



UNIVERSIDADE ESTADUAL DE
CAMPINAS

Instituto de Matemática, Estatística e
Computação Científica

ISAAC DAYAN BASTOS DA SILVA

Sistema de Alerta de Incêndio Florestal

Campinas

2017

Isaac Dayan Bastos da Silva

Sistema de Alerta de Incêndio Florestal

Tese apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Matemática Aplicada.

Orientador: João Frederico da Costa Azevedo Meyer

Coorientador: Laécio Carvalho de Barros

O arquivo digital corresponde à versão final da Tese defendida pelo aluno Isaac Dayan Bastos da Silva e orientada pelo Prof. Dr. João Frederico da Costa Azevedo Meyer.

Campinas

2017

Agência(s) de fomento e nº(s) de processo(s): CAPES

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

Si38s Silva, Isaac Dayan Bastos da, 1980-
Sistema de alerta de incêndio florestal / Isaac Dayan Bastos da Silva. –
Campinas, SP : [s.n.], 2017.

Orientador: João Frederico da Costa Azevedo Meyer.
Coorientador: Laécio Carvalho de Barros.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de
Matemática, Estatística e Computação Científica.

1. Sistemas fuzzy. 2. Aprendizado de máquina. 3. Incêndios florestais -
Aspectos ambientais. 4. Modelos classificadores. 5. Autômato celular. I. Meyer,
João Frederico da Costa Azevedo, 1947-. II. Barros, Laécio Carvalho de, 1954-.
III. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e
Computação Científica. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Forest fire warning system

Palavras-chave em inglês:

Fuzzy systems

Machine learning

Forest fires - Environmental aspects

Classifier models

Cellular automata

Área de concentração: Matemática Aplicada

Titulação: Doutor em Matemática Aplicada

Banca examinadora:

João Frederico da Costa Azevedo Meyer [Orientador]

Rodney Carlos Bassanezi

Marcos Eduardo Ribeiro do Valle Mesquita

Luciana Takata Gomes

Magda da Silva Peixoto

Data de defesa: 03-07-2017

Programa de Pós-Graduação: Matemática Aplicada

**Tese de Doutorado defendida em 03 de julho de 2017 e aprovada
pela banca examinadora composta pelos Profs. Drs.**

Prof(a). Dr(a). JOÃO FREDERICO DA COSTA AZEVEDO MEYER

Prof(a). Dr(a). RODNEY CARLOS BASSANEZI

Prof(a). Dr(a). MARCOS EDUARDO RIBEIRO DO VALLE MESQUITA

Prof(a). Dr(a). LUCIANA TAKATA GOMES

Prof(a). Dr(a). MAGDA DA SILVA PEIXOTO

As respectivas assinaturas dos membros encontram-se na Ata de defesa

À minha família

Agradecimentos

Primeiro agradeço a Deus por tudo que proporcionou até o presente momento. Agradeço a minha esposa, Ana Clara, pelo companheirismo do dia-a-dia, por todo apoio para a realização do doutorado, pois sem ela teria sido muito difícil, quase impossível. Por todo amor que ela me deu desde o dia que nos conhecemos e por ter me proporcionado ser pai de uma linda e inteligente filha. Agradeço à Ana Beatriz por ser minha filha e por iluminar todos os meus dias quando me chama de pai. Agradeço a minha mãe e meu pai por todo amor, educação, incentivo e orientação que sempre me deram, por serem os melhores pais do mundo e quero que saibam que esse título pertence a vocês também. Agradeço também as minhas irmãs pelo amor e pela ajuda que sempre me deram. Agradeço as minhas tias Telza, Telma e Tércia por todo amor e apoio que me deram para eu conseguir me estabelecer em Campinas e no período que estive em São Paulo. Agradeço ao meu orientador e amigo professor Joni que sempre me escutou, aconselhou e ajudou. Igualmente ao professor, coorientador e amigo prof. Laécio, pelos conselhos, orientações e ajuda. A vocês que me auxiliaram muito mesmo antes do ingresso no doutorado. Agradeço enormemente aos professores Rodney Bassanezi e Marcos Valle que, mesmo sem serem meus orientadores, me acolheram como orientando e me ajudaram muito para que eu alcançasse os resultados da minha tese. Aos professores Luciana Takata, Estevão Laureano, Petronio Pulino, Fernando Von Zuben, Fernando Gomide pelas orientações a mim dispensadas. Agradeço muito à professora Elaine Polleti pelas orientações no início da pesquisa e durante o estágio docente. Agradeço aos meus amigos da Unicamp, especialmente a Diego, Chrys, Douglas, Tatiana, Alex, Fábio, Arlison, Anna Ligia, Tiemi, Fernanda e Ana Carolina. Por todo o companheirismo durante esse período difícil, mas que a ajuda de vocês foi fundamental para que eu enfrentasse todas as adversidades. Aos meus amigos Maurício, Alírio, Bruno, Miro, Daniel Almeida, Marcelo, Vinícius, Elizardo, Miguel e Acácio, pela amizade de irmãos que sempre nos manterá próximos, independentemente da distância física. Em especial, agradeço ao meu amigo Sebastião pela ajuda durante fase mais difícil da pesquisa. Agradeço aos meus amigos Antonio Júnior, Paulo Roberto e Andrio. Agradeço, In Memoriam, ao meu amigo Tadeu Melo que me incentivou muito para que eu desenvolvesse esse projeto. Agradeço aos meus amigos Estácio, Ezacanael, Michael, Miguel, Guilherme, Luidhi, Andhello, Fortunarro e Cassiano por toda ajuda e orientação, principalmente nesse período de doutoramento. Agradeço também a todos os meus amigos que participaram do grupo Inova Ufac no ano de 2012. Agradeço a Maria Suely pela ajuda durante o doutorado. Agradeço a meu amigo e médico, doutor Rodrigo Ubiratan, pela força e orientação. À CAPES pela bolsa de estudos.

Resumo

O principal objetivo desta tese é apresentar um sistema de alerta antecipado de incêndio florestal. Este sistema é baseado em três subsistemas acoplados: um **classificador kNN fuzzy**, que a partir das variáveis distâncias para curso d'água e estrada mais próximos, altitude e tipologia florestal, gera o índice de risco de incêndio florestal, um algoritmo ***Subtractive Clustering*** e um **sistema fuzzy** que geram o perigo de incêndio florestal. De modo acoplado o risco e perigo de incêndio florestal podem simular a evolução espaço-temporal do índice de alerta de incêndio. A saída destes sistemas é entendida como pertinências de subconjuntos fuzzy. O alerta de incêndio florestal é utilizado para gerar simulações de propagação de incêndio a partir de metodologia de autômatos celulares. Propomos duas metodologias distintas para obtermos o perigo de incêndio florestal: a primeira consiste em utilizar um sistema fuzzy, que realiza métodos de aprendizagem supervisionada e não supervisionada, para fornecer uma série temporal de valores de perigo e que depende apenas de variáveis climáticas (umidade relativa do ar e precipitação); a segunda envolve um sistema dinâmico fuzzy que gera a mesma série temporal a partir de um sistema baseado em regras fuzzy com as mesmas variáveis de entrada do modelo de aprendizado de máquina. O índice de perigo gerado por esses modelos funciona também como regulador da dinâmica temporal do alerta de incêndio. Os modelos matemáticos são aplicados a dados reais geo-referenciados do Estado do Acre-Brasil.

Palavras-chave: Sistema Fuzzy; Aprendizado de Máquina; Incêndio Florestal; Modelos Classificadores; Autômatos Celulares

Abstract

The main purpose of this thesis is to present a fire forest early warning system. This system is based on three coupled subsystems: a fuzzy kNN classifier, which generates the forest fire risk index from the distances to the nearest watercourse and road, altitude and forest typology; a **Subtractive Clustering** algorithm and a **fuzzy system** that create the danger of forest fire. In a coupled way the risk and danger of forest fire can simulate the spatio-temporal evolution of the fire warning index. The forest fire warning is used to generate simulations of fire spread using cellular automata methodology. We propose two different methodologies to obtain the forest fire danger: the first is based upon a fuzzy system that performs supervised and unsupervised learning methods providing a time series of danger values and depends only on climatic variables (air relative humidity and rainfall); The second involves a dynamic fuzzy system which generates the same time series from a system based on fuzzy rules with the same input variables as the machine-learning model. The index of danger generated by these models also acts as a regulator of the temporal dynamics of the warning of fire. The mathematical models are applied to real georeferenced data from the State of Acre-Brazil.

Keywords: Fuzzy System; Machine Learning; Forest Fire; Classifier Models; Cellular Automata.

Lista de ilustrações

Figura 1 – Incêndio no Estado do Acre em maio de 2013 retirada de (GALVÃO, 2013).	15
Figura 2 – Diagrama da arquitetura do modelo geral que acopla um classificador kNN fuzzy, sistema fuzzy e autômatos celulares em que ρ é o risco de incêndio florestal, π_t é o perigo de incêndio florestal no tempo t , W é o alerta de incêndio florestal e q é o número de valores de perigo de incêndio passados que são utilizados a cada instante t	26
Figura 3 – Localização do Estado do Acre no Brasil na América Latina (INC, 2017).	28
Figura 4 – Evolução dos números de focos de calor ocorridos no Estado do Acre no período de 1988 a 2016 (INPE, 2011).	29
Figura 5 – Estrutura de um sistema baseado em regras fuzzy adaptada de (SILVA, 2005).	32
Figura 6 – Estrutura do sistema dinâmico fuzzy retirada de (SILVA, 2005)	34
Figura 7 – Diagrama da arquitetura de uma máquina de aprendizagem. Adaptado de (CHERKASSKY; MULIER, 1998).	35
Figura 8 – Diagrama em blocos da arquitetura do modelo de sistema fuzzy para o perigo de incêndio florestal.	44
Figura 9 – Centros dos agrupamentos obtidos pelo agrupador <i>Subtractive Clustering</i> e uma parte do conjunto de dados de treinamento.	45
Figura 10 – Função de pertinência de um subconjunto fuzzy centrado no ponto $(U, P) = (89.75, 12.67)$	46
Figura 11 – Percentual de focos de calor e perigo de incêndio florestal para o Estado do Acre em 2014.	47
Figura 12 – Ocorrência de focos de calor F_t e alerta climático α_t para o Estado do Acre em 2014.	47
Figura 13 – Curva ROC obtida da classificação realizada para avaliar o modelo de perigo de incêndio florestal.	48
Figura 14 – Variável linguística: umidade relativa do ar (%).	50
Figura 15 – Variável linguística: precipitação diária em mm.	50
Figura 16 – Diagrama em blocos da estrutura do sistema dinâmico fuzzy	51
Figura 17 – Variável linguística variação do perigo $\Delta(U_t, P_t)$	52
Figura 18 – Perigo de incêndio florestal (π^d) e percentual de focos de calor (dados reais).	53
Figura 19 – Alerta climático dinâmico (α_t^d) e ocorrência de focos de calor F_t	53
Figura 20 – Curva ROC associada à classificação utilizada para avaliar o desempenho do modelo de sistema dinâmico fuzzy.	54

Figura 21 – Diagrama em blocos da arquitetura do modelo de risco de incêndio florestal com o classificador kNN fuzzy.	58
Figura 22 – Mapas de ocorrência de focos de calor e risco de incêndio florestal do Estado do Acre em 2014.	63
Figura 23 – Diagrama em bloco da estrutura do modelo de alerta de incêndio florestal obtido por um classificador kNN fuzzy, agrupador e sistema fuzzy: ρ é o risco de incêndio florestal, π_t é o perigo de incêndio florestal no tempo t e q é o número de atrasos de tempo do perigo de incêndio.	66
Figura 24 – Mapa de Alerta de Incêndio Florestal para seis dias do ano de 2014. . .	67
Figura 25 – Diagrama da arquitetura do modelo de autômatos celulares.	69
Figura 26 – Vizinhança de Moore	70
Figura 27 – Diagrama em blocos que representa as condições para as mudanças de estado dos autômatos. W é o alerta de incêndio florestal da célula, π_t é o perigo de incêndio no instante t , p é o número aleatório com distribuição uniforme $[0, 1]$, c_1 é a condição de propagação e B é o limiar de extinção.	73
Figura 28 – Séries temporais de focos de calor de 2014, perigo de incêndio florestal (dois modelos distintos) e percentual de nós “queimando” do autômato celular	74
Figura 29 – Estados do autômato celular nas iterações 230, 240, 250 e 365 do ano de 2014.	75
Figura 30 – Focos de calor ocorridos em 2014 e as células “queimadas” do autômato celular.	76

Lista de tabelas

Tabela 1	– Restrições à somatória da FMA, de acordo com a precipitação do dia (SOARES; BATISTA, 2007)	22
Tabela 2	– Informações sobre os conjuntos fuzzy obtidos pelo agrupador <i>subtractive clustering</i> para a cidade de Rio Branco-AC, no período 2003-2013.	46
Tabela 3	– Classificação utilizada para avaliar a eficiência do modelo de perigo de incêndio florestal.	48
Tabela 4	– Base de regras utilizadas no sistema dinâmico fuzzy. Na primeira coluna estão as adjetivações que definem os números fuzzy correspondentes à umidade relativa do ar, na primeira linha estão os relativos à precipitação diária e cada posição da tabela informa o valor da variação do perigo $\Delta(U_t, P_t)$ correspondente.	50
Tabela 5	– Valores de umidade e parâmetro f_i associados aos 9 agrupamentos (fornecidos pelo <i>Subtractive Clustering</i>) juntamente com as pertinências relativas aos conjuntos fuzzy da variável linguística umidade do sistema fuzzy. Na coluna umidade encontram-se os valores das primeiras coordenadas dos centros dos agrupamentos obtidos no modelo de perigo de incêndio com aprendizado de máquina.	55
Tabela 6	– Ordem decrescente da densidade de focos de calor por área, tipologia florestal, número de focos de calor de 2002 a 2012, área total de cada tipologia e densidade de focos de calor por unidade de área.	61
Tabela 7	– Descrição dos códigos do Zoneamento Ecológico Econômico do Estado do Acre.	61
Tabela 8	– Estados das células do AC	70

Sumário

Introdução	14
1 Índices de Incêndio Florestal	17
1.1 Risco de Incêndio Florestal	17
1.2 Perigo de Incêndio Florestal	18
1.3 Alerta de Incêndio Florestal	19
1.4 Índices de Risco e Perigo de Incêndio Florestal: uma revisão de literatura	20
1.5 Sistema de Alerta Antecipado	24
1.6 Local de Estudo: o Estado do Acre-Brasil	28
2 Conceitos Básicos	30
2.1 Subconjuntos Fuzzy	30
2.2 Variáveis Linguísticas	31
2.3 Sistema Baseado em Regras Fuzzy	31
2.4 Sistema Dinâmico Fuzzy	33
2.5 Aprendizado de Máquina	34
2.6 Classificadores	35
2.6.1 Classificadores knn e knn fuzzy	36
2.6.2 Agrupador <i>Subtractive Clustering</i>	37
2.6.3 Curva ROC	38
3 Índice de Perigo de Incêndio Florestal	40
3.1 Caracterização das Variáveis	41
3.2 Perigo de Incêndio Florestal - versão com Sistemas de Aprendizado Supervisionado e Não-Supervisionado	42
3.2.1 Perigo de Incêndio Florestal para o Estado do Acre - Estudo 1	45
3.2.1.1 Resultados	45
3.3 Perigo de Incêndio Florestal através de Sistema Dinâmico Fuzzy	49
3.3.1 Perigo de Incêndio Florestal para o Estado do Acre - Estudo 2	52
3.3.1.1 Resultados	52
3.4 Avaliação Conjunta dos Modelos de Perigo de Incêndio Florestal	54
4 Índice de Risco de Incêndio Florestal	57
4.1 Mapa de risco de incêndio florestal para o Estado do Acre-Brasil	59
4.1.1 Caracterização das Variáveis de Entrada e Saída	59
4.1.2 Tratamento dos dados	61
4.1.3 Conjuntos de treinamento, validação e teste e o valor do parâmetro k	62
4.1.4 Resultados	62
5 Índice de Alerta de Incêndio Florestal e Simulações de Propagação	65
5.1 Evolução Temporal do Alerta de Incêndio Florestal	65

5.1.1	Evolução temporal do alerta de incêndio: estudo para o Estado do Acre	66
5.1.1.1	Resultados	66
5.2	Modelos de Propagação de Incêndio Florestal	67
5.3	Autômatos Celulares	68
5.3.1	Regras de Mudança de Estado	70
5.3.2	Estudo de Caso - 3: Estado do Acre	72
5.3.2.1	Resultados	73
6	Considerações Finais	77
	REFERÊNCIAS	80

Introdução

Incêndios florestais que ocorrem na Amazônia, como um todo, são o resultado de severas secas que ocorrem na região combinadas com práticas de queimadas para limpeza de clareiras, que muitas vezes se tornam incontroláveis. O impacto ambiental gerado pela queima afeta a qualidade do solo e do ar, a floresta fica cada vez mais propensa a novos incêndios, a biodiversidade é destruída, além de muitos outros efeitos negativos (ACRE, 2013).

O processo de desmatamento na região Amazônica é, quase sempre, realizado por meio do uso do fogo que é uma forma de baixo custo de remover plantas que tem potencial para competir com culturas agrícolas ou pastos e também para realizar a transformação da vegetação em cinzas. Os fatores sócio-econômicos são preponderantes para a utilização dessa prática de uso da terra que acaba gerando consequências não-intencionais tal como a queimada de enormes áreas de florestas tropicais, justamente porque o uso do fogo aumenta a inflamabilidade das paisagens Amazônicas e provoca um ciclo vicioso de degradação florestal. Nesse ciclo as florestas são substituídas por uma vegetação mais inflamável, que provoca o aumento do risco de incêndios (ACRE, 2010; NEPSTAD; MOREIRA; ALENCAR, 1999).

Os impactos gerados pela seca e pelos incêndios florestais têm grande influência no futuro das sociedades da Amazônia. A atividade humana no planeta induz mudanças climáticas que podem gerar o aumento da temperatura mas também alterar a distribuição de chuvas o que pode causar efeitos sobre as secas que podem ocorrer no futuro, tornando-as mais extensas e frequentes e pode deixar a sociedade cada vez mais vulnerável a eventos extremos (ACRE, 2010).

As estatísticas sobre a ocorrência de incêndios florestais no Brasil e em outros países mostram que a maioria dos incêndios tem origem em fontes que resultam, direta ou indiretamente, de atividades humanas (SOARES; BATISTA, 2007). Nepstad, Moreira e Alencar (1999) afirmam que a ocorrência de incêndios florestais não é um fenômeno recente na Amazônia porém ele é muito mais frequente na atualidade devido à ação antrópica.

Nepstad, Moreira e Alencar (1999) fazem uma classificação do fogo na Amazônia em três tipos principais:

- **Queimadas para desmatamento** são intencionais e estão relacionadas à queima e à derrubada da floresta;
- **Incêndios florestais rasteiros** são causados por queimadas que fogem ao controle e invadem as florestas; e

- **Queimadas e os incêndios em áreas já desmatadas** são o resultado do fogo que pode ser intencional ou acidental em pastagens, lavouras e capoeiras.

Durante a expansão da economia da borracha no fim do século XIX, o Estado do Acre desempenhou um papel relevante na história da região amazônica pela produtividade dos seringais presentes no local e pela riqueza natural dos rios acreanos. Em 2005 houve um período de estiagem longo e severo na Amazônia que contribuiu para que as atividades de queimadas começassem mais cedo do que o normal (entre os meses de agosto e setembro) e para a ocorrência de incêndios florestais nunca registrados nessa região e que comprometeram grandes áreas de floresta nativa (ACRE, 2010; PANTOJA; BROWN, 2009). As áreas mais atingidas do Acre foram as regionais do Alto e Baixo Acre (ACRE, 2010). A Figura 1 ilustra o combate a um incêndio ocorrido no Acre no ano de 2013. Não só o Acre, mas outros estados da região amazônica, tem sido objeto de atenção pela grande mídia brasileira em 2015 (BERTONI, 2015).



Figura 1 – Incêndio no Estado do Acre em maio de 2013 retirada de (GALVÃO, 2013).

As queimadas na Amazônia são monitoradas diariamente pelo Instituto Nacional de Pesquisas Espaciais (INPE) através de satélites meteorológicos que registram a localização de focos de calor (NEPSTAD; MOREIRA; ALENCAR, 1999). Os focos de calor, inclusive, são utilizados como indicadores de quando e onde as queimadas ocorrem na Amazônia e atualmente são utilizados no monitoramento de queimadas no Acre. Os dados sobre focos de calor são disponibilizados diariamente pelo Centro de Previsão do Tempo e Estudos Climáticos (CPTEC) do INPE (ACRE, 2010; INPE, 2011).

Existem incertezas associadas aos dados de focos de calor, pois o número total desses focos não é igual ao de queimadas. Diversos fatores podem influenciar na detecção dos focos de calor e, por isso, há a necessidade de que esses dados sejam interpretados e

de que haja uma validação de campo para reduzir a margem de erro (ACRE, 2010; INPE, 2011).

Uma das medidas preventivas fundamentais para prevenir e combater os incêndios florestais é a utilização de índices de risco e perigo de incêndio florestal (SOARES; BATISTA, 2007; NUNES; SOARES; BATISTA, 2008). Dessa forma, esta tese se propõe a criar modelos matemáticos de índices de perigo e risco de incêndio florestal que possam auxiliar na tarefa de combate e prevenção aos incêndios florestais.

Esta tese está organizada da seguinte forma: os índices associados a incêndio florestal são apresentados no Capítulo 1. No Capítulo 2 detalhamos conceitos básicos que são utilizados nos diversos modelos matemáticos apresentados nessa tese. O Capítulo 3 apresenta dois modelos matemáticos usados para obter o cálculo do perigo de incêndios envolvendo métodos de aprendizagem de máquinas supervisionados e não supervisionados e sistema dinâmico fuzzy. A modelagem de risco de incêndio florestal a partir de um classificador kNN fuzzy é apresentada no Capítulo 4. O Capítulo 5 detalha o modelo matemático do índice de alerta de incêndio florestal que se baseia na combinação dos índices de risco e perigo de incêndio florestal e apresenta, também, um modelo de propagação de incêndio florestal usando a metodologia de autômatos celulares. Em todos os capítulos apresentamos a metodologia e um caso de estudo para o Estado do Acre-Brasil. A tese termina com as conclusões no Capítulo 6.

1 Índices de Incêndio Florestal

O início do incêndio florestal, sua dispersão e intensidade estão associados com vários fatores ambientais entre eles destacamos: cobertura vegetal, relevo, condições climáticas, tipo de combustível. Cada um deles tem contribuição distinta para o grau de risco e perigo de incêndios locais (SOARES; BATISTA, 2007; FIORUCCI; GAETANI; MINCIARDI, 2008; NEPSTAD; MOREIRA; ALENCAR, 1999).

Em uma área florestal podem-se distinguir dois tipos de fatores determinantes do grau de perigo de incêndio: os de caráter permanente tais como material combustível, relevo e tipo de floresta e os variáveis, caracterizados pelas condições meteorológicas (NUNES; SOARES; BATISTA, 2008).

A literatura diferencia os conceitos de **índice de risco e perigo de incêndio florestal**. Enquanto o primeiro está associado à probabilidade de um incêndio se iniciar em função da presença e/ou atividades de agentes causadores, o segundo relaciona-se às características do material combustível (caráter permanente) e condições climáticas (caráter variável) que predispõem à ignição do incêndio ou representam dificuldade para a extinção dele (BROWN; DAVIS, 1973). Nas próximas seções, apresentamos mais detalhes sobre esses conceitos que são importantes para a compreensão dos modelos desenvolvidos nesta tese.

