trabalho, ou seja, caracterização de um sistema baseado em regras fuzzy e suas etapas.

A quinta seção, enfatiza os resultados obtidos do estudo por intermédio da aplicação do modelo aqui proposto, de forma aderente aos dados selecionadas para uso do modelo híbrido ANFIS.

Ao final, a sexta seção, finaliza o estudo trazendo uma breve descrição do que foi realizado, destacando os itens mais relevantes dos resultados obtidos e direciona quais são as perspectivas para futuros estudos.

1.1 Objetivo Geral

Desenvolver um índice de classificação de níveis de risco de queimadas para o território Brasileiro, por meio da análise de dados reais tendo como base variáveis relacionadas a fatores climáticos, dentre elas o dia sem chuva e precipitação (chuva).

1.2 Objetivos específicos

- Identificar e padronizar as fontes de dados para elaboração do índice;
- Desenvolver uma máquina de aprendizado com base em sistema de inferência neuro-fuzzy para classificação do risco de queimadas aplicada a cidade de Sorocaba; e
- Analisar e comparar os resultados obtidos pelo sistema de inferência, comparando com níveis de classificação pertencentes em outros índices aplicados no Brasil.

2. REVISÃO BIBLIOGRÁFICA

2.1 Proteção Florestal

Proteção florestal é a área da ciência florestal, conhecida como o ramo da silvicultura que objetiva proteger a floresta de seus inimigos, através do controle, prevenção e manejo de agentes.

A proteção florestal surge como elemento de monitoramento para identificar os possíveis impactos tratados pela ocorrência destes fenômenos. Os quais, podem ocorrer em decorrência de ações naturais, raios, vulcões ou por efeitos antrópicos, negligência, bem como de forma acidental ou intencional.

Em diversos estudos realizados por (TORRES et al., 2018), e (SILVA et al., 2016), são tratados os efeitos e impactos causados devido a presença do fogo, como a modificação do ambiente, perdas no âmbito social, ambiental, econômico e da biodiversidade da área afetada.

Nos meios agrícolas e florestais a utilização do fogo é tratado como um assunto polêmico e muitas vezes mal compreendido. Para alguns, este fenômeno é reconhecido com preocupação, representando a principal forma de destruição.

No entanto, em outras regiões, culturas e crenças a adoção da queima de forma controlada aplicada a vegetação específica é reconhecida como uma valiosa ferramenta nas ações de manejo agrícola e florestal.

Estudos ambientais, relacionados a presença do fogo no território Brasileiro não é um assunto recente, ao contrário, o mesmo começou a ganhar mais importância a partir do ano de 1963.

Em especial nas localidades do estado do Paraná, no distrito florestal de Monte Alegre, durante o período de inverno do referido ano, devido ao grande desastre ambiente, econômico e social provocado pela presença do fogo.

Naquele período, aproximadamente 2.000.000 ha de florestas sofreram com o poder do fogo, 20.000 ha de plantações de *Araucária angustifólia*, 500.00 ha de florestas primárias e 1.500.000 ha de florestas secundárias na localidade da Fazenda Monte Alegre estabelecida em 1943 foram consumidas pelo fogo. Alguns relatos da época destacam que 8.000 imóveis foram destruídos, 5.700 famílias desabrigadas, 110 pessoas morreram no estado do Paraná.

Atualmente, no Brasil, milhares de queimadas foram registradas nos últimos anos, estas consideradas como umas das principais formas de geração de impactos em nossas florestas, vida selvagem e meio ambiente.

As alterações climáticas e mudanças ambientais, proporcionam ao fogo uma fonte de energia natural cada vez mais devastadora, combinada com a velocidade do vento, umidade relativa e temperatura, ganham mais poder, aumentando sua intensidade e velocidade de destruição (NUNES et al., 2006).

2.2 Risco de Fogo

Ao pensarmos na maneira de estimar uma possível ocorrência de queimadas, bem como, a presença do risco de fogo, estamos fazendo referência aos princípios da estimativa meteorológica chamada risco de fogo aplicada a queima da vegetação.

A mesma queima pode ser fruto da falta de chuvas ou precipitações decorrente dos mais variados tipos de fatores climáticos que promovem a diminuição da umidade no solo (SOARES; BATISTA, 2007).

Para estimar ou prever o conceito do risco de fogo, devemos inicialmente considerar a contagem de dias seguidos sem chuva em um determinado local, para a quantidade máxima de 120 dias seguidos da não presença de chuva. Esta é uma forma básica e simples de tentar se estabelecer um fator que nos possibilite dizer, que em decorrência da falta de chuvas a possibilidade da ocorrência de uma queimada ou presença do risco de fogo poderá ser maior.

Considerando os mais diversos estudos relacionados a presença do Risco de Fogo, que permitiram estabelecer a concepção de indicadores relacionados a ocorrência de queimadas, torna-se necessário destacar alguns que serviram como fonte do saber na concepção do estudo aplicado a este trabalho, sendo eles: "*Comparação entre Métodos de Risco de Queimadas pelo Índice de Haines, Angström, Monte Alegre e o Hidrotermométrico Na Região de Viçosa-Mg no período de 01 A 30 de Setembro de 2003*" (LEMOS et al., 2012), e "*Desempenho dos Índices de Nesterov e Fórmula de Monte Alegre no Distrito De Lichinga, Norte de Moçambique*" (MBANZE, et al., 2017).

Ambos os estudos destacados, enfatizam como alguns dos mais importantes métodos de identificação do risco de queimadas podem ser aplicados nos mais variados territórios ao redor do mundo, baseados em fatores como a estabilidade atmosférica, temperatura do ar, umidade relativa, entre outros.

Para o território brasileiro, o método Fórmula de Monte Alegre surge como um dos mais relevantes e adequados a ser utilizado, em decorrência da utilização poucos fatores para

obter o resultado que classifique o risco de queimada no local de estudo ao qual o mesmo está sendo aplicado. Vale ressaltar que ao longo deste trabalho, este indicador será mais detalhado.

Ao considerarmos outros fatores que podem ser adotados aos métodos para identificar um risco de queimadas, existe um que está presente nos métodos apresentados, mais especificamente o fator que deve determinar o período de análise mensurado a através da quantidade de dias corridos.

Este fator pode ser considerado um elemento chave na maneira afim de calcular o risco de queimadas, adotando os dados obtidos ao longo de um período de dias que poderá variar entre 1 ou 30 dias no máximo.

2.3 Ocorrência de Queimadas

Nos últimos anos, as questões ambientais têm se tornado cada vez mais evidentes em todo o mundo devido à importância do tema aliada a busca pela preservação ambiental e maneiras para obter um desenvolvimento cada vez mais sustentável.

Com a consolidação da Agenda 21, a Organização das Nações Unidas, a ONU, estabeleceu em 2015 a Agenda 30, que se refere a um plano de ação para o planeta, para as pessoas e para o futuro, sendo composta de 17 Objetivos de Desenvolvimento Sustentável e 169 metas.

Dentre os seus 17 objetivos, destaca-se o objetivo 15 que trata como proteger, recuperar e promover o uso sustentável dos ecossistemas terrestres. Como também, gerir sustentabilidade das florestas, procurando combater a desertificação, afim de deter e reverter a degradação da terra e a perda de biodiversidade, amparado pelo item 15.1. Este enfatiza a necessidade de estabelecer a conservação, recuperação e uso sustentável de ecossistemas em especial florestas e degradação da terra até o ano de 2020 (ONU, 2015).

A ocorrência de queimadas bem como sua propagação apresenta uma forte relação com os fatores climáticos, os quais influenciam no desenvolvimento de focos de calor, especialmente em determinadas áreas ou regiões do Brasil. A necessidade de utilização de dados meteorológicos como fonte de análise, torna-se uma importante ferramenta para elaboração de um plano de prevenção.

Adotar as medidas preventivas, através da utilização de um índice capaz de classificar os níveis de risco de queimadas, de forma confiável é essencial para um planejamento mais eficiente das medidas de prevenção, adoção de ações rápidas e efetivas nas atividades de combate a queimada (NUNES et al., 2006, 2007, 2008).

Diversos fatores de classificação de risco de queimadas são baseados em fórmulas determinísticas, que por sua vez, necessitam da utilização de fatores climáticos como elementos para fundamentação de observações, além, de conhecimentos prévios que inúmeros pesquisadores utilizam (DUARTE et al., 2019).

Entretanto, esta combinação torna-se específica, a partir do momento em que se visa a possibilidade de criação de um índice para classificar possíveis níveis de risco de fogo em uma determinada área ou localidade.

O Centro de Previsão de Tempo e Estudos Climáticos do Instituto Nacional de Pesquisas Espaciais (CPTEC/INPE) utiliza diariamente diferentes satélites (sensores) com objetivo de produzir dados essenciais que possam ajudar a identificar a chamada radiação termal.

Com ela é possível reconhecer esta fonte de alta emissão de radiação como uma queimada ou foco de calor.

3. ÍNDICES DE CLASSIFICAÇÃO DE NÍVEIS DE RISCO DE FOGO

Como forma de possibilitar a utilização de uma fonte confiável para prevenção e planejamento eficiente visando estabelecer medidas de prevenção e combate, estudos começaram a ser realizado no Brasil em meados dos anos de 1960.

Naquela época, assim como hoje, entendia-se que adotar este tipo de recurso seria fundamental para criar ações rápidas e efetivas como fonte de aprendizado para as possíveis ocorrências de queimadas provocadas por diversas ações naturais ou não.

Pode-se dizer que a primeira classificação de níveis de risco de queimadas foi elaborada no Brasil a partir do ano de 1965, motivado pelo efeito provocado pelo grande incêndio de 1963 na Fazenda de Monte Alegre.

Inicialmente definiu-se um possível sistema criado pela empresa *Klabin S.A*, sendo está a mais afetada pelo incêndio de 1963, o qual tinha como objetivo principal servir como forma de proteção contra a ocorrência de e queimadas nas localidades do estado do Paraná. A partir das informações consolidadas pela Klabin, permitiu-se estabelecer um dos maiores bancos de dados composto por dados relacionados as ocorrências de e queimadas no Brasil.

Esta preciosa fonte de dados atualizada constantemente, proporcionou a criação de um dos mais utilizados e conhecidos indicadores relacionados ao perigo do risco de fogo, adotado em diversos países da América do Sul, bem como em Moçambique, conhecido como Fórmula de Monte Alegre (SOARES, 2007).

Sabendo da correlação entre as mais diversas variáveis meteorológicas e os graus de perigo (SOARES; BATISTA, 2007) e (TETTO, 2012) ressaltam que essa relação não pode ser considerada o elemento primordial, ao ponto de garantir eficiência na previsão de ocorrência do risco de fogo ou queimada.

Outros índices relacionados aos níveis de risco de incêndio ou queimadas estão presentes no Brasil além do Fórmula de Monte Alegre, dentre eles: de Angstron, Nesterov, Fórmula de Monte Alegre Alterada (FMA+) e Risco de Fogo Previsto (RF) (SOARES; BATISTA, 2007).

3.1 Índice Fórmula De Monte Alegre

Afim de ilustrar os diversos índices de classificação de níveis de risco de queimadas aplicados ao Brasil, foi escolhido como temática de estudo e observação o índice Fórmula de Monte Alegre, que apresenta grande destaque nos trabalhos e artigos acadêmicos estudados

para elaboração do estudo aplicado a este trabalho. TETTO (2012) afirma que o este índice foi o primeiro relacionado ao perigo do fogo desenvolvido no Brasil, de autoria de *Ronaldo Viana Soares* em 1972. Este apresenta um bom desempenho não só em regiões de clima semelhante, mas também em várias outras regiões, justificando-se em alguns casos apenas o ajuste das classes de perigo.

Conhecido como índice cumulativo, calculado através do uso das variáveis meteorológicas: Umidade Relativa do Ar (URi13h); e Total Diário de Precipitação Pluviométrica. A umidade relativa do ar, corresponde ao valor em porcentagem obtido obrigatoriamente no horário das 13hr referente a localidade que se deseja analisar, valor aplicado diretamente para computar o cálculo do perigo de fogo.

Por outro lado, o total diário de precipitação pluviométrica, é utilizado como fator restritivo, considerado como elemento de ajuste na somatória a ser realizada pelo FMA. Em outras palavras, o indicador referente ao total de precipitação poderá ser ajustado mediante a necessidade, afim de possibilitar uma classificação e identificação do risco de perigo de fogo com maior acurácia.

Como forma de classificar o risco de perigo de fogo para o dia atual, o índice Fórmula de Monte Alegre, leva em consideração a ocorrência de precipitação classificada através do indicador do volume pluviométrico mensurado em milímetros.

Estes valores são classificados, utilizados como parâmetro condicional, servindo como elemento de ajuste afim de obter o resultado referente ao FMAhoje. O resultado a ser obtido para o FMAhoje, será modificado caso o valor do volume da precipitação do dia anterior for maior que 2,5 milímetros. A Tabela 1 apresenta as faixas de classificação de valores em relação ao indicador da precipitação do dia anterior, como também, a modificação do cálculo de ajuste na identificação do risco de perigo de fogo.

Tabela 1 - Restrições na somatória da Fórmula de Monte Alegre

Precipitação mm dia^{-1}	**Modificação no cálculo**
< 2,5	Nenhuma, isto é, continuar o cálculo e a somatória como se não houvesse precipitação.
2,5 a 4,9	Abater 30% na FMA acumulada até a véspera e somar (100/H) do dia, isto é: $\mathbf{FMA_{hoje} = 0,7 * FMA_{ontem} + FMA_{i(hoje)}}$
5,0 a 9,9	Abater 60% na FMA acumulada até a véspera e somar (100/H) do dia, isto é: $\mathbf{FMA_{hoje} = 0,4 * FMA_{ontem} + FMA_{i(hoje)}}$
10,0 a 12,9	Abater 80% na FMA acumulada até a véspera e somar (100/H) do dia, isto é: $\mathbf{FMA_{hoje} = 0,2 * FMA_{ontem} + FMA_{i(hoje)}}$
> 12,9	Interromper o cálculo anterior (FMA = 0) e começar novo cálculo no dia seguinte

Fonte: Soares; Batista (2007).

SOARES (2007) também estabeleceu uma escala de interpretação para o indicador Fórmula de Monte Alegre, onde, uma vez calculado o índice de incêndio, os valores acumulados são classificados em diferentes graus de perigo de incêndio Tabela 2.

Tabela 2 - Graus de perigo de incêndio classificados pelo INPE

FMA	Grau de perigo
0 - 1,0	Nulo
1,1 - 3,0	Pequeno
3,1 – 8,0	Médio
8,1 – 20,0	Alto
> 20,0	Muito alto

Fonte: Soares; Batista (2007).

Os dados de focos de calor utilizados neste estudo são provenientes da plataforma Sistema de Monitoramento de Queimadas e Florestais por Satélite em Tempo Quase-Real, do Instituto Nacional de Pesquisas Espaciais (INPE, 2018).

3.2 Índice Risco de Fogo Observado (RFO)

O índice Risco de Fogo Observado (RFO), foi criado pelo Centro de Previsão de Tempo e Estudos Climáticos/Instituto Nacional de Pesquisas Espaciais (CPTEC/INPE), em decorrência da análise e registro das ocorrências de centenas de milhares de queimadas nos principais biomas (tipos de vegetação) presentes no território nacional, como também, em outros países com formam a América do Sul.

Sua primeira versão foi aplicada para estimar o risco de fogo observado, a qual está datada a partir do ano de 2002.

O RFO, desde então, passou a ser considerado como uma das principais ferramentas e fontes de aprendizado no que se refere ao Programa de Monitoramento de Queimadas e Florestais do INPE, podendo ser facilmente acessado e consultado através do endereço eletrônico: https://queimadas.dgi.inpe.br/queimadas/portal.

O mesmo posteriormente foi adicionado como componente da Ação 20V9 do Ministério da Ciência Tecnologia e Inovações (MCTIC) no atual Plano Plurianual do Governo Federal (PPA).

Este índice, apresenta como princípio o risco da queima da vegetação (presença de fogo) de acordo com a quantidade de dias sem chuva. O mesmo, leva em consideração fatores

relacionados à tipo do ciclo natural de desfolhamento da vegetação, temperatura máxima e umidade relativa mínima do ar, em adicional a possível presença do fogo na região de estudo (SETZER et al.; 2019).

Utiliza como referência para seu cálculo o denominado fator “Dias de Secura” ou “Secura”, que faz referência a um valor hipotético de dias sem nenhum registro de precipitação identificado nos últimos 120 dias na área de estudo.

Além dos fatores destacados anteriormente, a possibilidade da ocorrência de uma queimada, também pode ser gerada em sua grande maioria por inúmeras ações provocadas pelo Homem e não somente de forma natural, oriunda de raios ou outros fatores climáticos. Este também é um outro possível fator importante na consideração da possível presença ou não de uma queimada.

Para a definição do cálculo dos tipos de risco de fogo, a velocidade do vento não é considerada, como também a umidade do solo na realização dos possíveis cálculos, pois trata-se de uma variável decorrente do histórico da precipitação e não necessariamente se torna válida.

Por exemplo, ao analisarmos uma determinada área resultante de desmatamento, que apresenta uma vegetação em estado seco num período de meses, está referida área poderá sofre uma queimada num espaço de tempo pequeno de dias, mesmo tendo passado por um período de longas chuvas.

Este fato normalmente se deve ao subsolo desta área estar encharcado, em decorrência do acumulo de água em sua camada mais argilosa, tornando-se matéria orgânica no local, o que dificulta ainda mais o processo de escoamento da água presente. Outro ponto importante, que permite a ocorrência de uma queimada, está condicionado com a presença de forte insolação, combinada com as altas temperaturas do ar.

Afim de possibilitar a criação e utilização do risco de fogo observado, o CPTEC/INPE utilizaram como elemento chave os dados meteorológicos coletados ao longo dos dias nas centenas de estações de superfícies espalhadas pelo país.

Estes dados, são utilizados em um processo matemático chamado de interpolação, de modo a possibilitar que seja determinado o valor de uma função, dentro de um intervalo, a partir do conhecimento dos valores extremos desse intervalo.

A partir da combinação deste conjunto de dados, em especial os valores referentes aos dias de secura e precipitação, ambos já catalogados para o período e área de estudo observado, este sistema tem a possibilidade de estimar a possível presença de uma queimada, risco de fogo

ou somente um aumento na temperatura, mas que não representa a ocorrência destes fenômenos.

Para um período de alguns dias conhecidos o RFO possibilita que seja calculado outro índice conhecido como Risco de Fogo Previsto (RF), o qual é normalmente utilizado para intervalos entre 1 e 3 dias.

Para períodos acima de 3 dias, surge o Risco de Fogo Futuro (RFF) para algumas semanas respeitando o intervalo máximo de 4 semanas ou 30 dias, também calculado com base nos valores obtidos através do RFO.

Em ambos os cálculos de previsões, são utilizados satélites com resolução espacial de 25 km para o modelo numérico *Global Forecast System* (GFS). Este modelo permite identificar a precipitação, temperatura máxima e umidade relativa mínima do ar, fatores utilizados de forma obrigatória para identificar o RFO.

Da mesma forma que o índice FMA, o RF para 3 dias, também apresenta este comportamento, ou seja, necessita do uso de algumas fórmulas e dados coletados para sua concepção. Este índice, é resultante da análise do risco de fogo observado, com base no dia anterior, utiliza as variáveis meteorológicas Temperaturas Máximas Previstas e Umidades Mínimas (SISMANOGLU et al.; 2019).

3.3 Estudos Relacionados aos Indicadores de Risco de Fogo

A utilização de índices como indicadores de risco de fogo iniciou-se em meados dos anos de 1930 no Canadá e nos Estados Unidos. Estes índices são divididos em métodos não cumulativos e cumulativos. Os índices classificados como não cumulativos baseiam-se somente nas condições do tempo vigentes no dia. Já os cumulativos levam em consideração as condições climáticas de uma sucessão de dias.

Com objetivo de fundamentar a discussão sobre o tema, foi realizado um levantamento em bases de dados como Biblioteca Digital de Periódicos – Universidade Federal do Paraná (UFPR), Instituto de Pesquisas e Estudos Florestais (IPEF), Periódico Ciência Florestal (Universidade Federal de Santa Maria) e Science Direct.

Acerca de estudos que abordem, direta ou indiretamente, o controle e indicadores de risco de queimadas (*control and risk indicators of fires,* em inglês), abrangendo o período de 2008 a 2020, dos quais para elaboração deste projeto foram selecionados 8 trabalhos, conforme apresenta o Quadro 1.

Quadro 1 - Estudos sobre indicadores de risco de queimadas

ESTUDO	TÍTULO	AUTOR(ES)	TIPO	ORIGEM	ANO
01	Elaboração de um Fator de Risco de Florestais utilizando Lógica Fuzzy.	Isaac D. B. Silva, e Antonio C.F. Pontes Jr.	Artigo	Nacional	2008
02	Desempenho da Fórmula de Monte Alegre e da Fórmula de Monte Alegre Alterada no Distrito Florestal de Monte Alegre.	José R. S. Nunes, Ivone S. N. Fier, Ronaldo V. Soares, e Antonio Carlos Batista.	Artigo	Nacional	2009
03	Comparação entre Métodos de Risco de Queimadas pelo Índice de Haines, Angström, Monte Alegre e o Hidrotermométrico Na Região de Viçosa-Mg no período de 01 A 30 de Setembro de 2003.	Carlos F. Lemos, Flávio Justino, Hygor A. Rossoni.	Artigo	Nacional	2012
04	Perigo de incêndio florestal: aplicação da Fórmula de Monte Alegre e avaliação do histórico para Piracicaba, SP.	Clayton A. Alvares, Ítalo R. Cegatta, Lucas A. A. Vieira, Rafaela F. Pavani, Eduardo M. de Mattos, Paulo C. Sentelhas, José L. Stape,e Ronaldo V. Soares.	Artigo	Nacional	2014
05	Risco de ocorrência de queimada e de incêndio e as medidas de prevenções, em Belém – PA, no ano de 2015.	Maria do Carmo F. de Oliveira, José A. de S. Jr, Patrícia P. N. da Cruz, e José D. S. Filho.	Artigo	Nacional	2016
06	Desempenho dos Índices de Nesterov e Fórmula de Monte Alegre no Distrito De Lichinga, Norte de Moçambique.	Aires A. Mbanze, Antonio C. Batista, Alexandre F. Tetto, Axel M. Romero, e John Mudekwe.	Artigo	Internacional	2017
07	Fire danger index efficiency as a function of fuel moisture and fire behavior.	Fillipe T. P. Torres, Joyce M. N. Romeiro, Ana Carolina de A. Santos, Ricardo R. de O. Neto, Gumercindo S. Lima, e José C. Zanuncio.	Artigo	Internacional	2018
08	A wildfire warning system applied to the state of Acre in the Brazilian Amazon.	Laércio Silva, M.E. Valle, L.C. Barros, e J.F.C.A. Meyer.	Artigo	Internacional	2020

Fonte: Elaborado com base em *Silva e Pontes Jr (2008), Nunes, Fier, Soares e Batista (2009), Lemos, Justino e Rossoni (2012), Alvares, Cegatta, Vieira, Pavani, Mattos, Sentelhas, Stape e Soares (2014), Oliveira, Souza, Cruz e Souza (2016), Mbanze, Batista, Tetto, Romero e Mudekwe (2017), Torres, Romeiro, Santos, Neto, Lima e Zanuncio (2018), Silva, Valle, Barros e Meyes (2020).*

Através do levantamento realizado, podemos observar que cinco estudos nacionais e um internacional se relacionam diretamente com os indicadores de classificação de risco de fogo.

Desta relação, quatro estudos destacam de forma direta a presença do Índice Fórmula de Monte Alegre, e um estudo demonstra de forma prática a combinação da Lógica Fuzzy como ferramenta para elaboração de um índice de queimadas (SILVA et al., 2016).

Por fim, outro estudo apresenta uma comparação entre outros índices e métodos utilizados para classificar o risco de queimadas (MBANZE et al., 2017).

Os demais trabalhos abordam aspectos relacionados ao comportamento do fogo, ocorrência de queimadas em locais específicos do Brasil, mas não se preocupam em evidenciar o uso de possíveis indicadores de controle e identificação de riscos (TORRES et al., 2018).

4. TEORIA DOS CONJUNTOS FUZZY

A Teoria dos conjuntos fuzzy, proposta por Lofti Zadeh (1965), é aplicada para casos onde existe incerteza na classificação dos dados. Por meio desta teoria é possível representar as variáveis por conjuntos fuzzy, sendo possível que um elemento pertença a mais de um conjunto ao mesmo tempo.

Podemos considerar o seguinte exemplo aplicado a classificação da altura de uma árvore. Na lógica Aristotélica, ela pode ser descrita como pequena, alta ou média, pois esta lógica classifica os elementos considerando apenas as situações de pertencimento ou não pertencimento a um dado conjunto. No entanto, na lógica fuzzy a altura da árvore pode ter pertinência em um ou mais conjuntos.

Ao se comparar a altura de árvores adultas como por exemplo de uma pitangueira (Eugene uniflora), que chega a medir de 6 a 12 metros com uma paineira (Celia speciosa) que varia entre 15 e 30 metros, podemos dizer que a pitangueira é baixa em relação a paineira, mas ao comparar a paineira com um jequitibá-rosa (Cariniana legalis) uma das maiores árvores nativas brasileira que pode atingir até 50 metros de altura, neste caso a paineira pode ser considerada como baixa.

Esse tipo de comparação envolve ainda outras variáveis como a idade da planta e condições de cuidados, todas estas informações podem ser avaliadas por meio da lógica fuzzy, que possibilita que a paineira, neste caso, possa estar em mais de um intervalo de classificação (LORENZI, 2008).

A função de pertinência de um subconjunto fuzzy é definida através da utilização de uma função pré-estabelecida, em que seu intervalo varia de 0 a 1 representando desta forma o grau de pertinência de um elemento x de U de um subconjunto fuzzy F:

$$\mu_F: U \rightarrow [0,1] \qquad (1)$$

onde:

U = Conjunto clássico;

F = Subconjunto fuzzy.

O subconjunto fuzzy F é constituído por elementos x de um conjunto clássico U. O valor $\mu_F(x)$ indica o grau com que o elemento x de U está no conjunto fuzzy F.

As operações típicas para conjuntos clássicos como união e intersecção são também definidas para conjuntos fuzzy, que estão bem definidas por BARROS e BASSANEZI (2010).

4.1 Sistemas de Inferência Fuzzy

Segundo Barros e Bassanezi (2010) um sistema de inferência fuzzy é um processo computadorizado que trabalha do mesmo jeito que as nossas habituais tomadas de decisões.

Os sistemas de inferência fuzzy envolvem quatro módulos principais : fuzzificação, base de regras, inferência fuzzy e defuzzificação.

No primeiro módulo, as entradas do sistema são modeladas por conjuntos fuzzy, com seus respectivos domínios. Para essa etapa, é importante a presença de especialistas do fenômeno que será modelado para a construção das funções de pertinência, que descrevem o conjunto fuzzy. Após a entrada, os dados são submetidos a uma base de regras composta por proposições fuzzy também estabelecidas pelo especialista do fenômeno estudado. Essas regras fuzzy são elaboradas a partir de proposições condicionais, tais como "Se (estado) Então (resposta)", baseadas em variáveis linguísticas, e então, são modeladas por conjuntos fuzzy. Os conjuntos fuzzy que compõem o "estado" são chamados de antecedentes e os que compõem a "resposta", consequentes.

No módulo de inferência, cada proposição é traduzida para a forma matemática utilizando as técnicas fuzzy. Nesta etapa são definidas as t-normas, t-conormas e regras de inferência utilizadas para se obter a saída fuzzy que será utilizada pelo sistema.