1.1 Risco de Incêndio Florestal

Segundo Countryman (1966) o risco de incêndio para qualquer área está relacionado à quantidade de agentes de ignição disponíveis por fontes humanas e por causas naturais de fogo. Essas últimas estão totalmente relacionados a fatores ambientais e são, como exemplo, a combustão espontânea e o relâmpago. Já o número de agentes de ignição humanos que podem estar presentes são determinados pelo tipo de uso em uma área (atividade em curso), fatores ambientais ou sociais e econômicos (chamados de fatores “humanos”).

Já Allgöwer, Carlson e Wagtendonk (2003) realizam a análise de risco a partir de dois conceitos:

1. Probabilidade de ignição e propagação (relacionada à causa de incêndios, propriedades e condições do combustível) denominada como perigo de incêndio; e
2. Resultado e consequências denominados de vulnerabilidade a incêndio (devido a distribuição, composição do combustível).

Allgöwer, Carlson e Wagtendonk (2003) definem risco de incêndio como a combinação desses dois aspectos. As variáveis a serem consideradas para cada um desses aspectos dependem das escalas de tempo consideradas e que podem ser distintas.

Chuvieco, Allgöwer e Salas (2003), Fao (1986) definem risco de incêndio como:

- A chance de início do fogo pela presença e atividade de agentes causadores;
- Um elemento do perigo de incêndio em qualquer área; e
- Um número que tem relação ao potencial dos agentes de ignição presentes em uma determinada área (em um dia de avaliação).

Como as definições de risco de incêndios florestais são muito próximas e estão centradas na presença de agentes causadores de incêndio, procuramos criar, nesta tese, um modelo de risco de incêndio que, essencialmente, possui variáveis associadas direta ou indiretamente ao “fator humano” e ao relevo que possuem em comum uma variação pequena em uma grande escala de tempo.

1.2 Perigo de Incêndio Florestal

Perigo (“danger”) de incêndio, segundo Fao (1986), é definido como “o resultado, frequentemente expresso como um índice, de fatores constantes e variáveis que afetam a criação, disseminação e dificuldade de controle de incêndios e os danos que causam”. Essa terminologia é amplamente aceita pela comunidade florestal (CHUVIECO; KASISCHKE, 2007). Countryman (1966) afirma que o perigo de incêndio tem sido definido como o potencial de danos por incêndio.

O **Índice** de Perigo de Incêndio é um número relativo que indica a severidade do perigo de incêndio florestal determinado pelas condições de queima e outras variáveis de perigo de incêndio (CHUVIECO; ALLGÖWER; SALAS, 2003; FAO, 1986). O cálculo do perigo de incêndio deve ser feito através das propriedades físicas do mundo biótico e abiótico que é relacionado ao comportamento e à ocorrência de incêndios (ALLGÖWER; CARLSON; WAGTENDONK, 2003).

O processo de estimativa do perigo de incêndio refere-se ao “processo de avaliação e integração sistemática dos fatores individuais e combinados que determinam a facilidade de início e propagação de um incêndio, dificuldade de controle e alguns impactos imediatos e evidentes do fogo (por exemplo, queimadura da copa, profundidade de queimadura) com base na avaliação do risco de ignição, do ambiente de incêndio (isto é, combustíveis, clima, topografia) e valores em risco” (COUNTRYMAN, 1966).

Os sistemas de estimativa de perigo de incêndio fornecem índices numéricos (valores indiretos) em diversas escalas de tempo que estão associadas às condições físicas que predisõem à ignição ou a manutenção de condições favoráveis à propagação de incêndios. Esses sistemas são baseados em dados meteorológicos e modelos físicos, semi-físicos e modelos empíricos e requerem a avaliação de diversos fatores que estão relacionados às condições ambientais e à atividade humana. Como exemplo de condições ambientais utilizadas por sistemas de perigo de incêndio podemos citar: propriedades do material combustível, clima, topografia, entre outros (CHUVIECO *et al.*, 2003). Para Allgöwer, Carlson e Wagtendonk (2003) a obtenção de um perigo de incêndio efetivo é necessário conhecer a distribuição espacial e temporal do tipos e propriedades do combustível, variáveis sócio-econômicas, vegetação, possíveis fontes de calor e condições meteorológicas.

Pode-se utilizar os índices de perigo de incêndio para (SOARES; BATISTA, 2007; NUNES; SOARES; BATISTA, 2007; NUNES; SOARES; BATISTA, 2008):

1. Estabelecimento de zonas de risco, que possibilita um melhor planejamento das atividades de prevenção contra incêndios florestais;
2. Permissão para queimas controladas em períodos que possuem menores índices de perigo de incêndio;
3. Estimativa do comportamento do fogo com vistas a adoção de técnicas de combate ao fogo que sejam mais efetivas; e
4. Definição da estação de incêndios e a advertência pública do grau de perigo, informando à população sobre os cuidados a serem tomados em situações críticas.

O índice de perigo proposto nessa tese tem uma conotação diferente de vários outros trabalhos da literatura. Utilizamos esse índice para descrever o potencial de danos por incêndio analisado apenas pelas condições climáticas, pois essas variáveis possuem em comum uma rápida variação em um curto intervalo de tempo.

Na próxima seção apresentamos um novo índice chamado alerta de incêndio florestal que tem papel fundamental nesta tese e na modelagem de propagação de incêndio florestal realizada na Seção 5.3.

1.3 Alerta de Incêndio Florestal

O índice de alerta de incêndio florestal é uma proposta dessa tese que pretende acoplar as informações de risco e perigo de incêndio florestal para que as informações climáticas (que são de rápida variação no tempo) possam ser acopladas às informações que possuem pouca (ou nenhuma) variação temporal, tais como: tipologia florestal (que agrega

informação do “fator humano”), distâncias para estrada e curso d’água mais próximos, altitude e outras que podem ser incorporadas posteriormente ao modelo.

Para definir esse índice foi necessário adequar os modelos de índice e perigo de incêndio para que as saídas de ambos os modelos estivessem no intervalo $[0, 1]$. Dessa forma, utilizamos uma t-norma (mais detalhes no Capítulo 2) para acoplar esses modelos para, assim, obtermos um índice gerado com um modelo de variação espaço-temporal.

A seguir apresentamos uma revisão dos modelos matemáticos para a obtenção de índices (risco e perigo) associados a incêndio florestal propostos na literatura.

1.4 Índices de Risco e Perigo de Incêndio Florestal: uma revisão de literatura

Vários países propuseram, através de agências governamentais ou por meio de pesquisadores independentes, sistemas de informação sobre incêndios florestais baseados em índices de risco e perigo de incêndio florestal.

Desde 1968 o Serviço Florestal Canadense faz uso do Sistema Canadense de Avaliação do Perigo de Incêndio Florestal (CFFDRS) que é baseado em dois sistemas que utilizam equações empíricas: sistema Fire Weather Index (FWI) e Fire Behavior Prediction (FBP).

O *Fire Weather Index* (FWI) do Canadá possui como variáveis de entrada temperatura do ar, umidade relativa do ar, vento, chuva e o modelo é baseado em dois índices e três códigos computacionais: FFMC (*Fine Fuel Moisture Code*), DMC (*Duff Moisture Code*), DC (*Drought Code*), ISI (*Initial Spread Index*), BUI (*BuidUp Index*). Estes dois últimos índices auxiliam no cálculo da umidade de três tipos de combustíveis a partir de informações de temperatura, umidade relativa, velocidade do vento e precipitação em 24h. Mais detalhes sobre o modelo matemático podem ser encontrados em [Van Wagner \(1987\)](#).

Já o Fire Behavior Prediction acrescenta a informação de topografia e cobertura vegetal ao FWI e possui um modelo semi-físico elíptico de evolução de incêndio originado a partir de uma única fonte ([GROUP, 1992](#)).

Outro sistema que acopla diversos índices de risco de incêndio florestal é o Sistema Nacional de Avaliação de perigo de incêndio (NFDRS) desenvolvido nos Estados Unidos e que leva em consideração o efeito do teor de umidade do combustível, clima e do declive na taxa de espalhamento do fogo por um modelo de propagação semi-físico ([ROTHERMEL, 1972](#); [DEEMING; BURGAN; COHEN, 1977](#); [ANDERSON, 1982](#)).

O sistema de risco de incêndio florestal proposto por [Fiorucci, Gaetani e](#)

Minciardi (2008) foi baseado nos sistemas CFFDRS e NFDRS e envolve três modelos: Carga de Combustível, Umidade do Combustível e Potencial Espalhamento do Fogo. Esse modelo destaca-se dos demais por utilizar cálculo diferencial para estimar a umidade do combustível, porém a extensa quantidade de parâmetros a serem estimados o torna de difícil aplicação prática.

Pesquisadores propuseram índices de perigo a partir de equações empíricas em diversos países, tais como Portugal, através do Serviço Florestal Português (LOURENÇO, 1990), Espanha (NATURALEZA, 1993), França (SOL, 1990), Itália (PALMIERI; INGHI-LESI; SIANI, 1993), Finlândia (FFFI) (GOLDAMMER; BUCHHOLZ; RESCH, 2003) e Brasil (SOARES; BATISTA, 2007). Dados estatísticos foram utilizados para obter o índice de perigo de incêndios desenvolvido na Austrália por Mc Arthur (MCARTHUR; FORESTRY; BUREAU, 1966). Posteriormente, Noble, Bary e Gill (1980) utilizaram fórmulas matemáticas que se ajustaram aos dados utilizados por Mc Arthur.

Em Portugal, índice elaborado por Lourenço (1990) utiliza temperatura, umidade relativa do ar e velocidade do vento. O modelo é expresso pela equação:

$$IR_{LL} = T/U + V \quad (1.1)$$

onde IR_{LL} é o índice de risco de incêndio florestal, T em ($^{\circ}\text{C}$) é a temperatura do ar, U em percentual (%) é a umidade relativa do ar e V é o fator de correção que depende da direção e velocidade do vento; todas essas variáveis devem ser medidas no mesmo local e na mesma hora do dia.

O sistema de perigo de incêndio desenvolvido pelo Serviço Florestal Espanhol (NATURALEZA, 1993) tem como variáveis de entrada: temperatura do ar, umidade do ar, declividade, aspecto, iluminação, tipos de combustíveis, exposição do combustível e malha de rodovias. O sistema Icona leva em consideração três subíndices diferentes: *Weather Danger Index*, *Fuel Hazard Component* e *Fire Incidence Index* (simplificado por *Human Risk index*). Para modelar o tipo de vegetação o sistema utiliza o Behave (ROTHERMEL, 1972).

Na França, Sol (1990) apresentou um modelo que utiliza um conjunto de equações empíricas que dependem da velocidade do vento, da temperatura do ar, do solo, do ponto de orvalho, da umidade do solo e da pressão de vapor saturado a determinada temperatura.

O *Finish Forest Fire Index* – FFFI é um índice proposto por Goldammer, Buchholz e Resch (2003) na Finlândia que utiliza radiação solar, temperatura e umidade do ar, velocidade do vento e precipitação como variáveis de entrada. O modelo proposto utiliza a equação $DW = P - E + / - F_1$ onde DW é a mudança no conteúdo de umidade volumétrica, P é a precipitação, E é a evaporação e F_1 é o fluxo de água para dentro ou

fora da camada superficial. O modelo também utiliza a relação:

$$E = DE \cdot E_{pot} \quad (1.2)$$

onde DE é a eficiência de secagem e E_{pot} é o potencial de evaporação. O potencial de evaporação é calculado pela evaporação gerada pela fórmula de Penman-Monteith (PENMAN, 1948). Uma tabela é utilizada para relacionar a umidade do solo ao risco de incêndio.

Criada em 1972 a partir de pesquisas realizadas pelo professor Ronaldo Viana Soares da Universidade Federal do Paraná (UFPR), a Fórmula de Monte Alegre surgiu utilizando dados da região central do estado do Paraná-Brasil, mais especificamente a partir de dados meteorológicos e de ocorrência de incêndios na Fazenda Monte Alegre, no município de Telêmaco Borba. É o primeiro índice brasileiro de incêndio florestal e ainda atualmente um dos mais importantes (SOARES; BATISTA, 2007). Esse índice tem como variáveis a umidade relativa do ar, medida às 13h, e a precipitação. A equação básica para esse índice é:

$$FMA = \sum_{i=1}^n \frac{100}{H_i} \quad (1.3)$$

sendo n o número de dias do período de observação, H_i a umidade relativa do ar em porcentagem e FMA a Fórmula de Monte Alegre (*cf* (DA SILVA; PONTES JR., 2011)).

Sendo acumulativo, o índice está sujeito às restrições de precipitação, como mostra a Tabela 1.

Tabela 1 – Restrições à somatória da FMA, de acordo com a precipitação do dia (SOARES; BATISTA, 2007)

CHUVA DO DIA (mm)	MODIFICAÇÃO NO CÁLCULO DA FMA
$\leq 2,4$	Nenhuma
2,5 a 4,9	Abater 30% na FMA calculada na véspera e somar $(100/H)$ do dia.
5,0 a 9,9	Abater 60% na FMA calculada na véspera e somar $(100/H)$ do dia.
10,0 a 12,9	Abater 80% na FMA calculada na véspera e somar $(100/H)$ do dia.
$\geq 12,9$	Interromper o cálculo (FMA= 0) e recomeçar a somatória no dia seguinte.

Os índices de perigo de incêndio de McArthur para pastagens e florestas têm sido amplamente utilizados na Austrália tanto para a previsão de perigo de incêndio

quanto para a interpretação do comportamento do fogo (MCARTHUR; FORESTRY; BUREAU, 1966). Os índices descrevem o perigo de incêndio relativo do dia-a-dia em uma grande área, dependem da velocidade do vento, umidade relativa do ar e temperatura e também indicam a probabilidade de ignição, dificuldade de contenção e taxa de propagação esperada (NOBLE; BARY; GILL, 1980).

Noble, Bary e Gill (1980) apresentam um conjunto de equações para descrever os dados fornecidos por McArthur que fornece um método para descrever o sistema de previsão e serve para modelagem e processamento de máquina.

Técnicas estatísticas são geralmente aplicadas para obter tanto o risco quanto o perigo de incêndio florestal associados às probabilidades de ocorrência de incêndio (GOLDARAG; MOHAMMADZADEH; ARDAKANI, 2016; PREISLER *et al.*, 2004).

Preisler *et al.* (2004) definiram o risco de incêndio a partir de três probabilidades: a probabilidade de ocorrência de incêndio, a probabilidade condicional de um grande incêndio dado que houve ignição e a probabilidade incondicional de grande incêndio. Os autores ajustaram aos dados uma regressão logística não paramétrica espacial e temporalmente. Nesse trabalho, eles demonstram um método para avaliar o perigo de incêndio e de variáveis climáticas associadas aos incêndios e produzem mapas de probabilidades para prever o número total de incêndios e de grandes incêndios esperados em dada região e num determinado período de tempo.

Os modelos utilizados no trabalho de Goldarag, Mohammadzadeh e Ardakani (2016) foram regressão logística e redes neurais artificiais. As variáveis utilizadas nos modelos foram: temperatura do ar, duração da luz do sol, dia com índice de precipitação maior ou igual a 1, umidade relativa do ar, frequência de focos de calor, NDVI (*Normalized Difference Vegetation Index*), cobertura terrestre, altitude. Os dados espaciais foram obtidos de satélites e foi realizada uma interpolação estatística espacial desses dados. Os parâmetros dinâmicos foram medidos de 2001 a 2004 e 2005 foi o ano utilizado para avaliar os métodos.

O risco e o perigo de incêndio florestal são obtidos através do uso de fórmulas baseadas na combinação linear das variáveis disponíveis em mapas geo-referenciados manipulados pelo software GIS (SIVRIKAYA *et al.*, 2014; NUNES; SOARES; BATISTA, 2008).

Uma abordagem alternativa para a predição de risco de incêndio florestal é a utilização de sistemas de aprendizado de máquina que realizam o aprendizado a partir de dados disponíveis. Alonso-Betanzos *et al.* (2002), Goldarag, Mohammadzadeh e Ardakani (2016) utilizaram redes neurais *multilayer perceptron* para classificar a presença ou não de focos de calor em regiões de Espanha e Irã utilizando como entrada variáveis climáticas que foram extrapoladas/interpoladas no espaço a partir de posições de estações meteorológicas

e a saída da rede é um valor real no intervalo $[0, 1]$ que é convertido em forma de classe de risco de incêndio florestal.

Todos esses modelos citados acima escolhem, *a priori*, distribuições de probabilidade, equações ou funções que serão utilizadas para obter os índices de risco bem como de perigo de incêndio.

San-Miguel-Ayanz *et al.* (2011) propuseram uma classificação dos índices de incêndios florestais em índices de longo-prazo, curto-prazo e dinâmicos de acordo com a rapidez de variação no tempo das entradas dos modelos utilizados. Os índices de curto-prazo utilizam variáveis de entrada que são modificadas quase que continuamente, como é o caso de variáveis climáticas. Já os índices de longo-prazo dependem de variáveis que são pouco ou nada modificadas no curto e médio prazos, como exemplo, podemos citar o tipo de vegetação e topografia do local em estudo. Os índices dinâmicos agregam os dois tipos de variáveis.

A seguir apresentamos o sistema de alerta de incêndio florestal proposto nessa tese.

1.5 Sistema de Alerta Antecipado

O principal objetivo desta tese é apresentar um sistema de alerta antecipado de incêndio florestal. Este sistema é baseado em três sistemas acoplados para simular o comportamento espaço-temporal do alerta de incêndio. A saída destes sistemas é entendida como pertinências a subconjuntos fuzzy. Estas pertinências são então associadas para gerar o alerta de incêndio florestal. Esse último índice é utilizado para gerar simulações de propagação de incêndio florestal a partir de metodologia de autômatos celulares (ver Capítulo 5.3).

Propomos separar as variáveis espaciais para compor um índice de longo prazo denominado aqui **Risco de Incêndio Florestal** (ρ) e as variáveis temporais para o índice de curto prazo **Perigo de Incêndio Florestal** (π). Nós associamos esses índices para criar um novo índice denominado aqui **Alerta de Incêndio Florestal** (\mathbb{W}) que pode ser usado como ferramenta de alerta antecipado de incêndio florestal para tomada de decisões por parte do poder público.

Nós escolhemos usar um classificador k-vizinhos mais próximos fuzzy (kNN fuzzy) para obter o risco de incêndio, que irá fornecer um mapa espacial de risco de incêndio a partir de variáveis locais, tais como altitude, tipologia florestal e outros.

Propomos duas metodologias distintas e mutuamente excludentes para obtermos o perigo de incêndio florestal: a primeira consiste em utilizar um sistema fuzzy que usa aprendizagem supervisionada e não supervisionada fornecendo uma série temporal de

valores de perigo e que depende apenas de variáveis climáticas (umidade relativa do ar e precipitação) (DA SILVA *et al.*, 2016); a segunda envolve um sistema dinâmico fuzzy que gera uma série temporal de valores de perigo de incêndio a partir de um sistema baseado em regras fuzzy com as mesmas variáveis de entrada do primeiro modelo (DA SILVA; BARROS; MEYER, 2015). O índice gerado por esses modelos funciona também como regulador da dinâmica temporal do alerta de incêndio florestal.

Ambas as abordagens não requerem muitos parâmetros e os poucos que são utilizados podem ser estimados a partir de técnicas estatísticas, como a validação cruzada de 10 pastas (*10-fold cross validation*) (HAYKIN, 2001). Outra vantagem, exceto para o modelo de sistema dinâmico fuzzy, é que a relação entre as variáveis de entrada e o índice de saída não é definida *a priori* pelo pesquisador, mas é aprendida a partir dos dados disponíveis. Além disso, nenhuma distribuição de probabilidade nem fórmula empírica *a priori* são usadas para descrever essas variáveis.

Todas as variáveis de entrada utilizadas em todos os sistemas desta tese são de fácil obtenção e dispensam a utilização de experiências físicas/químicas ou fórmulas empíricas.

Finalmente, o uso de sistemas fuzzy ajuda a modelar a incerteza inerente a este problema complexo. Algumas desvantagens destes modelos são a necessidade de uma grande quantidade de dados disponíveis e de poucas variáveis de entrada.

O diagrama em blocos da Figura 2 apresenta a arquitetura dos modelos utilizados nessa tese.

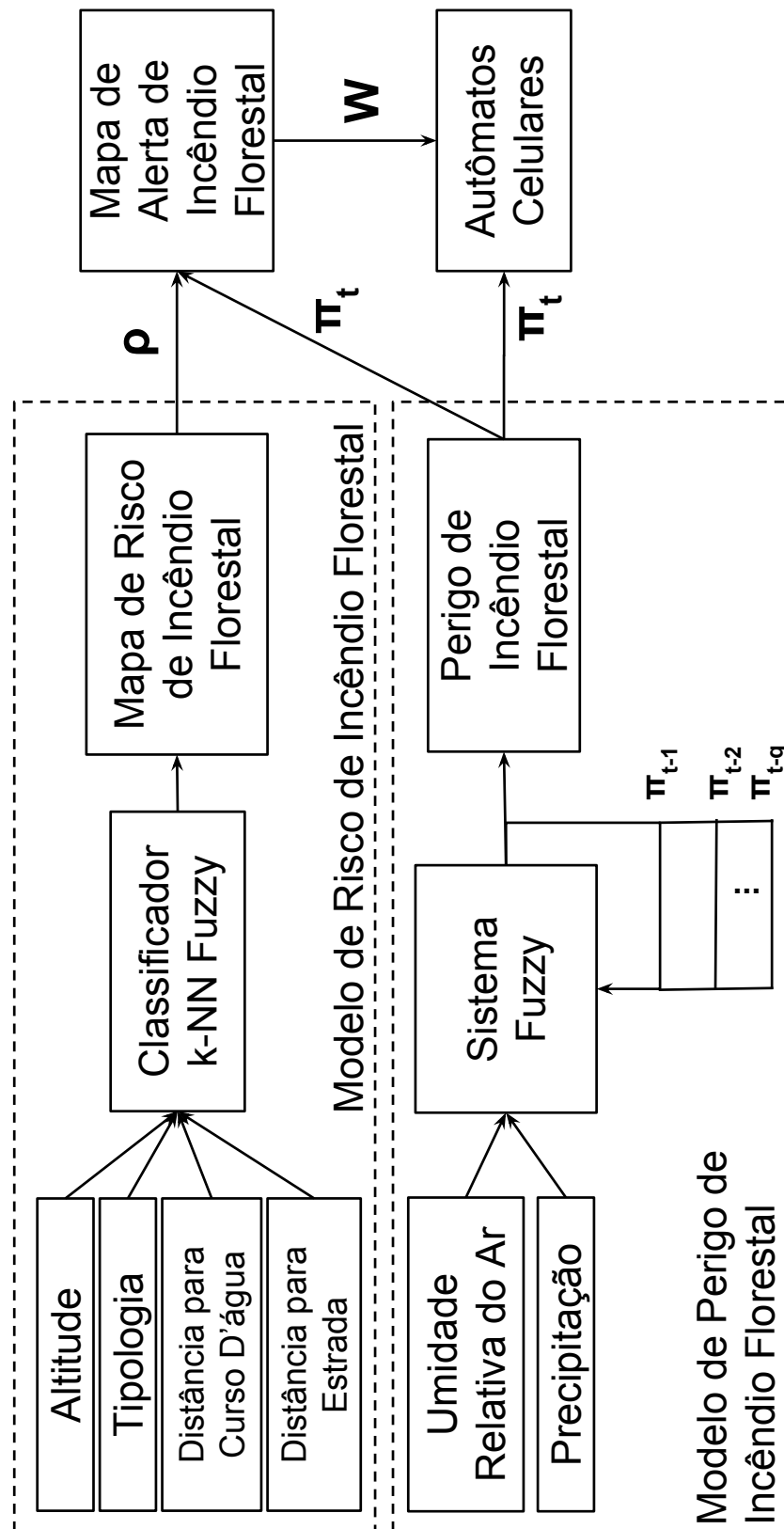


Figura 2 – Diagrama da arquitetura do modelo geral que acopla um classificador kNN fuzzy, sistema fuzzy e autômatos celulares em que ρ é o risco de incêndio florestal, π_t é o perigo de incêndio florestal no tempo t , \mathbb{W} é o alerta de incêndio florestal e q é o número de valores de perigo de incêndio passados que são utilizados a cada instante t .