Por fim, o último módulo, que representa a estrutura de um sistema de inferência fuzzy é denominada defuzzificação. Nesta etapa, é indicado um método para defuzzificar a saída e obter um número real que, finalmente, indicará a saída a ser adotada. Entre as diversas técnicas, a mais comumente utilizada é o método do centro de área também conhecido como método do centro de gravidade ou do centroide.

Os sistemas de inferência fuzzy do tipo Mamdani e do tipo Takagi-Sugeno estão entre os mais conhecidos. Os sistemas do tipo Mamdani (MAMDANI e ASSILIAN, 1995), adotam um método para o processo de decisão baseado em regras do tipo "SE A ENTÃO B", as quais tanto o antecedente quanto o consequente são valores de variáveis linguísticas, expressos por meio de conjuntos fuzzy. Nestes sistemas agregação do conjunto de regras é realizada através do operador união sobre todas as relações individuais. Baseado na regra de composição de

inferência max-mín, propõe uma relação fuzzy binária para modelar matematicamente a base de regras.

Por outro lado, identificando algumas dificuldades para a utilização do processo de decisão adotado no controlador de Mamdani, Takagi e Sugeno (1983) propuseram um método de tomada de decisão simplificado, baseado na lógica fuzzy, onde somente o antecedente das regras é formado por variáveis fuzzy e o consequente de cada regra é definido como uma função dos valores de entrada de cada regra. O resultado de cada regra é, portanto, um valor numérico (não um conjunto fuzzy), que assume como peso o valor da pertinência resultante do processamento do antecedente da regra. Esta determinação dispensa, neste caso, a definição de uma função de implicação específica. Por padrão a resposta final do sistema de Takagi-Sugeno é obtida pela média ponderada das respostas das regras individuais. Isto é, para este tipo de sistema não se faz necessário o processo de defuzzificação.

Os sistemas de Mamdani e Sugeno têm sofrido aprimoramentos ao longo do tempo e têm sido utilizados para diferentes aplicações. Estão disponíveis atualmente em diversos pacotes computacionais, com diferentes recursos para implementação dos seus componentes, como forma de facilitar o desenvolvimento de sistemas fuzzy para a solução de diferentes problemas de controle a partir da seleção de um desses tipos básicos.

4.2 Redes Neurais

As redes neurais são conhecidas como uma forma de computação não algorítmica que relembram a estrutura de processamento do cérebro humano, também chamada de conexionismo ou sistemas de processamento paralelo e distribuído. São formadas por unidades de processamento simples (nodos ou neurônio artificial) que computam determinadas funções matemáticas (normalmente não lineares).Estas unidades estão dispostas em uma ou mais camadas, interligadas por um grande número de conexões unidirecionais.As conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo, servindo como elementos de ponderação de acordo com as entradas recebidas para cada neurônio existente na rede.

O sistema de redes neurais realiza o processamento de informações, formada por constituintes estruturais, que se assemelham a neurônios biológicos, os quais realizam a transmissão das informações por meio da diferença de potencial elétrico na sua membrana (MONTANARI, 2016). Estes “neurônios” apresentam uma disposição que se associa a um algoritmo de aprendizagem, reconhecido como uma rede alimentada diretamente por múltiplas camadas,

dentro desta estrutura poderemos encontrar os nós chamados de ocultos, estes responsáveis em interagir com os dados de entrada e saída.Neste conjunto de múltiplas camadas, encontramos a primeira camada que representa a entrada dos neurônios de fonte, que interagem com os neurônios ocultos, inseridos nos padrões de ativação.

Maneira, a qual os dados de saída correspondem diretamente aos dados de entrada da segunda camada e assim sucessivamente, estabelecendo estas interligações.

SILVA (2016), destaca que uma das maneiras de trabalhar com redes neurais consiste no uso do perceptron, o qual, possui o objetivo de classificar os conjuntos de estímulos aplicados nas entradas, esta é a maneira como as redes neurais artificiais apresentam a capacidade de aprender mediante a apresentação de exemplos.

A partir, do uso de Algoritmo de Aprendizado, formado por um conjunto de procedimentos para adaptar os pesos das conexões entre as unidades de processamento, tendo como objetivo possibilitar as redes neuras artificiais “aprender” uma determinada função.

Para redes de múltiplas camadas o perceptron tem sido aplicado para resolução de problemas, submetido a treinamentos realizados através de um algoritmo baseado na regra de aprendizagem por correção de erros. A qual, consiste na propagação dos valores dos pesos sinápticos constantes, formada pelo vetor de entrada dos dados aplicado aos nós da rede, estabelecendo assim um efeito de propagação para ela, possibilitando um conjunto de saída.

Além disso, temos a reprogramação, operação que aplica aos pesos sinápticos a regra de correção de erros, subtraindo da rede a resposta real, produzindo assim um sinal de erro, enviado para parte de trás da rede.

A mudança de pesos sinápticos, é realizada para que a resposta real da rede consiga estar mais próxima cada vez da resposta desejada, realizando assim o treinamento da rede, através do uso do algoritmo conhecido como *back-propagation*.

Todavia, ao realizar a adoção de uma rede neural artificial, temos a capacidade de estabelecer uma possível solução a uma determinada tarefa, o que ocorre inicialmente por uma fase de aprendizagem.

4.3 Modelos Híbridos

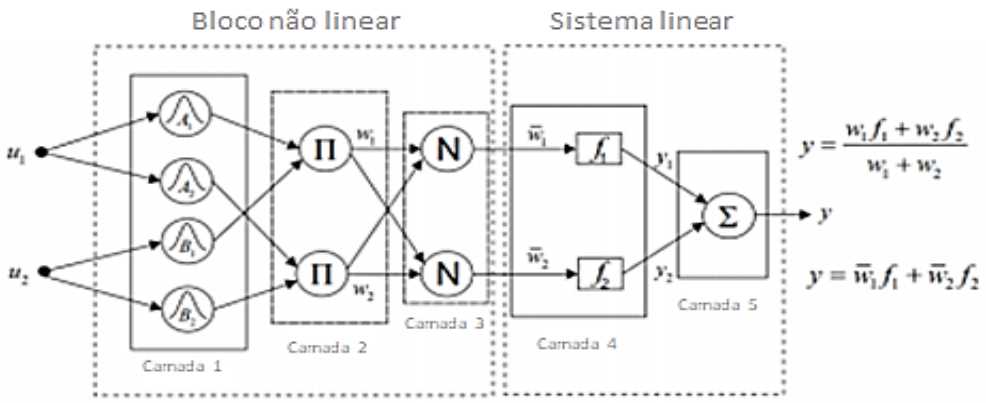
O modelo neuro-fuzzy é denominado como modelo híbrido, que estabelece uma relação entre o sistema de inferência fuzzy e as redes neurais, possibilitando que as informações, sejam treinadas, criando as bases de regras de forma automática.

Esta integração entre as diferentes técnicas: Redes Neurais, Algoritmos Genéticos e Lógica Fuzzy, nos permite criar um sistema computacional que incorpore o conhecimento humano, capaz de aprender e se adaptar ao ambiente de acordo com as incertezas e imprecisões, conhecido como ANFIS.

Apresenta a capacidade de aprendizagem que permite extrair as informações contidas em um conjunto de observações conhecidos como pares de entrada/saída, denominada como uma abordagem do tipo caixa preta, devido aos resultados não serem interpretáveis.

O ANFIS utiliza a metodologia proposta por Takagi e Sugeno, podendo ser dividida em até 5 camadas, conforme a Figura 1 (CUEVAS et al., 2017).

Figura 1 - Arquitetura padrão de um sistema ANFIS.



Fonte**:** Adaptada de: Cueva et al, (2017).

As camadas podem ser divididas em lineares e não lineares, a lógica fuzzy é responsável por dividir estes dois blocos, sendo às camadas anteriores a ela não lineares e após lineares (CUEVAS et al., 2017).

Camada 1: Corresponde a entrada dos valores. Estes definidos de acordo com o conjunto de variáveis de entrada (pares de entrada) estabelecido para o modelo que será submetido ao ANFIS. Esta camada é formada pela função de pertinência Ai, sendo Oi1 a função Ai, que especifica o grau de pertinência dos valores para satisfazer os termos linguísticos.

$$0_i^{\overline{1}} = u_{Ai}(x) \tag{2}$$

Camada 2: Composta pelo operador lógico fuzzy, que atribui graus de pertinência a cada uma das variáveis, que são descritas por “variáveis linguísticos”.

$$w^1 = o_i^1 \cdot O_i^2 \tag{3}$$

Camada 3: Calcula o grau de ativação normalizado de cada nó por meio da razão entre a aplicabilidade da i-ésima regra (wi) e o somatório das aplicabilidades.

$$O_i^3 = \overline{w}_i = \frac{wi}{\Sigma_J^H = 1} \omega_J \tag{4}$$

Camada 4: Multiplica a saída da 3ª camada com uma função fi, estabelecendo uma combinação linear para cada valor existente dentro do conjunto de pares de entrada. Os valores ($piqiri$) representam os parâmetros ajustáveis que determinam a resposta da função.

$$O_i^4 = \overline{w}_i f_i = \overline{w}(\rho_i u_1 + q_i u_2 + r_i) \tag{5}$$

Camada 5: Calcula a soma de todos os nós da camada anterior.

$$O_i^5 = \sum_i \overline{w} f_i = \frac{\sum_i \overline{w} f}{\sum_i w} \tag{6}$$

Podemos dizer que a aprendizagem do sistema ANFIS acontece através dos dois conjuntos de parâmetros, sendo estes envolvidos diretamente na fase de treinamento. Os parâmetros conhecidos como antecedentes, são tratados por constantes que caracterizam as funções de pertinências, e os parâmetros dos consequentes, são definidos com parâmetros lineares da saída do modelo de inferência.

Desta maneira, a aprendizagem do modelo ANFIS faz uso dos algoritmos do gradiente descendente para otimizar os parâmetros do antecedente e o algoritmo de mínimos quadrados de modo a determinar os parâmetros lineares dos consequentes, sendo assim, o aprendizado é realizado em duas etapas.

A possibilidade de interação entre as camadas existentes em um ANFIS, possibilita dizer que sistema de inferência apresenta as mesmas vantagens se comparado aos sistemas criados por especialistas (SAMER et al., 2021).

Sua utilização pode ser aplicada de duas formas distintas, ambas relacionadas à maneira como o sistema é treinado, o qual pode ser do tipo de grade conhecida em inglês como *grid*

partition, ou agrupamento subtrativo clusterizado conhecida em inglês como s*ub. clustering* (JANG, 1993).

Ao adotar a primeira forma, o sistema de inferência estabelecer uma divisão em relação aos dados por ele utilizados. Esta divisão é definida de maneira uniforme, estabelecendo o número e o tipo de função de pertinência para o conjunto de valores, como também, quais as regras serão envolvidas em cada possível combinação de entrada.

No entanto a segunda maneira, requer o uso de uma relação entre os dados de entrada e os saída utilizados pelo sistema. Está relação tem por objetivo criar os agrupamentos que deverão indicar as características pertencente ao sistema, permitindo assim identificar o chamado o centro ou centro de um sistema de regras fuzzy, posteriormente conhecido como Sugeno (LIMA, 2010).

5. METODOLOGIA E TRATAMENTO DOS DADOS

5.1 Área de estudo

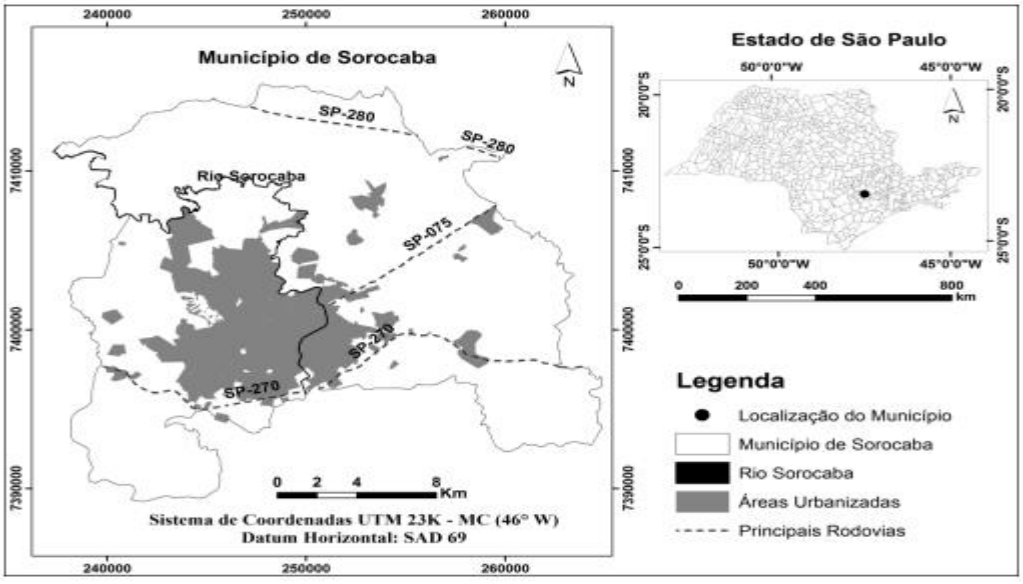
Pertencente ao interior do estado de São Paulo, a cidade de Sorocaba, destaca-se por ser a quarta maior população do interior paulista, e possui a maior população da região sudeste do estado, com aproximadamente 658 mil habitantes segundo o (IBGE, 2020), com área territorial de 450,38 km², encontra-se localizada a 87 km de distância da capital do Estado.

Apresenta um clima subtropical Cwa, segundo a classificação climática de Köppen, com variações de temperaturas classificadas como quentes próximas aos 37ºC no verão e 22ºC no inverno, bem como precipitações anuais próximas a 1297 mm (milímetros) (IBGE, 2020).

O relevo é classificado como ondulado, caracterizado pela presença de vertentes e altos de serra, altitude média de 632 metros em relação ao nível do mar. A maior altitude é de 1.028 metros, nas cabeceiras do rio Pirajibu, na Serra de São Francisco, próximo a Alumínio, a menor altitude 539 metros está no vale no Rio Sorocaba.

A vegetação natural original é formada pela mata atlântica, com locais de floresta ombrófila densa de montanha, composta por vegetação de cerrado e secundária em vários estágios de sucessão. A Figura 2 apresenta o Mapa da localização do município de Sorocaba no estado de São Paulo.

Figura 2 - Localização do município de Sorocaba no estado de São Paulo.



Fonte: Adaptado de: Lourenço et al., (2014).

A escolha da cidade de Sorocaba como área de estudo ocorreu com base na variação do número de risco de fogo apresentado nos últimos anos, em especial para o período definido para este trabalho.

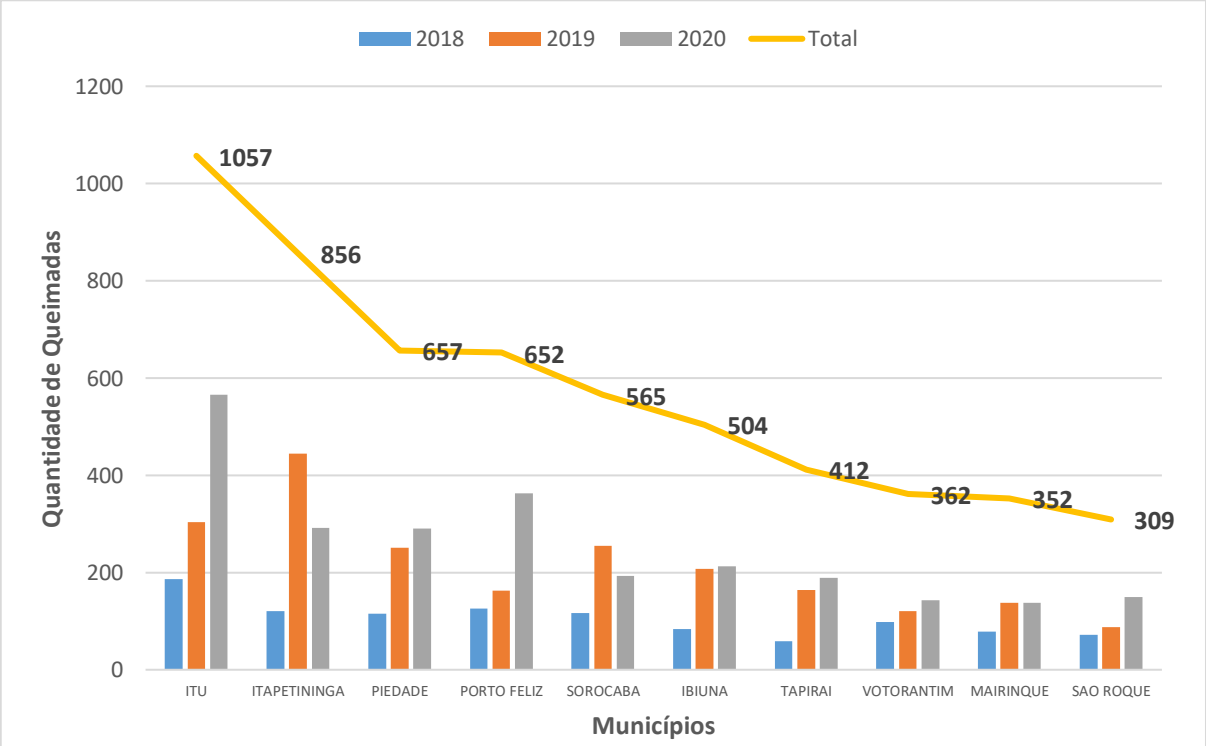
5.2 Desenvolvimento do Modelo

Foram considerados a utilização do histórico de dados pluviométricos combinados com dados espaciais referentes ao registro da presença do risco de fogo nas localidades relacionadas a área de estudo. Para atender as condições em relação aos dados pluviométricos, a presença da estação de captação de dados pluviométricos (chuvas), EDEN(Pirajibu) localizada no bairro do Eden, satisfez as necessidades que o estudo apresentava.

No que se refere ao histórico da presença de fogo, a cidade de Sorocaba apresenta uma variação interessante de focos de queimadas por ano, dentre os 27 municípios que forma a região metropolitana de Sorocaba, o município aparece na quinta posição em relação ao total de queimadas ocorridas entre anos de 2018, 2019 e 2020.

A Figura 3 destaca os 10 munícipios que mais apresentaram focos de queimadas para o respectivo período de anos.

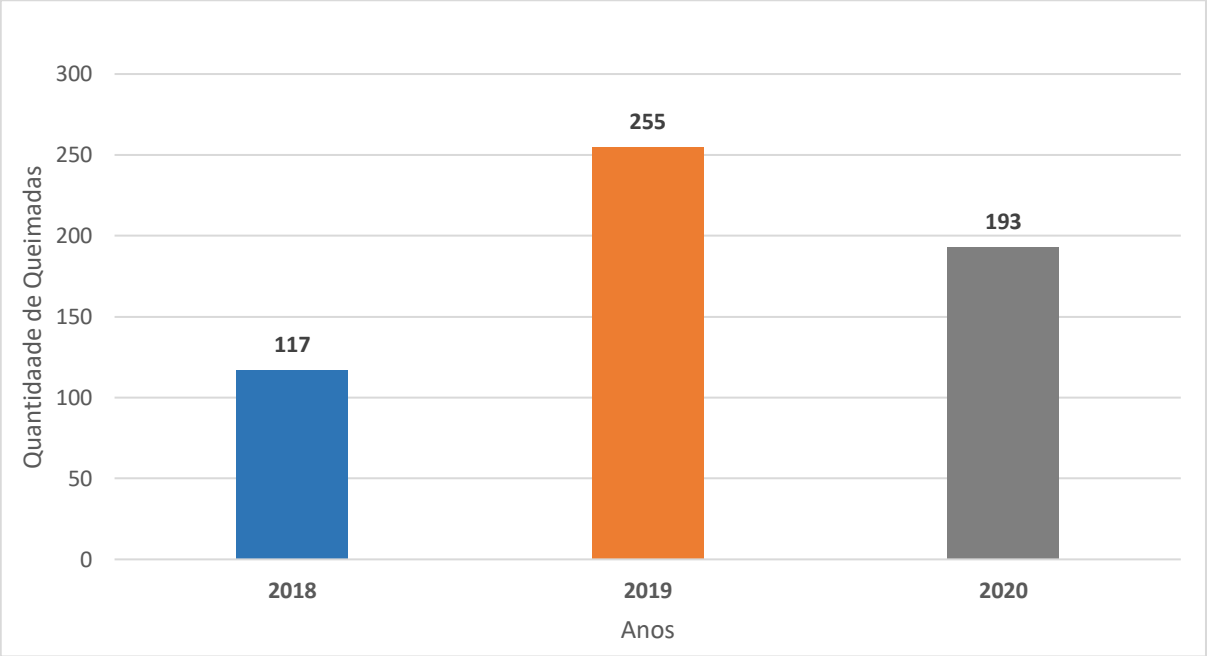
Figura 3 – Os 10 municípios com as maiores quantidades de queimadas.



Fonte: Elaboração própria.

A partir da visualização dos dados apresentados na Figura 3, observamos que a cidade de Sorocaba apresentou um total de 565 focos de queimadas entre os anos de 2018 e 2020. A Figura 4 apresenta a distribuição de queimadas para cada ano em relação a cidade de Sorocaba.

Figura 4 – Distribuição de queimadas por ano ocorridas no Município de Sorocaba.



Fonte: Elaboração própria.

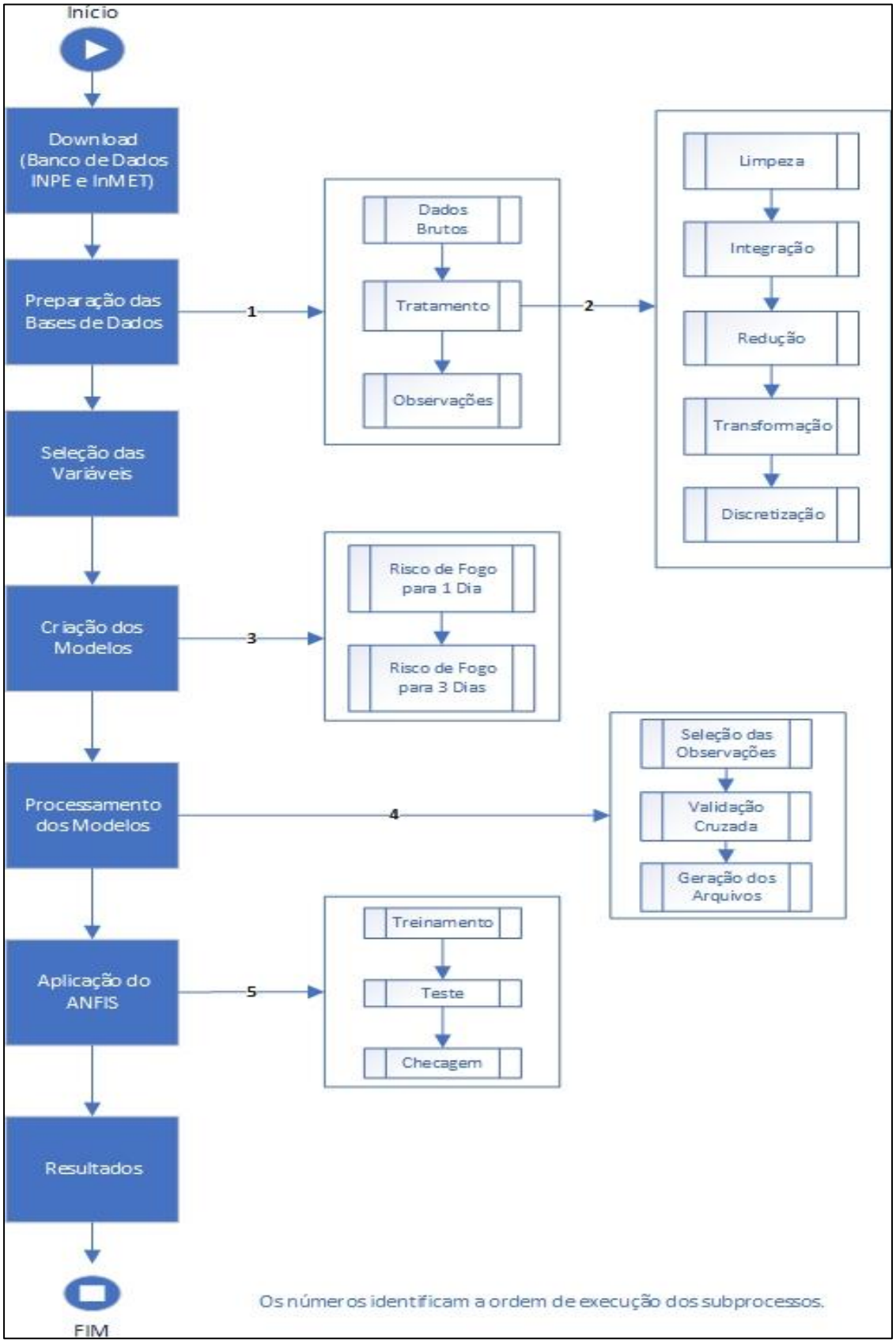
A relação completa do total de queimadas calculada para todos os 27 municípios que constituem a RMS, pode ser consultada nos Apêndices deste trabalho.

5.3 Construção do Índice de Classificação de Níveis de Riscos de Queimada no Brasil

O Índice de classificação de níveis de risco de queimadas no Brasil (ICRQB) foi desenvolvido, com base na análise da ocorrência do risco de fogo observado e precipitações relacionadas à cidade de Sorocaba, localizada no interior do estado de São Paulo. Através da combinação dos dados do Instituto Nacional de Pesquisas Especiais combinados com os dados do Instituto Nacional de Meteorologia.

Esta combinação permitiu estabelecer as condições necessárias na definição da forma de trabalho que permitisse a criação do índice. A Figura 5 destaca a sequência de processos e subprocessos relacionados à construção do ICRQB.

Figura 5 - Fluxograma – Construção do ICRQB.



Fonte: Elaboração própria.

5.4 Seleção das variáveis

A partir da compreensão dos elementos que formam a estrutura dos índices apresentados anteriormente, com base nos estudos destacados anteriormente no quadro 1, foi possível definir as algumas das variáveis de entrada selecionadas para este estudo, dentre elas: Dia sem chuva e Precipitação.

Em adicional, foi definida como a variável de saída o Risco de Fogo Observado, elemento presente no banco de dados de queimadas disponibilizado pelo INPE, como também, utilizada em diversos estudos relacionados a presença de queimadas ou risco de fogo.

Ao consideramos alguns estudos realizados por (LEMOS, 2012), a utilização das variáveis de entrada dia sem chuva e precipitação, combinadas com a variável risco de fogo observado, tornam-se elementos obrigatórios para a construção de um estudo voltado a presença da alta radiação provocada pelo fogo, sendo esta considerada queimada ou risco de fogo.

No entanto, a utilização de somente estas variáveis fosse realizada, poderíamos nos deparar com o resultado possivelmente enviesado, ou seja, algo não concreto, mediante a existência de valores nas preestabelecidas variáveis que retratam somente a ocorrência da queimada e precipitação ocorrida na área de estudo. Em adicional, a ausência da não ocorrência destes fenômenos não estavam identificadas.

Desta forma, tornou-se necessário estabelecer quais poderiam ser as outras variáveis de entrada, combinadas com as já definidas, tornando o processo de análise dos valores existentes nas bases de dados mais confiável.

Sendo assim, ao longo das pesquisas, análises e simulações realizadas novas variáveis foram reconhecidas, justamente para atender a questões relacionadas a não ocorrência de queimadas e precipitações, sendo elas presença da chuva e presença da queimada.

A escolha destas variáveis relacionadas à ocorrência ou não tanto da precipitação (chuva) como também das queimadas, estão relacionadas diretamente com os dados utilizados neste estudo.