A Figura 2 separa os modelos de perigo e risco de incêndio florestal em dois grandes blocos (em linhas tracejadas).

O bloco denominado **Modelo de Risco de Incêndio Florestal** possui a arquitetura do modelo de risco de incêndio cujas variáveis de entrada são: tipologia florestal, distâncias para estrada e curso d'água mais próximos e altitude. Estas variáveis possuem em comum o fato de terem uma baixa variação em um tempo muito longo, dessa forma elas podem ser consideradas constantes no período de tempo considerado para o aprendizado e validação desse modelo que é de 2003 a 2013.

A saída do classificador do tipo k-vizinhos mais próximos fuzzy utilizado é um mapa de risco de incêndio para o domínio estudado no qual a cada célula é atribuído um grau de risco que é definido como a pertinência ao conjunto fuzzy definido pela presença de focos de calor, como será detalhado no Capítulo 4.

O bloco denominado **Modelo de Perigo de Incêndio Florestal** possui a arquitetura do modelo de perigo de incêndio cujas variáveis de entrada são: umidade relativa do ar (em %) e precipitação (em mm) diárias. Estas variáveis possuem em comum o fato de terem uma rápida variação temporal.

Utilizamos dois modelos distintos (e mutuamente excludentes) para obtermos o perigo de incêndio florestal. O primeiro é um sistema dinâmico fuzzy que utiliza um sistema baseado em regras fuzzy e o segundo utiliza um sistema de aprendizado de máquina que acopla um agrupador do tipo *Subtractive Clustering* e um sistema fuzzy. A saída desse sistema é uma série temporal diária de valores perigo de incêndio que considera, além do valor do perigo no instante de tempo t , q valores atrasados no tempo. Esse modelo será detalhado no Capítulo 3.

Observando a Figura 2, percebemos que o bloco denominado **Mapa de Alerta de Incêndio Florestal** recebe informações de perigo (π_t) e risco (ρ) de incêndio florestal para formar um novo índice aqui chamado de **alerta de incêndio florestal** (\mathbb{W}). Esse índice possui uma variação espaço-temporal e consegue acoplar aqueles dois primeiros índices de uma forma que as características de ambos sejam preservadas. De forma resumida, o alerta de incêndio assumirá valores maiores em áreas de maior risco de incêndio e no período de tempo que o perigo de incêndio for maior também, o que é importante para o alerta antecipado de incêndios florestais.

A simulação de propagação de incêndios será realizada pelo bloco denominado **Autômatos Celulares**. Esse modelo utiliza as informações de alerta de incêndio para simular a evolução da frente de fogo para áreas mais críticas (maiores valores de alerta), porém o período de tempo no qual essa evolução acontece é determinado pelo perigo de incêndio.

Nesta tese escolhemos aplicar todos os modelos matemáticos desenvolvidos

nesta tese ao Estado do Acre-Brasil. A seguir, identificamos algumas características desse Estado brasileiro.

1.6 Local de Estudo: o Estado do Acre-Brasil

O Estado do Acre fica situado na região noroeste do Brasil, possui extensão de $164.123,712 \text{ km}^2$, e segundo o Censo do Instituto Brasileiro de Geografia e Estatística (IBGE) do ano de 2010, existiam 733.559 habitantes (BRASIL, 2013). Esse Estado possui 22 municípios e os dois maiores são Rio Branco e Cruzeiro do Sul. O município de Rio Branco é a capital com população de 336.038 habitantes e fica localizado no leste do Estado que é a região mais populosa do Estado. A segunda maior cidade é Cruzeiro do Sul que possui 78.507 habitantes e situa-se no extremo oeste do Estado (BRASIL, 2013). A Figura 3 ilustra a posição do Estado do Acre no Brasil.



Figura 3 – Localização do Estado do Acre no Brasil na América Latina (INC, 2017).

A ocorrência de queimadas em vários Estados da região amazônica, incluindo o Acre, foi objeto de atenção pela grande mídia brasileira em 2015 (BERTONI, 2015) e o total de focos de calor no Acre vem aumentando nos últimos anos e alcançando valores próximos aos do ano de 2005, no qual ocorreram fenômenos extremos de seca e incêndios florestais (INPE, 2011; PANTOJA; BROWN, 2009). A Figura 4 apresenta a evolução da quantidade total de focos de calor detectados no Estado do Acre no período de 1988 a 2016.

Em resumo, neste Capítulo 1, apresentamos os conceitos (a) de risco, (b) de perigo e (c) de alerta de incêndio florestal e uma sucinta revisão de literatura dos principais

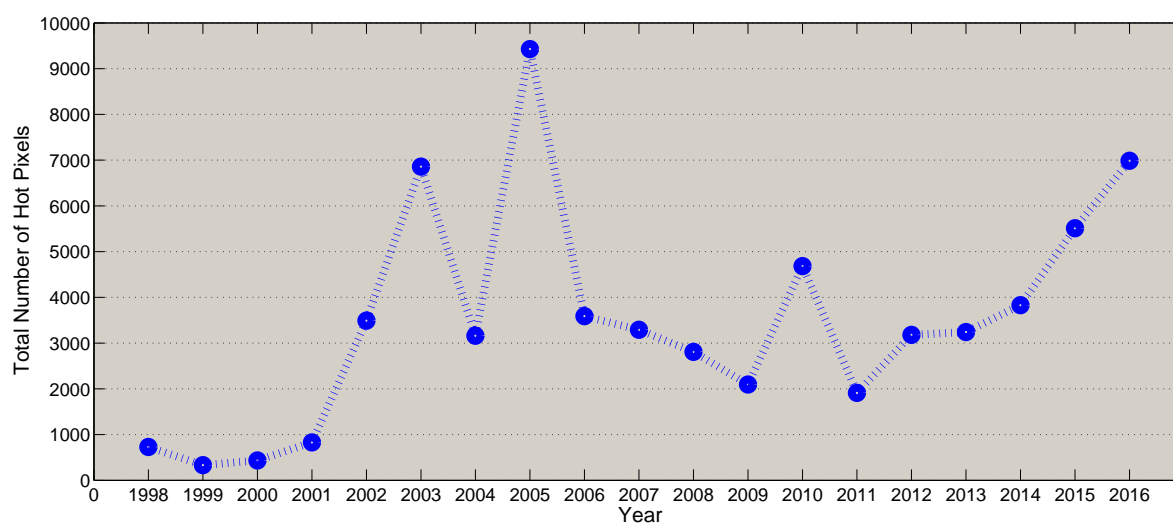


Figura 4 – Evolução dos números de focos de calor ocorridos no Estado do Acre no período de 1988 a 2016 (INPE, 2011).

sistemas e modelos matemáticos utilizados para obter índices de risco e perigo de incêndios. A Arquitetura dos modelos que foram utilizados nesta tese é apresentada na Figura 2 e também foram discutidos características do Estado do Acre, que foi o local escolhido para aplicar e verificar os modelos matemáticos criados, o que virá a seguir no texto.

A seguir, no Capítulo 2, apresentamos alguns conceitos básicos de instrumental necessário para o trabalho, incluindo, sobretudo, aspectos da teoria fuzzy.

2 Conceitos Básicos

Nesse capítulo apresentamos alguns conceitos básicos sobre conjuntos fuzzy, medidas de similaridade, curva ROC, classificadores baseados em aprendizado de máquina, entre outros, para uma melhor compreensão das técnicas utilizadas nessa tese.

2.1 Subconjuntos Fuzzy

Um subconjunto fuzzy A num universo de discurso U é caracterizado por uma função de pertinência $A : U \rightarrow [0, 1]$, onde $A(u)$ indica o grau de pertinência do elemento $u \in U$ ao subconjunto fuzzy A . Denotamos por $\mathbb{F}(U)$ a família de todos subconjuntos fuzzy de U . Além disso, dizemos que $A \in \mathbb{F}(U)$ é um subconjunto de $B \in \mathbb{F}(U)$, denotando essa inclusão por $A \subseteq B$, se $A(u) \leq B(u)$, $\forall u \in U$. Por fim, denotamos por \bar{A} o complemento padrão de um subconjunto fuzzy A , se $\bar{A}(u) = 1 - A(u)$, $\forall u \in U$ (VALLE; SOUZA, 2015).

Outra definição importante é a de norma triangular ou t-norma.

Definição 1. (*t-norma*). Um operador $t : [0, 1] \times [0, 1] \rightarrow [0, 1]$ é uma *t-norma* se satisfizer as seguintes condições (BARROS; BASSANEZI; LODWICK, 2017):

1. $t(1, x) = x$;
2. $t(x, y) = t(y, x)$;
3. $t(x, t(y, z)) = t(t(x, y), z)$;
4. se $x \leq u$ e $y \leq v$ então $t(x, y) \leq t(u, v)$;

Definição 2. Uma relação fuzzy E sobre $U_1 \times U_2 \times \dots \times U_n$ é qualquer subconjunto fuzzy de $U_1 \times U_2 \times \dots \times U_n$. Se o universo for formado de apenas dois conjuntos $U_1 \times U_2$ a relação é chamada de fuzzy binária sobre $U_1 \times U_2$ (BARROS; BASSANEZI; LODWICK, 2017).

Definição 3. Seja E uma relação fuzzy binária sobre U então para quaisquer x e y de U a relação fuzzy E é (DE BAETS; DE MEYER, 2005):

1. simétrica se $E(x, y) = E(y, x)$;
2. localmente reflexiva se $E(x, x) \geq E(x, y)$.

Outro conceito importante é o de medida de similaridade, também conhecida por equivalência fuzzy ou índice de igualdade. Na literatura existem diversas definições de medidas de similaridade (LAUREANO, 2014), nesta tese utilizamos a definição proposta por De Baets e De Meyer (2005):

Definição 4. *Medida de Similaridade é uma relação fuzzy binária simétrica e localmente reflexiva.*

Como exemplo, apresentamos três medidas de similaridade baseadas em distância: Medida do Complemento de Hamming e duas medidas propostas por [De Baets e De Meyer \(2005\)](#) descritas abaixo.

Dados dois vetores $x, y \in [0, 1]^n$ as medidas de similaridade S_h, S_1, S_2 entre x e y são dadas pelas fórmulas:

$$S_h(x, y) = 1 - \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (2.1)$$

$$S_1(x, y) = \frac{1}{n} \left[n + \sum_{i=1}^n (x_i \cdot y_i) - \max \left(\sum_{i=1}^n x_i, \sum_{i=1}^n y_i \right) \right] \quad (2.2)$$

$$S_2(x, y) = \frac{\sum_{i=1}^n [x_i \cdot y_i + (1 - x_i) \cdot (1 - y_i)]}{n + \sum_{i=1}^n (x_i \cdot y_i) - \min \left(\sum_{i=1}^n x_i, \sum_{i=1}^n y_i \right)} \quad (2.3)$$

Na próxima seção abordamos o conceito de variáveis linguísticas que é importante para a compreensão do modelo de perigo de incêndio florestal que foi elaborado a partir de um sistema baseado em regras fuzzy.

2.2 Variáveis Linguísticas

Para [Barros, Bassanezi e Lodwick \(2017\)](#) e [Cecconello \(2006\)](#), nas operações entre conjuntos fuzzy é comum utilizar-se as variáveis linguísticas, que podem ser qualitativas ou quantitativas, cujos valores assumidos são subconjuntos fuzzy. Variável linguística é um substantivo (qualitativo) e seus valores são adjetivos. Os conjuntos fuzzy representam os estados de uma variável linguística, em geral, são expressos por valores subjetivos como pequeno, muito, alto, etc.

Na próxima seção apresentaremos alguns conceitos sobre sistema baseado em regras fuzzy que utiliza valores assumidos por variáveis linguísticas e que serão modelados por conjuntos fuzzy.

2.3 Sistema Baseado em Regras Fuzzy

Um sistema baseado em regras fuzzy (SBRF), essencialmente, possui quatro componentes principais: um processador de entrada (ou fuzzificador), uma base de regras

fuzzy, um método de inferência fuzzy e um processador de saída (ou defuzzificador), gerando um número real como sua saída (PEDRYCZ; GOMIDE, 2007; BARROS; BASSANEZI; LODWICK, 2017). A Figura 5 ilustra um sistema baseado em regras fuzzy.

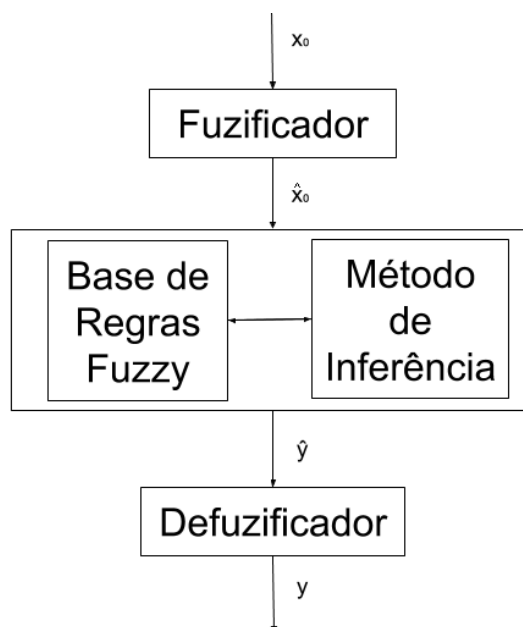


Figura 5 – Estrutura de um sistema baseado em regras fuzzy adaptada de (SILVA, 2005)

Nessa estrutura, o sistema faz a transformação de variáveis *crisp* (ou clássicas) em valores fuzzy utilizando as variáveis linguísticas. Para essa transformação, um conjunto de inferências fuzzy é utilizado para as tomadas de decisões e, após isso, haver um processo que permita representar um conjunto fuzzy por um número *crisp* (PEDRYCZ; GOMIDE, 2007; BARROS; BASSANEZI; LODWICK, 2017).

A fuzzificação é o processo no qual ocorre a conversão dos valores de entrada do sistema para conjuntos fuzzy, com as respectivas faixas de valores onde estão definidos. (PEDRYCZ; GOMIDE, 2007; PEIXOTO, 2005; MASSAD *et al.*, 2008; BARROS; BASSANEZI; LODWICK, 2017).

Já a base de regras fuzzy providencia uma maneira formal para representar as estratégias, objetivos e informações utilizadas por especialistas na área, através das suas experiências. O especialista auxilia na formulação de um conjunto de regras fuzzy, articulando as associações de entradas/saídas linguísticas (PEIXOTO, 2005; SHAW; SIMÕES, 1999; BARROS; BASSANEZI; LODWICK, 2017). Uma base de regras também pode ser construída a partir do uso de técnicas de aprendizado de máquina. (PEDRYCZ; GOMIDE, 2007)

O método de inferência utilizado no sistema baseado em regras fuzzy é o de Mamdani, que utiliza o operador OU (modelado pelo operador máximo \vee) e os operadores E e ENTÃO (modelados pelo operador mínimo \wedge) em cada regra individual da base de regras fuzzy. Para um sistema com mais de uma regra, tem-se a contribuição de cada

regra individualmente, de forma a se obter a saída fuzzy (PEDRYCZ; GOMIDE, 2007; BARROS; BASSANEZI; LODWICK, 2017).

A defuzzificação é o processo que permite a representação de um conjunto fuzzy por um valor real que melhor o represente (BARROS; BASSANEZI; LODWICK, 2017). O método do Centro de Gravidade ou Centróide é a técnica mais comum de defuzzificação. Pode ser compreendido como uma média ponderada, utilizando a função de pertinência $\mu_A(x)$ como o peso do valor x . As Equações 2.4 e 2.5 referem-se ao domínio discreto e domínio contínuo, respectivamente (BARROS; BASSANEZI; LODWICK, 2017).

$$\bar{z} = \frac{\sum_x \mu_A(x)x}{\sum_x \mu_A(x)}. \quad (2.4)$$

$$\bar{z} = \frac{\int_x \mu_A(x)x \, dx}{\int_x \mu_A(x) \, dx}. \quad (2.5)$$

Utilizamos o SBRF para definir o sistema dinâmico fuzzy que será utilizado nesta tese.

2.4 Sistema Dinâmico Fuzzy

Nesta tese utilizamos o sistema dinâmico fuzzy discreto que possui a forma:

$$\begin{cases} x_{k+1} = f(x_k) \\ x_0 \in \mathbb{R}^n \end{cases} \quad (2.6)$$

onde $f : \mathbb{R} \rightarrow \mathbb{R}$ quase linear, isto é,

$$f(x) = x + \Delta(x), \quad (2.7)$$

com $\Delta(x) \in \mathbb{R}^n$ obtido por um sistema baseado em regras fuzzy (SILVA, 2005).

A arquitetura do sistema dinâmico fuzzy 2.6 pode ser visualizada na Figura 6.

Na próxima seção, abordamos de modo resumido informações sobre aprendizado de máquina cuja compreensão é necessária para um melhor entendimento dos modelos matemáticos – desenvolvidos nos Capítulos 3 e 4 – que utilizam sistemas de aprendizado de máquina para obter os índices de perigo e risco de incêndio florestal, respectivamente.

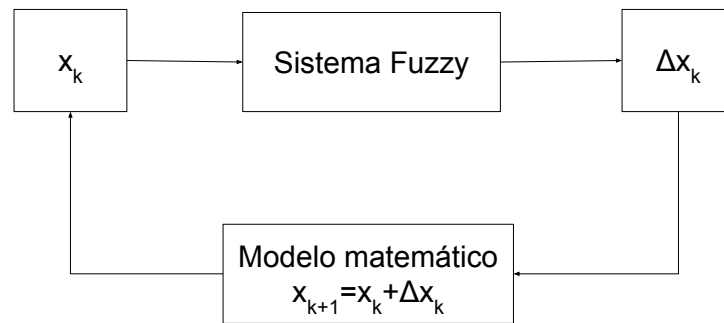


Figura 6 – Estrutura do sistema dinâmico fuzzy retirada de (SILVA, 2005)

2.5 Aprendizado de Máquina

A aprendizagem de máquina é um campo da inteligência artificial que cria algoritmos e técnicas que visam extrair padrões e regras computacionais. As técnicas de aprendizagem de máquina são amplamente aplicadas em reconhecimento de padrões e mineração de dados, utilizando algoritmos computacionais e métodos estatísticos para extrair conhecimento (ARTERO, 2009). Nesta tese utilizamos algoritmos que possuem essas características e buscam associar tanto variáveis climáticas quanto físicas à ocorrência de focos de calor, como exemplo temos o classificador kNN fuzzy e o *Subtractive Clustering* (ver Subseções 2.6.1 e 2.6.2).

Haykin (2001) afirma que existem muitas atividades associadas à noção de aprendizagem e não possuem definição precisa, pois dependem do ponto de vista do observador e do contexto no qual ele se encontra.

Como exemplo das múltiplas definições de aprendizagem, Cherkassky e Mulier (1998) definem, com uma visão mais voltada para a estatística, que aprender é o processo de estimar uma estrutura desconhecida ou dependência de um sistema a partir de um número finito de observações e de forma geral a aprendizagem abrange três componentes: gerador, sistema e máquina de aprendizagem, como pode ser visualizado na Figura 7.

O gerador é utilizado para obter vetores de entrada aleatórios a partir de uma distribuição de probabilidade fixa e desconhecida.

O sistema é responsável por produzir um valor de saída para cada vetor de entrada a partir de uma densidade condicional fixada e desconhecida, por fim, o objetivo da máquina de aprendizagem é criar um conjunto de funções $f(x, \omega)$, $\omega \in \Omega$, onde Ω é um conjunto de parâmetros abstratos usados apenas para indexar o conjunto de funções escolhidas *a priori* (CHERKASSKY; MULIER, 1998).

Um algoritmo de aprendizagem é um conjunto preestabelecido de regras bem-definidas para a solução do problema de aprendizagem (HAYKIN, 2001).

Na próxima seção abordaremos mais detalhes sobre o ato de classificar e sobre

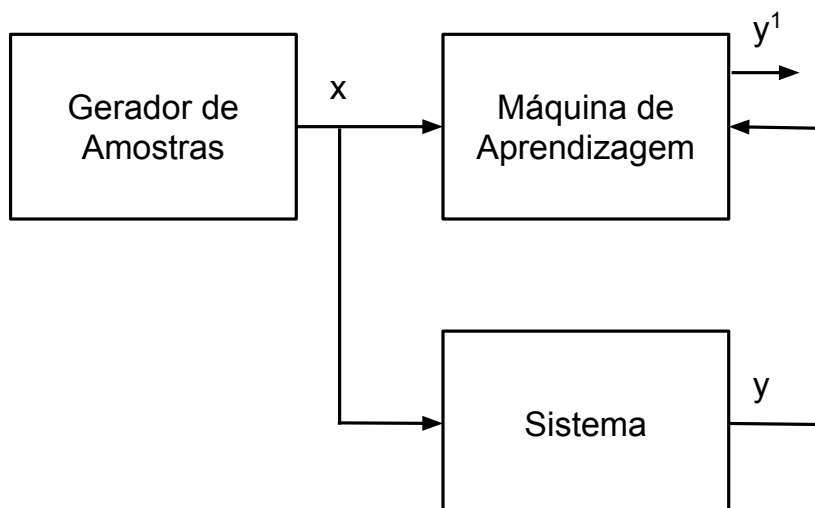


Figura 7 – Diagrama da arquitetura de uma máquina de aprendizagem. Adaptado de (CHERKASSKY; MULIER, 1998).

algoritmos classificadores que foram utilizados nesta tese para obter os índices de perigo e risco de incêndio florestal.

2.6 Classificadores

O ato de classificar é uma das atividades mais antigas do ser humano (ANDERBERG, 1973) e, com a quantidade de dados disponíveis crescendo cada vez mais, torna-se essencial classificá-los em um conjunto de categorias (XU; WUNSCH, 2005). De forma geral, os sistemas são supervisionados ou não supervisionados dependendo da disponibilidade ou não de dados rotulados (DUDA *et al.*, 2001; BISHOP, 1995; EVERITT; LANDAU; LEESE, 2001; CHERKASSKY; MULIER, 1998; JAIN; DUBES, 1988).

Na classificação supervisionada, realiza-se a modelagem de alguma função matemática $y = y(x, w)$, onde w é um vetor de parâmetros ajustáveis que realiza o mapeamento a partir de um conjunto de vetores de dados de entrada a um conjunto finito de rótulos de classes discretas. A classificação não supervisionada é caracterizada pelo fato de não haver dados rotulados disponíveis e possui o objetivo de encontrar agrupamentos naturais em um grande conjunto de dados que revelem padrões e que forneçam uma representação concisa do comportamento deles (EVERITT; LANDAU; LEESE, 2001; XU; WUNSCH, 2005; JAIN; DUBES, 1988; CHIU, 1997).

Dessa forma, dado um conjunto de dados, o objetivo de um algoritmo agrupador é criar subconjuntos disjuntos de modo que as amostras pertencentes a cada um desses subconjuntos tenham características comuns. Segundo Xu e Wunsch (2005) algoritmos agrupadores realizam a partição do conjunto de dados em uma quantidade de *clusters* que podem ser grupos, subconjuntos ou categorias. Não há consenso sobre a definição de *clusters*

(EVERITT; LANDAU; LEESE, 2001), que é descrito pela maioria dos pesquisadores, como sendo homogêneo internamente e tendo uma separação externa, ou seja, padrões que pertencem ao mesmo *cluster* devem ser mais similares entre si do que padrões de outros *clusters*, de acordo com algum conceito claro e significativo de similaridade (HANSEN; JAUMARD, 1997; JAIN; DUBES, 1988).

Em aprendizado de máquina, existem inúmeras propostas para tratamento desses problemas de classificação que se baseiam no uso dos mais diversos tipos de classificadores tais como: regressão logística, análise Discriminante, redes neurais artificiais, classificadores baseados em regras, árvores de decisão etc. Esses classificadores podem ser usados individualmente ou combinados em *ensembles* (COELHO, 2006; SEMOLINI, 2002).

Nessa tese utilizamos tanto algoritmos que realizam classificação supervisionada (knn fuzzy no Capítulo 4) e não supervisionada (*Subtractive Clustering* no Capítulo 3) e apresentamos conceitos básicos sobre esses classificadores na próxima subseção.

2.6.1 Classificadores knn e knn fuzzy

O algoritmo utilizado nessa tese para obter o risco de incêndio florestal (ver Figura 2) utiliza a regra dos k vizinhos mais próximos. Essa regra foi introduzida por Fukunaga e Hostetler (1975) e é muito utilizada em aplicações de classificação de padrões (WU; IANAKIEV; GOVINDARAJU, 2002), sistema de reconhecimento de fala (WHITE; FONG, 1975), classificação de Web site (KWON; LEE, 2003) entre outros. A abordagem utilizada em um determinado problema de classificação depende do conhecimento *a priori* que possuímos sobre ele. Se há o conhecimento estatístico completo sobre as distribuições de probabilidade conjunta das variáveis, pode-se obter uma classificação com probabilidade mínima de erro de classificação utilizando o método de Bayes (DUDANI, 1976). Agora se o conhecimento acerca do problema reside em n amostras classificadas corretamente, não existe procedimento de classificação que seja ótimo com respeito a todas as possibilidades de distribuições de probabilidade associadas ao conjunto de amostras (DUDANI, 1976).

Utilizamos nessa tese o método de classificação não-probabilístico que utiliza a regra dos k vizinhos mais próximos. Essa regra consiste em atribuir a classe para uma determinada amostra a partir da classe dos k vizinhos mais próximos. Uma vantagem de utilizar esse método é que não é necessário especificar, *a priori*, distribuições de probabilidade para as variáveis, mas sim, extrair a relação que existe entre elas a partir do conjunto de amostras disponíveis. Outra vantagem é que a medida que o valor de k vizinhos e n amostras tendem ao infinito tal que $k/n \rightarrow 0$, o erro de classificação do kNN tende ao erro de classificação mínimo do método de Bayes. Uma desvantagem consiste no elevado custo computacional quando o valor de n e k são altos, porém existem diversos trabalhos na literatura que buscam métodos para reduzir esse custo (WU; IANAKIEV;

GOVINDARAJU, 2002).

A métrica utilizada nos algoritmos elaborados nessa tese que utilizam a regra dos k vizinhos mais próximos é a distância de Mahalanobis que difere da distância Euclidiana por levar em consideração a correlação entre os conjuntos de dados. Dados dois vetores x e y a distância de Mahalanobis é dada pela equação (DE MAESSCHALCK; JOUAN-RIMBAUD; MASSART, 2000):

$$\|x - y\|^2 = (x - y)^T C^{-1} (x - y) \quad (2.8)$$

onde C é a matriz de covariância.

Apresentamos a seguir o pseudocódigo do algoritmo knn utilizado nesta tese.

Função: knearestneighbours

Descrição: O algoritmo classifica um determinado padrão na classe 0 (zero) ou 1 (um) de acordo com os k vizinhos mais próximos a ele.

Pseudo-código: 1. Inicialize k , $A_{n \times p}$ matriz com conjunto de dados e o vetor S de estados (0 ou 1) de todos os padrões de entrada.

2. Utilize uma função que encontra os k vizinhos mais próximos de todos os n padrões com uma determinada métrica.

3. Para cada padrão $i = 1, \dots, n$ é calculado a soma de todos estados dos k vizinhos mais próximos.

4. Se o resultado for menor que $(k + 1)/2$, ao estado do padrão analisado é atribuído o valor 0 e, caso contrário, é atribuído 1. fim

Para o knn fuzzy o pseudocódigo é idêntico com somente o acréscimo de uma função de pertinência que depende da distância entre os k vizinhos mais próximos, como descrito no Capítulo 4.

Na próxima subseção apresentamos uma técnica de agrupamento de dados conhecida como *Subtractive Clustering* utilizado para obter o perigo de incêndio florestal no Capítulo 3.

2.6.2 Agrupador *Subtractive Clustering*

O algoritmo *Subtractive Clustering* é utilizado para encontrar agrupamentos nos dados de umidade e precipitação utilizados pelo modelo cuja arquitetura é apresentada no bloco do Modelo de Perigo de Incêndio da Figura 2.

O objetivo desse algoritmo é separar um conjunto de dados em subconjuntos disjuntos que possuam alguma característica em comum. Cada subconjunto será denominado

de grupo que será definido por uma função gaussiana f dada por:

$$f(x) = \exp \left[-\frac{1}{2} \sum_{i=1}^n \left(\frac{x - C_i}{\sigma_i} \right)^2 \right]. \quad (2.9)$$

Portanto, precisamos que, para cada grupo, seja necessário definir o ponto $C = (C_1, C_2, \dots, C_n)$, chamado de centro, e as dispersões $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$ nas direções correspondentes às n variáveis do problema em questão.

Os algoritmos de agrupamento normalmente necessitam que o número de centros de grupos e suas posições iniciais sejam determinados *a priori*. Como exemplos, podemos citar o método de auto-organização de Kohonen (KOHONEN, 1990) e o algoritmo Fuzzy C Means (BEZDEK, 1981). A escolha do número de centros e suas localizações influenciam fortemente na qualidade da solução obtida (CHIU, 1997).

(YAGER; FILEV, 1994) propuseram um algoritmo, denominado método de montanha, para estimar a quantidade e a localização inicial dos centros dos agrupamentos. Seu método é baseado em criar uma malha no espaço de dados e calcular um valor potencial para cada ponto da malha baseado em suas distâncias aos pontos dos dados reais. Já Chiu (1997) propôs uma extensão desse método, chamado *Subtractive Clustering*, no qual cada ponto dos dados é considerado como um centro de agrupamento potencial, dispensando o uso de uma malha.

A primeira vantagem do *Subtractive Clustering*, reside no fato de que o número de “pontos de malha” a serem avaliados é igual ao número de pontos de dados, isso acontece de forma independente da dimensão problema. Ele também dispensa a necessidade de especificar uma resolução de malha, o que é uma vantagem, pois dispensa o conflito gerado entre precisão e complexidade computacional (CHIU, 1997).

Haja vista que serão utilizados classificadores para estimar o perigo e o risco de incêndio florestal nos Capítulos 3 e 4, abordamos, na próxima seção, conceitos sobre a teoria ROC que será utilizada nessa tese para avaliar a qualidade da classificação.

2.6.3 Curva ROC

A análise ROC (Receiver Operating Characteristic) originou-se a partir do desenvolvimento da teoria de decisão estatística entre 1950 e 1960 na psicologia e na avaliação de detecção de sinais de radar (METZ, 1986). Essa análise também foi utilizada para avaliar diagnósticos médicos (SWETS, 1979; HAJIAN-TILAKI, 2013). O desenvolvimento de metodologias de análise estatística dos dados ROC (pontos de uma curva ROC) foi impulsionado pelas várias aplicações da análise ROC à medicina (METZ, 1986; HANLEY; MCNEIL, 1982). Destacamos ainda que as curvas ROC são comumente usadas

para apresentar os resultados de problemas de decisão binária em aprendizado de máquina (DAVIS; GOADRICH, 2006).

Fazendo uma analogia ao diagnóstico de uma doença, para qualquer teste de diagnóstico é fixado um valor de corte para a variável em estudo que determina a classificação dos indivíduos em doente ou sadio. Valores de corte elevados conduzem a um teste pouco sensível e muito específico, em contraposição, valores de corte baixos invertem esse comportamento (CASTANHO *et al.*, 2007; BRAGA, 2000). Esse teste é avaliado pela comparação relativa da fração de verdadeiros positivos (*FVP*), fração de falsos positivos (*FFP*), fração de verdadeiros negativos (*FVN*) e fração de falsos negativos (*FFN*) (CASTANHO *et al.*, 2007; BRAGA, 2000; BRADLEY, 1997).

Após realizada a classificação, se *VP* é o número de verdadeiros positivos, *FP* de falsos positivos, *VN* verdadeiros negativos e *FN* de falsos negativos e conhecendo o número total de casos positivos *CP* e de casos negativos *CN*, podemos obter as frações citadas acima através das fórmulas (BRADLEY, 1997):

$$\begin{aligned} FVP &= \frac{VP}{CP} & FVN &= \frac{VN}{CN} \\ FFP &= \frac{FP}{CP} & FFN &= \frac{FN}{CN} \end{aligned}$$

Por definição, uma curva ROC no plano unitário é a representação gráfica dos pares (*FFP*, *FVP*), resultantes da variação do valor de corte ao longo de um eixo x (BRAGA, 2000).

Bradley (1997) investigou o uso da área sob a curva ROC (AUC do inglês *Area Under Curve*) como uma medida de desempenho para algoritmos de aprendizado de máquina e concluiu que a AUC tem sido uma das melhores maneiras de avaliar o desempenho de um classificador em um conjunto de dados quando uma avaliação de um número único é necessária ou um ponto de operação ainda não foi determinado.

Nesta tese utilizamos as áreas sob curvas ROC para avaliar quantitativamente os modelos de perigo de incêndio florestal cujos resultados foram interpretados como classificações binárias.

Neste Capítulo 2 apresentamos alguns conceitos básicos sobre subconjuntos fuzzy, sistemas baseados em regras fuzzy, sistema dinâmico fuzzy, medidas de similaridade, curva ROC e classificadores (baseados em aprendizado de máquina), para melhorar a compreensão de todos os modelos matemáticos utilizados nos capítulos posteriores e de ferramentas propostas para avaliar os resultados desses modelos.

No Capítulo 3 apresentamos resultados e bases conceituais para a modelagem do índice de perigo de incêndio florestal, como definido no Capítulo 1.

3 Índice de Perigo de Incêndio Florestal

Nesse capítulo apresentamos dois modelos matemáticos para o perigo de incêndio florestal (mutuamente excludentes) cuja arquitetura está detalhada no bloco **Modelo de Perigo de Incêndio Florestal** da Figura 2. Ambos possuem umidade relativa do ar e precipitação pluvial diária como entradas e a saída é um ponto de uma série temporal que indica o valor de perigo de incêndio π_t relativo ao dia t .

Segundo Flannigan e Wotton (2001) incêndios florestais estão fortemente ligados às condições meteorológicas e climáticas. O clima é um fator crítico para determinar o comportamento dos incêndios e o tamanho da área queimada por eles.

Tendo em vista a importância do clima para esse fenômeno, nesta seção apresentamos uma metodologia para estimar o perigo de incêndio florestal que depende apenas de variáveis climáticas.

A primeira modelagem (Seção 3.2) utiliza um sistema acoplado de agrupador e classificador que realiza aprendizado supervisionado e não-supervisionado para, através do uso dados históricos do clima, gerar a série temporal supracitada. Como a modelagem utiliza técnicas de aprendizado de máquina, não há vínculo com informações de pessoas, pois esse último modelo aprende as características dos dados históricos das variáveis de interesse (DA SILVA *et al.*, 2016).

Já segunda modelagem, apresentada na Seção 3.3, é realizada através de um sistema dinâmico fuzzy baseado em regras fuzzy (SBRF - ver Subseção 2.3) que foi elaborado a partir de informações dadas por especialistas da área de incêndios florestais, ou seja, a modelagem depende da qualidade dessas informações (DA SILVA; PONTES JR., 2011).

O objetivo de criar dois modelos distintos para resolver o mesmo problema reside no fato de que um fornece informações quantitativas importantes sobre alguns parâmetros da modelagem (versão com aprendizagem de máquina) e o outro, elaborado com sistema baseado em regras, possui características de agregar facilmente informações subjetivas que podem ser fornecidas por especialistas. A comparação entre os resultados desses modelos é realizada na Seção 3.4.

Na próxima seção abordamos a caracterização das variáveis utilizadas no modelo de perigo de incêndio florestal.

3.1 Caracterização das Variáveis

Os dados referente às variáveis climáticas utilizadas nesta tese foram obtidas do banco de dados do Instituto Nacional de Meteorologia (BRASIL, 2011). Foram solicitados dados de 1969 a 2015, porém percebemos que havia ausência de dados em determinados períodos de tempo.

Dessa forma, restringimos as séries temporais de dados da cidade de Rio Branco-AC relativas aos anos de 2003 a 2014, mesmo assim houve a necessidade de um pré-tratamento dos dados.

Em uma primeira tentativa de modelagem do perigo de incêndio florestal, utilizamos umidade relativa do ar (%), precipitação pluvial diária (mm) e temperaturas diárias máxima, mínima e média, porém percebemos que as variáveis de temperatura não tinham influência alguma no resultado do modelo de perigo de incêndio com sistema dinâmico fuzzy (Seção 3.3).

Isso ocorre por vários motivos como, por exemplo, a variação térmica que ocorre nessa cidade não é suficientemente diferente ao longo dos 365 dias do ano, por isso não interferiu no resultado obtido.

Assim, resolvemos descartar essas variáveis por agora e trabalhar apenas com umidade relativa do ar e precipitação.

A temperatura não foi testada com o modelo de perigo de incêndio que utilizou aprendizado de máquina (Seção 3.2). Essa e outras variáveis podem ser utilizadas nessa modelagem em trabalhos futuros.

A média de umidade relativa do ar diária de Rio Branco entre 2003 a 2014 foi de 84,64% e o desvio padrão foi 7,16. Esses valores apresentam uma característica “constante” e importante da região amazônica que é uma umidade alta – se comparada a outras regiões do Brasil.

A média da precipitação diária de Rio Branco entre 2003 a 2014 foi de 5,49 mm e o desvio padrão foi 12,86. Esse valor de desvio padrão maior do que o dobro da média apenas confirma o quão irregular é a ocorrência de chuvas quando observadas de forma diária.

Na próxima seção abordamos o modelo de perigo de incêndio florestal baseado em um modelo de aprendizado de máquina que possui como entradas a umidade relativa do ar e a precipitação.

3.2 Perigo de Incêndio Florestal - versão com Sistemas de Aprendizado Supervisionado e Não-Supervisionado

Sejam $U_t = U(t)$ e $P_t = P(t)$, respectivamente, a umidade relativa do ar e a precipitação pluvial no instante de tempo $t \in T = \{0, 1, 2, \dots, n\}$ em uma certa região. Além disso, seja $F_t = F(t)$ um indicador da ocorrência de focos de calor no instante de tempo $t \in T$, ou seja, $F_t = 1$ quando há ocorrência de focos de calor no tempo t e $F_t = 0$, caso contrário. Finalmente, suponhamos que dispomos um conjunto representativo $\{(U_t, P_t, F_t) : t \in T\}$ de dados históricos de uma certa região.

Com base no conjunto de dados históricos, determinamos o perigo de incêndio florestal π_t no instante de tempo t como uma função da umidade relativa do ar, U_t , e da precipitação pluvial, P_t , ou seja, $\pi_t \equiv \pi(U_t, P_t)$. Em outras palavras, o perigo de incêndio dependente apenas de variáveis climáticas bem como de dados históricos sobre a ocorrência ou não dos focos de calor (*hot pixels*) em cada instante t . Descrevemos a seguir uma metodologia para determinar o perigo de incêndio π_t .

Primeiramente, agrupamos os pontos (U_t, P_t) , $t \in T$, utilizando a técnica de aprendizado não-supervisionado *Subtractive Clustering*. A Subseção 2.6.2 apresenta mais detalhes sobre essa técnica. Vamos denotar por m o número de grupos, $C_i = (U_i, P_i)$, para $i = 1, \dots, m$, os centros dos grupos fornecidos pelo *Subtractive Clustering* aplicado aos dados históricos. Vamos denotar também por $\sigma = (\sigma_U, \sigma_P)$ as dispersões fornecidas pelo mesmo classificador para cada uma das variáveis: umidade do ar e precipitação, respectivamente.

Os centros dos grupos e as dispersões podem ser usados para determinar conjuntos fuzzy sobre o produto cartesiano da umidade relativa do ar pela precipitação pluvial. Formalmente, a pertinência de um ponto (U_t, P_t) ao conjunto fuzzy determinado pelo centro $C_i = (U_i, P_i)$ é dada pela equação 3.1.

$$\mu_i(U_t, P_t) = \exp \left[-\frac{1}{2} \left(\frac{(U_i - U_t)^2}{\sigma_U^2} + \frac{(P_i - P_t)^2}{\sigma_P^2} \right) \right], \quad \forall i = 1, 2, \dots, m. \quad (3.1)$$

Os subconjuntos fuzzy caracterizados pelas funções de pertinência μ_1, \dots, μ_m definem uma partição do universo de discurso obtido pelo produto cartesiano da umidade relativa do ar com a precipitação pluvial. Sobretudo, considerando a frequência de focos de incêndio dos dados históricos, atribuímos a cada um dos subconjuntos fuzzy um parâmetro

f_i , dado pela equação

$$f_i = \frac{\sum_{t \in A_i} \mu_j(U_t, P_t) F_t}{\sum_{t \in A_i} \mu_i(U_t, P_t)}, \quad \forall i = 1, \dots, m \quad (3.2)$$

em que os conjuntos A_i são dados por:

$$A_i = \left\{ t \in T : \mu_i(U_t, P_t) = \max_{k=1:m} \{ \mu_k(U_t, P_t) \} \right\}. \quad (3.3)$$

é o conjunto dos instantes de tempo para os quais (U_t, P_t) possuem a maior pertinência ao i -ésimo conjunto fuzzy.

Observe-se que, para qualquer $i \in \{1, \dots, m\}$, o parâmetro f_i é dado por uma média ponderada dos indicadores da ocorrência de focos de incêndio F_t para alguns valores de $t \in T$. Por conta disso, houve a necessidade de padronizar o parâmetro f_i de modo que o valor máximo seja 1. Além disso, uma vez que o fenômeno ocorre somente em condições de baixa umidade e precipitação, foi incorporado ao modelo um limiar $L > 0$, abaixo do qual os parâmetros f_i são eliminados. Assim, esse limiar L tem a função de selecionar quais conjuntos fuzzy (definidos pelo centro (U, P) e pela dispersão σ) contribuem para a ocorrência do fenômeno. Concluindo, definimos o *fator de perigo* do i -ésimo conjunto fuzzy, denotado por ϕ_i , como segue para todo $i = 1, \dots, m$:

$$\phi_i = \begin{cases} \frac{f_i}{\max_j \{f_j\}}, & \text{se } f_i \geq L, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.4)$$

Os fatores de perigo são usados para estimar o perigo de incêndio pontual $\tilde{\pi}_t$ de forma semelhante a (3.2). Precisamente, dadas a umidade relativa do ar U_t e a precipitação pluvial P_t num instante qualquer t , definimos $\tilde{\pi}_t \equiv \tilde{\pi}_t(U_t, P_t)$ por

$$\tilde{\pi}_t = \frac{\sum_{i=1}^m \mu_i(U_t, P_t) \phi_i}{\sum_{i=1}^m \mu_i(U_t, P_t)}, \quad (3.5)$$

em que $\mu_i(x_t)$ e ϕ_i representam, respectivamente, a pertinência de (U_t, P_t) e o fator de perigo do i -ésimo conjunto fuzzy.

Finalmente, admitimos que o perigo de incêndio não possui grandes variações entre os instantes de tempo t e $t + 1$. Em outras palavras, π_t depende não somente da umidade relativa do ar U_t e da precipitação pluvial P_t , no instante t , mas também depende k valores passados $\pi_{t-1}, \pi_{t-2}, \dots, \pi_{t-k}$ do perigo de incêndio. Formalmente, definimos o

perigo de incêndio atual $\pi_t \equiv \pi_t(U_t, P_t)$ pela média aritmética dos k valores anteriores com uma estimativa pontual $\tilde{\pi}_t$, ou seja,

$$\pi_t = \frac{1}{(k+1)} \left(\tilde{\pi}_t + \sum_{j=1}^k \pi_{t-j} \right). \quad (3.6)$$

O diagrama da Figura 8 ilustra a arquitetura do modelo proposto para o perigo de incêndio florestal que é o mesmo apresentado no bloco **Modelo de Perigo de Incêndio Florestal** na Figura 2.

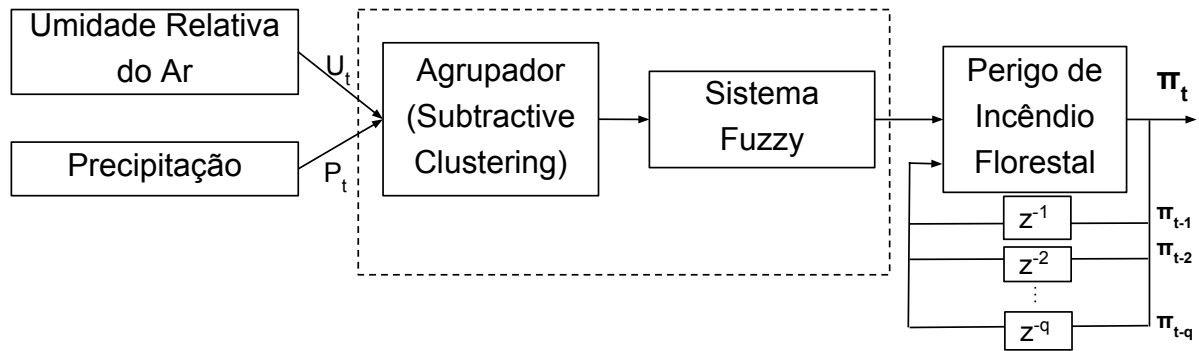


Figura 8 – Diagrama em blocos da arquitetura do modelo de sistema fuzzy para o perigo de incêndio florestal.

Destacamos a importância da definição do perigo de incêndio florestal a partir da média aritmética. Sem ela, o modelo apenas fornece valores de perigos de incêndio muito próximos de zero ou muito próximos de um, de forma alternada. Esse fato acontece, provavelmente, em virtude da adoção da variável de saída que é considerada no modelo ser **ocorrência de foco de calor** (binária), ao invés de expressar a **quantidade total de focos de calor**. Além disso, para que o modelo seja uma ferramenta de alerta antecipado de ocorrências de incêndios florestais, é importante que ele utilize as informações de perigo de incêndio florestal dos dias anteriores ao dia t para caracterizar uma janela de tempo correspondente ao período mais seco do ano.

Para avaliar o desempenho do modelo, definimos o **Alerta Climático** α_t da seguinte forma

$$\alpha_t = \begin{cases} 1, & \text{se } \pi_t > 0, \\ 0, & \text{c. c..} \end{cases} \quad (3.7)$$

Na próxima subseção, esse modelo será aplicado ao Estado do Acre-Brasil e lá será discutido como o **Alerta Climático** α_t é utilizado para avaliarmos quantitativamente o desempenho do modelo proposto.