As bases de dados não apresentavam em suas observações dados relacionados a não ocorrência destes dois fenômenos naturais, ou seja, existiam somente as observações disponíveis em cada uma delas afirmava que evidenciaram uma queimada ou precipitação (chuva) ocorrida na cidade de Sorocaba para cada dia entre os anos de 2018, 2019 e 2020.

A criação destas variáveis foi condicionada a necessidade ofertar aos modelos implementados para estudo aqui proposto uma forma de auxílio para garantir maior

assertividade na definição do resultado a ser obtido, em consideração todas as possibilidades para cada um dos dias em análise, sabendo que para a mesma data poderíamos ter a presença dos dois fenômenos, somente um deles ou nenhum deles identificados.

Outra variável importante também foi incorporada ao estudo, denominada: acumulado da precipitação para três dias anteriores. Sua utilização está intrinsecamente relacionada à forma utilização dos índices Fórmula de Monte Alegre e Hidrotermométrico. Ambos apresentam características similares ao ICRQB, utilizam dados cumulativos e são aplicados ao território brasileiro. A Tabela 3 descreve as características das variáveis selecionadas para posterior utilização aos modelos elaborados para este trabalho.

Tabela 3 - Descrição das variáveis selecionadas

Nome da Variável	**Descrição**	**Faixa de Valores**	**Unidade**	**Tipo**
DiaSemChuva	Representa o número de dias sem chuva.	0 até 120	Dias	Entrada
PrecipitaçãodoDiaAnterior	Representa o valor da precipitação do dia anterior.	0 até 120	Milímetros	Entrada
AcumuladoPrecipitaçãoTrêsDiasAnteriores	Representa o valor total acumulado de precipitação para o período de três dias anteriores.	0 até 120	Milímetros	Entrada
PresençaChuva	Representa a identificação da presença ou não da chuva.	0 ou 1	Verdadeiro/Falso	Entrada
PresençaQueimada	Representa a identificação da ocorrência ou não de uma queimada.	0 ou 1	Verdadeiro/Falso	Entrada
RiscoFogoObservado	Representa o valor relacionado ao Risco de Fogo Observado para 1 ou 3 dias.	0.1 até 1	Decimais	Saída

Fonte: Elaboração própria.

5.5 Base de dados

Os dados apresentados neste trabalho, advém do banco de queimadas do INPE, disponibilizado pelo seu Programa Queimadas, para a cidade de Sorocaba de acordo com o período dos dias entre 01/01/2018 e 31/12/2020.

A forma de obtenção dos dados para ambos os bancos de dados, pode ser feita diretamente através do download dos arquivos, de acordo com período entre 30 ou 365 dias, preferencialmente no formato de valores separados por vírgula.

A Tabela 4 descreve a representação de cada variável existente no banco de dados de queimadas do INPE.

Tabela 4 - Descrição das variáveis Banco de Dados de Queimadas do INPE

Nome da Variável	Descrição
Datahora	Apresenta o horário de referência da passagem do satélite segundo o fuso horário de Greenwich (GMT);
Satelite	Identifica o nome do algoritmo utilizado e referência ao satélite provedor da imagem;
Pais / Estado / Município	Entidades político-geográficas onde foram identificadas a presença de focos de incêndio no território brasileiro;
Bioma	Nome do Bioma segundo a referência do Instituto Brasileiro de Geografia e Estatísticas (IBGE);
DiaSemChuva	Número de dias sem chuva até a detecção do foco de calor, os quais são mensurados dentro da faixa de valores entre 0 (zero) e 120 (cento e vinte);
Precipitação	Valor da precipitação acumulada no dia até o momento da detecção do foco de calor; os quais são mensurados dentro da faixa de valores entre 0 (zero) e 120 (cento e vinte);
RiscoFogoObservado	Representa o valor correspondente ao Risco de Fogo Observador para o dia da detecção do foco de calor, os quais são identificados dentro da faixa de valores entre 0.1 (zero ponto um) e 1 (um);
Latitude / Longitude	Apresenta as coordenadas geográficas do centro do pixel que contém um possível foco de incêndio ou presença de calor acima de 47º C.

Fonte: INPE – Programa Queimadas.

Em conjunto com os dados de queimadas do INPE, utilizou-se também o banco de dados de precipitações diárias mantidas pelo INMET, coletados pela estação pluviométrica

Pirajibu, localizada no bairro do Eden pertencente a cidade de Sorocaba. A Tabela 5 descreve as variáveis existentes no banco de dados do INMET, selecionadas para o estudo.

Tabela 5 - Descrição das variáveis Banco de Dados de Precipitações

Nome da Variável	Descrição
Dias de Chuvas	Total de Dias de chuvas ocorridos para o mês referente a data da precipitação.
Total Precipitação Mês	Total do valor da Precipitação Acumulada calculada ao longo do mês de acordo com a data da precipitação.
Precipitação Do Dia	Valor referente ao volume de precipitação identificado para a data da precipitação.

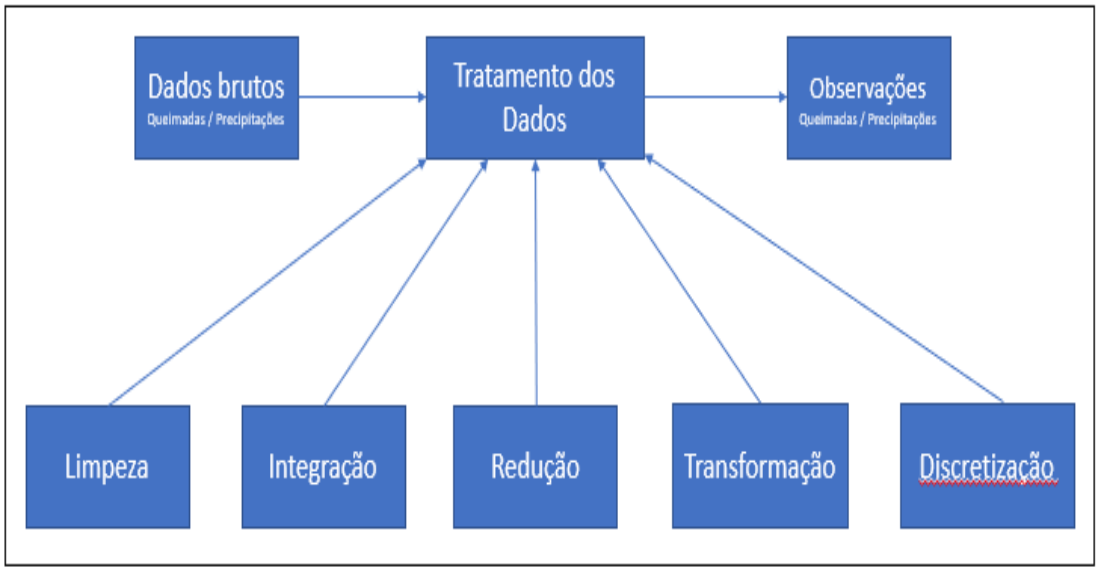
Fonte: INMET.

5.6 Preparação da Base de Dados

Sabendo da necessidade de identificar somente os dados de queimadas relacionados para a cidade de Sorocaba armazenados no banco de dados do INPE, como também, combinar os dados de precipitação existentes no banco de dados do INMET, foi necessário realizar um processo de preparação das respectivas bases de dados para posterior utilização.

A Figura 6 traz uma visão resumida do processo de preparação das bases de dados, um breve resumo sobre cada etapa está descrito na sequência.

Figura 6 - Etapas do processo de preparação das bases de dados.



Fonte: Elaboração própria.

- Dados brutos: Os dados existentes no banco de dados de queimados do INPE, apresentaram registros para os focos de queimadas relacionados a todos os estados e municípios do território brasileiro, totalizando para o período entre os anos de 2018 e 2020 mais de 11 milhões.

  Tornou-se necessário realizar uma seleção dos dados, estabelecendo um filtro para que fosse identificado somente aqueles relacionados a área de estudo, resultando no valor de aproximadamente de 560 linhas de dados selecionadas.

- Tratamento dos Dados:
  - Limpeza: para entrada de valores ausentes, remoção de ruídos e correção de inconsistências;
  - Integração: para unir os dados relacionados às queimadas e precipitações, formando uma fonte de dados única. Nesta tarefa foram consideradas também a não ocorrência de queimadas e precipitações;
  - Redução: para reduzir o conjunto de dados a ser analisado, considerando por exemplo a ocorrência da queimada ou precipitação de acordo com a data e não a data e hora;
  - Transformação: para estabelecer uma padronização dos dados, estabelecendo um formato específico para o estudo; e
  - Discretização: para redução da quantidade de valores contínuos existente em algumas variáveis.

- Observações: Resultado do tratamento de dados. A partir deste resultado, temos o conjunto de observações de 1286 identificadas, as quais estão totalmente preparadas para serem aplicadas aos modelos criados para o estudo.

Após a do processo de preparação das bases de dados, foi possível identificar os dados relacionados as queimadas e seus respectivos risco de fogo, como também, os dados referentes às precipitações diárias para os anos de 2018, 2019 e 2020.

5.7 Validação Cruzada

A validação cruzada é uma técnica para avaliar a capacidade de generalização de um modelo, a partir de um conjunto de dados, amplamente empregada em problemas onde o

objetivo da modelagem é a predição. Tem como objetivo estimar o quão preciso é este modelo na prática, ou seja, o seu desempenho e assertividade para um novo conjunto de dados (LING, 2019).

Uma das maneiras de fazer a divisão desses dados é usando o método *holdout*, o qual consiste em dividir os dados em 70-30 de maneira aleatória, ou seja, 70% dos dados para a fase de treinamento e 30% para a fase de teste ou checagem.

A desvantagem de usar essa técnica, está condicionada a possibilidade de selecionar uma porção de dados para treino, teste ou checagem que são muito parecidas tendo uma boa avaliação do modelo. Nesse caso, porém, quando colocamos o modelo em produção os dados novos são muito diferentes dos dados já conhecidos pelo modelo, o que nos gera péssimos resultados, praticamente tornando inutilizável este conjunto de dados (WONG, 2017).

Outra possibilidade para se evitar problemas relacionados a aleatoriedade dos dados, consiste na criação de grupos, estes formados por conjuntos de dados, como forma de divisão. Esta alternativa surge como técnica que visa eliminar a aleatoriedade, reduzir a variância, afim de estabelecer um resultado mais robusto. A única desvantagem apresentada por esta técnica, está diretamente relacionada a possíveis impactos no que se refere ao tempo de processamento, mais especificamente ao desempenho (LING, 2019).

Na possibilidade de a base de dados selecionada apresentar uma grande quantidade (milhões) de observações a serem analisadas, o tempo de processamento demandado pela validação cruzada poderá se tornar um pouco mais demorado, proporcionando um possível impacto ou demora no que se refere a apresentação dos resultados.

Para este trabalho, foi utilizada a técnica de validação cruzada k-*fold*, onde temos um número de k (*folds = 5*), ou seja, 5 (cinco) que representa a quantidade de agrupamentos de dados criados, considerando 4 grupos combinados que formam um único grupo selecionado para a fase de treinamento de dados, e grupo restante para a fase de checagem.

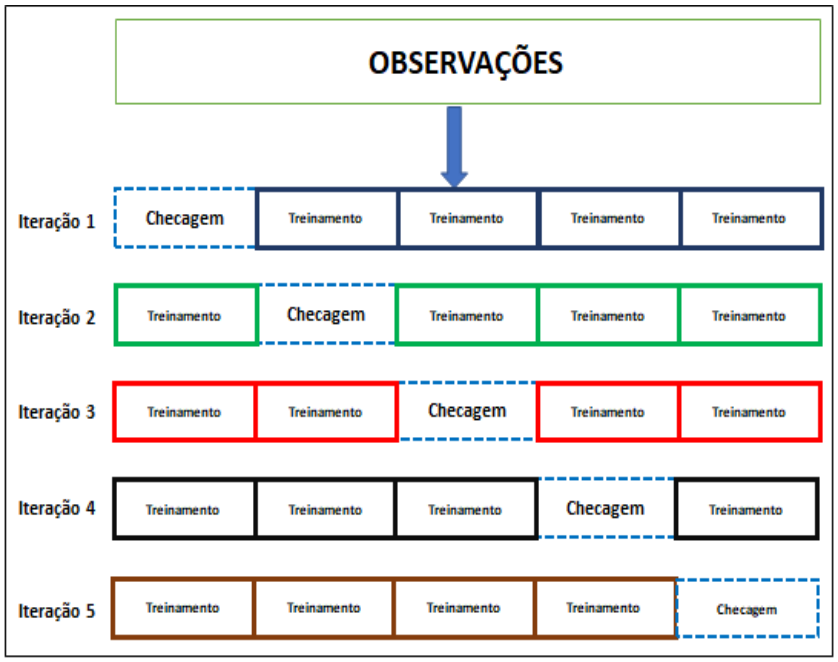
Do total de 1286 observações identificadas na fase de preparação dos dados, 900 foram distribuídas em 5 (cinco) grupos compostos obrigatoriamente por 180 observações, as quais foram selecionadas de forma aleatória, não se repetindo observações já selecionadas em outros grupos no momento da composição do grupo atual.

A definição da quantidade de números de iterações a serem utilizadas, foi estabelecida tendo como base a quantidade de *folds*, neste caso 5-*folds*, com objetivo de permitir o processamento, análise e apresentação dos resultados.

Para posterior comparação entre os valores obtidos referentes aos grupos de treinamentos, para com os valores de cada grupo de checagem, realizando está comparação ao

término de cada iteração. A Figura 7 ilustra o conjunto de grupos de treinamentos e checagens definidos para serem submetidos ao uso da validação cruzada 5-*folds*.

Figura 7 - Validação Cruzada 5-folds.



Fonte**:** Elaboração própria.

Vale ressaltar, que cada grupo foi nomeado como *TrainingGroup* (grupo de treinamento), acrescentando-se uma combinação numérica que representa a junção dos grupos que formam o respectivo grupo de treinamento. A Figura 8 apresenta um exemplo desta nomenclatura.

Figura 8 – Nomenclatura – TrainingGroup.



Fonte: Elaboração própria.

Podemos dizer que o TrainingGroup2345 representa a junção dos grupos de treinamento: 2, 3, 4 e 5 totalizando o número 720 observações. A Tabela 6 descreve todos os grupos de treinamentos criados para serem aplicados na validação cruzada.

Tabela 6 - Grupos de Treinamento

Nome do Grupo	Número dos Grupos Combinados	Número do Grupo Selecionado para Checagem
TrainingGroup2345	2, 3, 4 e 5	1
TrainingGroup1345	1, 3, 4 e 5	2
TrainingGroup1245	1, 2, 4 e 5	3
TrainingGroup1234	1, 2, 3 e 5	4
TrainingGroup1235	1, 2, 3 e 4	5

Fonte: Elaboração própria.

5.8 Modelos para Predição do Risco de Fogo

Como todo estudo, diversas suposições e análises foram realizadas até se encontrar a melhor maneira de implementação dos modelos, a fim de conseguir obter o resultado mais robusto, nem sempre o mais preciso.

Após a preparação da base de dados e seleção das variáveis pertinentes ao resultado a ser atingido pelo estudo, vários modelos foram elaborados, treinados e testados, para conseguir identificar os mais assertivos no que se refere a média de erros absolutos, como também, no tempo de processamento em relação a quantidade de observações selecionadas.

Chegou-se a definição da utilização de 6 modelos distintos, com o propósito de predizer a ocorrência do risco de fogo observado para 1 (um) ou 3(dias), que apresentaram os melhores valores de médias de erros ao longo das fases de treinamento, teste e checagem, em conjunto com o um tempo de processamento para obtenção dos resultados.

Estes modelos foram nomeados através de uma sequência numérica, como por exemplo: Modelo 1, Modelo 2, Modelo 3 e assim respectivamente até o Modelo 6.

Os modelos 1, 2 e 3 tem a finalidade de apresentar como resultado a predição do Risco de Fogo para 1 dia, por outro lado os modelos: 4, 5 e 6 estão designados a resultar a predição para 3 dias. De maneira prática, os modelos apresentam diferenças em suas estruturas, ou seja, não possuem a mesma quantidade de variáveis de entrada.

Esta diferença foi estabelecida com intuito de identificar as diferenças no comportamento de cada modelo, em adicional com variáveis nos possíveis resultados.

A igualdade existente em cada modelo se refere a variável de saída, todos possuem a mesma quantidade, mais precisamente uma única variável. As Tabelas 7, 8, 9, 10, 11 e 12, descrevem a estrutura de cada modelo.

Tabela 7 – Descrição das variáveis de entrada e saída selecionadas no Modelo 1

Nome da Variável	Faixa de Valores	Representa	Tipo
DiaSemChuva	0 até 120	Dias	Entrada
PrecipitaçãoDiaAnterior	0 até 120	Milímetros	Entrada
AcumuladoPrecipitaçãoTrêsDiasAnteriores	0 até 120	Milímetros	Entrada
PresençaChuva	0 e 1	Verdade/Falso	Entrada
PresençaQueimada	0 e 1	Verdade/Falso	Entrada
RiscoFogoObservadaParaUmDia	0,1 até 1	Nível do Risco	Saída

Fonte: Elaboração própria.

Tabela 8 - Descrição das variáveis de entrada e saída selecionadas no Modelo 2

Nome da Variável	Faixa de Valores	Indicador	Tipo
DiaSemChuva	0 até 120	Dias	Entrada
PrecipitaçãoDiaAnterior	0 até 120	Milímetros	Entrada
AcumuladoPrecipitaçãoTrêsDiasAnteriores	0 até 120	Milímetros	Entrada
RiscoFogoObservadaParaUmDia	0,1 até 1	Nível do Risco	Saída

Fonte: Elaboração própria.

Tabela 9 - Descrição das variáveis de entrada e saída selecionadas no Modelo 3

Nome da Variável	Faixa de Valores	Representa	Tipo
DiaSemChuva	0 até 120	Dias	Entrada
PrecipitaçãoDiaAnterior	0 até 120	Milímetros	Entrada
RiscoFogoObservadaParaUmDia	0,1 até 1	Nível do Risco	Saída

Fonte: Elaboração própria.

Tabela 10 - Descrição das variáveis de entrada e saída selecionadas no Modelo 4

Nome da Variável	Faixa de Valores	Representa	Tipo
DiaSemChuva	0 até 120	Dias	Entrada
PrecipitaçãoDiaAnterior	0 até 120	Milímetros	Entrada
AcumuladoPrecipitaçãoTrêsDiasAnteriores	0 até 120	Milímetros	Entrada
PresençaChuva	0 e 1	Verdade/Falso	Entrada
PresençaQueimada	0 e 1	Verdade/Falso	Entrada
RiscoFogoObservadaParaTrêsDia	0,1 até 1	Nível do Risco	Saída

Fonte: Elaboração própria.

Tabela 11 - Descrição das variáveis de entrada e saída selecionadas no Modelo 5

Nome da Variável	Faixa de Valores	Representa	Tipo
DiaSemChuva	0 até 120	Dias	Entrada
PrecipitaçãoDiaAnterior	0 até 120	Milímetros	Entrada
AcumuladoPrecipitaçãoTrêsDiasAnteriores	0 até 120	Milímetros	Entrada
RiscoFogoObservadaParaTrêsDia	0,1 até 1	Nível do Risco	Saída

Fonte: Elaboração própria.

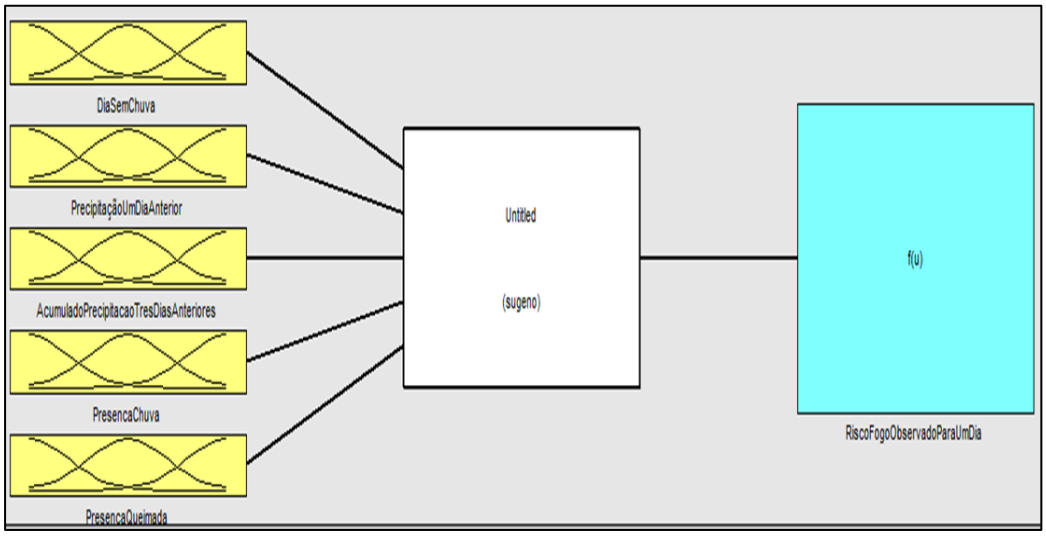
Tabela 12 - Descrição das variáveis de entrada e saída selecionadas no Modelo 6

Nome da Variável	Faixa de Valores	Representa	Tipo
DiaSemChuva	0 até 120	Dias	Entrada
PrecipitaçãoDiaAnterior	0 até 120	Milímetros	Entrada
RiscoFogoObservadaParaTrêsDia	0,1 até 1	Nível do Risco	Saída

Fonte: Elaboração própria.

A Figura 9 exibe a estrutura definida pelo modelo híbrido ANFIS para os modelos 1 e 4, através da ferramenta *AnfisEdit* existente no software MatLab. Estes modelos apresentam um conjunto de 5 (cinco) variáveis de entrada e 1 (uma) variável de saída, considerado de elaboração mais complexa se comparado com os demais modelos.

Figura 9 - Estrutura criada para os Modelos 1 e 4.



Fonte: Elaboração própria.

5.9 Dados Brutos – Extração, Transformação e Carga

Uma das principais atividades relacionadas a manipulação de grandes volumes de dados, a ações de Extração, Transformação e Carga de dados, conhecidas pela sigla em inglês ETL (*Extract, Transformation and Load*), não poderia deixar de ser aplicada. Ainda mais para um estudo que apresenta uma volumetria considerável de milhões de linhas dados ao logo dos anos de monitoramento realizados pelo INPE e INMET.

Como destacado anteriormente, através da opção exportar dados disponível no programa de queimadas, de forma indireta estamos fazendo uso das operações de ETL, realizando a extração do período de dados, para que o INPE faça o processo de transformar este conteúdo em um arquivo com extensão *.csv*, posteriormente o mesmo conteúdo armazenado no arquivo e liberado para download.

A partir do momento em que realizamos o download deste arquivo, estamos iniciando o processo de ETL de forma direta, processamos a extração desta massa de dados, e agora temos a possibilidade de realizar as devidas transformações (*tratamentos*) aplicadas ao conjunto de dados selecionado.

Por meio dos tratamentos a serem realizados, poderemos permitir que os conjuntos de dados brutos possam se transformar em dados reconhecíveis, ou seja, interpretáveis, compreensíveis, desta forma, estaremos criando uma fonte de conhecimento disponível para ser estudada.

Por fim, as ações relacionadas a carga de dados para uma ferramenta computacional, seja ela criada pelos mais diversos fornecedores de software, consiste na passagem dos dados que estão sendo extraídos, tratados e agora armazenados em um novo local lógico, como por exemplo um: arquivo texto ou arquivo de banco de dados.

Inicialmente foram utilizados recursos computacionais tradicionais, como o *Microsoft Excel* e o editor de texto *Notepad* para o acesso e leitura dos respectivos arquivos, mas que infelizmente não se mostraram ferramentas capazes de manipular volumes de dados acima de 1.600.000 (um milhão e seiscentas mil linhas de dados). Outra ferramenta analisada, foi ambiente de desenvolvimento integrado *R-Studio*, mas também apresentou algumas limitações técnicas, tais como, a limitação para se realizar a leitura e carregamento de dados em memória ou para um *dataframe* acima do valor de 8 *gigabytes*.

Desta forma, tornou-se necessário a adoção de um ambiente com maior poder de processamento computacional relacionado as questões de leitura e gravação de dados, para atender estes necessidades e permitir aplicar os procedimentos necessários para se alcançar os

objetivos deste trabalho. O Sistema Gerenciador de Banco de Dados (SGBD) *Microsoft SQL Server* desenvolvido pela Microsoft, foi definido como o ambiente computacional, o qual deverá desempenhar as funções de armazenar, processar e ofertar os dados a serem utilizados pelo modelo híbrido posteriormente implementado neste trabalho.

5.10 Tratamento dos Dados

O Microsoft SQL Server, surge como a opção adequada ao uso do ETL, apresenta a capacidade adotada no acesso a diversas com fontes de dados das mais variadas origens, tamanhos e formatos, possibilitando para usuários dos mais diversos níveis de conhecimento e experiência com ferramentas ou ambientes de bancos de dados realizar o acesso, leitura e manipulação de dados.

Para o escopo de ferramentas utilizadas para este trabalho, foram escolhidas a utilização do *Microsoft SQL Server 2019 Express Edition* para o ambiente de banco de dados, em conjunto com o *Microsoft SQL Server Management Studio* para o ambiente de desenvolvimento e ETL.

Sabendo quais as ferramentas vamos utilizar, em para qual cenário estão sendo aplicadas, podemos começar a colocar em prática a técnica de ETL, realizar o acesso ao arquivo de dados selecionado como fonte de demonstração denominado: Fonte de Dados – Queimadas – 2018.csv, composto em seu conteúdo por mais de 1.800.000 (um milhão e oitocentos mil linhas) de focos de incêndio queimadas identificadas.

Antes de avançarmos, gostaria de enfatizar que todas ações e configurações aplicadas a técnica de *ETL*, podem ser realizadas diretamente na interface de desenvolvimento por linha de comando existente no Management Studio.

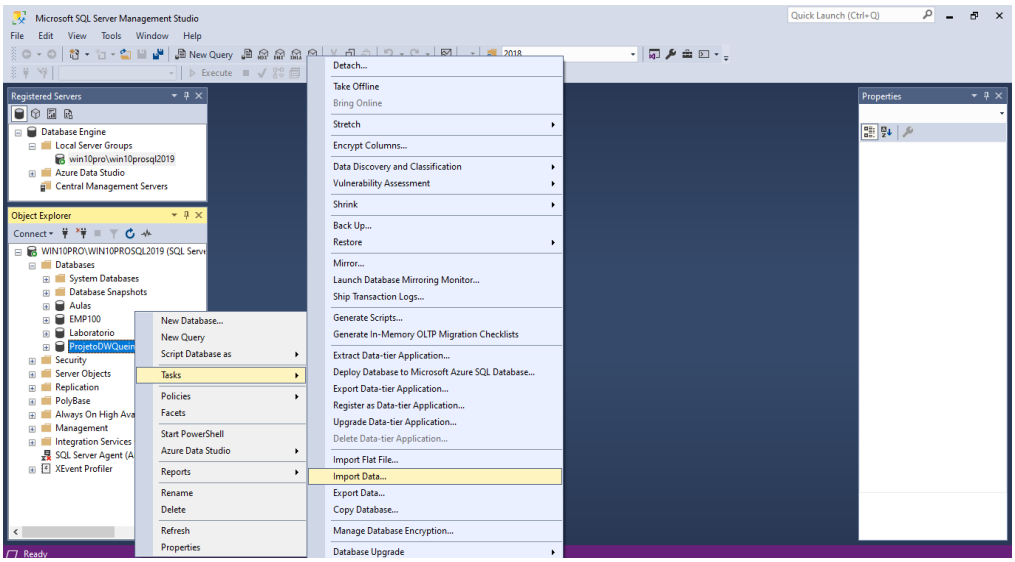
5.11 Limpeza e Integração

Começamos nossa operação de ETL, pela ação de extração (leitura) dos dados contidos no arquivo fonte de dados – queimadas – 2018, este selecionado para a demonstração. dentro do ambiente de desenvolvimento Microsoft SQL Server Management Studio.