3.2.1 Perigo de Incêndio Florestal para o Estado do Acre - Estudo 1

Nesta subseção apresentamos um estudo do modelo de aprendizado de máquina para obtermos o perigo de incêndio florestal para o Estado do Acre-Brasil. Para tanto, considerou-se um conjunto de dados contendo informações da umidade relativa do ar, precipitação pluvial e número de focos de calor na região da capital do estado, Rio Branco, nos anos de 2003 a 2014 (BRASIL, 2011; INPE, 2011).

Precisamente, os dados de 2003 a 2013 foram usados para treinamento enquanto que os dados de 2014 foram usados como conjunto de teste. A partir dos dados de treinamento, foram obtidos, através da técnica *Subtractive Clustering*, nove conjuntos fuzzy no produto cartesiano da umidade do ar pela precipitação pluvial, cujas coordenadas U_i e P_i dos centros estão mostradas na Tabela 2 e a Figura 9 apresenta graficamente esses centros e uma parcela dos dados de treinamento utilizados. Ilustramos a função de pertinência de um dos nove conjuntos fuzzy na Figura 10 que apresenta também parte dos dados de treinamento e os outros centros.

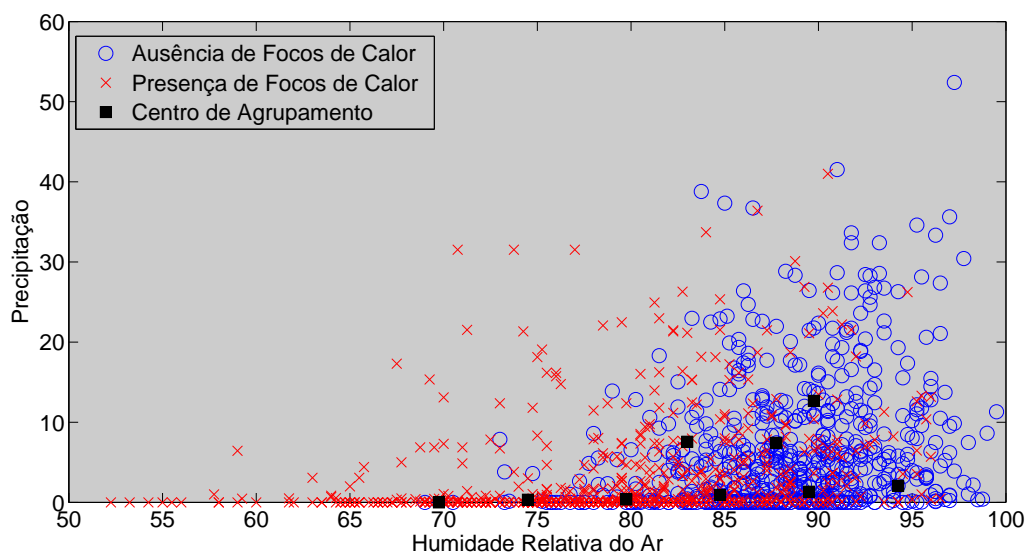


Figura 9 – Centros dos agrupamentos obtidos pelo agrupador *Subtractive Clustering* e uma parte do conjunto de dados de treinamento.

3.2.1.1 Resultados

A Tabela 2 também apresenta o fator de perigo ϕ_i para cada um dos conjuntos fuzzy. O parâmetro L foi determinado através de uma busca exaustiva considerando os valores $L \in \{0.1; 0.2; \dots; 0.9; 1\}$, em virtude do baixo custo computacional e de ausência de informações complementares de especialistas e da literatura. Adotamos o limiar $L = 0.3$ em (3.4), pois o mesmo maximiza o desempenho médio do modelo utilizando a técnica de validação cruzada com 10 pastas. O valor desse parâmetro depende da região em estudo e

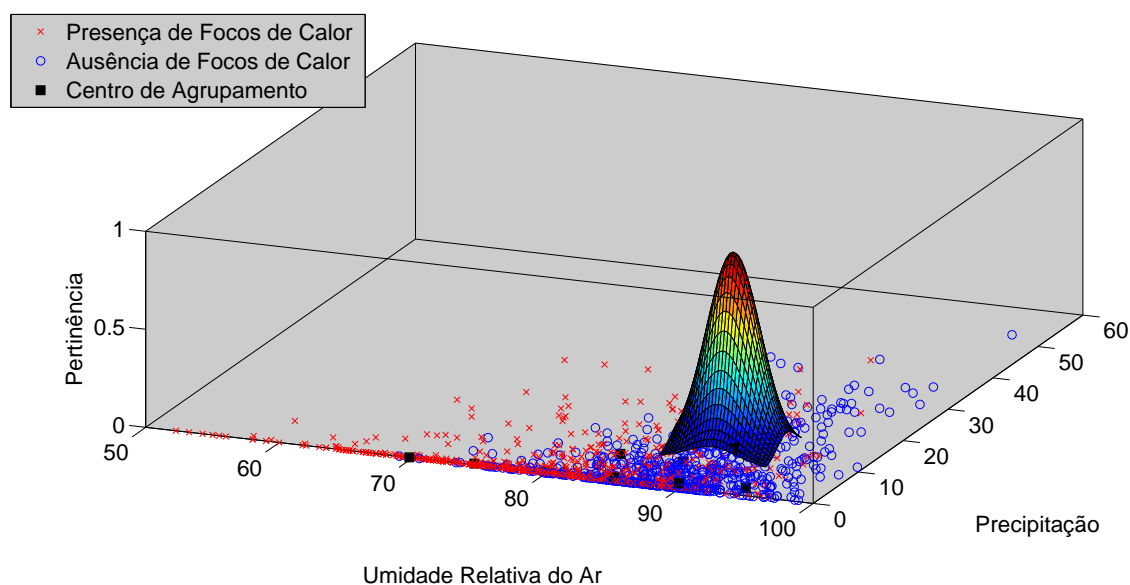


Figura 10 – Função de pertinência de um subconjunto fuzzy centrado no ponto $(U, P) = (89.75, 12.67)$.

poderia ser determinado de diversas formas, tais como: informações de um especialista, busca exaustiva ou obtendo informações da literatura específica.

Tabela 2 – Informações sobre os conjuntos fuzzy obtidos pelo agrupador *subtractive clustering* para a cidade de Rio Branco-AC, no período 2003-2013.

Grupos	Umidade (%)	Precipitação (mm)	f_i	ϕ
7	94,25	2,07	0,00	0,00
9	89,75	12,67	0,09	0,00
3	89,50	1,33	0,11	0,00
5	87,75	7,43	0,21	0,00
1	84,75	0,93	0,21	0,00
8	83,00	7,57	0,64	0,64
2	79,75	0,37	0,69	0,69
4	74,50	0,33	0,89	0,89
6	69,75	0,00	1,00	1,00

Observamos, na Tabela 2, que os conjuntos fuzzy que possuem centros com valores relativamente baixos de umidade e precipitação (para a região estudada) estão associados a maiores fatores de perigo ϕ_i . Isso concorda com a literatura específica (SOARES; BATISTA, 2007; NEPSTAD; MOREIRA; ALENCAR, 1999), a qual informa que o período de maior perigo de incêndio florestal é quando temos ausência de chuvas e, como consequência, a umidade relativa do ar diminui significativamente.

O número k de atrasos considerados para o cálculo do perigo de incêndio florestal também foi determinado por uma busca exaustiva combinada com a validação cruzada com 10 pastas. Foram testados os valores para $k \in \{1, \dots, 10\}$ e o valor ótimo foi

de $k = 3$.

A simulação realizada com o conjunto de teste considerou os dados climáticos e de focos de calor de Rio Branco-AC para os 365 dias do ano de 2014. As Figuras 11 e 12 ilustram os resultados obtidos.

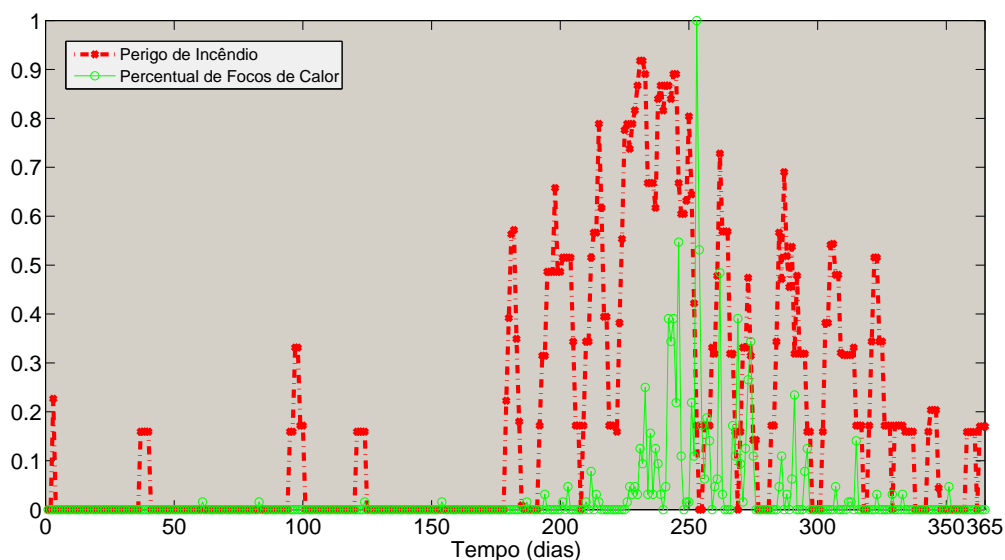


Figura 11 – Percentual de focos de calor e perigo de incêndio florestal para o Estado do Acre em 2014.

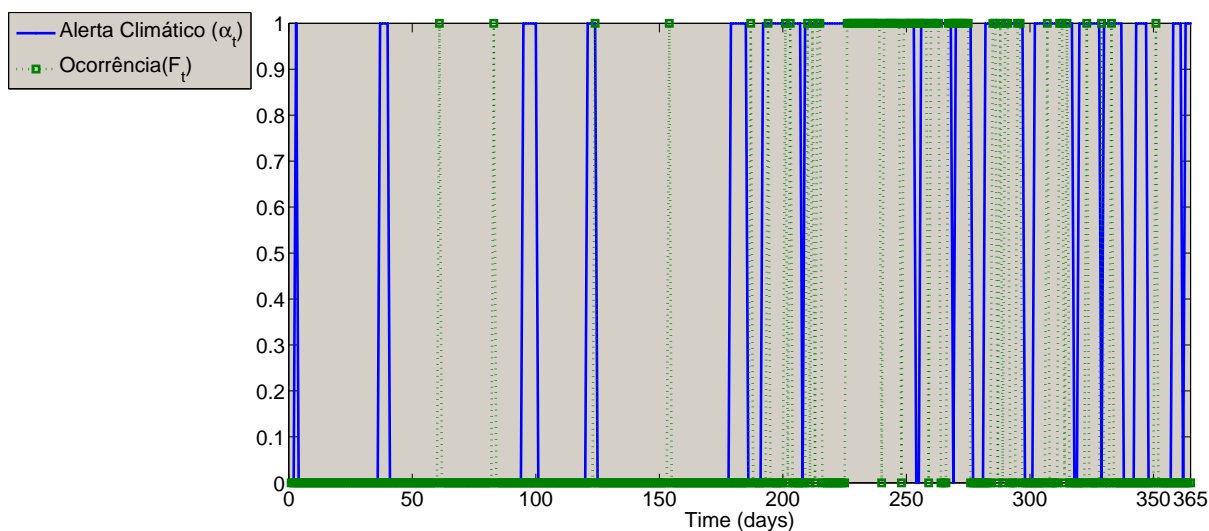


Figura 12 – Ocorrência de focos de calor F_t e alerta climático α_t para o Estado do Acre em 2014.

Pode-se verificar na Figura 11 que o perigo de incêndio permanece nulo, com algumas exceções, durante o primeiro semestre do ano que é um período bastante chuvoso na região norte. Porém, a partir do mês de junho, as chuvas começam a ficar escassas e a umidade começa a diminuir, o que provoca a ocorrência de focos de calor.

O perigo de incêndio florestal começa a aumentar significativamente no mesmo período em que aparecem os focos de calor e, mesmo atingindo alguns valores nulos no período de seca, esse índice permanece positivo e alto durante a ocorrência dos focos de calor.

Para avaliar quantitativamente a estimativa do perigo de incêndio (π_t), consideramos os resultados obtidos como uma classificação binária na qual, para cada instante de tempo t , o verdadeiro positivo (tp) é caracterizado por $\alpha_t = 1$ (perigo é positivo) e $F_t = 1$, (houve a ocorrência de foco de calor). De forma análoga, o verdadeiro negativo (tn) ocorre quando $\alpha_t = 0$ e $F_t = 0$. O falso negativo (fn) ocorre quando $\alpha_t = 0$ e $F_t = 1$ e, por fim, para o falso positivo (fp) temos $\alpha_t = 1$ e $F_t = 0$. A Tabela 3 resume essas ideias.

Tabela 3 – Classificação utilizada para avaliar a eficiência do modelo de perigo de incêndio florestal.

Classificação	Alerta Climático (α_t)	Ocorrência F_t
Verdadeiro Positivo (tp)	1	1
Verdadeiro Negativo (tn)	0	0
Falso Positivo (fp)	1	0
Falso Negativo (fn)	0	1

Após serem preditos os valores de perigo de incêndio florestal para o conjunto de teste (ano de 2014), obteve-se taxa de acerto igual a 0.6712 e a área abaixo da curva ROC (do inglês *Receiver Operator Characteristic*) de 0.8562. Mais detalhes sobre curva ROC podem ser encontrados na Subseção 2.6.3. A Figura 13 apresenta a curva ROC obtida.

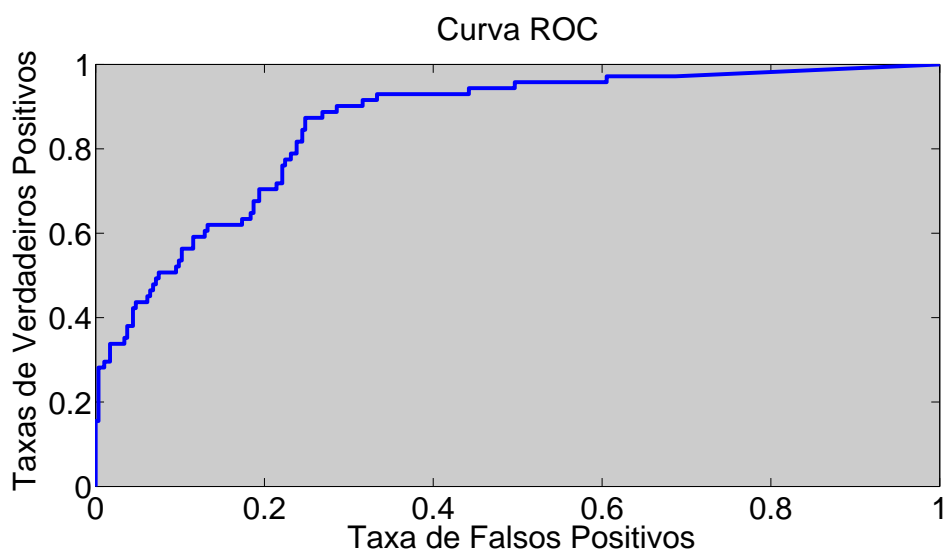


Figura 13 – Curva ROC obtida da classificação realizada para avaliar o modelo de perigo de incêndio florestal.

Na próxima seção, discutimos um modelo matemático alternativo para o perigo de incêndio florestal. Esse modelo utiliza um sistema dinâmico fuzzy que utiliza um sistema

baseado em regras fuzzy (SBRF) para gerar a série temporal de valores de perigo de incêndio somente a partir de informações climáticas.

3.3 Perigo de Incêndio Florestal através de Sistema Dinâmico Fuzzy

A ideia principal dessa seção é desenvolver uma modelagem baseada em informações subjetivas de especialistas de incêndio florestal acerca da relação existente entre a umidade relativa do ar e a precipitação pluvial diária.

Além disso, também foram utilizados dados históricos das variáveis existentes para embasar as informações obtidas. Esses dados subjetivos foram traduzidos em variáveis linguísticas – modeladas por conjuntos fuzzy – e uma base de regras fuzzy.

Tanto a partir das informações dadas por especialistas quanto observando as séries temporais diárias de umidade relativa do ar, precipitação e focos de calor procuramos identificar valores de umidade e precipitação que estavam ocorrendo no início, meio e fim do período de ocorrência dos focos de calor.

Identificamos valores dessas variáveis de entrada que impedem a ocorrência do fenômeno. Esses valores foram utilizados para ajustar o suporte dos conjuntos fuzzy que compõem as variáveis linguísticas de entrada.

A base de regras apresentada na Tabela 4 foi obtida a partir de informações qualitativas tais como: “Quando a umidade está muito alta e a precipitação está muito alta não há ocorrência de incêndios florestais”.

Apresentamos mais detalhes sobre subconjuntos fuzzy, variáveis linguísticas, sistema de base de regras fuzzy e sistema dinâmico fuzzy no Capítulo 2.

Os valores do suporte dos números fuzzy para todas as variáveis linguísticas do sistema fuzzy foram ajustados a partir de informações fornecidas por especialistas e das informações de umidade, precipitação e focos de calor da cidade de Rio Branco-AC nos anos de 2003 a 2013. Para o sistema dinâmico serão utilizadas as variáveis umidade relativa do ar e precipitação pluvial que assumem os estados: MUITO BAIXA, BAIXA, MÉDIA, ALTA, MUITO ALTA, conforme Figuras 14 e 15.

O método de inferência utilizado no sistema baseado em regras fuzzy é o de Mamdani e a técnica de defuzzificação utilizada foi o método do Centro de Gravidade ou Centróide (ver Seção 2.3) (GOMIDE; PEDRYCZ, 1998; BARROS; BASSANEZI; LODWICK, 2017).

A Tabela 4 apresenta a base de regras do sistema baseado em regras fuzzy utilizado nesta tese (adaptada de (DA SILVA; PONTES JR., 2011)). Essas regras foram elaboradas a partir da informação de um especialista em incêndio florestal do Acre e de análise subjetiva de dados históricos tanto climáticos quanto de focos de calor.

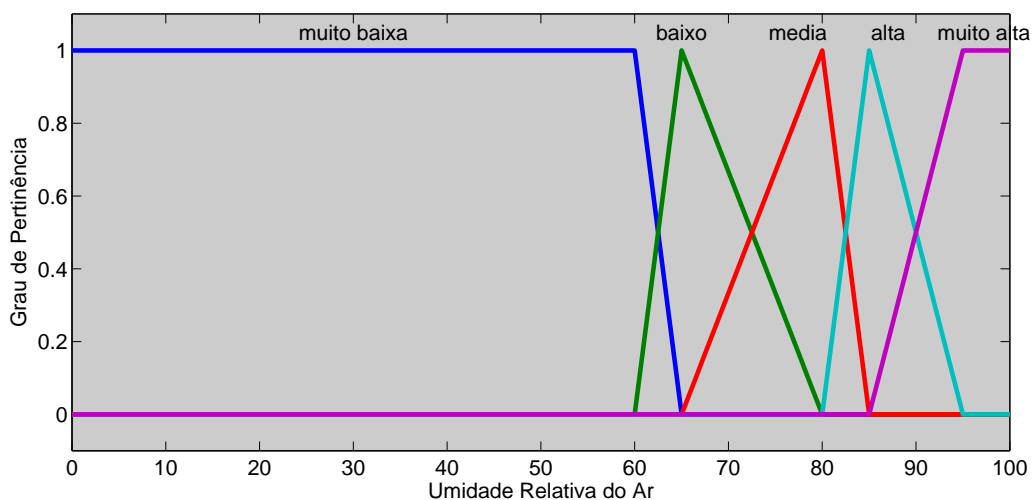


Figura 14 – Variável linguística: umidade relativa do ar (%).

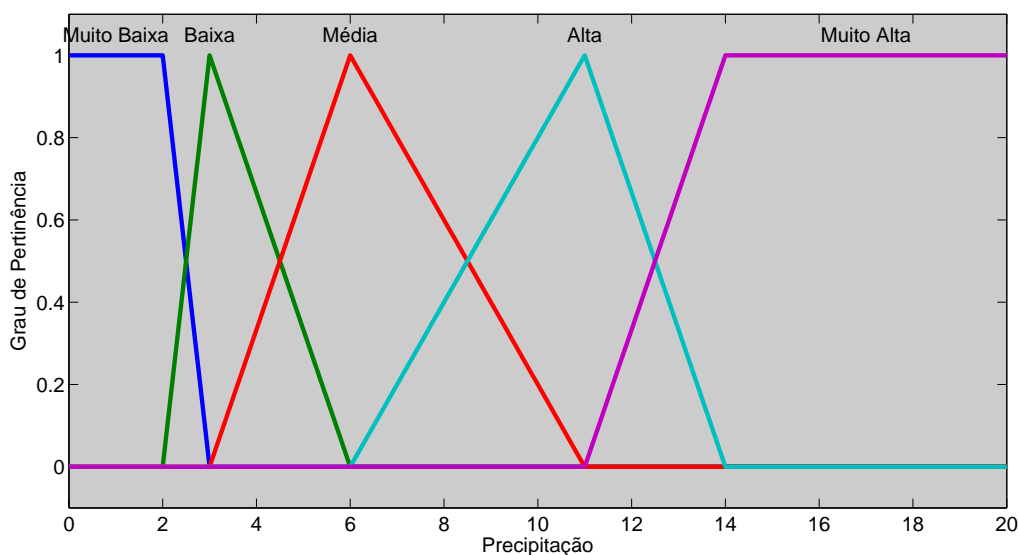


Figura 15 – Variável linguística: precipitação diária em mm.

Tabela 4 – Base de regras utilizadas no sistema dinâmico fuzzy. Na primeira coluna estão as adjetivações que definem os números fuzzy correspondentes à umidade relativa do ar, na primeira linha estão os relativos à precipitação diária e cada posição da tabela informa o valor da variação do perigo $\Delta(U_t, P_t)$ correspondente.

Umidade \ Precipitação	Precipitação				
	muito baixa	baixa	média	alta	muito alta
muito baixa	Alta Pos.	Alta Pos.	Média Pos.	Média Neg.	Baixa Neg.
baixa	Alta Pos.	Média Pos.	Baixa Pos.	Média Neg.	Média Neg.
média	Média Pos.	Média Pos.	Média Neg.	Alta Neg.	Alta Neg.
alta	Média Neg.	Média Neg.	Alta Neg.	Alta Neg.	Alta Neg.
muito alta	Alta Neg.	Alta Neg.	Alta Neg.	Alta Neg.	Alta Neg.

Essas regras se aplicam ao comportamento do fenômeno dos incêndios florestais para o clima amazônico. Para aplicar esse modelo a outras regiões do Brasil e do mundo, tanto a base de regras quanto os suportes dos conjuntos fuzzy das variáveis linguísticas devem ser reajustadas.

Nesse modelo também utilizamos uma média móvel na série temporal à medida ela vai sendo construída. Isso é feito a partir de uma média aritmética entre a variação do perigo $\Delta(U_t, P_t)$ do dia t com os valores de perigo de incêndio de n atrasos da série temporal. O valor de n pode ser obtido a partir do uso de alguma técnica estatística aplicada aos dados da série histórica como, por exemplo a *10-fold cross validation* aqui adotada.