Será utilizada a opção Tarefas (*Tasks*) em conjunto com a opção Importar Dados (*Import Data)*, ambas disponíveis a partir do momento em que clicamos com o botão da direita do mouse em um dos bancos de dados **ModeloHibrido.** Sendo este ambiente, criado exclusivamente à ser utilizado como ambiente de banco de dados aplicado a esta parte do estudo

realizado neste trabalho. A Figura 10 ilustra a opção Import Data existente no Management Studio através menu Tasks, obtido ao se clicar com o botão da direita do mouse no baco de dados selecionado.

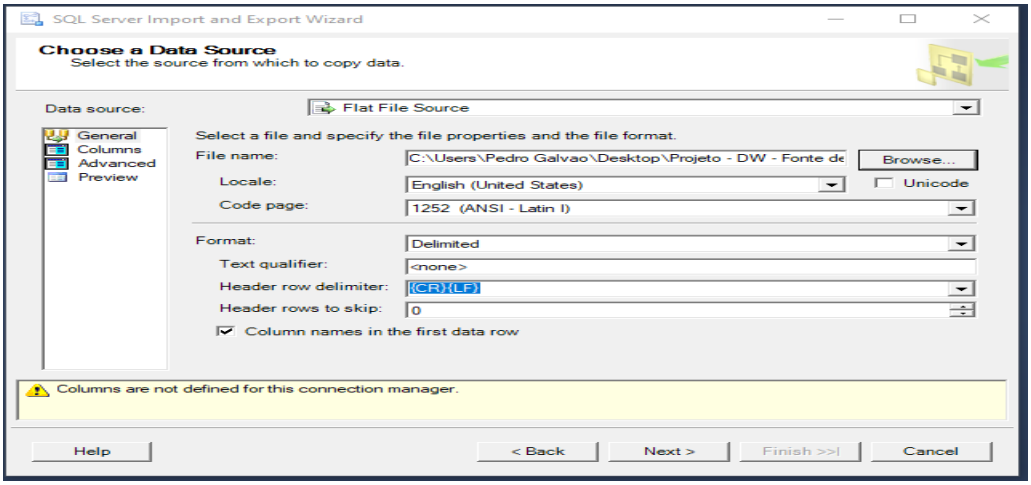
Figura 10 - Menu Tasks – Import Data.



Fonte: Elaboração própria.

Após clicar na opção import data, a ferramenta *SQL Server Import and Export Wizard* será aberta. Nosso próximo passo será definir a origem de dados (*data source*) definindo a mesma como arquivo de origem simples (*Flat File Source)*, conforme a Figura 11.

Figura 11 - Origem de dados.



Fonte: Elaboração própria.

5.12 Redução e Transformação

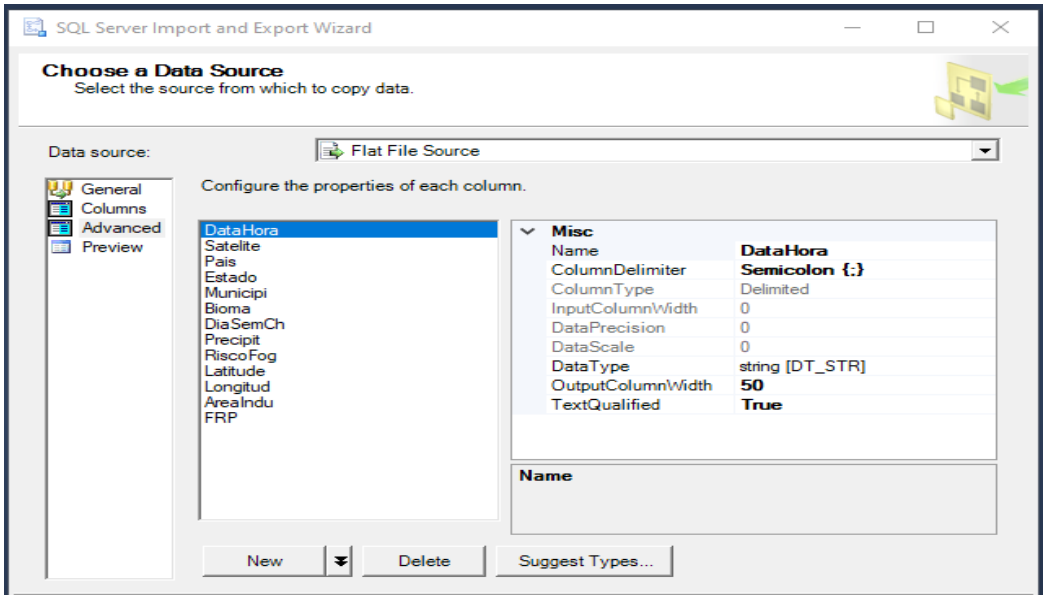
Quando estamos trabalhando com grandes volumes de dados, obtidos de fontes externas, o tratamento ou para muitos a limpeza e padronização, consiste num conjunto de ajustes necessários para que seja possível se estabelecer o formato mais próximo ao que cada dado poderá apresentar de significado ou valor.

Devemos definir o formato de apresentação de um dado relacionado a data e hora, se o mesmo deverá ser apresentado como data, ou hora, bem como, se este dado deverá ser decomposto em outros dados, como por exemplo: “dia, mês, ano, hora, minuto, segundos”, como forma de se criar uma hierarquia de dados.

Enfim o tratamento de dados requer diversos cuidados e atenções afim de resultar um novo dado padronizado da melhor maneira possível. Através da ferramenta importação de dados existente no Microsoft SQL Server, temos a capacidade de aplicar de forma mais ágil e gráfica os procedimentos envolvidos nas possíveis limpezas e tratamentos.

Dentre elas, a partir da utilização da opção Avançada (*Advanced*), o assistente irá nos apresentar a lista de variáveis existentes em nosso arquivo, para cada uma das variáveis teremos a nossa disposição a possibilidade de configurar suas respectivas características, como: Nome, Tipo de Dados, Tamanho, Precisão. A Figura 12 traz as configurações disponíveis relacionadas ao tratamento de dados.

Figura 12 - Tratamento de Dados.



Fonte: Elaboração própria.

5.13 Carregamento de Dados (Observações)

O carregamento de dados representa a ação de leitura dos dados existente contidos diretamente no arquivo texto, posteriormente tratados pela ferramenta de importação de dados, os quais serão armazenados no respectivo banco de dados **ModeloHibrido**. De forma mais especifica uma nova estrutura reconhecida como uma tabela, formada pelas variáveis e seus dados, onde cada variável representa um atributo (coluna), e cada linha de dados como um registro lógico.

Começamos informando o destino de armazenamento para nossos dados, através da tela Escolha um Destino (*Choose a Destination*). Para o cenário implementado neste estudo, será utilizando o servidor de banco de dados criado para os ambientes de testes e homologação, logo após escolher o servidor, devemos selecionar em qual banco de dados iremos depositar a massa de dados que estão sendo carregados neste momento.

A partir da confirmação destas configurações o Microsoft SQL Server, definirá o banco de dados **ModeloHibrido** como nossa estrutura base para o armazenamento e manipulação de dados. Por fim, o nome da tabela que será criada para receber os dados referentes aos focos de queimadas que estamos selecionando para serem analisados deverá ser informado. Nesta demonstração foi escolhido o nome **Queimadas2018** para a nova tabela.

As configurações necessárias ao carregamento dos dados estão definidas, podemos finalizar a ferramenta de importação, deixar para que o próprio Microsoft SQL Server faça o seu papel, ler todos os dados contidos no arquivo fornecido pelo INPE e resultar no armazenamento para a tabela Queimadas2018.

O mesmo procedimento foi realizado para os arquivos contendo o volume diário de precipitações para os meses correspondentes aos anos 2018, 2019 e 2020 disponibilizado pelo INMET.

5.14 Ambiente de Desenvolvimento para Computação Técnica MatLab

O ambiente de desenvolvimento para computação técnica MatLab, foi criado por Cleve Moler, presidente do departamento de ciências da computação na Universidade do Novo México em 1970.

O MATLAB surgiu como uma solução acadêmica com objetivo de ajudar os alunos a trabalhar com processamento e aprendizagem relacionado ao desenvolvimento de código sem

ter a necessidade de utilizar uma linguagem de programação, neste caso a linguagem Fortran, amplamente difundida naquela época.

O nome MATLAB vem de *Matrix Laboratory*. Originalmente desenvolvido para prover um acesso amigável ao tratamento de vetores e matrizes, formado por um conjunto de bibliotecas e funções matemáticas, que permitem auxiliar na geração de gráficos e manipulação de dados. Composto por uma vasta coleção de caixas de ferramentas (toolboxes) aplicadas as áreas específicas como: Equações Diferencias Ordinárias, Estatística, Processamento de Imagens e Sinais, Comunicações, Design de Controle, Teste e Medição, Modelagem e Análise Financeira e Biologia Computacional.

A linguagem e o ambiente de programação MATLAB permitem ainda que o usuário possa escrever suas próprias bibliotecas, assim, o usuário poderá enriquecer a linguagem, incorporando a ela novas funções.

Vale ressaltar que a versão MATLAB R2010b foi à escolhida como ambiente de desenvolvimento para o estudo apresentado neste trabalho. Através da licença para uso acadêmico vinculada a Universidade Estadual Paulista "*Júlio de Mesquita Filho*" – UNESP.

5.15 Implementando o Modelo Híbrido ANFIS para Classificação do Risco de Fogo

Ao se desenvolver um índice que nos auxilie na classificação de níveis de risco de fogo, deve-se levar em consideração possíveis adequações para uma determinada região onde este será aplicado, está é a forma mais coerente de fazer a predição em relação ao histórico de ocorrências das queimadas na localidade (SOARES et al., 2012).

Para a implementação de um ANFIS que permite fazer uso dos dados disponibilizados pelo INPE como forma para classificar os possíveis níveis de risco, foram consideradas as variáveis meteorológicas dia sem chuva e precipitação um dia anterior. Ambas combinadas com a variável indicadora risco de fogo observado, formando o conjunto de entradas e saídas aplicadas ao algoritmo ANFIS, tendo como o objetivo estabelecer a concepção do Índice de Classificação de Riscos de Queimadas do Brasil reconhecido pela sigla ICRQB.

Este índice será alimentado pelas entradas dados referentes aos dias sem chuva identificado pela faixa de valores entre 0 (zero) e 120 (cento e vinte), em conjunto com os valores da precipitação mensurados na unidade de medida em milímetros identificados pela faixa de valores entre 0 (zero) e 100 (cem).

Os valores referentes ao risco de fogo, os quais estão definidos dentro da faixa de valores entre 0(zero) e 1(um), foram calculados e obtidos com base média para o período de 3 dias.

O modelo híbrido selecionado para o estudo foi o configurado para ser processado pelo método de clusterização subtrativo aplicado ao algoritmo ANFIS, considerado como o mais adequado para concepção final do ICRQB, levando-se em consideração as características do volume de dados selecionados, e agilidade para o processamento de séries históricas de dados, que este método apresenta. O clusterização subtrativo, retorna os centros dos clusters computados e estima o número de clusters nos dados de entrada, possibilitando obter resultados de médias de erros absoluto de forma mais rápida.

O mesmo foi implementado através do software MatLab, em conjunto a linguagem M, a quantidade de épocas utilizadas durante a variação dos raios de influência foi igual a 100 (cem).

A saída apresentada pelo ANFIS terá como objetivo evidenciar o quanto este o ICRQB poderá ser capaz de ajudar a classificar os possíveis níveis de risco de fogo que poderá ocorrer em período de dias, mediante a análise e aprendizados que o modelo híbrido concebido pelo ANFIS foi capaz de elaborar.

5.16 Carregando dados no MatLab

O processo de carregamento de dados selecionados para serem aplicados ao modelo híbrido ANFIS, foi realizado através da utilização do comando *Select* (coletar, selecionar e apresentar), sendo este um dos mais conhecidos pelos profissionais da área de banco de dados.

O funcionamento deste comando, bem como, a estrutura de código implementada para gerar a porção de dados que serão armazenados nos arquivos utilizados para o treinamento "*training*" e "checking" (ambos posteriormente requeridos para o funcionamento do ANFIS), não serão abordados neste momento. Vale ressaltar que existem diversas formas de realizar este procedimento, ou seja, coletar os dados desejados e armazenar diretamente em um arquivo texto.

Como forma de facilitar a compreensão deste procedimento, será utilizada a operação de copiar os dados retornados pelo Microsoft SQL Server e depois colar diretamente em um novo arquivo no editor de texto Notepad. Algumas considerações devem ser destacadas para que procedimento realizado manualmente possa resultar nos arquivos que serão posteriormente utilizados no MatLab, dentre as quais destaco:

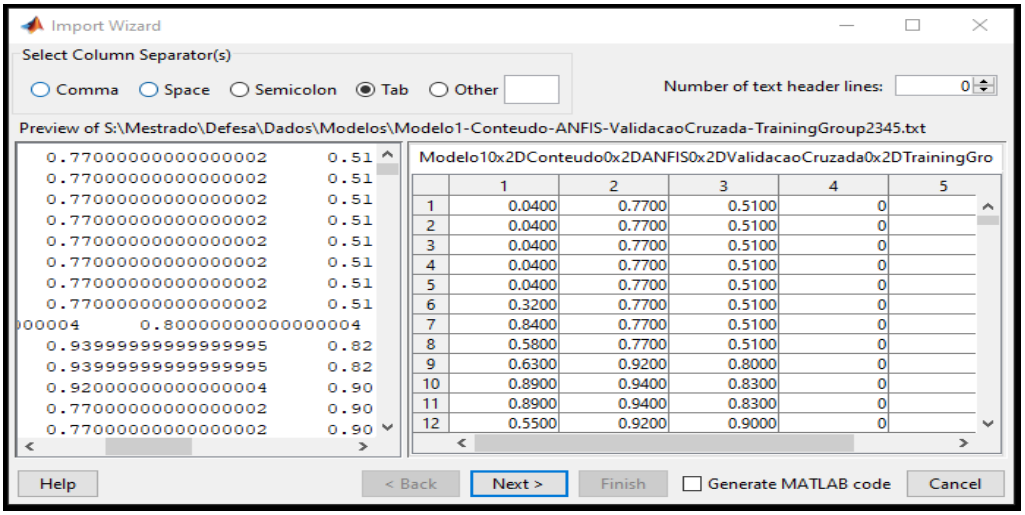
- A massa de dados selecionada para o procedimento de treinamento, deve corresponder a 70% do total de dados selecionados, ou outros 30% são utilizados para o procedimento de checagem ou teste;
- O nome dos arquivos fora definido como training e checking, justamente como forma de evidenciar em qual momento serão utilizados pelo ANFIS;
- A extensão .txt (Text File) comumente utilizada para arquivos em formato de texto simples (estrutura de arquivos que não apresentam formatação ou uso de cores) somente foi definida como maneira de facilitar o uso e obtenção de dados para as mais diversas ferramentas de leitura e gravação de dados no formato texto, isso inclui o Matlab, que tradicionalmente trabalha com a extensão .dat (General Data File), ou simplesmente arquivos de dados;
- Ao salvar ambos os arquivos no editor de texto Notepad, deve-se observar o valor existente na opção Enconding, que por padrão deverá apresentar o valor UTF-8 (Formato de Transformação Unicode ou Unicode Transformation Format); e
- Desta forma, caracteres especiais, símbolos e acentuações serão reconhecidos de forma fácil e limpa para qualquer ferramenta de manipulação de dados, sendo assim, ao realizar a ação de salvar os dados contidos no Notepad mantive o valor padrão para opção enconding na tela salvar.

Com os arquivos de treinamento e checagem criados, podemos caminhar para fazer uso dos mesmos dentro do MatLab, realizando a leitura dos respectivos valores de dia sem chuva, precipitação um dia anterior e risco de fogo observado.

Será utilizada a ferramenta Importação de Dados (*Import Data)* existente na *Toolbox Workspace,* disponível no ambiente de desenvolvimento computacional MatLab. Sua forma de utilização é bem simples, muito parecida com o processo de abertura de um arquivo através do próprio MatLab, como também, similar ao processo de importação de um arquivo .txt ou .csv no Microsoft Excel.

Ao clicar no botão será aberta a tela que nos selecionar o arquivo desejado e realizar o processo de leitura do mesmo, o que difere neste momento é a próxima tela que será apresentada. Esta, permite ao MatLab realizar o reconhecimento do tipo de arquivo que está sendo lido, identificar sua estrutura e possibilitar que o conteúdo seja definido e formatado para logo na sequência ser adicionado ao *Workspace* como uma variável. A Figura 13 apresenta a tela de definição do layout e estrutura do conteúdo disponível no arquivo training.txt.

Figura 13 - Ferramenta Import Wizard.



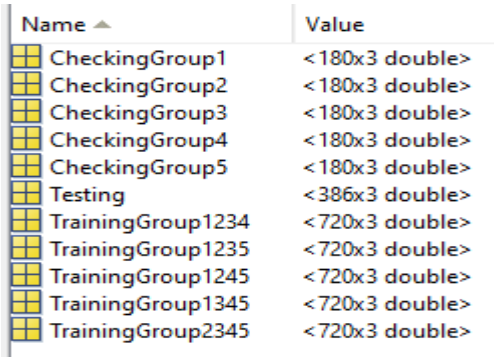
Fonte: Elaboração própria.

Com base na Figura 13, observamos que o assistente de importação de dados disponibilizado pelo MatLab, conseguiu realizar o acesso ao conteúdo do arquivo. Em nosso processo de importação foram carregados os 5 (cinco) arquivos definidos para as fases de treinamento, checagem e teste de dados.

Por fim, a próxima tela apresentada pelo assistente vai permitir escolher em qual estrutura reconhecida pelo MatLab nossos dados serão alocados, deixe selecionada a opção *Create variables preview*, para que seja criada uma variável definida com base no nome do arquivo que estamos importando.

Na sequência as respectivas variáveis serão adicionadas ao toolbox *Workspace* dentro do ambiente do MatLab, conforme ilustra a Figura 14.

Figura 14 - Relação de variáveis criadas na Toolbox Workspace.



Fonte: Elaboração própria.

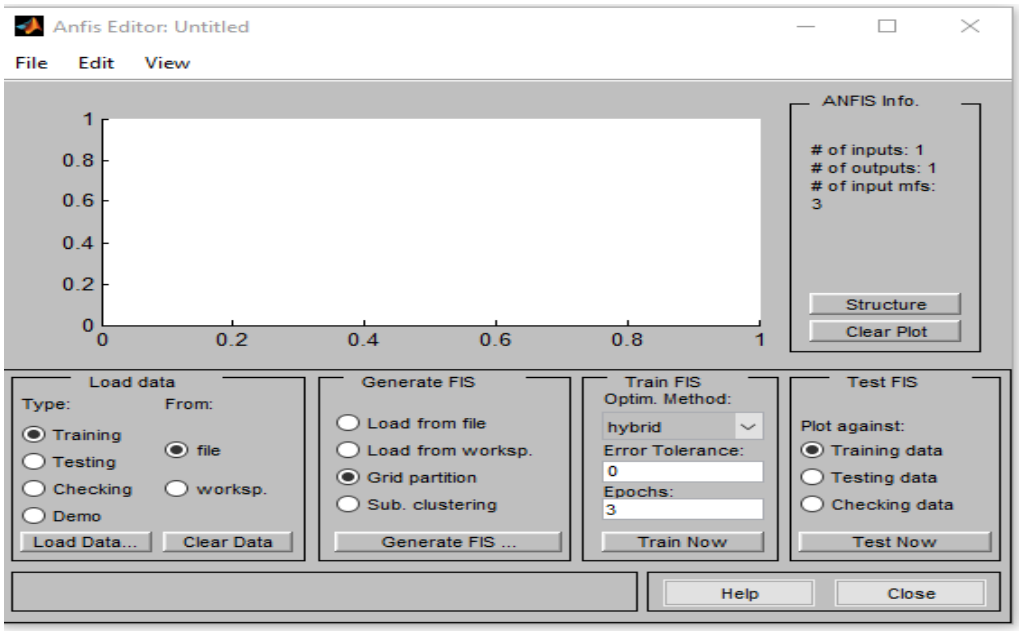
Nos anexos deste trabalho, consta a versão final do código-fonte desenvolvido através da linguagem Transact-SQL existente no Microsoft SQL Server. Este código foi implementado como objetivo de demonstrar como podemos automatizar todo processo de preparação da base de dados, em adicional, aplicar a técnica de validação cruzada diretamente em uma base de dados relacional.

Além disso, como podemos realizar a criação dos arquivos de treinamento, checagem e testes, para posterior utilização no software MatLab.

5.17 Utilizando a ferramenta AnfisEdit para materializar o ANFIS

A utilização do ANFIS aplicado ao ambiente de desenvolvimento computacional existente no MatLab, baseia-se no uso da ferramenta *Anfis Editor*, a mesma pode ser acessada pelo comando *AnfisEdit.* A Figura 15 apresenta o Anfis Editor inicializado dentro do MatLab.

Figura 15 - AnfisEdit carregado dentro do ambiente MatLab.



Fonte: Elaboração própria.

Ao observar a Figura 15, notamos que o editor anfis é formado por uma estrutura organizada em quatro fases: *Load Data, Generate FIS, Train FIS e Test FIS*, será realizado uma breve abordagem sobre estas quatro fases a seguir:

- *Load Data*: Conjunto de opções que nos permitem escolher qual será a forma de carregamento dos dados, disponível através de arquivos ou variáveis, em adicional definir qual ação existente no ANFIS estes dados serão utilizados, para este estudo como já foi destacado faremos uso dos dados para *Training* e *Checking*;
- *Generate FIS*: Conjunto de opções utilizadas para materialização de um novo *Fuzzy Inference System* (FIS), em suma, permitir que sistema de inferência fuzzy possa se tornar algo real. Para a geração de um FIS, podemos escolher carregar um arquivo já existente, carregar dados de um FIS alocados no *workspace*, em adicional definir qual será a técnica para aprendizado dos dados: *Grid Partition* ou *Sub. Clustering*;
- *Train FIS*: Através deste conjunto de opções, temos a possibilidade de realizar a ação de treinamento dos nossos dados, escolhendo qual será o método de otimização deste treinamento que por padrão vem definido como híbrido (*hybrid*), a quantidade de tolerância de erros durante o processamento, e a quantidade de época conhecida como quantidade de vezes que o treinamento será executado; e
- *TEST FIS*: Através desta opção, temos a possibilidade de realizar o teste dos dados selecionados para análise do algoritmo ANFIS.

Nos anexos deste trabalho, consta a versão final do código-fonte desenvolvido através da linguagem M existente no software MatLab. Este código foi implementado como objetivo de demonstrar como podemos automatizar o carregamento dos dados armazenados nos arquivos de treinamento, checagem e teste, definindo cada conjunto como uma variável única e exclusiva para ser utilizada dentro da ferramenta *AnfisEdit*.

5.18 Realizando o Treinamento dos Dados (Training)

Para a utilização do ANFIS devemos realizar preferencialmente a ação de treinamento combinada com a checagem ou teste dos dados.

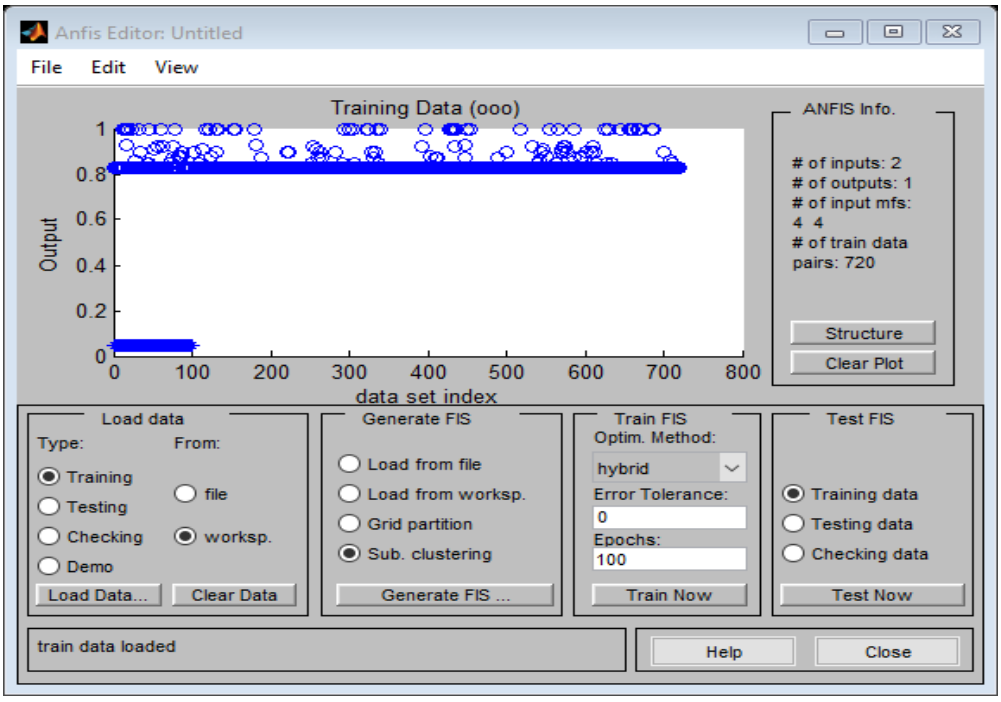
A ação de treinamento de dados é considerada obrigatória, ela deverá prover ao ANFIS as fontes de geração de estímulos que possam produzir ao longo do seu processamento percepções reconhecidas por ele, produzindo assim novos conhecimentos, tornando cada vez o sistema de inferência mais inteligente.

O treinamento de dados realizado para este estudo, apresenta um conjunto de configurações que foram definidas após a realização de diversas simulações, sendo elas:

- Load Data: Workspace declarando o nome da variável Training;
- Generate FIS: Sub. Clustering;
- Train FIS:
    - Optim. Method: Hybrid;
    - Error Tolerrance: 0; e
    - Epochs: 100 (cem) indica a quantidade de vezes em que o treinamento será realizado; e
- Test FIS: Training Data.

As Figuras 16, 17, 18 e 19 apresentadas respectivamente, evidenciam o comportamento do *AnfisEdit* ao longo do processamento de cada fase.

Figura 16 - Visualização dos dados de treinamento carregados.

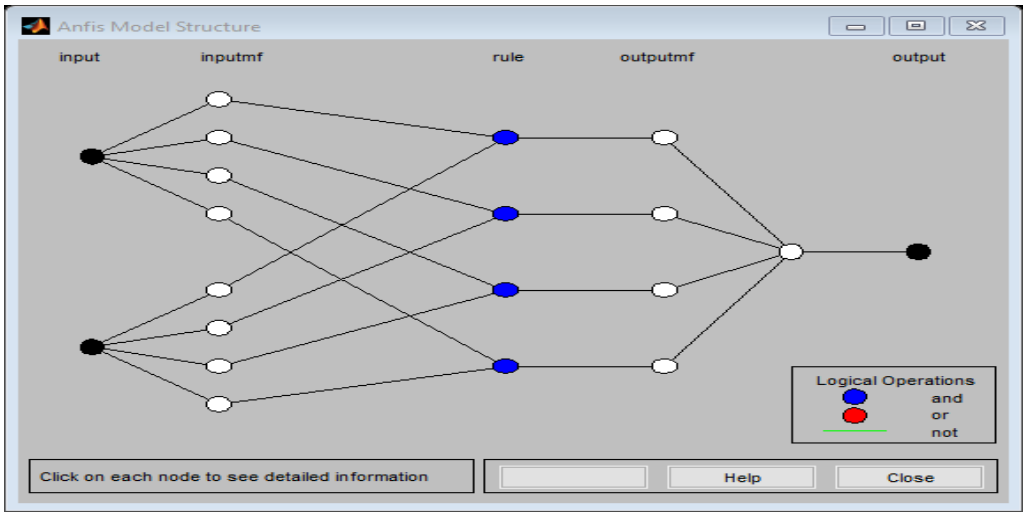


Fonte: Elaboração própria.