A Equação 3.8 é utilizada para obter o perigo de incêndio florestal π_t^d no instante t . Ela consiste em uma variação do sistema dinâmico fuzzy proposto por (SILVA, 2005).

$$\pi_t^d = \min \left\{ 1, \max \left[0, \left(\frac{1}{n+1} \sum_{i=1}^n \pi_{t-i}^d + \frac{1}{n+1} \Delta(U_t, P_t) \right) \right] \right\}. \quad (3.8)$$

Aqui, n é o número de atrasos no tempo, $\Delta(U_t, P_t)$ é obtido por um sistema baseado em regras fuzzy, no qual as variáveis linguísticas de entrada são a umidade relativa do ar U_t e precipitação pluvial total diária P_t no instante t . A saída do sistema é a Variação do Perigo $\Delta(U_t, P_t)$ conforme indicado na Figura 17. A Figura 16 ilustra a arquitetura do modelo do sistema dinâmico proposto. Essa arquitetura também faz parte da Figura 2 no bloco Modelo de Perigo de Incêndio Florestal.

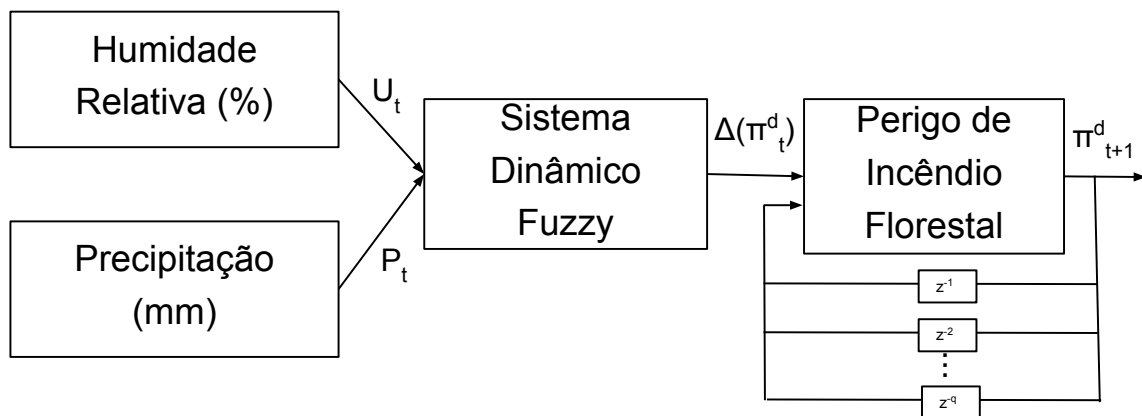


Figura 16 – Diagrama em blocos da estrutura do sistema dinâmico fuzzy

A modificação proposta na definição de sistema fuzzy é justificada pelo fato de que é desejável que o perigo de incêndio florestal seja um valor pertencente ao intervalo $[0, 1]$ a cada instante de tempo t . Como podem ocorrer valores de variação do perigo negativos ($\Delta(U_t, P_t) < 0$), possivelmente ocorreriam valores negativos de perigo de incêndio.

Além disso, sem utilizar a função mínimo, o valor de perigo de incêndio também poderia crescer de forma inadequada para uma possível normalização necessária.

Adotou-se a condição inicial do sistema dinâmico (perigo de incêndio florestal inicial) como sendo $\pi_0^d = 0$ pelo fato de que no primeiro dia do ano o Acre está em pleno período chuvoso e, salvo raras exceções, nesse período não há ocorrências de focos de calor. A partir daí, para cada dia t , $\Delta(U_t, P_t)$ e a média com n atrasos no tempo são calculados.

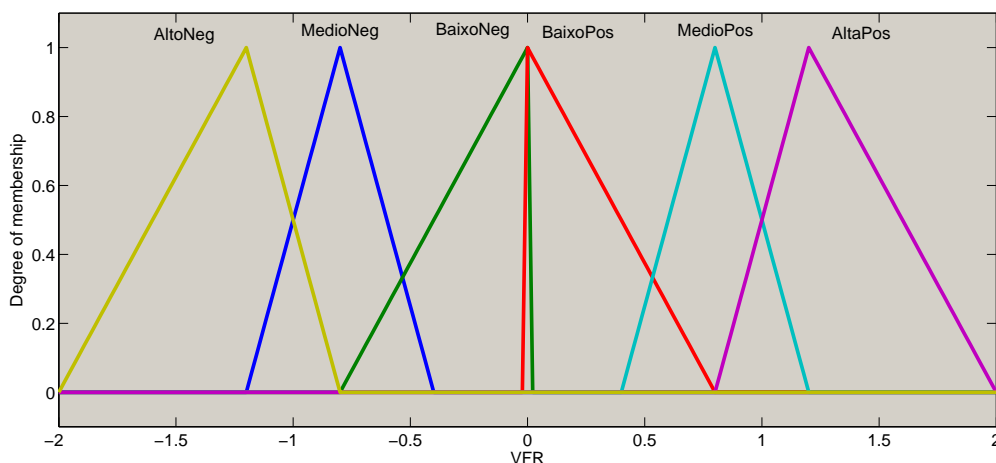


Figura 17 – Variável linguística variação do perigo $\Delta(U_t, P_t)$

Para avaliar o desempenho do modelo, utilizamos uma classificação binária utilizando a ocorrência de focos de calor F_t e o **Alerta Climático Dinâmico** α_t^d definido da seguinte forma:

$$\alpha_t^d = \begin{cases} 1, & \text{se } \pi_t^d > 0, \\ 0, & \text{c. c..} \end{cases} \quad (3.9)$$

A qualidade da modelagem será indicada pela área abaixo da curva ROC associada à classificação realizada. Na próxima subseção, aplicamos o modelo aos dados referentes à cidade de Rio Branco-AC do ano de 2014.

3.3.1 Perigo de Incêndio Florestal para o Estado do Acre - Estudo 2

Da mesma forma que no primeiro modelo apresentado na Subseção 3.2.1, o modelo de perigo de incêndio florestal utilizando sistema dinâmico fuzzy será avaliado a partir de dados de umidade relativa do ar, precipitação diária e ocorrência de focos de calor de Rio Branco-AC em 2014.

3.3.1.1 Resultados

As Figuras 18 e 19 apresentam o resultado das simulações.

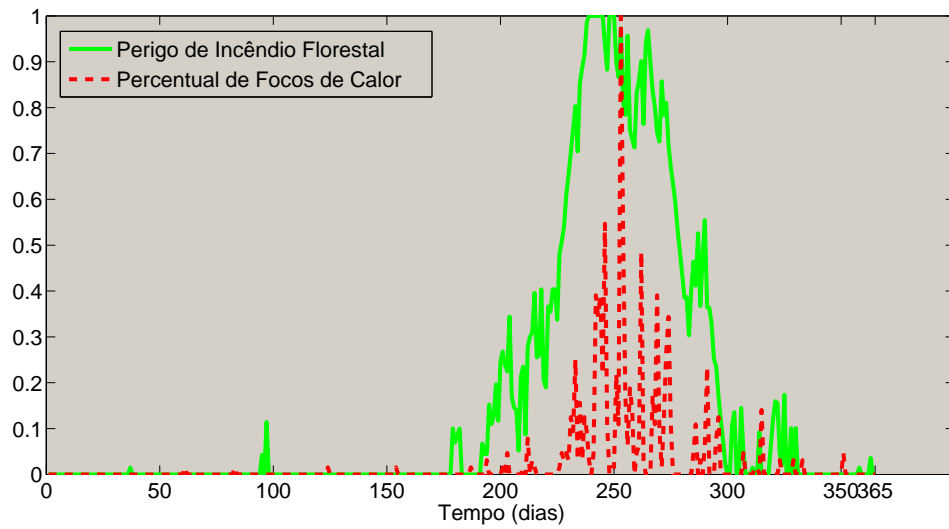


Figura 18 – Perigo de incêndio florestal (π^d) e percentual de focos de calor (dados reais).

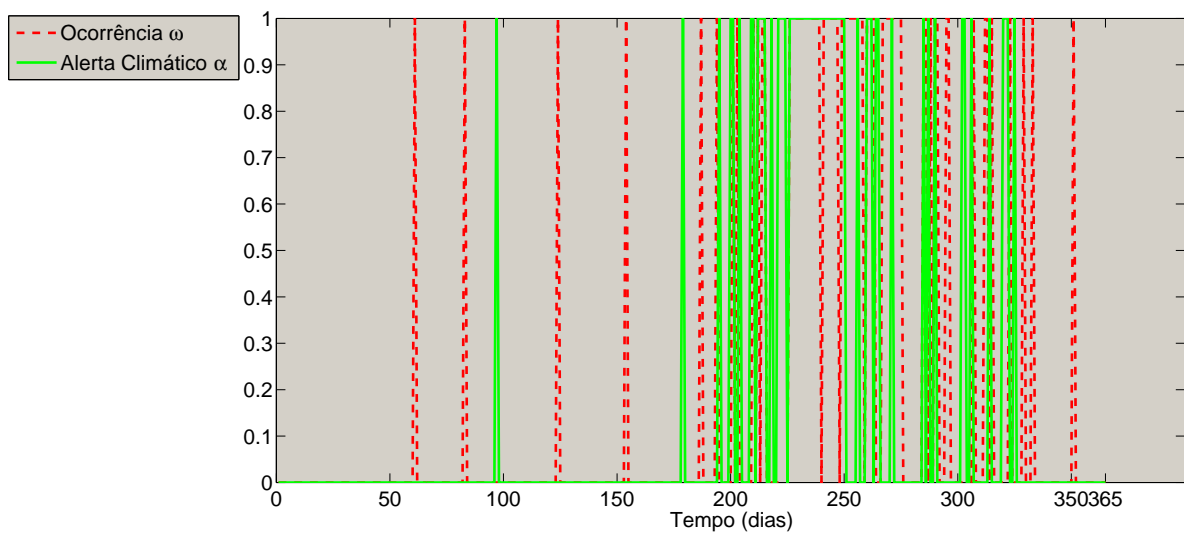


Figura 19 – Alerta climático dinâmico (α_t^d) e ocorrência de focos de calor F_t .

A Figura 19 apresenta o Alerta Climático Dinâmico (α_t^d) e a Ocorrência de Focos de Calor F_t que são duas variáveis binárias utilizadas na avaliação do modelo de perigo de incêndio florestal (com sistema dinâmico fuzzy) através da mesma classificação descrita na Tabela 3.

De forma análoga à Subseção 3.2.1, iremos utilizar uma classificação binária para avaliar o desempenho do modelo de perigo de incêndio com sistema dinâmico fuzzy. Dessa forma, iremos considerar o alerta climático dinâmico α_t^d e a ocorrência de focos de calor F_t para, a partir da Tabela 3, realizar a necessária classificação. O desempenho desse modelo será medido pela área abaixo da curva ROC (AUC) apresentada na Figura 20. Obtivemos o valor de $AUC = 0.8706$.

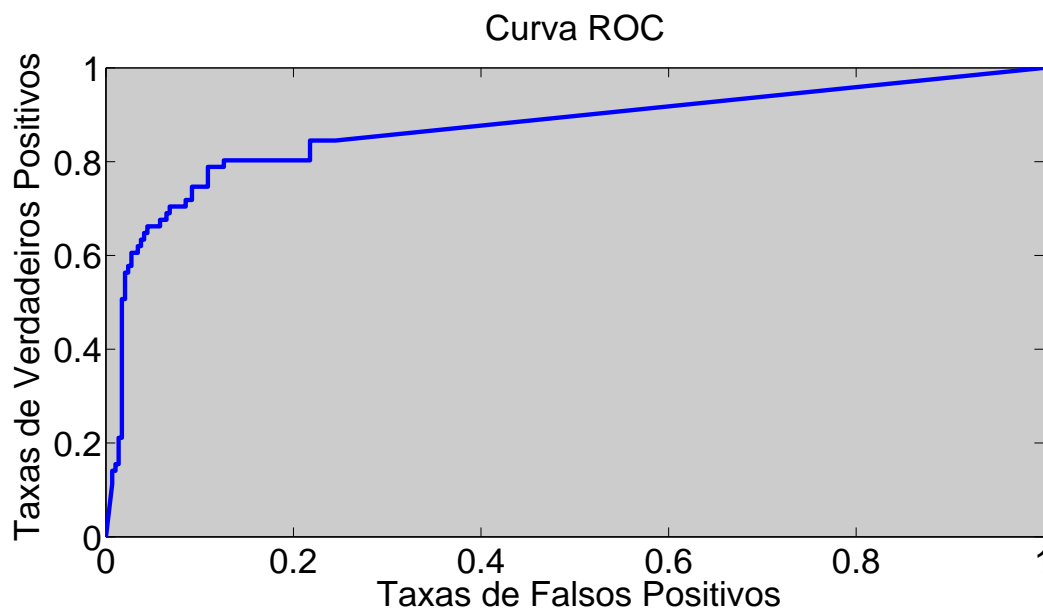


Figura 20 – Curva ROC associada à classificação utilizada para avaliar o desempenho do modelo de sistema dinâmico fuzzy.

Considerando as informações fornecidas pelos dois modelos de perigo de incêndio florestal quando aplicados aos dados reais do Estado do Acre, mostraremos na Seção 3.4 uma certa “equivalência” entre essas informações, apesar de terem sido elaborados a partir de técnicas diferentes.

3.4 Avaliação Conjunta dos Modelos de Perigo de Incêndio Florestal

Apesar dos modelos de perigo de incêndio descritos acima terem sido elaborados a partir de ferramentas matemáticas distintas, ambos apresentaram parâmetros semelhantes de forma quantitativa.

A Tabela 5 apresenta os valores de umidade obtidos como primeira coordenada dos 9 centros de agrupamentos fornecidos pelo *Subtractive Clustering*, o parâmetro f_i associado a cada agrupamento e as pertinências relativas aos conjuntos fuzzy que compõem a variável linguística umidade do sistema fuzzy.

Embora estejam disponíveis as demais pertinências, com vistas à melhor compreensão optamos por apresentar a Tabela 5 de modo mais “enxuto” e, assim, esperamos de mais fácil compreensão.

Como exemplo, o agrupamento 6 (seis) possui a primeira coordenada com valor de 69,75 mm de umidade e fator de perigo igual a 1 (um), o mesmo valor de umidade possui pertinência igual a 0,98 ao subconjunto fuzzy associado ao adjetivo **baixa** da variável linguística umidade.

Tabela 5 – Valores de umidade e parâmetro f_i associados aos 9 agrupamentos (fornecidos pelo *Subtractive Clustering*) juntamente com as pertinências relativas aos conjuntos fuzzy da variável linguística umidade do sistema fuzzy. Na coluna umidade encontram-se os valores das primeiras coordenadas dos centros dos agrupamentos obtidos no modelo de perigo de incêndio com aprendizado de máquina.

Grupo	Modelo com Aprendizado de Máquina		Modelo com Sistema Dinâmico fuzzy	
	Umidade	f_i	Maior Pertinência	Adjetivo
1	84,75	0,21	0,63	Alta
2	79,75	0,69	0,95	Média
3	89,50	0,11	0,73	Alta
4	74,50	0,89	0,55	Baixa
5	87,75	0,21	0,97	Alta
6	69,75	1,00	0,98	Baixa
7	94,25	0,00	0,85	Muito Alta
8	83,00	0,64	0,40	Média
9	89,75	0,09	0,70	Alta

Ao analisar a Tabela 5 pode-se perceber que valores de umidade associados aos conjuntos fuzzy “baixa” possuem maiores valores de parâmetro f_i que os associados aos conjuntos fuzzy “alta”. Valores de umidade que possuem maior pertinência ao conjunto fuzzy “baixa” também possuem maior parâmetro f_i em relação àqueles que possuem menor pertinência. Por um determinado aspecto, entendemos que o modelo de sistema de aprendizado de máquina fornece uma caracterização quantitativa da relação entre a variável umidade e o índice de perigo de incêndio que está sendo caracterizado pelo parâmetro f_i .

A mesma conclusão não podemos afirmar em relação à precipitação. Isso ocorre porque a precipitação possui caráter mais aleatório e como acontecem medições de chuva durante o período mais seco do ano, os dados reais possuem muito ruído o que prejudica os algoritmos de aprendizados de dados.

Nesse Capítulo 3 realizamos duas abordagens (mutuamente exclusivas) para obtermos o perigo de incêndio florestal. Esse índice possui a umidade relativa do ar e precipitação pluvial diária como variáveis de entrada e a saída é o perigo de incêndio florestal no dia t (π_t). A sequência de valores de perigo de incêndio é chamada de série temporal que foi obtida na Seção 3.2 por um sistema de aprendizado de máquina supervisionado e não supervisionado e na Seção 3.3 utilizamos o sistema dinâmico discreto 3.8 baseado em um Sistema Baseado em Regras Fuzzy (Seção 2.3) para obter a série temporal de perigo de incêndio. Nas Seções 3.2.1 e 3.3.1 aplicamos os modelos de perigo de incêndio a dados relativos à cidade de Rio Branco-AC e avaliamos esses modelos utilizando a área abaixo da curva ROC associada à classificação proposta na Tabela 3.

No próximo capítulo introduziremos a noção de risco de incêndio florestal (ρ)

que difere da modelagem realizada até agora pelo fato de considerar variáveis que não são modificadas (ou são muito pouco modificadas) com o passar do tempo. Utilizamos um classificador do tipo kNN fuzzy que nos fornece um mapa de risco de incêndio florestal estático no tempo, mas que será de fundamental importância para a definição do índice de alerta de incêndio florestal (W) realizada no Capítulo 5.

4 Índice de Risco de Incêndio Florestal

Neste capítulo apresentaremos uma metodologia, baseada na teoria dos conjuntos fuzzy, que fornece um índice ρ , chamado *risco de incêndio florestal*, relacionado à chance de um incêndio se iniciar. A arquitetura desse modelo está detalhada no bloco **Modelo de Risco de Incêndio Florestal** que é uma parte do diagrama geral apresentado na Figura 2.

Em termos matemáticos, deseja-se encontrar uma função $\rho : \Omega \rightarrow [0, 1]$ tal que $\rho(x) \in [0, 1]$ represente o risco de incêndio florestal associado ao ponto $x \in \Omega \subseteq \mathbb{R}^n$. Nesse trabalho, considerou-se um conjunto com informações geo-referenciadas $\Omega \subseteq \mathbb{R}^4$ cujos elementos são vetores cujas coordenadas representam: altitude (y_1), tipologia (y_2), distância para curso d'água (y_3) e distância para a estrada (y_4) mais próximos, respectivamente. Essas variáveis foram escolhidas por apresentarem uma forte relação com a ocorrência de incêndios e são utilizadas por vários trabalhos da área específica de incêndios florestais (NEPSTAD; MOREIRA; ALENCAR, 1999; BATISTA; OLIVEIRA; SOARES, 2002; BATISTA *et al.*, 2014). As coordenadas geográficas (i, j) de $x = (y_1, y_2, y_3, y_4)$ não são utilizadas para a obtenção de $\rho(x)$, porém elas são utilizadas para obtermos um *mapa de risco de incêndio florestal*, ou seja, para a posição (i, j) do mapa, obtemos o risco de incêndio florestal calculando $\rho(x)$ no ponto de coordenada (i, j) .

Vamos admitir que podemos determinar o mapa de risco usando um conjunto representativo de dados $\mathcal{T} = \{(x^r, \omega_1^r) : r = 1, \dots, n\} \subseteq \Omega \times \{0, 1\}$, chamado conjunto de treinamento. Aqui, $x^r = (y_1, y_2, y_3, y_4) \in \Omega$ contém informações geo-referenciadas enquanto $\omega_1^r \in \{0, 1\}$ indica a classe de x^r e representa a ocorrência ou não de focos de calor em x^r , para todo $r = 1, \dots, n$ em um certo intervalo de tempo.

Determinamos o risco de incêndio a partir de \mathcal{T} e do classificador binário do tipo k -vizinhos mais próximos fuzzy (kNN fuzzy, do inglês *k-nearest neighbours fuzzy*). As duas classes em questão representam a ausência (classe ω_0) ou presença (classe ω_1) de foco de calor e serão entendidas como subconjuntos fuzzy. Dessa forma, o risco de incêndio florestal é definido como sendo a pertinência de $x^r \in \Omega$ ao subconjunto fuzzy que define a presença de focos de calor. O diagrama apresentado na Figura 21 (bloco **Modelo de Risco de Incêndio Florestal** na Figura 2) ilustra a metodologia proposta.

De um modo geral, seja $\mathbb{V} = \{x^1, x^2, \dots, x^k\}$ uma k vizinhança do vetor $x^r = (y_1, y_2, y_3, y_4)$, o classificador kNN (*k-nearest neighbours*) identifica a classe de um determinado vetor x^r a partir do voto majoritário dos k vizinhos mais próximos de x^r , usando como base conjunto \mathcal{T} (HAYKIN, 2001).

A versão fuzzy desse classificador atribui a x^r um grau de pertinência a cada

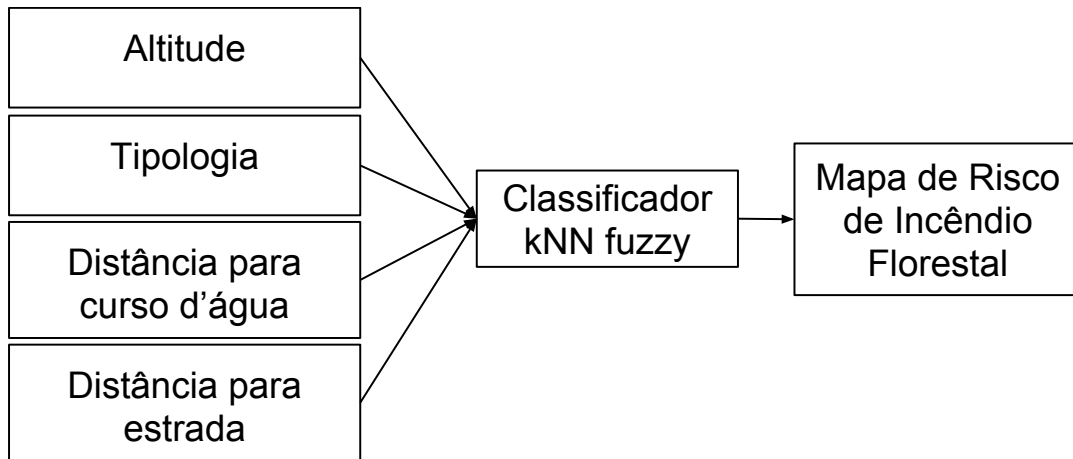


Figura 21 – Diagrama em blocos da arquitetura do modelo de risco de incêndio florestal com o classificador kNN fuzzy.

classe considerada. O grau de pertinência é uma função da distância de x^r aos elementos $x^l \in \mathbb{V}$ com $l = 1, \dots, k$ e das pertinências desses vizinhos nas classes possíveis (KELLER; GRAY; GIVENS JR., 1985). Mais informações sobre o classificador kNN e sua versão fuzzy pode ser encontradas na Subseção 2.6.1.

No problema proposto, consideramos a pertinência de x^r no conjunto dos pontos que define a presença de foco de calor. Então dado um vetor $x^r \in \Omega$, o risco de incêndio florestal $\rho(x^r)$ é definido pela Fórmula 4.1 que foi adaptada de Keller, Gray e Givens Jr. (1985).

$$\rho(x^r) = \frac{\sum_{j=1}^k \omega_1^j (1/\|x^r - x^j\|)^2}{\sum_{j=1}^k (1/\|x^r - x^j\|)^2}, \quad (4.1)$$

onde $\|\cdot\|$ é a norma de Mahalanobis (ver Subseção 2.6.1). Essa métrica foi utilizada para decorrelacionar as variáveis de entrada e, em relação à distância euclidiana, obtivemos uma melhora qualitativa no mapa de risco de incêndio. Note que as pertinências atribuídas a x^r são influenciadas pelo inverso do quadrado das distâncias dos vizinhos mais próximos e por suas respectivas pertinências a classe que define a presença focos de calor.

Por fim, o número de vizinhos k pode ser fornecido *a priori* ou estimado usando alguma técnica estatística. Nesse trabalho, propomos determinar k usando a técnica de validação cruzada com 10 pastas (*10-fold cross validation*) (HAYKIN, 2001). Em palavras, os dados disponíveis são divididos 10 subconjuntos disjuntos. Cada um dos 10 subconjuntos é usado uma vez para validação enquanto que os 9 subconjuntos restantes são usados como conjunto de treinamento parcial. O conjunto de treinamento parcial é

usado para determinar quais pontos foram atingidos por focos de calor e o desempenho dessa classificação é avaliado usando o conjunto de validação. Adota-se o valor de k que produz o menor erro considerando a média aritmética das 10 partições. O mapa de risco de incêndio final ρ é determinado usando todo o conjunto de dados \mathcal{T} .

Para avaliar quantitativamente o mapa de risco de incêndio florestal foi necessário encontrar uma medida de similaridade que tivesse a característica de captar graus de “proximidade” entre o risco de incêndio predito para células do domínio, que é uma variável que assume valores em um conjunto fuzzy, e o estado de presença ou não de foco de calor na célula (variável binária).

Destacamos três medidas de similaridade, todas baseadas em distância, que conseguiram capturar esses aspectos qualitativos da modelagem que são: complemento da distância relativa de Hamming (S_h) – que é um exemplo de medida de similaridade forte – e S_1 e S_2 que são medidas que fazem parte de uma classe de medidas de similaridades estudadas por De Baets, De Meyer e Naessens (2001). Mais informações sobre medida de similaridade e as fórmulas utilizadas podem ser encontradas na Seção 2.1. Os resultados obtidos para os dados relativos ao Estado do Acre são apresentados na Seção 4.1.

O mapa de risco de incêndio construído com a metodologia apresentada até este ponto será essencial para a construção do índice de alerta de incêndio florestal (\mathbb{W}), detalhado na Seção 5.1.

É apresentada a seguir a metodologia descrita acima em um estudo para o Estado do Acre, Brasil, na qual é obtido um mapa de risco de incêndio florestal utilizando os dados geo-referenciados com respeito aos anos de 2003 a 2014.

4.1 Mapa de risco de incêndio florestal para o Estado do Acre-Brasil

Nessa seção utilizamos os dados geo-referenciados do Estado do Acre para gerar o mapa de risco de incêndio florestal a partir da metodologia apresentada no início desse capítulo.

As variáveis utilizadas nesse modelo são caracterizadas na próxima subseção.

4.1.1 Caracterização das Variáveis de Entrada e Saída

Para gerarmos o banco de dados com informações de entrada e saída do modelo de risco de incêndio florestal foi necessário criar uma malha de retângulos (células) que recobre o domínio desejado. Para cada célula associamos o valor que as variáveis de entrada assumem no centro geométrico (centróide) do retângulo. Somente a variável de saída será quantificada a partir da presença ou ausência de foco de calor em cada célula do domínio.

Os dados das variáveis **tipologia florestal**, distância para **curso d'água** e para **estrada** mais próximos foram obtidos a partir dos mapas geo-referenciados do Zoneamento Ecológico Econômico fase 2, fornecidos pela Fundação de Tecnologia do Acre - FUNTAC e utilizando o software Arcgis versão 12.5.

As distâncias para curso d'água e estrada de uma determinada célula foram calculadas do centróide da célula à linha que corresponde ao curso d'água ou estrada mais próximos daquela célula.

Essas duas variáveis foram escolhidas em virtude de terem uma forte relação com a presença humana.

A variável altitude foi obtida de imagens de satélite disponibilizadas em [Miranda \(2014\)](#) e a altitude associada a uma determinada célula é a altitude do centróide da célula que é obtida a partir do cruzamento da imagem que contém informações de altitude com o mapa de centróides de todas as células.

Após esse procedimento, procuramos descobrir quais eram as altitudes dos pontos nos quais ocorreram focos de calor no Estado do Acre no período de 2002 a 2012.

A altitude média dos pontos nos quais ocorreram focos de calor nesse período é 202,65 metros e o desvio padrão desse mesmo conjunto de dados é 20,94 metros.

Esse valor relativamente pequeno de desvio padrão contribuiu para que essa variável fosse escolhida como entrada para o modelo já que, na imagem de satélite, o valor máximo da altitude foi 752 m e mínimo é 108 m.

O mapa de tipologia florestal é constituído de polígonos e cada um desses polígonos está associado a uma tipologia florestal. A tipologia que será associada a uma determinada célula é a tipologia do polígono ao qual o centróide dessa célula pertence.

A Tabela 6 apresenta os atributos da tipologia florestal existentes no banco de dados obtido associados à soma da área total em Hectares (HA) da respectiva tipologia, ao número total de focos de calor que ocorreram em cada tipologia no período de 2002 a 2012 e à densidade de focos de calor por unidade de área.

A variável tipologia florestal foi quantificada a partir da ordem decrescente da densidade de focos de calor por unidade de área identificada na Tabela 6.

A Tabela 7 apresenta o significado das siglas de tipologias florestal.

Utilizamos essa densidade pra caracterizar quais tipologias têm mais influência para a ocorrência do fenômeno. Dessa forma, a quantidade de focos de calor que ocorre em uma tipologia e a área total que essa tipologia possui no domínio irão determinar na quantificação dessa variável.

A variável de saída, ocorrência de foco de calor, foi obtida de mapas geo-

Ordem	Tipologia	Focos de Calor	Área Total (HA)	Focos/Área
1	Áreas Antropizadas	21811	1721156,99	0,0127
2	FAP - Aluvial + Vs	209	19746,96	0,0106
3	FAP - Aluvial	4426	840905,86	0,0053
4	FAP + FD	10044	2312959,56	0,0043
5	FAB - Aluvial	643	160369,93	0,0040
6	FAP	1248	405883,87	0,0031
7	FAB + FD	817	348755,98	0,0023
8	FAB + FAP	9163	5613623,72	0,0016
9	FAP - Aluvial + Pab	49	36891,26	0,0013
10	Campinaranas	6	5244,78	0,0011
11	FAB + FAP + FD	3345	3392057,19	0,0010
12	FABD	1305	1477398,19	0,0009
13	FD	33	44279,31	0,0007
14	FD -submontana	0	74271,93	0,0000
15	FAP +Pab	0	9356,3788	0,0000

Tabela 6 – Ordem decrescente da densidade de focos de calor por área, tipologia florestal, número de focos de calor de 2002 a 2012, área total de cada tipologia e densidade de focos de calor por unidade de área.

Sigla	Tipologia
Vs	Vegetação secundária
Aluvial	Tipologia em área aluvial
FAP	Floresta Aberta com Palmeiras
FAB	Floresta Aberta com Bambu
FD	Floresta Densa
FABD	Floresta Aberta com Bambu Dominante
Pab	Formações Pioneiras

Tabela 7 – Descrição dos códigos do Zoneamento Ecológico Econômico do Estado do Acre.

referenciados disponíveis no banco de queimadas do Instituto de Pesquisas Espaciais - Inpe (INPE, 2011).

Reunimos os mapas de focos de calor de 2003 a 2013 em apenas um que foi associado à malha de células e se houve ocorrência de pelo menos um foco de calor em uma determinada célula a variável de saída é igual a um e, caso contrário, ela assume o valor zero.

Na próxima seção apresentamos o tratamento realizado nos dados para que fosse possível utilizar o classificador kNN fuzzy para obtermos o risco de incêndio florestal.

4.1.2 Tratamento dos dados

Após a geração dos dados a partir dos mapas geo-referenciados percebemos que o número de padrões na classe ω_0 (ausência de focos de calor) é muito maior que

os da classe ω_1 (presença de focos de calor), por isso o conjunto de dados teve que ser balanceado, ou seja, obtivemos um outro conjunto de dados com quantidade igual para as duas classes, esses padrões foram escolhidos de forma totalmente aleatória para evitar possíveis distorções nas informações.

Em virtude da ocorrência de focos de calor ser diferente entre as regiões leste e oeste do Estado do Acre, separamos o conjunto balanceado em dois conjuntos que aqui são denominados: esquerda (oeste do Estado do Acre) e direita (leste do Estado do Acre) e representados usando sub-índices e e d , respectivamente.

Para cada um desses dois conjuntos, as características foram normalizadas e passaram a ter média zero e variância um. Isso foi necessário pelo fato da amplitude numérica de algumas variáveis ser muito maior do que outras.

4.1.3 Conjuntos de treinamento, validação e teste e o valor do parâmetro k

Realizamos, também, a separação do conjunto total de dados da seguinte forma: 85% foram usados como conjunto de treinamento \mathcal{T} , isto é, para sintetizar o mapa de risco de incêndio florestal (ρ). Os 15% restantes formaram o conjunto de teste \mathcal{A} , que foi usado para avaliar a metodologia proposta.

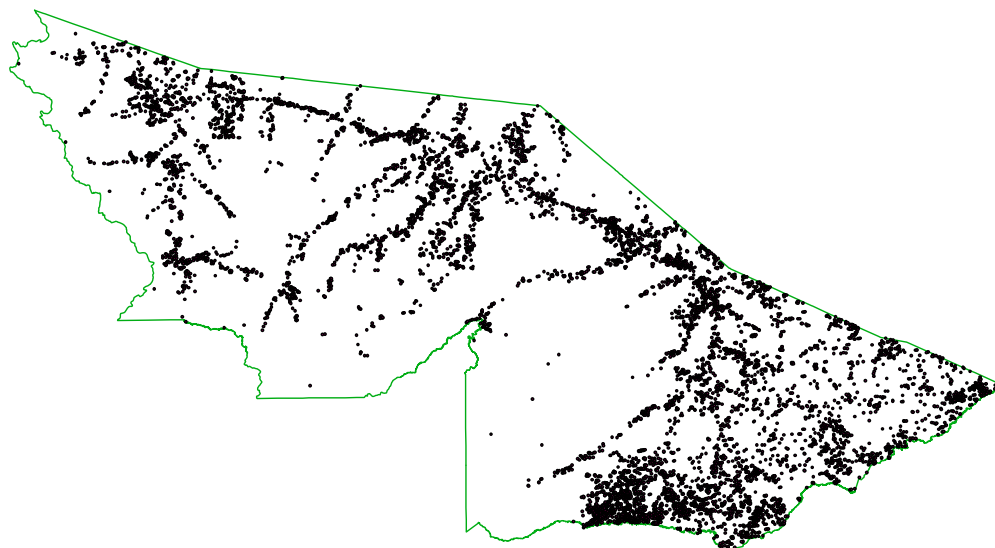
Conforme sugerido anteriormente, o número k de vizinhos mais próximos foi determinado usando a técnica de validação cruzada com dez pastas sobre o conjunto \mathcal{T} . Especificamente, para cada valor de $k \in \{1, 3, 5, \dots, 47, 49\}$, calculou-se a média do *Hamming Loss* entre os 10 conjuntos de validação. O menor *Hamming Loss* médio para o conjunto da direita foi 16.67%, obtido tomando $k_d = 29$ de vizinhos mais próximos. Para o conjunto da esquerda, o número $k_e = 21$ de vizinhos mais próximos forneceu o menor *Hamming Loss* médio de 12.32%.

Finalmente, os valores de $k_e = 21$ e $k_d = 29$, combinados com os conjuntos de treinamento \mathcal{T}_e e \mathcal{T}_d , foram utilizados para determinar os mapas de risco ρ_e e ρ_d para as regiões oeste e leste do Estado do Acre. O *Hamming Loss* calculado sobre todos os dados de teste ($\mathcal{A}_e \cup \mathcal{A}_d$) foi 16.41%. Como a taxa de acerto (acurácia) do kNN fuzzy foi de 83.59%, pode-se prever, para os dados disponíveis, de forma satisfatória, a ocorrência ou não de um foco de calor para o período considerando as informações espaciais de um ponto qualquer do domínio.

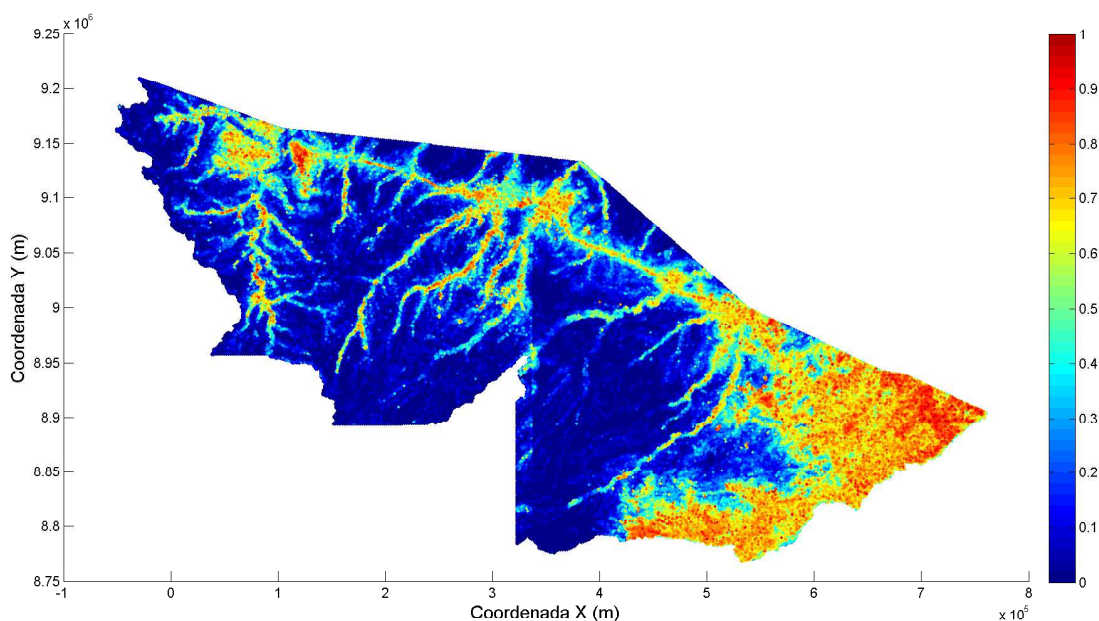
4.1.4 Resultados

A Figura 22 ilustra a predição do risco de incêndio para todas as células do domínio considerado e o mapa com todos os focos de calor ocorridos no ano de 2014. Note ainda que o padrão de risco de incêndio florestal acompanha o padrão de ocorrência de

focos de calor, tanto nas áreas mais urbanas do estado (lado leste do estado e região ao redor de Cruzeiro do Sul) quanto na área rural (centro-oeste).



a) Mapa de ocorrência de focos de calor do Estado do Acre em 2014 (INPE, 2011).



b) Mapa do risco de incêndio florestal do Estado do Acre em 2014.

Figura 22 – Mapas de ocorrência de focos de calor e risco de incêndio florestal do Estado do Acre em 2014.

Para avaliar quantitativamente essa “proximidade” entre os dois mapas utilizamos três medidas de similaridades descritas no início deste capítulo. Essas medidas de similaridade comparam o vetor que contém a informação binária de ocorrência de focos de calor do ano de 2014 para todas células do Estado do Acre (fenômeno real) e o vetor que possui estimativa do risco de incêndio gerado pelo classificador kNN fuzzy. Obtivemos os

valores de $S_h = 0.8291$, $S_1 = 0.8452$ e $S_2 = 0.8427$ o que caracteriza uma alta similaridade entre os vetores.

Esse mapa de risco de incêndio florestal obtido será essencial para a construção do índice de alerta de incêndio florestal (\mathbb{W}), detalhado na Seção 5, de forma conjunta com os modelos de perigo de incêndio florestal descritos no Capítulo 3.

Nesse Capítulo 4 abordamos o modelo matemático para obtermos o risco de incêndio florestal. Esse índice incorpora a altitude, a tipologia florestal, as distâncias para curso d'água e estrada mais próximos como variáveis de entrada e a saída é o mapa de incêndio florestal. Utilizamos o classificador kNN fuzzy para obter o risco de incêndio $\rho(x_i, y_j)$ para cada ponto $(x_i, y_j) \in \Omega$ como a pertinência ao subconjunto fuzzy associado à classe ω_1 que representa a ocorrência de focos de calor na célula (x_i, y_j) . Na Seção 4.1 aplicamos o modelo de risco de incêndio a dados relativos ao Estado do Acre-Brasil e avaliamos esse modelo utilizando três medidas de similaridade.

A seguir, no Capítulo 5, definimos o alerta de incêndio florestal (\mathbb{W}) que tem como objetivo agregar as características de todos os índices discutidos até agora. Esse alerta será utilizado para simular a propagação de incêndio florestal com a metodologia de autômatos celulares.

5 Índice de Alerta de Incêndio Florestal e Simulações de Propagação com Autômatos Celulares

5.1 Evolução Temporal do Alerta de Incêndio Florestal

O risco de incêndio florestal obtido no Capítulo 4 é estático e depende apenas de entradas que, ou não variam com o tempo, ou que podem variar muito pouco ao longo de vários anos. O perigo de incêndio florestal (Seção 3.2) depende do clima local que possui uma rápida variação diária.

Para que um índice seja útil como alerta antecipado de ocorrência de incêndio florestal, é importante que ele seja uma função do tempo e das coordenadas espaciais.

A ideia central para criar um índice que satisfaça a essa condição foi combinar os índices de risco e perigo de incêndio florestal (definições desses índices são descritas nas Seções 1.1 e 1.2) de forma apropriada para que o novo índice obtido – aqui chamado de **Alerta de Incêndio Florestal** (\mathbb{W}) – possa agregar as variáveis físicas e climáticas do local considerado e, assim, gerar um mapa de alerta dinâmico que evolua no tempo e no espaço. A arquitetura desse modelo está detalhada no bloco **Mapa de Alerta de Incêndio Florestal** na Figura 2.

Para combinar o risco e o perigo de incêndio florestal e obter o alerta de incêndio florestal \mathbb{W} utilizamos a t-norma de Lukasiewski (t_L) e, então $\mathbb{W}(x, t)$ no ponto $x \in \Omega$ e no tempo $t \in [0, T]$ é dado pela Equação 5.1.

$$\mathbb{W}(x, t) = \max \{0, \rho(x) + \pi_t - 1\} = t_L(\rho(x), \pi_t). \quad (5.1)$$

Essa t-norma foi escolhida, em particular, para que os pontos do domínio Ω associados a risco de incêndio florestal nulo permaneçam com alerta de incêndio florestal nulo durante a simulação temporal. A definição de t-norma é realizada na Subseção 2.1. Além disso, dados dois pontos $x, y \in \Omega$, se $\rho(x) > \rho(y)$ então para todo $t \in [0, T]$ teremos $\mathbb{W}(x, t) > \mathbb{W}(y, t)$, logo, áreas que possuem maior risco de incêndio florestal, permanecerão com valor de alerta de incêndio florestal maior durante as simulações.

Outra característica interessante dessa definição é que o alerta de incêndio florestal atinge o seu maior valor em regiões que possuem risco de incêndio ρ máximo e no instante $t_m \in [0, T]$ tal que o perigo de incêndio π_{t_m} também é máximo. Esse fato

é importante para a modelagem do fenômeno, pois os incêndios florestais na Amazônia somente acontecem, com raras exceções, no período mais seco do ano e as áreas que são mais atingidas são aquelas que possuem características específicas e um histórico de ocorrências desse fenômeno (NEPSTAD; MOREIRA; ALENCAR, 1999).

Na próxima subseção será apresentado um estudo evolutivo do alerta de incêndio florestal para o Estado do Acre-Brasil.

5.1.1 Evolução temporal do alerta de incêndio: estudo para o Estado do Acre

Utilizaremos nesta subseção o mapa de risco de incêndio florestal e a série temporal de valores de perigo de incêndio florestal preditos nos Capítulos (3) e (4) para o Estado do Acre-Brazil no ano de 2014. Com essas informações e utilizando a t-norma de Lukasiewski obtivemos uma sequência de mapas de alerta de incêndio florestal que, ao serem reunidos, geram um vídeo da evolução desse alerta durante esse ano.

O diagrama em blocos da Figura 23 apresenta a arquitetura do modelo de alerta de incêndio florestal. Essa arquitetura é a parte da Figura 2 que contempla apenas os blocos de risco, perigo e alerta de incêndio florestal.

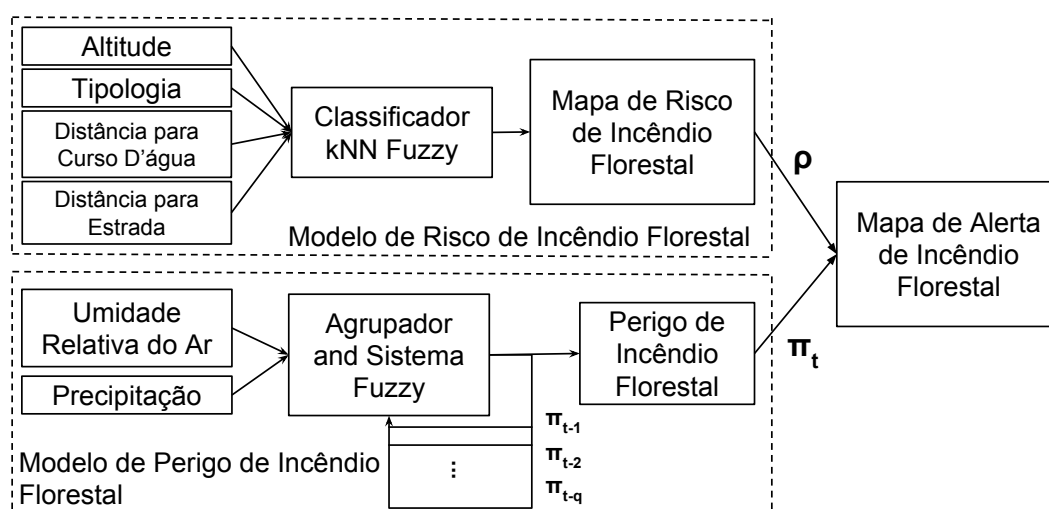


Figura 23 – Diagrama em bloco da estrutura do modelo de alerta de incêndio florestal obtido por um classificador kNN fuzzy, agrupador e sistema fuzzy: ρ é o risco de incêndio florestal, π_t é o perigo de incêndio florestal no tempo t e q é o número de atrasos de tempo do perigo de incêndio.

5.1.1.1 Resultados

A Figura 24 mostra seis cenários correspondentes a seis dias da evolução desse índice.