Observe que a Figura 16, exibe a área de superfície preenchida por bolinhas que simbolizam os dados carregados pelo *AnfisEdit* que posteriormente serão utilizados no treinamento. A cor azul identifica os elementos de dados carregados e reconhecidos pelo AnfisEdit, a parte superior direita da Figura 16, apresenta o resumo ANFIS Info, um sumário que evidencia o número de entradas e saídas e o número de funções de pertinência utilizada

pela concepção da estrutura do sistema de inferência fuzzy. Ao clicar no botão Structure exibido logo abaixo da ANFIS Info, temos a visualização da estrutura criada para este FIS, conforme a Figura 17.

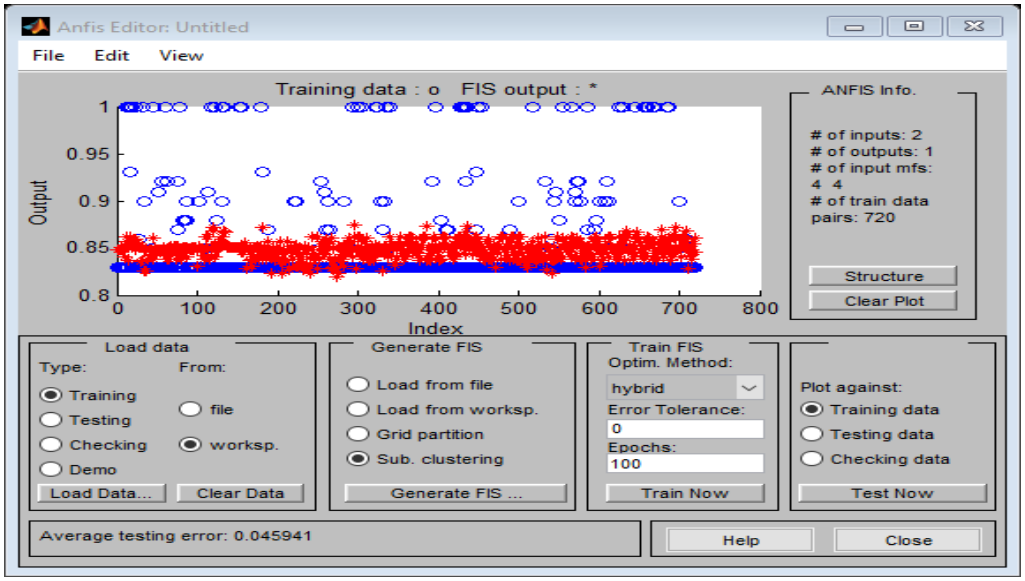
Figura 17 - Visualização da estrutura criada para o FIS.



Fonte: Elaboração própria.

Com a estrutura definida, podemos materializar o sistema de inferência fuzzy, utilizando o botão *Generate FIS*, e depois o botão *Train Now*, de acordo com a Figura 17.

Figura 18 - Épocas de treinamento de dados realizada.



Fonte: Elaboração própria.

A Figura 18 acima destaca o resultado gerado pelo *AnfisEdit* após o processamento do *train now*, as estrelinhas referencias através da cor azul, representam cada uma das épocas de treinamento executada.

Durante o período de execução e processamento de cada época, a ferramenta *Anfis Edit*or faz uso da janela de comandos (*Command Window*) existente na área de trabalho do MatLab, para registrar um conjunto de dados oriundos do processo de training que visam orientar como o ANFIS está trabalhando com os dados ofertados a ele.

Estes dados gerados pelo training, são reconhecidos como um sumário da estrutura que compõem o Anfis Info, composto por alguns parâmetros, como:

- Número de Nós (*Number of nodes*);
- Números de parâmetros linear (*Number of linear parameters*); e
- Número de pares de treinamento (*Number of training data pairs*); entre outros.

Ao navegarmos até a janela de comandos, analisarmos os parâmetros apresentados pelo Anfis Info, conseguimos compreende como os nossos dados se entrelaçam com a estrutura do ANFIS.

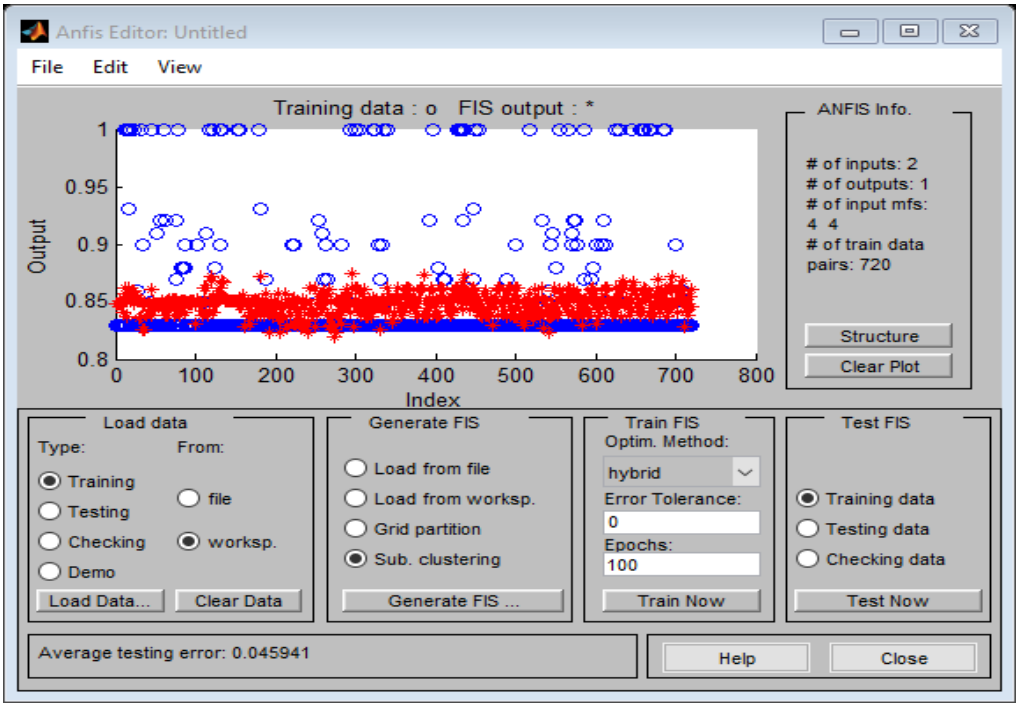
A última parte do processo de treinamento dos dados, consiste na realização da checagem ou teste, executada através do botão testar agora (t*est now)*. Há ação testar agora, faz uso do conjunto de dados carregados na fase de *load data*, em combinação com a estrutura do FIS criada na fase *generate fis*, e dos valores de treinamento obtidos para cada época ao término da fase *train now*.

Em resumo, a possibilidade de definir as configurações, combinadas com os elementos gerados pelo ANFIS, permitem que o treinamento de dados possa percorrer um caminho viável para sua realização.

O teste de dados deve ser considerado como um momento importante do que tange a análise dos mesmos, ao ser executado as percepções capturadas no seu processamento deverão ser apresentadas na área de superfície de plotagem de gráficos do *Anfis Editor*, destacando o resultado do conjunto de dados recém treinamento.

Por padrão, os elementos são plotados de duas formas diferentes para facilitar nossa compreensão, utilizando um conjunto de bolas azuis quem representam os dados treinados, mesclados com as estrelas vermelhas que identificam a saída do sistema de inferência fuzzy. A Figura 19 ilustra o retorno da execução do *Test Now*.

Figura 19 - Retorno da execução do Test Now.



Fonte: Elaboração própria.

5.19 Realizando a Checagem dos Dados (Checking)

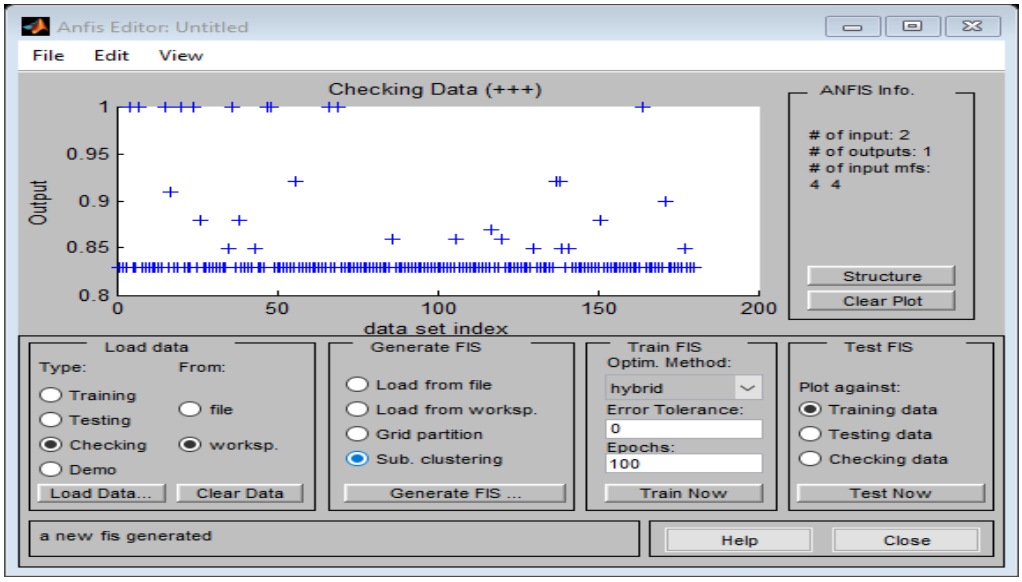
A fase de checagem de dados tem como objetivo, evidenciar o quanto nosso modelo híbrido, em conjunto com o sistema de inferência fuzzy está conseguindo aprender ao longo de cada ciclo de processamento, ou seja, avaliar se os resultados apresentados na fase de treinamento podem ser comparados com os resultados posteriormente obtidos ao término da fase de checagem.

Através do conjunto de variáveis de entrada entrelaçadas com a variável de saída, nosso sistema de inferência fuzzy adaptado com as redes neurais artificias deverá sensibilizar seus respectivos conjuntos de nós, afim, de ser estabelecer uma forma de leitura do conjunto de dados que estão sendo processados nesta fase comparados com os valores processados na fase de treinamento, verificando se estes dados, seguem ou apresentam características similares que permitam reconhecer o mesmo padrão ou comportamento.

As Figuras 20, 21 e 22 exibem respectivamente os resultados que o *Anfis Editor* retornou após o processamento de cada uma das operações executadas pela modelo híbrido mediante ao conjunto de dados selecionado para a fase de checagem.

Através da Figura 20 visualizamos na cor azul o símbolo “+” que identifica os dados de checagem carregados.

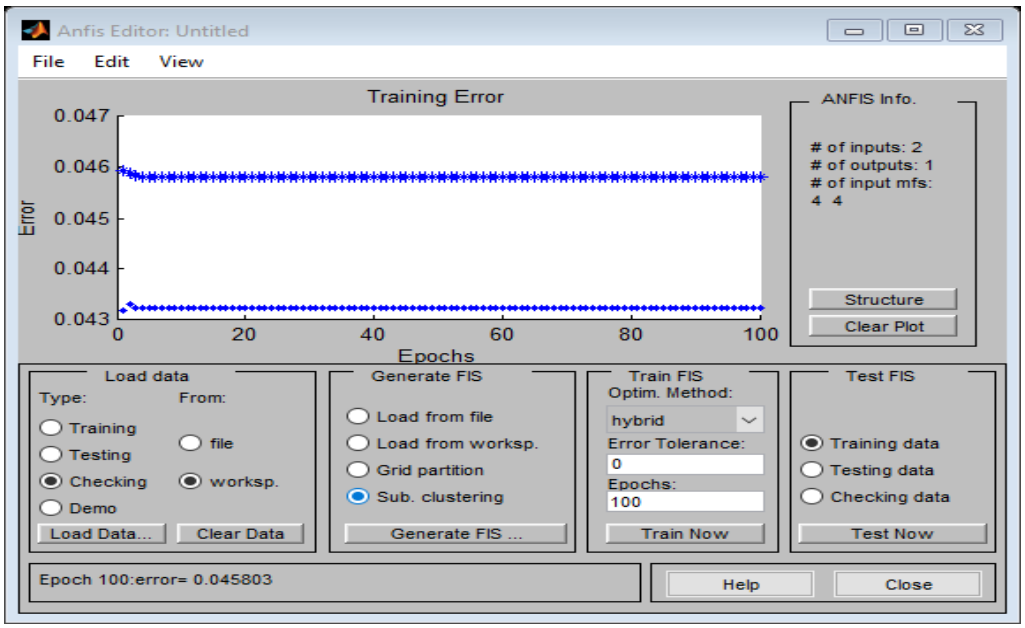
Figura 20 - Visualização dos dados de checagem carregados.



Fonte: Elaboração própria.

A Figura 21 traz o término da execução da fase de treinamento e os resultados obtidos.

Figura 21 - Época de treinamento de dados realizada.



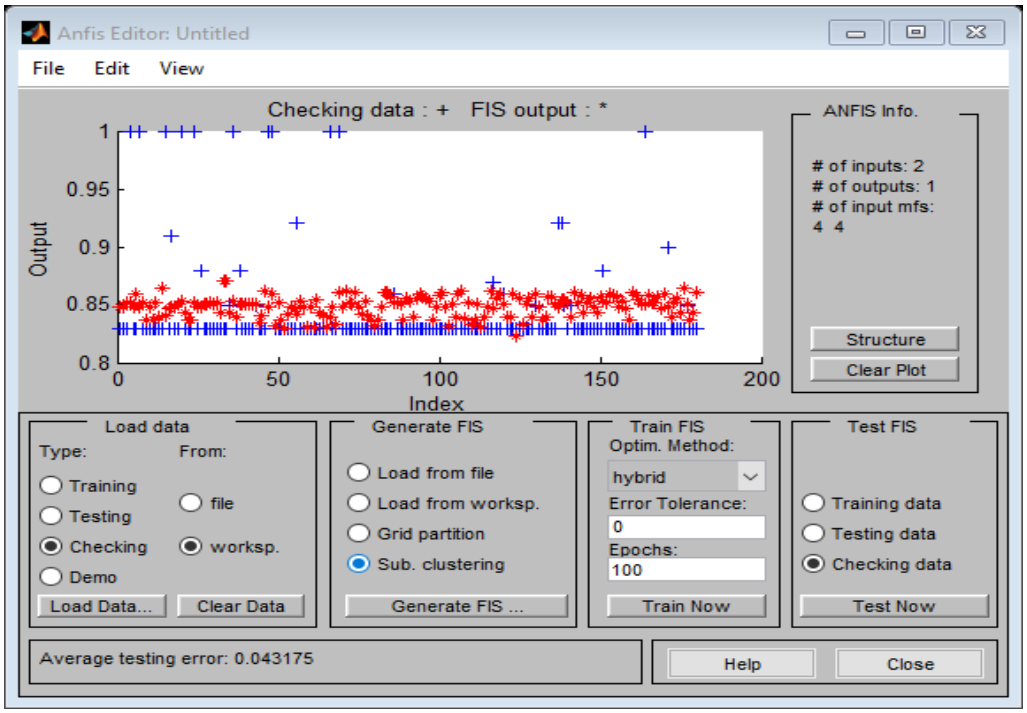
Fonte: Elaboração própria.

De maneira similar ao *train now* realizado na fase de treinamento, o *Anfis Editor* nos retorna na área de superfície o resultado do processamento realizado por ele ao longo da execução cíclica de cada época, conforme apresentou a Figura 21.

Conseguimos avistar a existência de dois elementos representados na cor azul, o asterisco exibido na parte inferior da área superfície que indica o valor do erro observado, na parte superior visualizamos o asterisco um pouco menor, que representa a variação do erro observado a cada época processada.

Esta distância entre cada elemento de acordo com sua época de treinamento, tem como objetivo principal mostrar as mudanças nos indicadores de percepção de erros obtidos na execução de cada época. A Figura 22 mostra o resultado obtido após a execução do *Test FIS*.

Figura 22 - Test FIS realizado.



Fonte: Elaboração própria.

A Figura 22 destaca os elementos plotados pelo modelo híbrido na área de superfície existente dentro *Anfis Editor*, ressaltando na cor azul o sinal de mais "+" que representa a saída obtida após a fase de checagem de dados. Estes mesclados com o asterisco na cor vermelha compreendem a saída identificada pelo sistema de inferência fuzzy vinculada com o modelo hibrido.

6. RESULTADOS E DISCUSSÃO

Para o estudo aplicado a este trabalho, foi selecionado o método de agrupamento subtrativo (*Sub. Clustering*). Este método apresentou melhores resultados relacionados ao tempo de processamento de dados, no que se refere a sua agilidade para ler e processar séries históricas que possuem mais de quatro dados de entrada (MILAN et al., 2021).

A escolha desta técnica se deve a forma de aprendizagem que a estrutura ANFIS envolve na fase de seleção de variáveis. Através da determinação do número de funções de pertinências por variável e a obtenção de um conjunto de regras fuzzy. Quando não se conhece "a priori" quantos agrupamentos deve haver para um determinado conjunto de dados, o agrupamento subtrativo é um algoritmo rápido e robusto para saber este número (AZAD et al., 2019).

Ainda, esta técnica permite a localização do centro do agrupamento, sendo as funções de pertinências e as regras obtidas a partir destes centros de agrupamento e, portanto, com estas informações é possível gerar um sistema de inferência fuzzy do tipo Takagi-Sugeno que modela o comportamento dos dados.

O método do algoritmo de agrupamento subtrativo considera um conjunto de N amostras (vetores de observações) de dados. As mesmas, definidas em um espaço de dimensão m + n (no problema da identificação do sistema, m é o número de entradas e n é o número de saídas), as quais são normalizadas em cada uma das dimensões, de modo que os dados estejam limitados por um hipercubo unitário (CHIU, 1994; ABRAHAM et al., 2005).

Após a partição do espaço de entrada são geradas as funções de pertinência, neste caso funções com o formato gaussiano devido à escolha do método subtrativo. A seguir, os parâmetros das funções de pertinência de entrada e saída são otimizados com um algoritmo híbrido como forma de alcançar o mapeamento das entradas para com as saídas identificadas.

A utilização do modelo híbrido ANFIS através do algoritmo agrupamento subtrativo foi realizado mediante ao uso do software de linguagem computacional MatLab. Através da caixa de ferramentas denominada *AnfisEdit*.

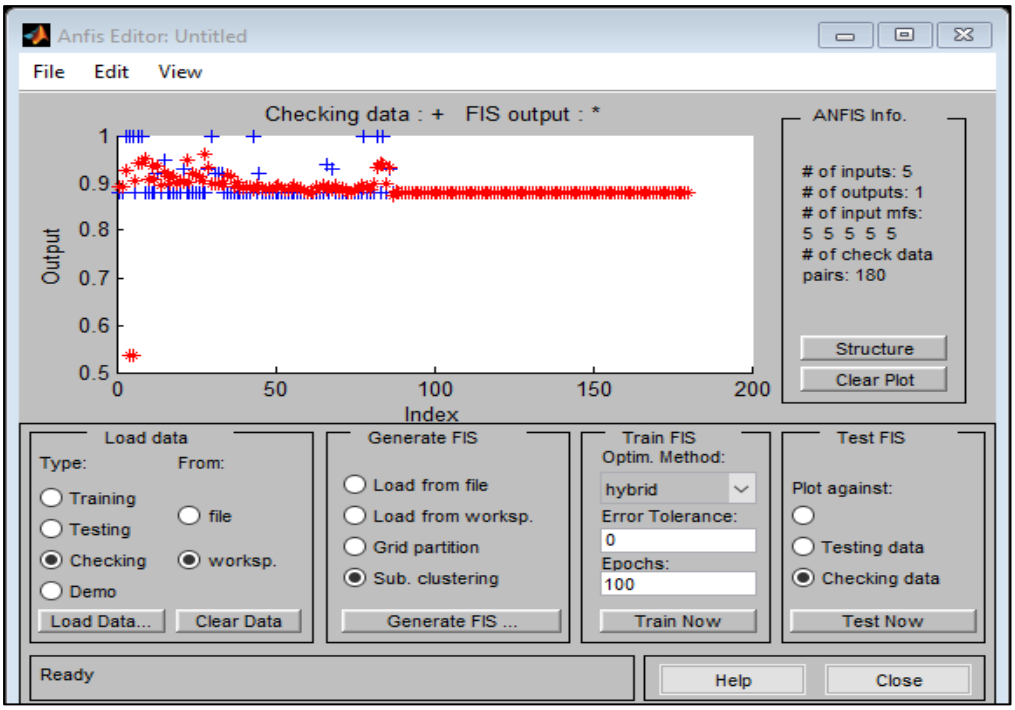
A partir do processamento das 1286 observações identificadas durante a fase de preparação dos dados, 386 observações (30% do total de observações) foram selecionadas de forma aleatória para compor o conjunto de dados condicionados à fase de treinamento.

As outras 900 observações deslocaram-se para uso no processo de validação cruzada, a qual estabeleceu os 5 (cinco) grupos compostos por 180 observações também selecionadas de forma aleatória sem repetição.

Estes grupos foram submetidos às fases de treinamento e checagem de dados para cada um dos 6 (seis) modelos. Os quais foram processados através do modelo híbrido ANFIS configurado para o FIS aplicado ao método de agrupamento subtrativo, processado ao longo de 100 épocas.

Esta caixa de ferramentas, pode ser utilizada também para trabalhar com o outro algoritmo particionamento em grid. A Figura 23 ilustra um dos modelos processados através da ferramenta *Anfis Editor*.

Figura 23 - Ferramenta AnfisEdit.



Fonte: Elaboração própria.

A Tabela 13 destaca os resultados obtidos pelos modelos elaborados na predição da presença do risco de fogo para 1 (um) dia, para as fases treinamento e checagem de dados.

Tabela 13 - Resumo dos Resultados – Modelos período de 1 dia

Modelos	**Variáveis**		**Média dos Erros Absolutos**	**Desvio Padrão**
	Entrada	Saída		
Modelo 1	5	1	0,02	0,05
Modelo 2	3	1	0,03	0,01
Modelo 3	2	1	0,03	0,01

Fonte: Elaboração própria.

Através dos resultados apresentados na Tabela 13 nota-se que o modelo 1 formado por 5 (cinco) variáveis de entrada apresentou o menor valor de média de erros absolutos, mas o maior valor de desvio padrão se comparado com os outros dois modelos. Esta diferença se deve a quantidade de dados de entradas e seus respectivos intervalores de valores.

A Tabela 14 apresenta os resultados referentes aos modelos elaborados para predição da presença do risco de fogo para o período máximo de 3 dias, valores obtidos ao término da execução das fases de treinamento e checagem, de acordo com as definições da validação cruzada 5-*fold*.

Tabela 14 - Resumo dos Resultados – Modelos período de 3 dias

Modelos	**Variáveis**		**Média dos Erros Absolutos**	**Desvio Padrão**
Modelo 4	5	1	0,05	0,18
Modelo 5	3	1	0,05	0,01
Modelo 6	2	1	0,04	0,00

Fonte: Elaboração própria.

De posse dos resultados apresentados nas Tabela 13 e 14, observamos que entre os modelos criados para predizer o risco de fogo observado para 1 dia, o Modelos 3 aparece como uma boa alternativa.

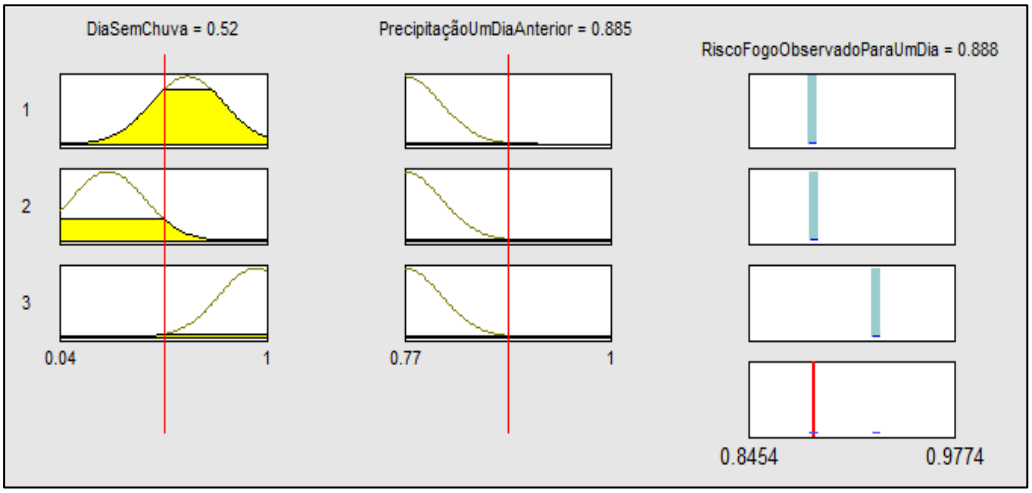
Este modelo destaca-se por sua simplicidade, possui uma pequena estrutura formada pelas variáveis de entrada: dia sem chuva e precipitação do dia anterior, constituídas por dados existentes diretamente na base dados do INPE e INMET, característica que torna sua implementação mais fácil, rápida e simples.

Ao contrário do que acontece nos Modelos 1 e 2 que possuem estruturas mais complexas, formados por algumas variáveis de entrada que não existiam inicialmente na base de dados disponíveis para este trabalho. Dentre elas, a variável: acumulado precipitação para três dias anteriores, gerada através do cálculo da soma do volume de precipitação dos três dias anteriores a data de análise do risco de fogo observado.

Mediante a estes requisitos destacados anteriormente, como também, de acordo com os valores relacionados a média de erros absolutos e desvio padrão, o Modelo 3 se destaca ainda mais. Destaque que permite a este modelo, se tornar o modelo híbrido indicado ao objetivo deste trabalho, que visa criar um índice capaz de estabelecer uma forma de classificar os níveis de risco de queimadas no Brasil.

A Figura 24 destaca a relação de funções de pertinências definidas para o Modelo 3.

Figura 24 - Funções de pertinência existentes no Modelo 3.



Fonte: Elaboração própria.

Observando a Figura 23, podemos notar que a segunda variável de entrada precipitação para um dia anterior praticamente utilizou uma única função pertinência, ao contrário da primeira variável de entrada dia sem chuva, a qual está utilizando para este conjunto de valores apresentados 3 funções de pertinências, do tipo gaussiana.

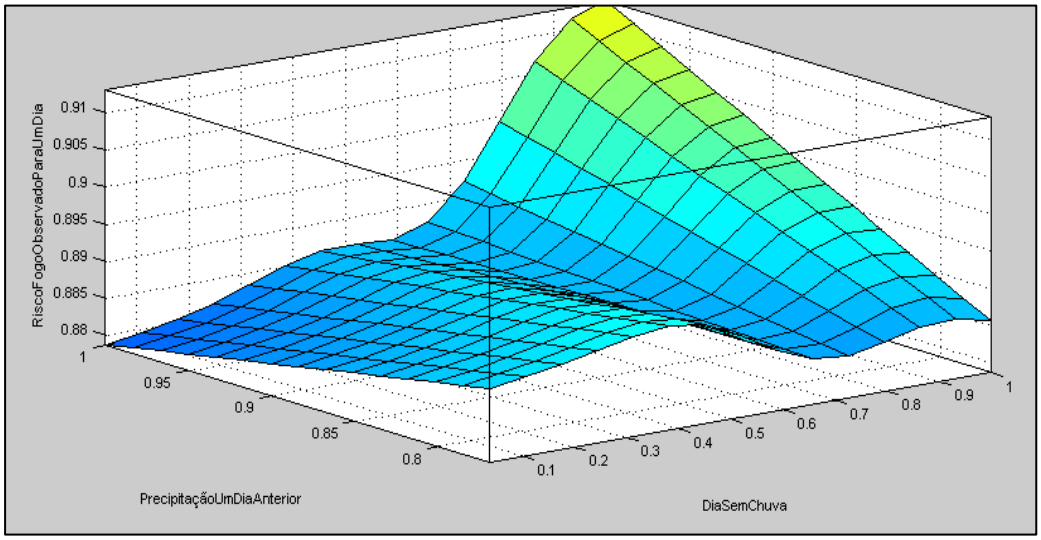
Este comportamento se torna evidente quando os valores existentes em cada linha de observação selecionadas pelo ANFIS apresentam intervalos de valores muito próximo, ou seja, a diferença entre estes respectivos valores é muito pequena, fazendo com que a mesma função de pertinência aparentemente venha a ser utilizada.

No que se refere a variável de saída risco de fogo observado, a mesma, utiliza a função da saída linear do tipo Takagi-Sugeno, possibilitando estabelecer um padrão de reconhecimento e apresentação de cada linha de dados de saída reconhecida pelo modelo.

A Figura 25 exibe o gráfico de superfície gerado pelo modelo híbrido ANFIS de acordo com a combinação entre algumas variáveis de entrada dia sem chuva e precipitação existentes no Modelo 3.

Por padrão este gráfico pode ser consultado ao final de uma das fases de treinamento, checagem ou teste. Ao visualizamos o gráfico de superfície, temos a possibilidade de identificar os valores que permitem formar a chamada saída do sistema, ou seja, como os dados de entrada combinados com os dados de saída permitiram a elaboração de um resultado exibido de forma visual.

Figura 25 - Gráfico de Superfície - Dia sem chuva e Precipitação.



Fonte: Elaboração própria.

Podemos notar, o quanto mais próximo do número 1 a variável dia sem chuva se aproxima, maior a superfície do gráfico se aproxima de valores mais altos realizados a variável de saída risco fogo observado.

Da mesma maneira que o modelo 3 se destacou entre o conjunto de modelos implementados para predizer a presença do risco de fogo para um dia, os modelos elaborados para 3 dias também apresentaram um possível candidato a ser utilizado como modelo híbrido aplicado ao índice, mais especificamente o Modelo 6.

Este, apresentou o valor para média de erros absolutos próximo a 4%, formado pelo mesmo conjunto de variáveis de entrada se comparado ao Modelo 3.

Logo após a execução das fases de treinamento e checagem, todos os modelos foram submetidos a fase a avaliação do sistema de inferência fuzzy, que possui a finalidade de garantir se os resultados retornados pelo modelo híbrido ANFIS estão próximos aos chamados valores reais.

O processo de avaliação do sistema de inferência fuzzy, foi realizada mediante ao uso da ferramenta de linha de comando *EvalFIS*, existente no software MatLab. Nos anexos deste trabalho está disponibilizada a versão final do código-fonte desenvolvido através da linguagem M existente no software MatLab que demonstrar como podemos realizar o procedimento de avaliação do modelo 3 criado neste trabalho.

A partir da utilização da função *EvalFIS*, nosso sistema estará sendo submetido a uma validação de sua estrutura e principalmente das regras estabelecidas para seu aprendizado.

Ao final do processamento da avaliação do sistema, podemos obter o resultado desta operação, que consiste basicamente na apresentação da relação de valores do conhecimento obtido pelo modelo, tendo como base os valores de entrada combinado com os valores de saída, criando uma lista denominada como saídas calculadas ou lista de resultados de aprendizado.

Resumidamente, a lista de resultados de aprendizados criada logo após a avaliação do sistema de inferência fuzzy pertencente a cada modelo, foi posteriormente combinada com a relação de observações também existente nos modelos destacados já apresentados neste trabalho.

Através desta combinação entre os chamados dados reais e os dados calculados, foi possível comparar de forma concreta os resultados de saídas que o sistema de inferência está retornando. Em suma, ter a certeza que o modelo realmente conseguiu aprender com base nos conjuntos de dados a ele ofertados.

A Tabela 15 ilustra uma pequena porção de dados selecionados de forma aleatória que representam os resultados obtidos após o procedimento de avaliação do Modelo 3.

Tabela 15 - Variável risco de fogo observado combinada com a saída calculada

RiscoFogoObservado (Saída Real)	**RiscoFogoObservado (Saída Calculada)**	**Diferença (Módulo)**
0.84	0.88	0.04
0.85	0.89	0.04
0.87	0.89	0.02
0.88	0.89	0.01
0.89	0.89	0.00
0.90	0.88	0.02
0.91	0.89	0.02
0.94	0.89	0.05
0.95	0.89	0.06
1	0.89	0.11

Fonte: Elaboração própria.

Os valores da variável de saída risco de fogo observado se referem a predição do risco de fogo para um dia. Estes representam dados reais, combinados com a saída calculada resultante da avaliação do FIS pertencente ao modelo 3.

Esta comparação, indica que o modelo híbrido ANFIS, mostrou-se eficaz para a elaboração do índice aqui proposto, permitindo a agregação de diversas variáveis, na criação de regras condicionais e parâmetros de forma automática.

Através da combinação das variáveis de entrada dia sem chuva e precipitação, em conjunto com a definição da variável de saída risco fogo observado para 1 ou 3 dias, foi possível atingir o objetivo final deste estudo, que visa estabelecer uma forma de classificar os níveis de risco de fogo.

Esta classificação foi elaborada com base na mesma definição criada pelo INPE em 2002 e atualizada em 2019, empregado nos atuais indicadores de classificação de níveis de risco de queimadas ou risco de fogo para 1 ou 3 dias. Tanto para o indicador risco de fogo observado, como também, para o risco de fogo previsto.

A Tabela 16 apresenta os níveis de classificação adaptados de acordo com as definições do INPE, considerando a mesma nomenclatura e faixas de valores.

Tabela 16 - Níveis de Classificação de Risco de Fogo

Nível de Classificação	Intervalos
Mínimo	[0, 0,15]
Baixo	(0,15, 0,40]
Médio	(0,40, 0,70]
Alto	(0,70, 0,95]
Crítico	(0,95, 1]

Fonte: Adaptado de: INPE, 2019.

A partir da definição dos níveis de classificação e faixas de valores, foi aplicado o uso da matriz de confusão com objetivo de identificar para cada nível a quantidade de observações pertencentes a ela.

Através da utilização da matriz de confusão foi possível verificar o grau de assertividade dos resultados, em adicional, compreender quais os valores estão relacionados as classes estatísticas: sensibilidade e especificidade.

De forma sucinta a classe estatística sensibilidade nos ajuda a mensurar a capacidade de minimizar os chamados valores falsos negativos, reconhecida como a habilidade de identificar de forma correta as observações que o modelo também identificou corretamente dentro de cada nível de classificação (LEVER et al., 2016).

Por outro lado, a classe estatística especificidade tem o objetivo de mensurar os chamados valores negativos existentes no conjunto de valores retornados pelo modelo também como negativos, denominada como a capacidade de diferenciar corretamente as observações que o modelo identificou pertencente a uma classificação, mas que na verdade podem pertencer à outra (LEVER et al., 2016).

A Tabela 17 apresenta a matriz de confusão concebida a partir do uso dos níveis de classificação, com base no dados existentes no arquivo de teste aplicado ao Modelo 3.

Tabela 17 - Matriz de Confusão

	Mínimo	**Baixo**	**Médio**	**Alto**	**Crítico**
Mínimo	0	0	0	0	0
Baixo	0	0	0	0	0
Médio	0	0	0	0	0
Alto	0	6	21	313	0
Crítico	0	2	6	28	10

Fonte: Elaboração própria.

A matriz de confusão, foi elaborada utilizando as 386 observações selecionadas para a fase de teste do modelo 3, referentes aos anos de 2018, 2019 e 2020, correlacionadas a cidade de Sorocaba.

Através dos valores existentes na matriz de confusão, foram realizados os cálculos para identificar as classes estatísticas sensibilidade e especificidade. Os valores de cada classe, estão definidos para intervalo entre os números 0 (zero) e 1 (um), reconhecida como forma padrão de apresentação destes valores estabelecido pela matriz de confusão.

A Tabela 18 traz os valores correlacionados entre o nível de classificação e a classe estatística.

Tabela 18 - Resultados das classes estatísticas Sensibilidade e Especificidade

	Mínimo	**Baixo**	**Médio**	**Alto**	**Crítico**
Sensibilidade	0	0	0	0.92	1
Especificidade	1	1	1	0.40	0.90

Fonte: Elaboração própria.

De posse do conjunto de valores apresentados na Tabela 18, podemos realizar uma análise com objetivo de examinar os valores retornados oriundos da execução da matriz de confusão, relacionados ao nível de classificação crítico.

Nesta análise, podemos evidenciar que a classe estatística especificidade está bem próxima ao valor 1 (um), e a sensibilidade atingiu o valor 1 (um). De forma geral, o valor médio para classe sensibilidade apresenta está próximo ao valor 0,38 e a classe especificidade o valor médio obtido foi de 0,86.

Os valores médios 0,38 para a classe estatística sensibilidade e 0,86 para classe estatística especificidades foram calculados logo após a construção da matriz de confusão. Foram considerados o conjunto de valores retornados por estas classes estatísticas pertencentes a cada nível de classificação, desta forma, realizou-se a soma de todos os valores, e posteriormente a divisão de acordo com o total de níveis de classificação apresentadas para serem utilizadas na construção da matriz de confusão.

Podemos dizer que os resultados apresentados pela matriz confusão evidenciam que Modelo 3, demonstra uma boa capacidade em conseguir identificar de forma específica um valor que estava atribuído a uma classificação mas na verdade deveria pertencer a outra, capacidade que diminuição a presença dos chamados valores negativos dentro do conjunto de resultados analisados.

Em suma, a capacidade de acerto na identificação das observações que o modelo híbrido identificou como verdadeiro pode ser considerada como correta em relação a sua sensibilidade para os níveis de classificação alto e crítico com base no conjunto de dados utilizados para o construção da matriz de confusão.

Como também, a especificidade que apresenta o valor próximo a 1 (um), a qual representa a baixa existência de valores negativos para o níveis de classificação mínimo, baixo, médio e crítico. Sendo assim, considerando que todas as observações identificadas pelo modelo híbrido mais especificamente o Modelo 3, estão categorizadas de forma correta.

7. CONSIDERAÇÕES FINAIS

As queimadas têm sido intensificadas nas últimas décadas em todo território brasileiro, fenômeno que tem ocasionado à emissão de gases e poluentes na atmosfera, contribuindo para alterações na qualidade do ar e no clima. Fatores o que tem influenciado a qualidade de vida da população, colaborando para a ocorrência de problemas de saúde em inúmeros seres humanos.

Estes, ocasionando diversas alterações nas características do solo e morte de seres vivos, podem interferir no equilíbrio do ambiente e provocar danos ambientais. Assim, as queimadas causam consequências na superfície e na atmosfera.

O estudo e o monitoramento do risco do fogo tem sido importante para o mapeamento das áreas afetadas ao longo do território brasileiro, com intuito de estimar as alterações ambientais, identificar os impactos no solo e cobertura da terra.

Esta é uma maneira de coletar, armazenar e fornecer dados que podem ser utilizados para planejamento de ações, como forma de tentar evitar futuros impactos ambientais, econômicos e sociais, visando estabelecer alternativas para mitigar estas e outras ações prejudiciais, oriundas da ação provocada por este tipo de fenômeno.

Cada vez mais, torna-se importante a ocorrência de estudos que visam obter novos conhecimentos relacionados aos fatores que possam ajudar na identificação da ocorrência de uma nova queimada, nas mais diversas áreas do conhecimento (MOHAMMED et al., 2020).

Propondo a construção de uma nova fonte de aprendizado que poderá ajudar na tomada de decisão frente aos problemas ambientais e de saúde pública, que objetivam criar ações planejadas e políticas públicas capazes de diminuir seus possíveis riscos e impactos.

De acordo com os modelos apresentados neste estudo, foram identificadas maneiras de avaliação da presença do risco de queimada ou fogo para o intervalo entre 1 ou 3 dias aplicadas de forma única e específica ao município de Sorocaba, localizada no interior do Estado de São Paulo.

Compreender os comportamentos, regras e condições que podem provocar a ocorrência de fenômenos naturais, foi um dos fatores predominante para se estabelecer a uma forma de análise inicial em relação ao conjunto de dados selecionados para este estudo. Os quais serviram como combustível para alimentar as possibilidades de aprendizados que poderiam ser descobertos.

A preparação da base de dados, possibilitou definir de forma mais clara, qual seria o cenário a ser trabalhado, ou seja, de posse dos dados tratados, padronizados e prontos para serem reconhecidos como observações, o caminho de aplicação do modelo híbrido ANFIS estava definido.

Os resultados atingidos através do processamento dos seis modelos selecionados para este estudo, mostram que a combinação entre o conhecimento humano e o aprendizado de máquina se tornam uma ferramenta extremamente poderosa.

Em suma, os resultados apresentados por este estudo, tendem a tornar-se uma valiosa fonte de conhecimento, ofertada para a sociedade de maneira geral, empresas e instituições de ensino ou pesquisa, no que tange a definição de programas e políticas públicas relacionadas a ações de prevenção, conservação e combate a áreas de queimadas com a presença constante de fogo.

Este estudo limitou na utilização de dados espaciais disponibilizados pelo INPE relacionados as queimadas e risco de fogo, os quais foram combinados com os dados pluviométricos disponibilizados pelo INMET, ambos relacionados a cidade de Sorocaba, para estabelecer a predição da presença da queimada para o período máximo de três dias.

Para trabalhos futuros pode-se analisar a comparação entre os resultados aqui obtidos, para com outros índices de classificação de níveis de riscos de queimadas utilizados no Brasil, como também na predição do risco de queimada ou fogo para períodos superiores da 30 dias.

8. REFERÊNCIAS

ABRAHAM, A. **Adaptation of Fuzzy Inference System Using Neural Learning, in: Fuzzy Systems Engineering: Theory and Practice** (Eds.: N. Nedjah, L. de Macedo Mourelle), Springer Verlag, Berlin, pp. 53-83, 2005.

AZAD, A.; MANOOCHEHRI, M.; KASHI, H.; FARZIN, S.; KARAMI, H.; NOURANI, V.; SHIRI, J. **Comparative evaluation of intelligent algorithms to improve adaptive neuro-fuzzy inference system performance in precipitation modelling**. J. HYDROL. 571, 214–224. https://doi.org/10.1016/j.jhydrol.2019.01.062, 2019.

BARROS, L.C. ; BASSANEZI, R.C. **Tópicos de Lógica Fuzzy e Biomatemática**. 2.ed. Campinas: Comissão de Publicações IMECC, 2010.

BOND, W. J.; WOODWARD, F. I.; MIDGLEY, G. F. **The global distribution of ecosystems in a world without fire**. New Phytologist, v. 165, p. 525-538, 2005.

CHIU, S.L. **Fuzzy model identification based on cluster estimation.** Journal of Intelligent & fuzzy systems, v. 2, p. 267–278, 1994.

CUEVAS, E.; DÍAZ, P.; ZALDÍVAR, D.; CISNERO, P.M. **Nonlinear system identification based on ANFIS-Hammerstein model using Gravitational search algorithm. Applied Intelligence.** v. 48, p. 182-203. Springer Nature. http://dx.doi.org/10.1007/s10489-017-0969-1, 2017.

DUARTE, M.L.; SILVA, T.A. **Avaliação do desempenho de três algoritmos na classificação de uso do solo a partir de geotecnologias gratuitas.** Revista de Estudos Ambientais, v. 21, p. 6-16. http://dx.doi.org/10.7867/1983-1501.2019v21n1p6-16, 2019.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Estimativa Populacional.** https://ibge.gov.br/cidades-e-estados/sp/sorocaba.html, 2021.

JANG, J.S.R. **ANFIS Adaptive-Network-based Fuzzy Inference System.** https://doi.org/10.1109/21.256541, 1993.

LEMOS, C.F.; JUSTINO, F.B.; COSTA, L.C. **Distribuição espacial do índice de Haines para Minas Gerais por análise da média atmosfera.** Revista Brasileira de Agropecuária Sustentável. https://periodicos.ufv.br/rbas/article/download/2667/1151, 2012.

LEVER, J.; KRZYWINSKI, M; ALTMAN, N. **Classification evaluation.** Nature Methods, v.13, p.541–542, 2016.

LIMA, F. **Estimador Neuro-Fuzzy de Velocidade Aplicado ao Controle Vetorial sem Sensores de Motores de Indução Trifásicos.** https://www.teses.usp.br/teses/disponiveis/3/3143/tde-20092011150232/publico/Tese_Fabio_Lima.pdf, 2010.

LING, H.; QIAN, C.; KANG, W.; LIANG, C.; CHEN, H. **Combination of Support Vector Machine and K-Fold cross validation to predict compressive strength of concrete in marine environment.** https://doi.org/10.1016/j.conbuildmat.2019.02.071, 2019.

LORENZI, Harri. **Árvores brasileiras: Manual de identificação e cultivo de plantas arbóreas do Brasil**, v.1, 5.ed. Nova Odessa, SP: Instituto Plantarum, p.384, 2008.

LOURENÇO, R. W.; SILVA, D. C. C.; SALES, J. C. A. **Elaboração de uma metodologia de avaliação de fragmentos de remanescentes florestais como ferramenta de gestão e planejamento ambienta. Revista Eletrônica Capital Cientifico.** Ambiência Guarapuava (PR) v. 10, n. 3, p. 685 – 698. ISSN 1808 - 0251. https://revistas.unicentro.br/index.php/ambiencia/article/view/2782/2372, 2014.

MAMDANI, E. H. e ASSILIAN S. **An experiment in linguistic synthesis with a fuzzy logic controller**. Int. J. Man-Machine Studies, vol. 7, p. 1-13, 1995.

MALMSTRÖM, A. **The importance of measuring fire severity – Evidence from microarthropod studies**. Forest Ecology and Management, v. 260, p. 62-70. 2010.

MBANZE, A.A; BATISTA, A.A; TETTO, A.F; ROMERO, A.M; MUDEKWE, J. **Desempenho dos Índices de Nesterov e Fórmula de Monte Alegre no Distrito de Lichinga, Norte de Moçambique.** Ciência Florestal, Santa Maria, v. 27, n. 2, p. 687-696. https://periodicos.ufsm.br/cienciaflorestal/article/view/27753, 2017.

MILAN, S.G.; ROOZBAHANI, A.; AZAR, A.N.; JAVADI, S. **Development of adaptive neuro fuzzy inference system – Evolutionary algorithms hybrid models (ANFIS-EA) for prediction of optimal groundwater exploitation.** https://doi.org/10.1016/j.jhydrol.2021.126258, 2021.

MONTANARI, T. **Histologia - Texto, atlas e roteiro de aulas práticas: Tecido Nervoso.** 3. ed. Porto Alegre: Universidade Federal do Rio Grande do Sul. http://www.ufrgs.br/livrodehisto/pdfs/4Nervoso.pdf, 2016.

MOHAMMED, A. J.; AURORA, G. V.; ANTONIO, F. S.; Samira, C.; ABDELOUHAB, Z. **A hybrid neuro-fuzzy inference system-based algorithm for time series forecasting applied to energy consumption prediction.** https://doi.org/10.1016/j.apenergy.2020.114977, 2020.

NUNES, J.R.S; FIER, I.S.N; SOARES, R.V; BATISTA, A.C. **Desempenho da Fórmula de Monte Alegre (FMA) e da Fórmula de Monte Alegre Alterada (FMA+) no Distrito Florestal de Monte Alegre.** https://revistas.ufpr.br/floresta/article/view/17827, 2012.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS (ONU). **Transformando Nosso Mundo: A Agenda 2030 para o Desenvolvimento Sustentável**. Nova Iorque: ONU. https://nacoesunidas.org/wp-content/uploads/2015/10/agenda2030-pt-br.pdf, 2015.

SAMER, M.; JABER, M.; ANDREW, C.; JENNIFER, C.; SIMON, P. **Robust data expansion for optimized modelling using adaptive neuro-fuzzy inference systems.** https://doi.org/10.1016/j.eswa.2021.116138, 2021.

SANTOS, J. F.; SOARES, R. V.; BATISTA, A. C. **Perfil dos incêndios florestais no Brasil em áreas protegidas no período de 1998 a 2002**. Revista Floresta, Curitiba, v. 36, n. 1, p. 93–100, 2003.

SETZER, W. A.; SISMANOGLU, R.; SANTOS, dos M. G. J. **Método do cálculo do risco de fogo do programa do INPE - versão 11, junho/2019.** http://queimadas.dgi.inpe.br/~rqueimadas/documentos/RiscoFogo_Sucinto.pdf, 2019.

SILVA, I.D.B.; PONTES, A.C.F.JR. **Elaboração de um Fator de Risco de Florestais utilizando Lógica Fuzzy.** 2011. https://www.ime.unicamp.br/~biomat/bio21_art9.pdf, 2011.

SILVA, I.N.; SPATTI, D.H.; FLAUZINO, R.A. **Redes Neurais Artificiais para engenharia e ciências aplicadas.** São Paulo: Artliber, 2016.

SOARES, R. V.; BATISTA, A. C. **Incêndios florestais: controle, efeito e usos do fogo.** Curitiba, Universidade Federal do Paraná, p.250, 2007.

SOUZA, de C.V.P. **Fuzzy neural networks and neuro-fuzzy networks: A review the main techniques and applications used in the literature**, v. 92. https://doi.org/10.1016/j.asoc.2020.106275, 2020.

SUGENO, M. **An introductory survey of fuzzy control**. Information Sciences 36, p. 59-83, 1983.

TETTO, A. F. **Comportamento histórico dos incêndios florestais na Fazenda Monte Alegre no período de 1965 a 2009**. p. 115, Setor de Ciências Agrárias, Universidade Federal do Paraná. https://acervodigital.ufpr.br/bitstream/handle/1884/27402/R%20-%20T%20-%20TETTO,%20ALEXANDRE%20FRANCA.pdf?sequence=1&isAllowed=y, 2012.

TORRES, F.T.P.; ROMEIRO, J.M.N.; SANTOS, A.C.A.; NETO, R.R.O.; LIMA, G.S.; ZANUNCIO, J.C. **Fire danger index efficiency as a function of fuel moisture and fire behavior.** https://www.sciencedirect.com/science/article/pii/S0048969718308775, 2018.

WHITLOCK, C.; SHAFER, S. L; MARLON, J. **The role of climate and vegetation change in shaping past and future fire regimes in the northwestern US and the implications for ecosystem management**. Forest Ecology and Management, v. 178, p. 5-21, 2003.

WONG, T.T. **Parametric methods for comparing the evaluated by k-fold cross validation on multiple data sets**. https://doi.org/10.1016/j.patcog.2016.12.018, 2017.

ZADEH, L. A. **Fuzzy sets**. Information and Control, vol. 8, p. 338-353, 1965.

ZADEH, L. A. **Outline of a new approach to the analysis of complex systems and decision processes**. IEEE Transactions on Systems, Man, and Cybernetics, vol. 1, p. 28-44, 1973.

9. APÊNDICE

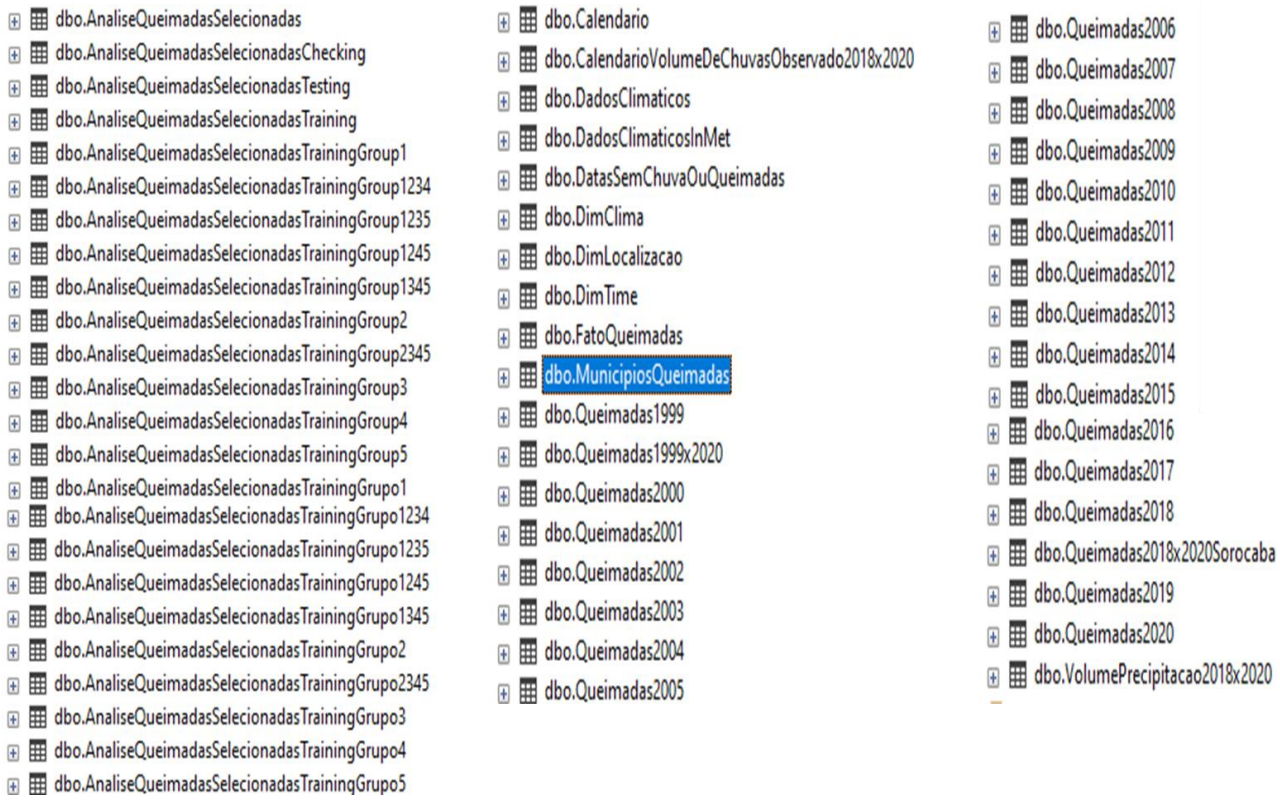
9.1 Figura 1 - Fonte de Dados – Amostra – Dados do INPE.

	A	B	C	D	E	F	G	H	I	J	K	L
1	DataHora;Satelite;Pais;Estado;Municipi;Bioma;DiaSemCh;Precipit;RiscoFog;Latitude;Longitud;AreaIndu;FRP											
2	2018/01/01 21:00:00;NOAA-12;Brasil;Maranhao;BACURI;Amazonia;-2;-2;777.7;-1.685;-45.2817;;											
3	2018/01/02 21:00:00;NOAA-12;Brasil;Parana;PRUDENTOPOLIS;Mata Atlantica;-2;-2;777.7;-24.885;-51.34;;											
4	2018/01/02 21:00:00;NOAA-12;Brasil;Mato Grosso do Sul;LAGUNA CARAPA;Mata Atlantica;-2;-2;777.7;-22.915;-54.965;;											
5	2018/01/02 21:00:00;NOAA-12;Brasil;Mato Grosso do Sul;LAGUNA CARAPA;Mata Atlantica;-2;-2;777.7;-22.9131;-54.955;;											
6	2018/01/02 21:00:00;NOAA-12;Brasil;Mato Grosso do Sul;LAGUNA CARAPA;Mata Atlantica;-2;-2;777.7;-22.9117;-54.9467;;											
7	2018/01/02 21:00:00;NOAA-12;Brasil;Rio Grande do Sul;URUGUAIANA;Pampa;-2;-2;777.7;-29.41;-56.6967;;											
8	2018/01/02 21:00:00;NOAA-12;Brasil;Rio Grande do Sul;URUGUAIANA;Pampa;-2;-2;777.7;-29.4117;-56.7031;;											
9	2018/01/02 21:00:00;NOAA-12;Brasil;Mato Grosso do Sul;RIO VERDE DE MATO GROSSO;Pantanal;-2;-2;777.7;-18.6381;-55.235;;											
10	2018/01/03 21:00:00;NOAA-12;Brasil;Santa Catarina;CELSO RAMOS;Mata Atlantica;-2;-2;777.7;-27.6581;-51.2467;;											
11	2018/01/03 21:00:00;NOAA-12;Brasil;Santa Catarina;CONCORDIA;Mata Atlantica;-2;-2;777.7;-27.1531;-52.0781;;											
12	2018/01/03 21:00:00;NOAA-12;Brasil;Santa Catarina;IRACEMINHA;Mata Atlantica;-2;-2;777.7;-26.845;-53.3231;;											
13	2018/01/03 21:00:00;NOAA-12;Brasil;Santa Catarina;MAREMA;Mata Atlantica;-2;-2;777.7;-26.8067;-52.6967;;											
14	2018/01/03 21:00:00;NOAA-12;Brasil;Rio Grande do Sul;TOROPI;Mata Atlantica;-2;-2;777.7;-29.4367;-54.2431;;											
15	2018/01/03 21:00:00;NOAA-12;Brasil;Rio Grande do Sul;URUGUAIANA;Pampa;-2;-2;777.7;-29.51;-56.8367;;											
16	2018/01/03 21:00:00;NOAA-12;Brasil;Rio Grande do Sul;ENCRUZILHADA DO SUL;Pampa;-2;-2;777.7;-30.71;-52.8481;;											
17	2018/01/03 21:00:00;NOAA-12;Brasil;Rio Grande do Sul;ENCRUZILHADA DO SUL;Pampa;-2;-2;777.7;-30.71;-52.8567;;											
18	2018/01/03 21:00:00;NOAA-12;Brasil;Rio Grande do Sul;QUEVEDOS;Mata Atlantica;-2;-2;777.7;-29.4467;-54.165;;											
19	2018/01/03 21:00:00;NOAA-12;Brasil;Parana;TURVO;Mata Atlantica;-2;-2;777.7;-24.7781;-51.3717;;											
20	2018/01/03 21:00:00;NOAA-12;Brasil;Parana;TURVO;Mata Atlantica;-2;-2;777.7;-24.87;-51.3717;;											
21	2018/01/03 21:00:00;NOAA-12;Brasil;Parana;TURVO;Mata Atlantica;-2;-2;777.7;-24.78;-51.38;;											
22	2018/01/03 21:00:00;NOAA-12;Brasil;Parana;CANDIDO DE ABREU;Mata Atlantica;-2;-2;777.7;-24.8131;-51.14;;											

9.2 Figura 2 - Fonte de Dados – Amostra – Dados do INMET.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	EstacaoCodigo	NivelConsistencia	SubBaciaCodigo	Estado	Municipio	Responsavel	EstacaoTipo	EstacaoNome	Latitude	Longitude	Data	Origemdosdados	Maxima	Total	DiaMaxima	DiasDeChuva	Dia0
2	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Jan-18	Pluviômetro	30	172	18	15	
3	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Feb-18	Pluviômetro	8	16	13	3	
4	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Mar-18	Pluviômetro	20	98	19	10	
5	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Apr-18	Pluviômetro	8	12	3	2	
6	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	May-18	Pluviômetro	0	0	1	0	
7	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Jun-18	Pluviômetro	2	3	11	2	
8	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Jul-18	Pluviômetro	8	39	17	9	
9	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Aug-18	Pluviômetro	10	26	25	7	
10	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Sep-18	Pluviômetro	20	74	18	5	
11	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Jan-20	Pluviômetro	7	0	4	0	
12	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Feb-20	Pluviômetro	19	96	10	15	
13	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Mar-20	Pluviômetro	2	3	6	2	
14	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Apr-20	Pluviômetro	0	0	1	0	
15	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	May-20	Pluviômetro	4	10	22	4	
16	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Jun-20	Pluviômetro	4	11	5	5	
17	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Jul-20	Pluviômetro	4	11	25	5	
18	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Aug-20	Pluviômetro	10	40	19	9	
19	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Sep-20	Pluviômetro	9.8	39.8	19	8	
20	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Oct-20	Pluviômetro	8	41	10	9	
21	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Nov-20	Pluviômetro	6	16	8	5	
22	2347059	Bruto	62	SP	SOROCABA	ANA	Pluviométrica	EDEN(PIRAJIBU)	-23:25:13	-047:24:16	Dec-20	Pluviômetro	12	55.8	27	11	

9.3 Figura 3 - Banco de Dados – Modelo Híbrido – Tabelas.



9.4 Figura 4 - Modelos - Arquivos – TrainingGroup, Testing e CheckingGroup.

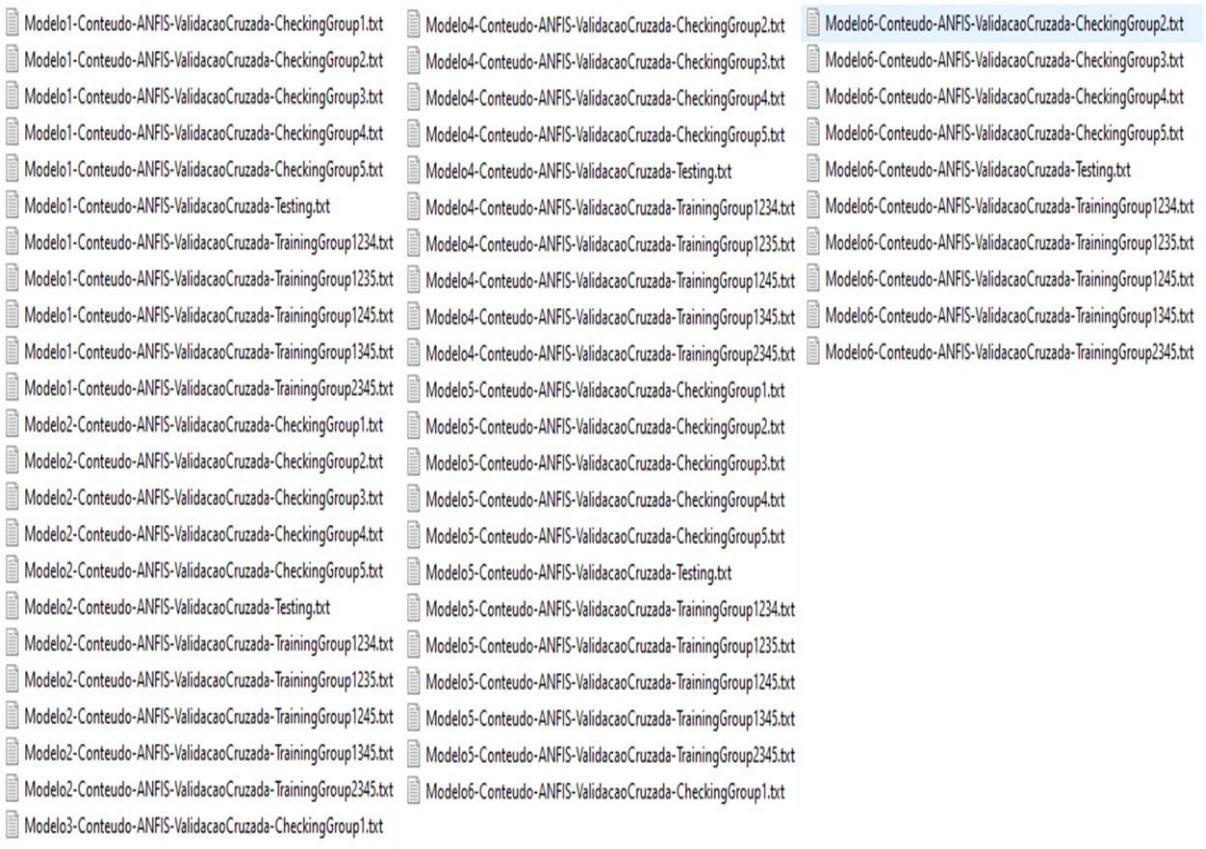
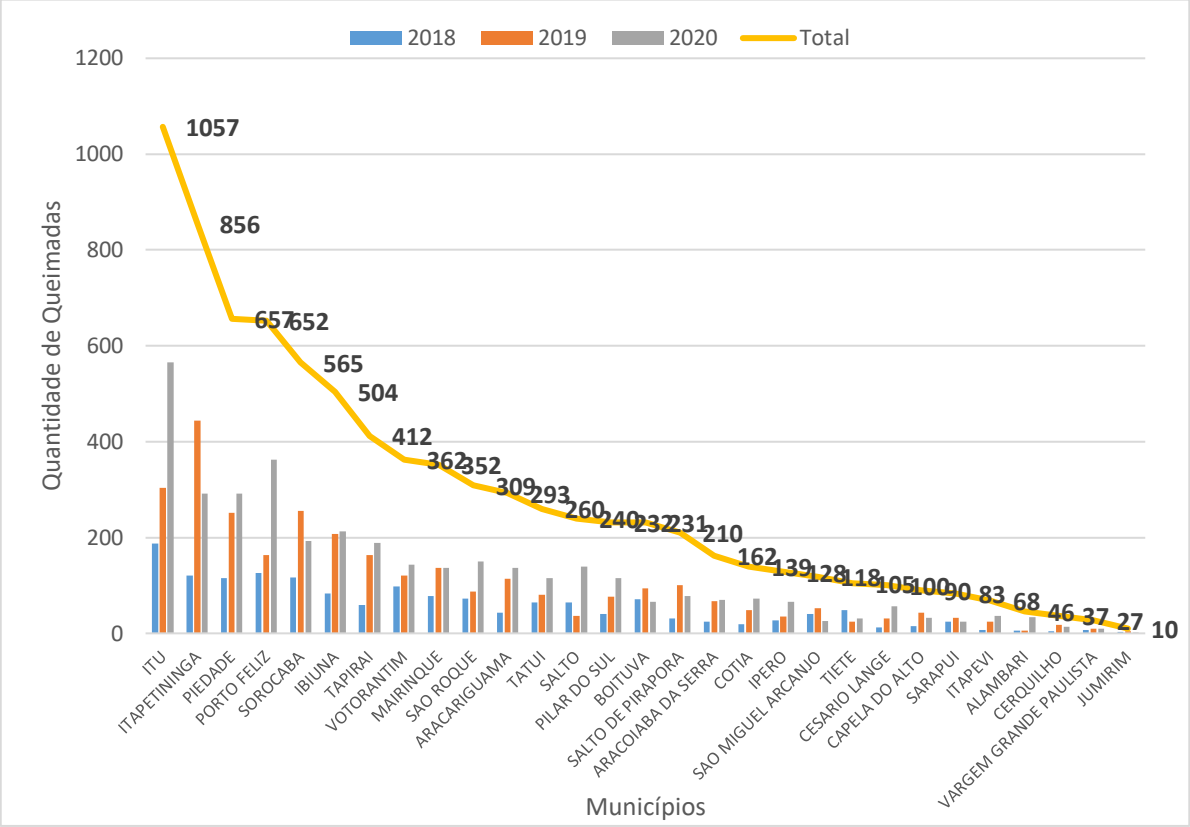


Figura 5 – Gráfico – RMS - Queimadas por Município.



10. ANEXOS

10.1 Tabela 1 - Base dados – Amostra aleatória de observações selecionadas para o estudo.

DiaSemChuva	PrecipitacaoUmDiaAnterior	RiscoFogoObservadoParaUmDia
0.04	0.77	0.88
0.04	0.77	0.88
0.04	0.77	0.88
0.04	0.77	0.88
0.04	0.77	0.88
0.04	0.77	1
0.04	0.77	1
0.04	0.77	0.88
0.99	0.77	0.88
0.17	0.77	0.88
0.17	0.77	0.88
1	0.77	0.88
0.11	0.77	0.88
0.11	0.77	0.88
0.32	0.77	0.88
0.67	0.77	0.88
0.11	0.77	0.88
0.79	0.77	0.88
0.26	0.77	0.88
0.84	0.77	0.88
0.36	0.77	1
0.91	0.77	0.88
0.95	0.77	1
0.55	0.77	0.95
0.58	0.77	0.94
0.58	0.77	0.94
0.63	0.92	0.88
0.74	0.77	0.88

10.2 Script 1 – Microsoft SQL Server - Criação – Modelo 3.

```
-- Acessando o Banco de Dados --
Use ProjetoDWQueimadas
Go

-- Verificando a existência das tabelas --
If Exists (Select name From Sys.tables Where Name = 'AnaliseQueimadasSelecionadas' Or N
ame = 'AnaliseQueimadasSelecionadasTrainingGroup1' Or Name = 'AnaliseQueimadasSeleci
onadasTrainingGroup2345')
Begin
 Drop Table AnaliseQueimadasSelecionadas
 Drop Table AnaliseQueimadasSelecionadasTesting
 Drop Table AnaliseQueimadasSelecionadasTraining
 Drop Table AnaliseQueimadasSelecionadasTrainingGroup1
 Drop Table AnaliseQueimadasSelecionadasTrainingGroup2
 Drop Table AnaliseQueimadasSelecionadasTrainingGroup3
 Drop Table AnaliseQueimadasSelecionadasTrainingGroup4
 Drop Table AnaliseQueimadasSelecionadasTrainingGroup5
 Drop Table AnaliseQueimadasSelecionadasTrainingGroup2345
 Drop Table AnaliseQueimadasSelecionadasTrainingGroup1345
 Drop Table AnaliseQueimadasSelecionadasTrainingGroup1245
 Drop Table AnaliseQueimadasSelecionadasTrainingGroup1235
 Drop Table AnaliseQueimadasSelecionadasTrainingGroup1234
End

-- Criando a Tabela AnaliseQueimadasSelecionadas --
Create Table AnaliseQueimadasSelecionadas
(CodigoAnaliseQueimadasSelecionadas Int Identity(1,1) Primary Key Clustered,
 DataAnaliseQueimadasSelecionadas Date,
 DiaSemChuva Float Not Null,
 PrecipitacaoUmDiaAnterior Float,
 DataRiscoFogoObservadoTresDiasPraFrente Date,
 RiscoFogoObservadoTresDiasPraFrente Float)
Go

-- Criando a Tabela AnaliseQueimadasSelecionadasTesting --
Create Table AnaliseQueimadasSelecionadasTesting
(CodigoAnaliseQueimadasSelecionadasTesting Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTesting Date,
 DiaSemChuvaTesting Float Not Null,
 PrecipitacaoUmDiaAnteriorTesting Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTesting Date,
 RiscoFogoObservadoTresDiasPraFrenteTesting Float)
Go
```

```
-- Criando a Tabela AnaliseQueimadasSelecionadasTraining --
Create Table AnaliseQueimadasSelecionadasTraining
(CodigoAnaliseQueimadasSelecionadasTraining Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTraining Date,
 DiaSemChuvaTraining Float Not Null,
 PrecipitacaoUmDiaAnteriorTraining Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTraining Date,
 RiscoFogoObservadoTresDiasPraFrenteTraining Float)
Go

-- Criando a Tabela AnaliseQueimadasSelecionadasTraining --
Create Table AnaliseQueimadasSelecionadasTrainingGroup1
(CodigoAnaliseQueimadasSelecionadasTrainingGroup1 Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTrainingGroup1 Date,
 DiaSemChuvaTrainingGroup1 Float Not Null,
 PrecipitacaoUmDiaAnteriorTrainingGroup1 Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTrainingGroup1 Date,
 RiscoFogoObservadoTresDiasPraFrenteTrainingGroup1 Float)
Go

-- Criando a Tabela AnaliseQueimadasSelecionadasTraining --
Create Table AnaliseQueimadasSelecionadasTrainingGroup2
(CodigoAnaliseQueimadasSelecionadasTrainingGroup2 Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTrainingGroup2 Date,
 DiaSemChuvaTrainingGroup2 Float Not Null,
 PrecipitacaoUmDiaAnteriorTrainingGroup2 Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTrainingGroup2 Date,
 RiscoFogoObservadoTresDiasPraFrenteTrainingGroup2 Float)
Go

-- Criando a Tabela AnaliseQueimadasSelecionadasTraining --
Create Table AnaliseQueimadasSelecionadasTrainingGroup3
(CodigoAnaliseQueimadasSelecionadasTrainingGroup3 Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTrainingGroup3 Date,
 DiaSemChuvaTrainingGroup3 Float Not Null,
 PrecipitacaoUmDiaAnteriorTrainingGroup3 Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTrainingGroup3 Date,
 RiscoFogoObservadoTresDiasPraFrenteTrainingGroup3 Float)
Go

-- Criando a Tabela AnaliseQueimadasSelecionadasTraining --
Create Table AnaliseQueimadasSelecionadasTrainingGroup4
(CodigoAnaliseQueimadasSelecionadasTrainingGroup4 Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTrainingGroup4 Date,
 DiaSemChuvaTrainingGroup4 Float Not Null,
 PrecipitacaoUmDiaAnteriorTrainingGroup4 Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTrainingGroup4 Date,
 RiscoFogoObservadoTresDiasPraFrenteTrainingGroup4 Float)
Go
```

```
-- Criando a Tabela AnaliseQueimadasSelecionadasTraining --
Create Table AnaliseQueimadasSelecionadasTrainingGroup5
(CodigoAnaliseQueimadasSelecionadasTrainingGroup5 Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTrainingGroup5 Date,
 DiaSemChuvaTrainingGroup5 Float Not Null,
 PrecipitacaoUmDiaAnteriorTrainingGroup5 Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTrainingGroup5 Date,
 RiscoFogoObservadoTresDiasPraFrenteTrainingGroup5 Float)
Go

-- Criando a Tabela AnaliseQueimadasSelecionadasTraining2345 -
- CheckingTrainingGroup1 --
Create Table AnaliseQueimadasSelecionadasTrainingGroup2345
(CodigoAnaliseQueimadasSelecionadasTrainingGroup2345 Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTrainingGroup2345 Date,
 DiaSemChuvaTrainingGroup2345 Float Not Null,
 PrecipitacaoUmDiaAnteriorTrainingGroup2345 Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTrainingGroup2345 Date,
 RiscoFogoObservadoTresDiasPraFrenteTrainingGroup2345 Float)
Go

-- Criando a Tabela AnaliseQueimadasSelecionadasTraining1345 -- -
- CheckingTrainingGroup2 --
Create Table AnaliseQueimadasSelecionadasTrainingGroup1345
(CodigoAnaliseQueimadasSelecionadasTrainingGroup1345 Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTrainingGroup1345 Date,
 DiaSemChuvaTrainingGroup1345 Float Not Null,
 PrecipitacaoUmDiaAnteriorTrainingGroup1345 Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTrainingGroup1345 Date,
 RiscoFogoObservadoTresDiasPraFrenteTrainingGroup1345 Float)
Go

-- Criando a Tabela AnaliseQueimadasSelecionadasTraining1245 -- -
- CheckingTrainingGroup3 --
Create Table AnaliseQueimadasSelecionadasTrainingGroup1245
(CodigoAnaliseQueimadasSelecionadasTrainingGroup1245 Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTrainingGroup1245 Date,
 DiaSemChuvaTrainingGroup1245 Float Not Null,
 PrecipitacaoUmDiaAnteriorTrainingGroup1245 Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTrainingGroup1245 Date,
 RiscoFogoObservadoTresDiasPraFrenteTrainingGroup1245 Float)
Go
```

```
-- Criando a Tabela AnaliseQueimadasSelecionadasTraining1235 -- -
- CheckingTrainingGroup4 --
Create Table AnaliseQueimadasSelecionadasTrainingGroup1235
(CodigoAnaliseQueimadasSelecionadasTrainingGroup1235 Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTrainingGroup1235 Date,
 DiaSemChuvaTrainingGroup1235 Float Not Null,
 PrecipitacaoUmDiaAnteriorTrainingGroup1235 Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTrainingGroup1235 Date,
 RiscoFogoObservadoTresDiasPraFrenteTrainingGroup1235 Float)
Go

-- Criando a Tabela AnaliseQueimadasSelecionadasTraining1234 -- -
- CheckingTrainingGroup5 --
Create Table AnaliseQueimadasSelecionadasTrainingGroup1234
(CodigoAnaliseQueimadasSelecionadasTrainingGroup1234 Int Primary Key Clustered,
 DataAnaliseQueimadasSelecionadasTrainingGroup1234 Date,
 DiaSemChuvaTrainingGroup1234 Float Not Null,
 PrecipitacaoUmDiaAnteriorTrainingGroup1234 Float,
 DataRiscoFogoObservadoTresDiasPraFrenteTrainingGroup1234 Date,
 RiscoFogoObservadoTresDiasPraFrenteTrainingGroup1234 Float)
Go

-- Limpando as Tabelas --
If ((Select Count(CodigoAnaliseQueimadasSelecionadas) From AnaliseQueimadasSeleciona
das) = 1 Or (Select Count(CodigoAnaliseQueimadasSelecionadasTesting) From AnaliseQuei
madasSelecionadasTesting) = 1)
Begin
 Truncate Table AnaliseQueimadasSelecionadas
 Truncate Table AnaliseQueimadasSelecionadasTesting
 Truncate Table AnaliseQueimadasSelecionadasTraining
 Truncate Table AnaliseQueimadasSelecionadasTrainingGroup1
 Truncate Table AnaliseQueimadasSelecionadasTrainingGroup2
 Truncate Table AnaliseQueimadasSelecionadasTrainingGroup3
 Truncate Table AnaliseQueimadasSelecionadasTrainingGroup4
 Truncate Table AnaliseQueimadasSelecionadasTrainingGroup5
 Truncate Table AnaliseQueimadasSelecionadasTrainingGroup2345
 Truncate Table AnaliseQueimadasSelecionadasTrainingGroup1345
 Truncate Table AnaliseQueimadasSelecionadasTrainingGroup1245
 Truncate Table AnaliseQueimadasSelecionadasTrainingGroup1235
 Truncate Table AnaliseQueimadasSelecionadasTrainingGroup1234
End
Go
```

```
-- Realizando a Análise e Inserindo na Tabela AnaliseQueimadasSelecionadas --
- Indicador 1 - Registro de Queimadas(INPE) para o Município de Sorocaba entre os anos de 2018 e 2020 --
Insert Into AnaliseQueimadasSelecionadas (DataAnaliseQueimadasSelecionadas, DiaSemChuva, PrecipitacaoUmDiaAnterior, DataRiscoFogoObservadoTresDiasPraFrente, RiscoFogoObservadoTresDiasPraFrente)
Select Q.DataHora As Data,
       Q.DiaSemChuva,
       (Select ValorPrecipitacao From VolumePrecipitacao2018x2020
        Where DataPrecipitacao = DateAdd(Day,-1,Q.DataHora)) As PrecipitacaoUmDiaAntes,
       DateAdd(Day,3,Q.DataHora) As DataRiscoFogoObservadoTresDiasPraFrente,
       IsNull(R.RiscoFogoObservadoTresDiasPraFrente,0) As RiscoFogoObservadoTresDiasPraFrente
From Queimadas2018x2020Sorocaba Q Outer Apply (Select Top 1 IsNull(RiscoFogo,0) From Queimadas2018x2020Sorocaba
                                               Where DataHora = DateAdd(Day,+3,Q.DataHora)) As R(RiscoFogoObservadoTresDiasPraFrente)
Order By Data Asc
Go

- Indicador 2 - Registro de Precipitação(ANA) para o Município de Sorocaba entre os anos de 2018 e 2020 --
Insert Into AnaliseQueimadasSelecionadas (DataAnaliseQueimadasSelecionadas, DiaSemChuva, PrecipitacaoUmDiaAnterior, DataRiscoFogoObservadoTresDiasPraFrente,RiscoFogoObservadoTresDiasPraFrente)
Select V.DataPrecipitacao As Data,
       (Select Abs(Count(CodigoVolumePrecipitacao)-1) From VolumePrecipitacao2018x2020
        Where DataPrecipitacao >= (Select Convert(Char(10),DateAdd(dd,-(Day(V.DataPrecipitacao)-1),V.DataPrecipitacao),23)
                                   From VolumePrecipitacao2018x2020
                                   Where DataPrecipitacao = V.DataPrecipitacao)
        And DataPrecipitacao <=V.DataPrecipitacao
        And ValorPrecipitacao =0) As DiaSemChuva,
       (Select Sum(ValorPrecipitacao) From VolumePrecipitacao2018x2020
        Where DataPrecipitacao = DateAdd(Day,-1,V.DataPrecipitacao)) As PrecipitacaoUmDiaAntes,
       DateAdd(Day,3,V.DataPrecipitacao) As DataRiscoFogoObservadoTresDiasPraFrente,
       IsNull(R.RiscoFogoObservadoTresDiasPraFrente,0) As RiscoFogoObservadoTresDiasPraFrente
From VolumePrecipitacao2018x2020 V Outer Apply (Select Top 1 RiscoFogo From Queimadas2018x2020Sorocaba
                                                Where DataHora = DateAdd(Day,+3,V.DataPrecipitacao)) As R(RiscoFogoObservadoTresDiasPraFrente)
Where V.ValorPrecipitacao > 0
Order By Data Asc
Go
```

```
- Indicador 3 - Registro de Precipitação(ANA) e Registro de Queimadas(INPE) para o Municí
pio de Sorocaba entre os anos de 2018 e 2020 --
Insert Into AnaliseQueimadasSelecionadas (DataAnaliseQueimadasSelecionadas, DiaSemCh
uva, PrecipitacaoUmDiaAnterior, DataRiscoFogoObservadoTresDiasPraFrente,RiscoFogoOb
servadoTresDiasPraFrente)
Select Distinct V.DataPrecipitacao As Data,
       Q.DiaSemChuva,
       (Select ValorPrecipitacao From VolumePrecipitacao2018x2020
        Where DataPrecipitacao = DateAdd(Day,-
1,V.DataPrecipitacao)) As PrecipitacaoUmDiaAntes,
       DateAdd(Day,3,V.DataPrecipitacao) As DataRiscoFogoObservadoTresDiasPraFrente,
       IsNull(R.RiscoFogoObservadoTresDiasPraFrente,0) As RiscoFogoObservadoTresDias
PraFrente
From VolumePrecipitacao2018x2020 V Inner Join Queimadas2018x2020Sorocaba Q
                                    On V.DataPrecipitacao = Q.DataHora
                                   Outer Apply (Select Top 1 RiscoFogo From Queimad
as2018x2020Sorocaba
                                                Where DataHora = DateAdd(Day,+3,V.
DataPrecipitacao)) As R(RiscoFogoObservadoTresDiasPraFrente)Where V.ValorPrecipitaca
o >0
Order By Data Asc
Go

- Indicador 4 - Sem Registro de Precipitação(ANA) e Sem Registro de Queimadas(INPE) par
a o Município de Sorocaba entre os anos de 2018 e 2020 --
Insert Into AnaliseQueimadasSelecionadas (DataAnaliseQueimadasSelecionadas, DiaSemCh
uva, PrecipitacaoUmDiaAnterior, DataRiscoFogoObservadoTresDiasPraFrente,RiscoFogoOb
servadoTresDiasPraFrente)
Select C.DataDatasSemChuvaOuQueimadas,
       (Select Count(CodigoVolumePrecipitacao)+1 From VolumePrecipitacao2018x2020
        Where DataPrecipitacao >= (Select Convert(Char(10),DateAdd(dd,-
(Day(V.DataPrecipitacao)-1),V.DataPrecipitacao),23)
                                   From VolumePrecipitacao2018x2020
                                   Where DataPrecipitacao = V.DataPrecipitacao)
         And DataPrecipitacao <=V.DataPrecipitacao
         And ValorPrecipitacao =0) As DiaSemChuva,
       V.ValorPrecipitacao,
       DateAdd(Day,3,V.DataPrecipitacao) As DataRiscoFogoObservadoTresDiasPraFrente,
       IsNull(R.RiscoFogoObservadoTresDiasPraFrente,0) As RiscoFogoObservadoTresDias
PraFrente
From DatasSemChuvaOuQueimadas C Inner Join VolumePrecipitacao2018x2020 V
                                    On V.DataPrecipitacao = DateAdd(Day,-
1,C.DataDatasSemChuvaOuQueimadas)
                                   Outer Apply (Select Top 1 RiscoFogo From Queimad
as2018x2020Sorocaba
                                                Where DataHora = DateAdd(Day,+3,V.
DataPrecipitacao)) As R(RiscoFogoObservadoTresDiasPraFrente)
Order By C.DataDatasSemChuvaOuQueimadas
Go
```

```
-- Normalizando os Dados --
;With CTENormalizacaoDeDados (Codigo, DiaSemChuva, PrecipitacaoUmDiaAnterior, Ris
coFogoObservadoTresDiasPraFrente)
As
(
Select CodigoAnaliseQueimadasSelecionadas,
        Round(CUME_DIST() Over(Order By DiaSemChuva),2) As DiaSemChuva,
        Round(CUME_DIST() Over(Order By PrecipitacaoUmDiaAnterior),2) As Precipitaca
oUmDiaAnterior,
        Round(CUME_DIST() Over(Order By RiscoFogoObservadoTresDiasPraFrente),2) As
 RiscoFogoObservadoTresDiasPraFrente
From AnaliseQueimadasSelecionadas)
Update AnaliseQueimadasSelecionadas
Set DiaSemChuva = C.DiaSemChuva,
    PrecipitacaoUmDiaAnterior = C.PrecipitacaoUmDiaAnterior,
    RiscoFogoObservadoTresDiasPraFrente = C.RiscoFogoObservadoTresDiasPraFrente
From AnaliseQueimadasSelecionadas A Inner Join CTENormalizacaoDeDados C
                                        On A.CodigoAnaliseQueimadasSelecionadas = C.C
odigo
Go

--
Selecionando 30% dos registros existentes na Tabela AnaliseQueimadasSelecionadas para ser
em inseridos na Tabela AnaliseQueimadasSelecionadasTesting --
Insert Into AnaliseQueimadasSelecionadasTesting(CodigoAnaliseQueimadasSelecionadasTe
sting, DataAnaliseQueimadasSelecionadasTesting,
                                                DiaSemChuvaTesting, PrecipitacaoUmDi
aAnteriorTesting,
                                                DataRiscoFogoObservadoTresDiasPraFre
nteTesting ,RiscoFogoObservadoTresDiasPraFrenteTesting)
Select Top 30 Percent CodigoAnaliseQueimadasSelecionadas, DataAnaliseQueimadasSeleci
onadas, DiaSemChuva, PrecipitacaoUmDiaAnterior,
                        DataRiscoFogoObservadoTresDiasPraFrente, RiscoFogoObservadoTre
sDiasPraFrente
From AnaliseQueimadasSelecionadas
Order By NewId()
Go

- Armazenando o restantes dos registros existentes na Tabela AnaliseQueimadasSelecionadas
 para serem inseridos na Tabela AnaliseQueimadasSelecionadasTraining --
Insert Into AnaliseQueimadasSelecionadasTraining(CodigoAnaliseQueimadasSelecionadasTr
aining, DataAnaliseQueimadasSelecionadasTraining,
                                                DiaSemChuvaTraining, PrecipitacaoUm
DiaAnteriorTraining,
                                                DataRiscoFogoObservadoTresDiasPraFr
enteTraining, RiscoFogoObservadoTresDiasPraFrenteTraining)
Select  CodigoAnaliseQueimadasSelecionadas, DataAnaliseQueimadasSelecionadas, DiaSem
Chuva, PrecipitacaoUmDiaAnterior,
        DataRiscoFogoObservadoTresDiasPraFrente ,RiscoFogoObservadoTresDiasPraFrent
e
```

```
From AnaliseQueimadasSelecionadas
Where CodigoAnaliseQueimadasSelecionadas Not In (Select CodigoAnaliseQueimadasSelec
ionadasTesting From AnaliseQueimadasSelecionadasTesting)
Order By NewId()
Go

-- Realizando a Separação dos dados para cada Group de Training -- Group1 --
Insert Into AnaliseQueimadasSelecionadasTrainingGroup1(CodigoAnaliseQueimadasSelecio
nadasTrainingGroup1, DataAnaliseQueimadasSelecionadasTrainingGroup1,
                                                    DiaSemChuvaTrainingGroup1, Pr
ecipitacaoUmDiaAnteriorTrainingGroup1,
                                                    DataRiscoFogoObservadoTresDia
sPraFrenteTrainingGroup1,
                                                    RiscoFogoObservadoTresDiasPra
FrenteTrainingGroup1)
Select Top (Select Count(CodigoAnaliseQueimadasSelecionadasTraining)/5 From AnaliseQu
eimadasSelecionadasTraining)
            CodigoAnaliseQueimadasSelecionadasTraining, DataAnaliseQueimadasSelecion
adasTraining, DiaSemChuvaTraining, PrecipitacaoUmDiaAnteriorTraining,
            DataRiscoFogoObservadoTresDiasPraFrenteTraining, RiscoFogoObservadoTres
DiasPraFrenteTraining
From AnaliseQueimadasSelecionadasTraining
Order By NewId()
Go

-- Realizando a Separação dos dados para cada Group de Training -- Group2 --
Insert Into AnaliseQueimadasSelecionadasTrainingGroup2(CodigoAnaliseQueimadasSelecio
nadasTrainingGroup2, DataAnaliseQueimadasSelecionadasTrainingGroup2,
                                                    DiaSemChuvaTrainingGroup2, Pr
ecipitacaoUmDiaAnteriorTrainingGroup2,
                                                    DataRiscoFogoObservadoTresDia
sPraFrenteTrainingGroup2,
                                                    RiscoFogoObservadoTresDiasPra
FrenteTrainingGroup2)
Select Top (Select Count(CodigoAnaliseQueimadasSelecionadasTraining)/5 From AnaliseQu
eimadasSelecionadasTraining)
            CodigoAnaliseQueimadasSelecionadasTraining, DataAnaliseQueimadasSeleciona
dasTraining, DiaSemChuvaTraining, PrecipitacaoUmDiaAnteriorTraining,
            DataRiscoFogoObservadoTresDiasPraFrenteTraining, RiscoFogoObservadoTres
DiasPraFrenteTraining
From AnaliseQueimadasSelecionadasTraining
Where CodigoAnaliseQueimadasSelecionadasTraining Not In (Select CodigoAnaliseQueima
dasSelecionadasTrainingGroup1 From AnaliseQueimadasSelecionadasTrainingGroup1)
Order By NewId()
Go
```

```
-- Realizando a Separação dos dados para cada Group de Training -- Group3 --
Insert Into AnaliseQueimadasSelecionadasTrainingGroup3(CodigoAnaliseQueimadasSelecio
nadasTrainingGroup3, DataAnaliseQueimadasSelecionadasTrainingGroup3,
                                                    DiaSemChuvaTrainingGroup3, Pr
ecipitacaoUmDiaAnteriorTrainingGroup3,
                                                    DataRiscoFogoObservadoTresDia
sPraFrenteTrainingGroup3,
                                                    RiscoFogoObservadoTresDiasPra
FrenteTrainingGroup3)
Select Top (Select Count(CodigoAnaliseQueimadasSelecionadasTraining)/5 From AnaliseQu
eimadasSelecionadasTraining)
            CodigoAnaliseQueimadasSelecionadasTraining, DataAnaliseQueimadasSeleciona
dasTraining, DiaSemChuvaTraining, PrecipitacaoUmDiaAnteriorTraining,
            DataRiscoFogoObservadoTresDiasPraFrenteTraining, RiscoFogoObservadoTres
DiasPraFrenteTraining
From AnaliseQueimadasSelecionadasTraining
Where CodigoAnaliseQueimadasSelecionadasTraining Not In (Select CodigoAnaliseQueima
dasSelecionadasTrainingGroup1 From AnaliseQueimadasSelecionadasTrainingGroup1)
And  CodigoAnaliseQueimadasSelecionadasTraining Not In (Select CodigoAnaliseQueimada
sSelecionadasTrainingGroup2 From AnaliseQueimadasSelecionadasTrainingGroup2)
Order By NewId()
Go

-- Realizando a Separação dos dados para cada Group de Training -- Group4 --
Insert Into AnaliseQueimadasSelecionadasTrainingGroup4(CodigoAnaliseQueimadasSelecio
nadasTrainingGroup4, DataAnaliseQueimadasSelecionadasTrainingGroup4,
                                                    DiaSemChuvaTrainingGroup4, P
recipitacaoUmDiaAnteriorTrainingGroup4,
                                                    DataRiscoFogoObservadoTresDi
asPraFrenteTrainingGroup4,
                                                    RiscoFogoObservadoTresDiasPra
FrenteTrainingGroup4)
Select Top (Select Count(CodigoAnaliseQueimadasSelecionadasTraining)/5 From AnaliseQu
eimadasSelecionadasTraining)
            CodigoAnaliseQueimadasSelecionadasTraining, DataAnaliseQueimadasSelecion
adasTraining, DiaSemChuvaTraining, PrecipitacaoUmDiaAnteriorTraining,
            DataRiscoFogoObservadoTresDiasPraFrenteTraining, RiscoFogoObservadoTres
DiasPraFrenteTraining
From AnaliseQueimadasSelecionadasTraining
Where CodigoAnaliseQueimadasSelecionadasTraining Not In (Select CodigoAnaliseQueima
dasSelecionadasTrainingGroup1 From AnaliseQueimadasSelecionadasTrainingGroup1)
And  CodigoAnaliseQueimadasSelecionadasTraining Not In (Select CodigoAnaliseQueimada
sSelecionadasTrainingGroup2 From AnaliseQueimadasSelecionadasTrainingGroup2)
And  CodigoAnaliseQueimadasSelecionadasTraining Not In (Select CodigoAnaliseQueimada
sSelecionadasTrainingGroup3 From AnaliseQueimadasSelecionadasTrainingGroup3)
Order By NewId()
Go
```

```
-- Realizando a Separação dos dados para cada Group de Training -- Group5 --
Insert Into AnaliseQueimadasSelecionadasTrainingGroup5(CodigoAnaliseQueimadasSelecio
nadasTrainingGroup5, DataAnaliseQueimadasSelecionadasTrainingGroup5,
                                                    DiaSemChuvaTrainingGroup5, Pr
ecipitacaoUmDiaAnteriorTrainingGroup5,
                                                    DataRiscoFogoObservadoTresDia
sPraFrenteTrainingGroup5,
                                                    RiscoFogoObservadoTresDiasPra
FrenteTrainingGroup5)
Select Top (Select Count(CodigoAnaliseQueimadasSelecionadasTraining)/5 From AnaliseQu
eimadasSelecionadasTraining)
            CodigoAnaliseQueimadasSelecionadasTraining, DataAnaliseQueimadasSeleciona
dasTraining, DiaSemChuvaTraining, PrecipitacaoUmDiaAnteriorTraining,
            DataRiscoFogoObservadoTresDiasPraFrenteTraining, RiscoFogoObservadoTres
DiasPraFrenteTraining
From AnaliseQueimadasSelecionadasTraining
Where CodigoAnaliseQueimadasSelecionadasTraining Not In (Select CodigoAnaliseQueima
dasSelecionadasTrainingGroup1 From AnaliseQueimadasSelecionadasTrainingGroup1)
And  CodigoAnaliseQueimadasSelecionadasTraining Not In (Select CodigoAnaliseQueimada
sSelecionadasTrainingGroup2 From AnaliseQueimadasSelecionadasTrainingGroup2)
And  CodigoAnaliseQueimadasSelecionadasTraining Not In (Select CodigoAnaliseQueimada
sSelecionadasTrainingGroup3 From AnaliseQueimadasSelecionadasTrainingGroup3)
And  CodigoAnaliseQueimadasSelecionadasTraining Not In (Select CodigoAnaliseQueimada
sSelecionadasTrainingGroup4 From AnaliseQueimadasSelecionadasTrainingGroup4)
Order By NewId()
Go

-- Carregando os dados no arquivo de Testing --
AnaliseQueimadasSelecionadasTesting - DiaSemChuva, PrecipitacaoUmDiaAnterior, Acum
uladoPrecipitacaoTresDiasAnteriores, Chuva, Queimada --
Declare @BCPComando Varchar(500)
Set @BCPComando='bcp "Select DiaSemChuvaTesting, PrecipitacaoUmDiaAnteriorTesting,
 RiscoFogoObservadoTresDiasPraFrenteTesting From ProjetoDWQueimadas..AnaliseQueim
adasSelecionadasTesting" queryout "S:\Mestrado\Defesa\Dados\Modelos\Modelo6-
Conteudo-ANFIS-ValidacaoCruzada-Testing.txt" -c -UTF8 -T -
SWIN10PRO\WIN10PROSQL2019'
Exec master..xp_cmdshell @BCPComando, No_Output
Print 'Arquivo: Modelo6-Conteudo-ANFIS-ValidacaoCruzada-Testing.txt criado.'
Go

-- Carregando os dados no arquivo de CheckingGroup1 -- Tabela TrainingGroup1 --
- AnaliseQueimadasSelecionadasTraining - DiaSemChuva, PrecipitacaoUmDiaAnterior, Acu
muladoPrecipitacaoTresDiasAnteriores, Chuva, Queimada, RiscoFogoObservadoTresDiasPra
Frente --
Declare @BCPComando Varchar(500)
Set @BCPComando='bcp "Select DiaSemChuvaTrainingGroup1, PrecipitacaoUmDiaAnterio
rTrainingGroup1, RiscoFogoObservadoTresDiasPraFrenteTrainingGroup1 From ProjetoDW
Queimadas..AnaliseQueimadasSelecionadasTrainingGroup1" queryout "S:\Mestrado\Defesa\
Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup1.txt" -c -
UTF8 -T -SWIN10PRO\WIN10PROSQL2019'
```

```
Exec master..xp_cmdshell @BCPComando, No_Output
Print 'Arquivo: Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup1.txt criado.'
Go

-- Montando e Inserindo o conteúdo para o TrainingGroup2345 -
- Checking TrainingGroup1 --
Insert Into AnaliseQueimadasSelecionadasTrainingGroup2345
Select * From AnaliseQueimadasSelecionadasTrainingGroup2
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup3
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup4
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup5
Go


-- Carregando os dados no arquivo de TrainingGroup2345 --
- AnaliseQueimadasSelecionadasTraining - DiaSemChuva, PrecipitacaoUmDiaAnterior, Acu
muladoPrecipitacaoTresDiasAnteriores, Chuva, Queimada, RiscoFogoObservadoTresDiasPra
Frente --
Declare @BCPComando Varchar(500)
Set @BCPComando='bcp "Select DiaSemChuvaTrainingGroup2345, PrecipitacaoUmDiaAnt
eriorTrainingGroup2345, RiscoFogoObservadoTresDiasPraFrenteTrainingGroup2345 From
ProjetoDWQueimadas..AnaliseQueimadasSelecionadasTrainingGroup2345" queryout "S:\Me
strado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-
TrainingGroup2345.txt" -c -UTF8 -T -SWIN10PRO\WIN10PROSQL2019'
Exec master..xp_cmdshell @BCPComando, No_Output
Print 'Arquivo: Modelo6-Conteudo-ANFIS-ValidacaoCruzada-
TrainingGroup2345.txt criado.'
Go

-- Carregando os dados no arquivo de CheckingGroup2 -- Tabela TrainingGroup2 --
- AnaliseQueimadasSelecionadasTraining - DiaSemChuva, PrecipitacaoUmDiaAnterior, Acu
muladoPrecipitacaoTresDiasAnteriores, Chuva, Queimada, RiscoFogoObservadoTresDiasPra
Frente --
Declare @BCPComando Varchar(500)
Set @BCPComando='bcp "Select DiaSemChuvaTrainingGroup2, PrecipitacaoUmDiaAnterio
rTrainingGroup2, RiscoFogoObservadoTresDiasPraFrenteTrainingGroup2 From ProjetoDW
Queimadas..AnaliseQueimadasSelecionadasTrainingGroup2" queryout "S:\Mestrado\Defesa\
Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup2.txt" -c -
UTF8 -T -SWIN10PRO\WIN10PROSQL2019'
Exec master..xp_cmdshell @BCPComando, No_Output
Print 'Arquivo: Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup2.txt criado.'
Go
```

```
-- Montando e Inserindo o conteúdo para o TrainingGroup1345 - Checking TrainingGroup2 -
Insert Into AnaliseQueimadasSelecionadasTrainingGroup1345
Select * From AnaliseQueimadasSelecionadasTrainingGroup1
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup3
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup4
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup5
Go

-- Carregando os dados no arquivo de TrainingGroup1345 --
- AnaliseQueimadasSelecionadasTraining - DiaSemChuva, PrecipitacaoUmDiaAnterior, Acu
muladoPrecipitacaoTresDiasAnteriores, Chuva, Queimada, RiscoFogoObservadoTresDiasPra
Frente --
Declare @BCPComando Varchar(500)
Set @BCPComando='bcp "Select DiaSemChuvaTrainingGroup1345, PrecipitacaoUmDiaAnt
eriorTrainingGroup1345, RiscoFogoObservadoTresDiasPraFrenteTrainingGroup1345 From
ProjetoDWQueimadas..AnaliseQueimadasSelecionadasTrainingGroup1345" queryout "S:\Me
strado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-
TrainingGroup1345.txt" -c -UTF8 -T -SWIN10PRO\WIN10PROSQL2019'
Exec master..xp_cmdshell @BCPComando, No_Output
Print 'Arquivo: Modelo6-Conteudo-ANFIS-ValidacaoCruzada-
CheckingGroup1345.txt criado.'
Go

-- Carregando os dados no arquivo de CheckingGroup3 -- Tabela TrainingGroup3 -------
AnaliseQueimadasSelecionadasTraining - DiaSemChuva, PrecipitacaoUmDiaAnterior, Acum
uladoPrecipitacaoTresDiasAnteriores, Chuva, Queimada, RiscoFogoObservadoTresDiasPraF
rente --
Declare @BCPComando Varchar(500)
Set @BCPComando='bcp "Select DiaSemChuvaTrainingGroup3, PrecipitacaoUmDiaAnterio
rTrainingGroup3, RiscoFogoObservadoTresDiasPraFrenteTrainingGroup3 From ProjetoDW
Queimadas..AnaliseQueimadasSelecionadasTrainingGroup3" queryout "S:\Mestrado\Defesa\
Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup3.txt" -c -
UTF8 -T -SWIN10PRO\WIN10PROSQL2019'
Exec master..xp_cmdshell @BCPComando, No_Output
Print 'Arquivo: Modelo6-Conteudo-ANFIS-ValidacaoCruzada-TrainingGroup3.txt criado.'
Go
```

```
-- Montando e Inserindo o conteúdo para o TrainingGroup1245 -  Checking TrainingGroup3 -
Insert Into AnaliseQueimadasSelecionadasTrainingGroup1245
Select * From AnaliseQueimadasSelecionadasTrainingGroup1
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup2
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup4
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup5
Go

-- Carregando os dados no arquivo de TrainingGroup1245 --
- AnaliseQueimadasSelecionadasTraining - DiaSemChuva, PrecipitacaoUmDiaAnterior, Acu
muladoPrecipitacaoTresDiasAnteriores, Chuva, Queimada, RiscoFogoObservadoTresDiasPra
Frente --
Declare @BCPComando Varchar(500)
Set @BCPComando='bcp "Select DiaSemChuvaTrainingGroup1245, PrecipitacaoUmDiaAnt
eriorTrainingGroup1245, RiscoFogoObservadoTresDiasPraFrenteTrainingGroup1245 From
ProjetoDWQueimadas..AnaliseQueimadasSelecionadasTrainingGroup1245" queryout "S:\Me
strado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-
TrainingGroup1245.txt" -c -UTF8 -T -SWIN10PRO\WIN10PROSQL2019'
Exec master..xp_cmdshell @BCPComando, No_Output
Print 'Arquivo: Modelo6-Conteudo-ANFIS-ValidacaoCruzada-
TrainingGroup1245.txt criado.'
Go

-- Carregando os dados no arquivo de CheckingGroup4 -- Tabela TrainingGroup4 --
- AnaliseQueimadasSelecionadasTraining - DiaSemChuva, PrecipitacaoUmDiaAnterior, Acu
muladoPrecipitacaoTresDiasAnteriores, Chuva, Queimada, RiscoFogoObservadoTresDiasPra
Frente --
Declare @BCPComando Varchar(500)
Set @BCPComando='bcp "Select DiaSemChuvaTrainingGroup4, PrecipitacaoUmDiaAnterio
rTrainingGroup4, RiscoFogoObservadoTresDiasPraFrenteTrainingGroup4 From ProjetoDW
Queimadas..AnaliseQueimadasSelecionadasTrainingGroup4" queryout "S:\Mestrado\Defesa\
Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup4.txt" -c -
UTF8 -T -SWIN10PRO\WIN10PROSQL2019'
Exec master..xp_cmdshell @BCPComando, No_Output
Print 'Arquivo: Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup4.txt criado.'
Go

-- Montando e Inserindo o conteúdo para o TrainingGroup1235 - Checking TrainingGroup4 -
Insert Into AnaliseQueimadasSelecionadasTrainingGroup1235
Select * From AnaliseQueimadasSelecionadasTrainingGroup1
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup2
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup3
Union All
Select *From AnaliseQueimadasSelecionadasTrainingGroup5
Go
```

```
-- Carregando os dados no arquivo de TrainingGroup1235 ---
- AnaliseQueimadasSelecionadasTraining - DiaSemChuva, PrecipitacaoUmDiaAnterior, Acu
muladoPrecipitacaoTresDiasAnteriores, Chuva, Queimada, RiscoFogoObservadoTresDiasPra
Frente --
Declare @BCPComando Varchar(500)
Set @BCPComando='bcp "Select DiaSemChuvaTrainingGroup1235, PrecipitacaoUmDiaAnt
eriorTrainingGroup1235, RiscoFogoObservadoTresDiasPraFrenteTrainingGroup1235 From
ProjetoDWQueimadas..AnaliseQueimadasSelecionadasTrainingGroup1235" queryout "S:\Me
strado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-
TrainingGroup1235.txt" -c -UTF8 -T -SWIN10PRO\WIN10PROSQL2019'
Exec master..xp_cmdshell @BCPComando, No_Output
Print 'Arquivo: Modelo6-Conteudo-ANFIS-ValidacaoCruzada-
TrainingGroup1235.txt criado.'
Go

-- Carregando os dados no arquivo de CheckingGroup5 -- Tabela TrainingGroup5 ---
- AnaliseQueimadasSelecionadasTraining - DiaSemChuva, PrecipitacaoUmDiaAnterior, Acu
muladoPrecipitacaoTresDiasAnteriores, Chuva, Queimada, RiscoFogoObservadoTresDiasPra
Frente --
Declare @BCPComando Varchar(500)
Set @BCPComando='bcp "Select DiaSemChuvaTrainingGroup5, PrecipitacaoUmDiaAnterio
rTrainingGroup5, RiscoFogoObservadoTresDiasPraFrenteTrainingGroup5 From ProjetoDW
Queimadas..AnaliseQueimadasSelecionadasTrainingGroup5" queryout "S:\Mestrado\Defesa\
Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup5.txt" -c -
UTF8 -T -SWIN10PRO\WIN10PROSQL2019'
Exec master..xp_cmdshell @BCPComando, No_Output
Print 'Arquivo: Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup5.txt criado.'
Go

-- Montando e Inserindo o conteúdo para o TrainingGroup1234 -
- Checking TrainingGroup5 --
Insert Into AnaliseQueimadasSelecionadasTrainingGroup1234
Select * From AnaliseQueimadasSelecionadasTrainingGroup1
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup2
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup3
Union All
Select * From AnaliseQueimadasSelecionadasTrainingGroup4
Go
```

```
-- Carregando os dados no arquivo de TrainingGroup1234 ---
- AnaliseQueimadasSelecionadasTraining - DiaSemChuva, PrecipitacaoUmDiaAnterior, Acu
muladoPrecipitacaoTresDiasAnteriores, Chuva, Queimada, RiscoFogoObservadoTresDiasPra
Frente --
Declare @BCPComando Varchar(500)
Set @BCPComando='bcp "Select DiaSemChuvaTrainingGroup1234, PrecipitacaoUmDiaAnt
eriorTrainingGroup1234, RiscoFogoObservadoTresDiasPraFrenteTrainingGroup1234 From
ProjetoDWQueimadas..AnaliseQueimadasSelecionadasTrainingGroup1234" queryout "S:\Me
strado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-
TrainingGroup1234.txt" -c -UTF8 -T -SWIN10PRO\WIN10PROSQL2019'
Exec master..xp_cmdshell @BCPComando, No_Output
Print 'Arquivo: Modelo6-Conteudo-ANFIS-ValidacaoCruzada-
TrainingGroup1234.txt criado.'
Go
```

10.3 Script 2 – MatLab – Carregamentos dos Modelos.

```
%Limpando a Command Window
clc

%Limpando o Workspace
clear

% Import the file Testing
Testing = importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-Testing.txt');

% Import the file CheckinGroup
CheckingGroup1 = importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup1.txt');

CheckingGroup2 = importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup2.txt');

CheckingGroup3 = importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup3.txt');

CheckingGroup4 = importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup4.txt');

CheckingGroup5 = importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-CheckingGroup5.txt');

% Import the file TrainingGroup
TrainingGroup2345 = importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-TrainingGroup2345.txt');
```

```
TrainingGroup1345 = importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-TrainingGroup1345.txt');

TrainingGroup1245 = importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-TrainingGroup1245.txt');

TrainingGroup1235 = importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-TrainingGroup1235.txt');

TrainingGroup1234 = importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-TrainingGroup1234.txt');

%Carregando do AnfisEdit
AnfisEdit
```

10.4 Script 3 – MatLab – Avaliação dos Modelos.

```
%Limpando a Command Window
clc

%Limpando o Workspace
clear

%Importando o arquivo CheckinGroup
Testing= importdata('S:\Mestrado\Defesa\Dados\Modelos\Modelo6-Conteudo-ANFIS-ValidacaoCruzada-Testing.txt');

%Definindo a leitura das variáveis de entrada
FISInputData = Testing (:,1:2);

%Realizando a leitura do FIS
FISRead = readfis('S:\Mestrado\Defesa\Dados\Modelos\FIS\Modelo6-Conteudo-ANFIS-FIS-Sugeno.fis');

%Realizando a avaliação do FIS
FISOutput = evalfis(FISInputData,FISRead);

%Escrevendo o resultado de saída para um arquivo em Excel
xlswrite('S:\Mestrado\Defesa\Dados\Modelos\FIS\Modelo6-Conteudo-ANFIS-EvalFIS-Saida.xls',FISOutput);
```