Uma observação importante sobre a evolução do alerta de incêndio florestal é que ele é zero em praticamente todo período do ano e somente começa a aumentar

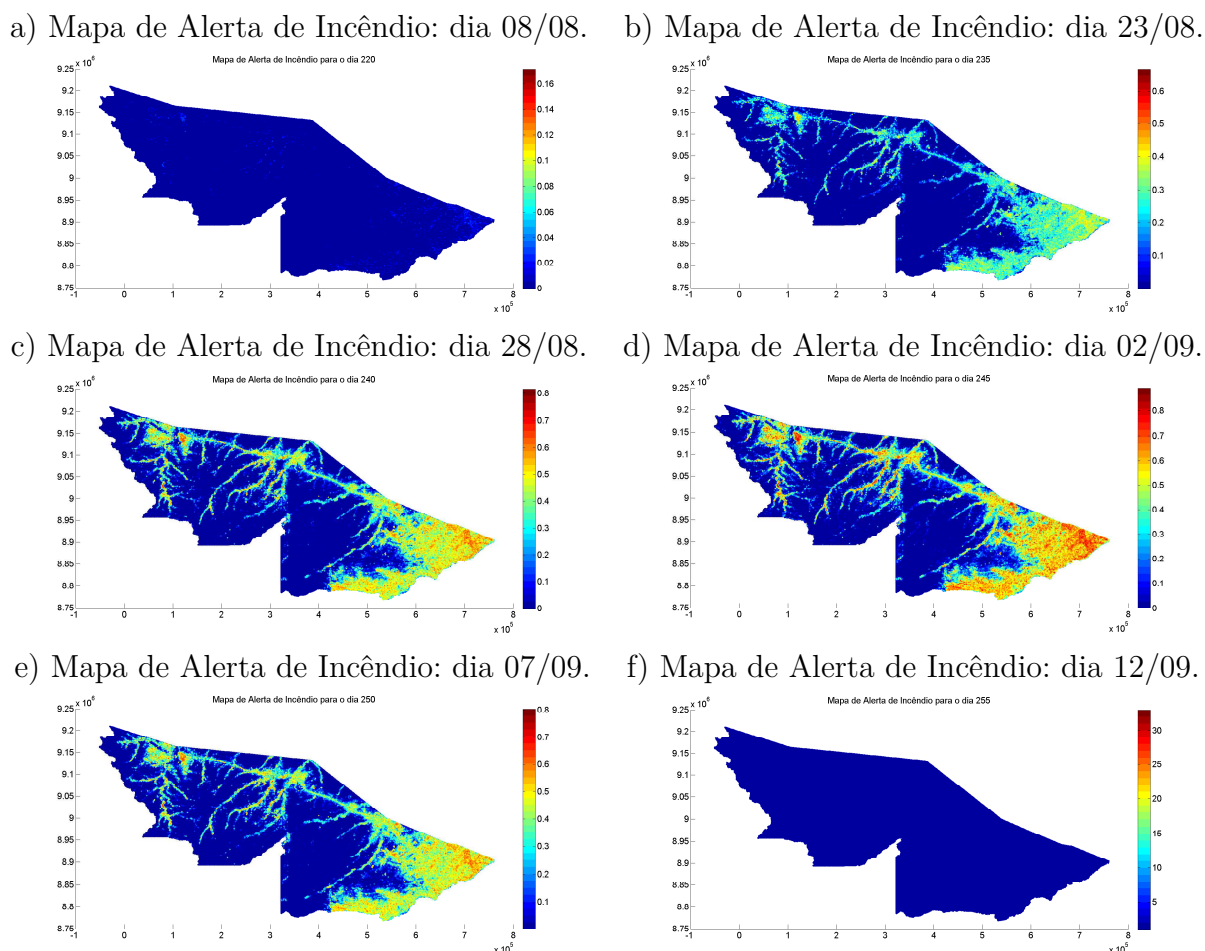


Figura 24 – Mapa de Alerta de Incêndio Florestal para seis dias do ano de 2014.

no período mais seco do ano que, para a região amazônica, é entre os meses de junho a novembro (NEPSTAD; MOREIRA; ALENCAR, 1999). Esse comportamento concorda com a real de ocorrência de focos de calor para a região estudada.

O índice de alerta de incêndio florestal herda do risco e do perigo de incêndio florestal todas as características importantes na modelagem do fenômeno: assume valores maiores em locais mais propensos a incêndio (maiores valores de risco de incêndio) e somente no período do ano no qual há ocorrência de focos de calor (maiores valores de perigo de incêndio). Esses fatos decorrem diretamente da definição do alerta a partir da t-norma entre o risco e o perigo de incêndio conforme detalhado na seção anterior.

Na próxima seção abordamos a definição de modelos matemáticos para incêndio florestal e informações sobre modelos que utilizam a metodologia dos autômatos celulares.

5.2 Modelos de Propagação de Incêndio Florestal

Os modelos matemáticos de incêndios florestais são geralmente compostos por um conjunto de equações cuja solução fornece valores numéricos para a evolução

espaço-temporal de uma ou mais variáveis (PASTOR *et al.*, 2003).

Os métodos de construção dos simuladores de incêndios florestais podem ser divididos em duas categorias, a que utiliza um sistema de grade regular cujas técnicas mais utilizadas (ALBRIGNT; MEISNER, 1999) são a percolação por ligação e o autômato celular e a que é baseada no plano contínuo possui como técnica mais usada a propagação de onda elíptica (PASTOR *et al.*, 2003).

Na técnica de simulação de incêndio florestal de percolação por ligação, a paisagem é representada por uma malha de divisões quadradas, triangulares ou hexagonais e a cada célula serão associadas as características ambientais correspondentes. Existe uma probabilidade específica (ajustada por um modelo matemático baseado em dados históricos de incêndio) de ignição e propagação associada a cada célula que rege o espalhamento do fogo de uma célula para suas vizinhas (TAVARES *et al.*, 2015; PASTOR *et al.*, 2003).

A técnica de simulação de automato celular é caracterizada por um conjunto de regras (baseadas em modelos matemáticos e semi empíricos de comportamento do fogo) que determina o estado da célula e a propagação do fogo para a vizinhança (DA SILVA; BARROS; MEYER, 2015; LOUZADA; FERREIRA JR., 2008).

O que caracteriza a propagação de ondas elípticas é a modelagem de uma frente de fogo que se desloca em uma meio contínuo e desenha um perímetro que é composto de um número finito de segmentos. Esse modelo é baseado no princípio de propagação das ondas de Huygens, cada vértice é considerado como ponto de ignição de um pequeno incêndio, que avança em forma elíptica em condições ambientais homogêneas seguindo algum critério de propagação que é definido por um modelo matemático empírico, semi-empírico ou teórico (ANDERSON *et al.*, 1982).

Na próxima seção, apresentamos a proposta inicial de um simulador de propagação baseado em autômatos celulares, porém os valores armazenados nas células dos autômatos são referentes ao risco de incêndio estimado no Capítulo 4 e as regras de propagação não são definidas por modelos semi empíricos, mas sim, definidas a partir do alerta de incêndio florestal e do perigo de incêndio, que determina a dinâmica temporal do autômato.

5.3 Autômatos Celulares

Os Autômatos Celulares (AC) são representações matemáticas para sistemas físicos com espaço, tempo e quantidades físicas discretizados (WOLFRAM, 1994).

Um autômato celular consiste em uma malha (ou “vetor”) regular e uniforme com uma variável discreta em cada sítio (“célula”). O estado de um autômato celular é completamente especificado pelos valores das variáveis em cada célula. A vizinhança

de uma célula com coordenadas (x_i, y_i) é definida como todos os sítios imediatamente adjacentes e, às vezes inclui ela própria. A evolução de um autômato é realizada em passos de tempo discretos e o valor de uma variável em todas as células é atualizado simultaneamente com base nos valores das variáveis em células na sua vizinhança no passo de tempo anterior dependendo de um conjunto de regras locais definidas (WOLFRAM, 1994).

Neste modelo de AC as coordenadas (x_i, y_j) das células são associadas somente a localizações geográficas, pois as variáveis físicas e climáticas (as mesmas utilizadas nesta tese) que interferem no funcionamento do AC estão incorporadas ao risco, perigo e alerta de incêndios.

O AC é composto por um conjunto de células com determinados valores, que interagem entre si em função de uma coleção finita de condições pré-definidas. Os estados (valores) das células são alterados conforme um conjunto de regras de transição, que depende da vizinhança (às vezes da própria célula também), ou seja, das células em torno da célula que será atualizada. Assim, o AC é composto de três partes: uma estrutura (“lattice” (tipo da rede de contato), ou seja, a geometria da célula (formato)), uma vizinhança e uma regra de transição local. De forma geral, a regra de transição de estados é imposta de forma paralela e sincronizada a todas as células. Uma configuração inicial de autômato, aparentemente simples, pode produzir resultados em que a conjuntura da matemática dos estados apresentará um alto nível de complexidade (WOLFRAM, 1994).

A arquitetura do modelo de autômatos celulares proposto nesta tese pode ser visualizada na Figura 25. Essa figura sintetiza a arquitetura proposta na Figura 2.

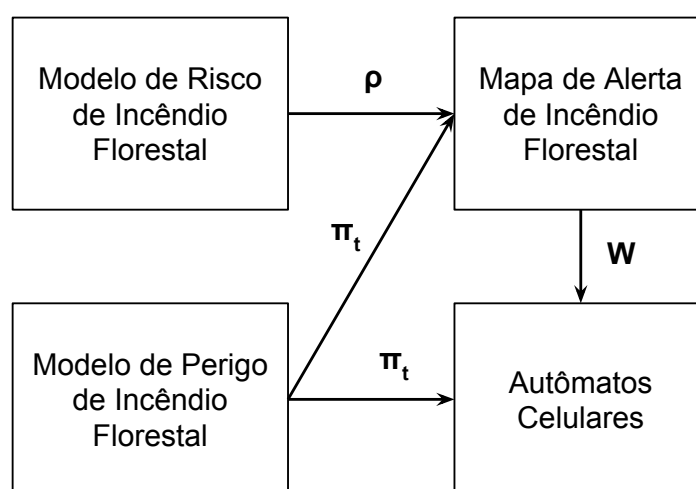


Figura 25 – Diagrama da arquitetura do modelo de autômatos celulares.

Pode-se representar a composição de um AC como sendo $U = (\mathbb{L}, \mathbb{Q}, \mathbb{R}, f)$, onde \mathbb{L} é a malha de células, \mathbb{Q} o conjunto de estados que cada célula assume, \mathbb{R} a vizinhança de uma determinada célula e f a regra ou função de transição local (ADAMATZKY, 1994).

A malha do AC utilizado nesta tese é bidimensional e retangular e cada célula $x \in \mathbb{L}$ será identificada por coordenadas (x_i, y_i) no plano xy .

Os valores que as células assumem são conhecidos como estados e o conjunto de estados que uma célula pode assumir é representado por \mathbb{Q} . Para o AC que irá simular a propagação de incêndio consideramos $\mathbb{Q} = \{0, 1, 2, 3\}$ com significados descritos na Tabela 8.

Tabela 8 – Estados das células do AC

0	Vazia
1	Vegetação
2	Queimando
3	Queimada

Para cada célula do autômato foi considerada vizinhança de Moore, ou seja, cada célula identificada por (x_i, y_j) possui oito vizinhos (x_{i+1}, y_{j+1}) , (x_{i+1}, y_{j-1}) , (x_{i-1}, y_{j+1}) , (x_{i-1}, y_{j-1}) , (x_{i+1}, y_j) , (x_{i-1}, y_j) , (x_i, y_{j-1}) , (x_i, y_{j+1}) .

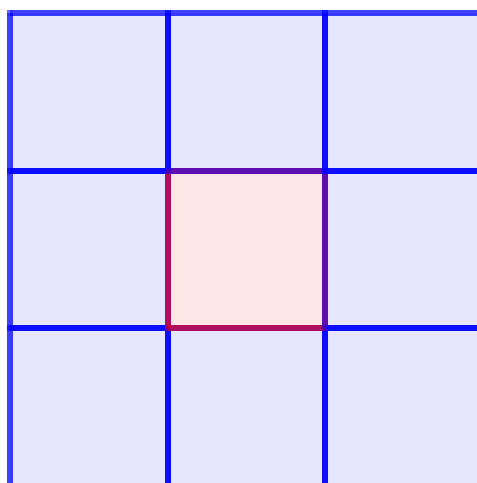


Figura 26 – Vizinhança de Moore

O estado da célula $z = (x_i, y_j)$ no instante t é indicado por z_t e o estado da vizinhança por $E(z_t) = (w_t^1, \dots, w_t^8)$, pois na vizinhança de Moore cada célula possui oito vizinhos. A célula z_t tem o seu estado alterado no próximo instante dependendo do estado das células vizinhas: $z_{t+1} = f(E(z_t)) = f(w_t^1, \dots, w_t^8)$ (ADAMATZKY, 1994).

5.3.1 Regras de Mudança de Estado

Para definir as regras de mudança de estado das células do autômato celular utilizamos o alerta de incêndio florestal $\mathbb{W}(x_i, y_j, n)$, definido na Seção 5.1, que pode ser calculado para todas as células com coordenadas $(x_i, y_j) \in \Omega$ e nos instantes $n \in \{1, \dots, T\}$.

A mudança do estado “vegetação” para o estado “queimando” de uma determinada célula (x_i, y_j) do autômato somente ocorre quando o perigo de incêndio florestal for positivo, ou seja, $\pi_t > 0$.

Além da condição $\pi_t > 0$, se uma célula (x_i, y_j) encontra-se no estado queimando, o “fogo” será propagado para outra célula $(u_i, v_j) \in E(x_i, y_j)$ pertencente à vizinhança da primeira quando o alerta de incêndio florestal $\mathbb{W}(u_i, v_j, n)$ for maior que c_1 . Isto caracteriza a **condição de propagação**:

$$\mathbb{W}(u_i, v_j, n) > c_1, \quad (5.2)$$

onde c_1 é uma constante que está associada à dificuldade de propagação do fogo em virtude da combinação do risco e perigo de incêndio da célula na vizinhança.

Além da **condição de propagação**, apenas um percentual de células vizinhas serão “incendiadas”, de forma aleatória, a partir da escolha de um parâmetro limiar. A escolha desse percentual é feita com um parâmetro aleatório $p((x_i, y_j))$ para cada célula, obtido por uma distribuição uniforme $U[0, 1]$, e por um parâmetro c_2 , que neste trabalho será denominado de **limiar de alcance**. Dessa forma, as células vizinhas incendiadas devem cumprir a condição $p((x_i, y_j)) < c_2$. A cada iteração, quando se tem $\pi_t > 0$, o algoritmo realiza também a “queima” aleatória de um percentual das células que satisfazem a condição 5.2.

Dadas as condições do problema, as informações disponíveis e as características matemáticas da modelagem optamos por utilizar a distribuição uniforme $[0, 1]$.

O fenômeno dos incêndios ocorrem, com algumas exceções, em um período do ano específico (período menos chuvoso), por isso existe a necessidade de realizar a mudança de estado de um percentual dos nós no estado “vegetação” para o estado “queimando” de forma aleatória quando o perigo de incêndio florestal assume um valor diferente de zero, sinalizando que o clima encontra-se propício à ocorrência de incêndios.

Para cada célula (x_i, y_j) , a mudança do estado “queimando” para o estado “queimado” é feita quando $p((x_i, y_j)) < B(\pi(n))$ onde $p((x_i, y_j))$ é um parâmetro aleatório com distribuição uniforme $U[0, 1]$ associado a célula (x_i, y_j) e $B(\pi(n))$ é um parâmetro chamado **Limiar de Extinção** que depende do perigo de incêndio no instante n . Esse parâmetro B , definido como uma função de $\pi(n)$, é dado pela Fórmula 5.3.

$$B(n+1) = \begin{cases} B(n) - c_3 * \pi(n) & \text{se } \pi(n) > 0 \\ B(n) + c_3 & \text{se } \pi(n) = 0 \end{cases} \quad (5.3)$$

onde c_3 é um parâmetro que define uma taxa de redução no parâmetro B , de uma iteração para a seguinte, em virtude de π ser positivo. Se obtivermos $B(n^*) < 0$ em algum instante

n^* então adota-se $B(n^*) = 0$.

Quanto maior o valor de B , mais rapidamente as células que estão no estado “queimando” mudam para “queimado”, isso representa uma extinção mais rápida do incêndio o que justifica a dependência que B possui em relação ao perigo de incêndio. Um valor positivo desse índice indica que as condições climáticas estão propícias ao início de novos incêndios e que eles estão ocorrendo e tendem a se prolongar, se não houver algum controle externo.

Agora a medida que as chuvas começam a ocorrer, o perigo de incêndio começar diminuir e existe a tendência de focos de incêndio se apagarem, ou seja, B deve parar de aumentar o seu valor, propiciando assim que o “fogo” das células comece a se apagar mais rapidamente.

O fenômeno dos incêndios ocorrem, com algumas exceções, em um período do ano específico (período menos chuvoso), por isso existe a necessidade de realizar a mudança de estado de um percentual dos nós no estado “vegetação” para o estado “queimando” de forma aleatória quando o perigo de incêndio florestal ultrapassa determinado valor *threshold* que será denominado c_4 , como relatado acima, sinalizando que o clima encontra-se propício à ocorrência de incêndios.

O percentual de células que são queimadas aleatoriamente é definido por $c_4 * \pi(n)$. Esse limiar de queima limita a quantidade de nós que serão forçados a mudar de estado, pois $\pi(n)$ atinge o valor 1 em determinada iteração e a quantidade de nós que seriam “queimados” iriam superar muito a quantidade daqueles atingidos pela “difusão” do processo.

A Figura 27 apresenta o diagrama em blocos com as condições para uma determinada célula mudar do estado vegetação para queimando e depois para queimado.

Os parâmetros c_1 , c_2 , c_3 e $c_4 \in [0, 1]$ foram determinados pelas simulações computacionais com os dados relativos ao Estado do Acre na Subseção 5.3.2.

5.3.2 Estudo de Caso - 3: Estado do Acre

As simulações de incêndio florestal para o Estado do Acre foram realizadas com uma malha de 226.267 pontos distribuídos na região em estudo e com dados climáticos e de focos de calor registrados no ano de 2014. Para cada célula do autômato foi realizada a predição do risco incêndio florestal com a metodologia indicada no Capítulo 4.

Os dados climáticos utilizados foram umidade e precipitação do ano de 2014 para a cidade de Rio Branco-AC (foram considerados os mesmos para o Estado do Acre inteiro) e, a partir delas, calculou-se o perigo de incêndio florestal pela metodologia aplicada no Capítulo 3. Cada passo de tempo do autômato corresponde a um dia (24h).

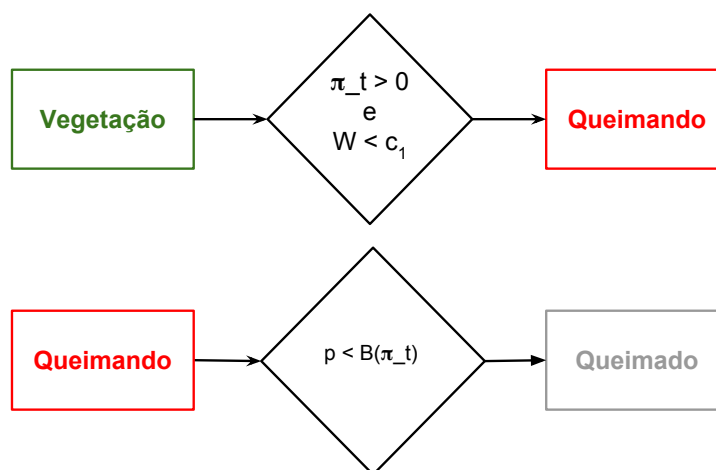


Figura 27 – Diagrama em blocos que representa as condições para as mudanças de estado dos autômatos. W é o alerta de incêndio florestal da célula, π_t é o perigo de incêndio no instante t , p é o número aleatório com distribuição uniforme $[0, 1]$, c_1 é a condição de propagação e B é o limiar de extinção.

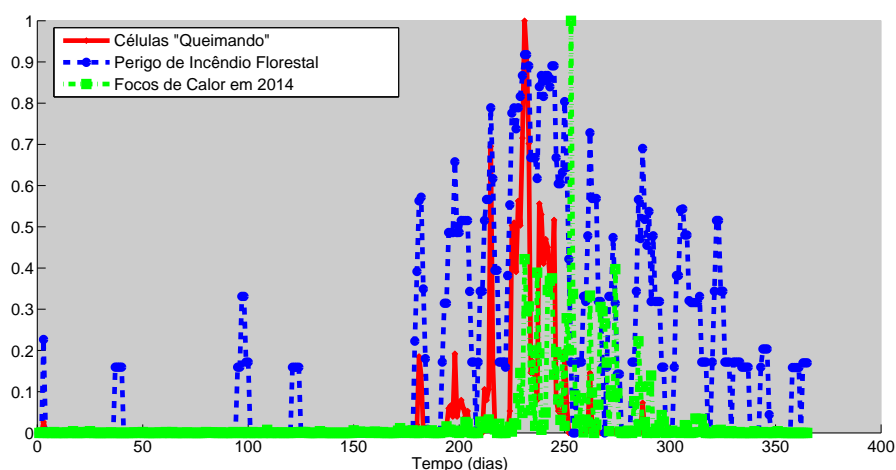
5.3.2.1 Resultados

Para determinar os parâmetros c_1 , c_2 e $c_3 \in [0, 1]$ citados na Seção 5.3, considerou-se o conhecimento prévio das células em que ocorreram os focos de calor no Estado do Acre durante o ano de 2013. A partir disso, gerou-se simulações com autômato celular para valores de $c_1, c_3 \in [0, 1]$ e $c_2 \in [0, 0.1]$ de forma exaustiva, considerando dois parâmetros constantes e variando o terceiro. Os parâmetros que foram escolhidos são aqueles que maximizam o percentual de focos de calor “atingidos” pelas células no estado queimando sem entretanto “queimarem” todas as células da região oeste do acre que possui risco de incêndio florestal muito maior que as demais.

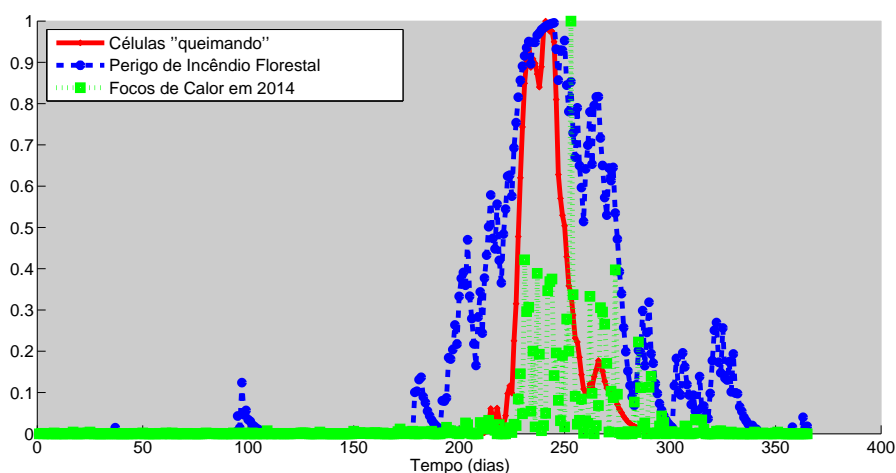
Para esse trabalho, encontramos $c_1 = 0.2$ (**condição de propagação** (5.2)), $c_2 = 0.05$ (**limiar de alcance**) e $c_3 = 0.1$ (taxa de decrescimento do **limiar de extinção** B), obtivemos também $c_4 = 0.4$ a partir de simulações realizadas com os dados climáticos e de focos de calor relativos ao ano de 2013, haja vista que o conjunto que deseja-se realizar as simulações de propagação e estimativa dos índices é relativo ao ano de 2014.

Com o risco e o perigo de incêndio florestal, o autômato celular consegue simular a propagação de incêndios com a aplicação das regras de transição de estados vistas na Subseção 5.3.1. Dessa forma, durante a simulação, somente células com risco de incêndio florestal positivo podem mudar para o estado queimando e só o fazem no período de tempo (perigo de incêndio acima de um determinado limiar) que coincide com a ocorrência de incêndios no Estado do Acre. Os resultados das simulações para os dois modelos de perigo de incêndio discutidos no Capítulo 3 podem ser verificados na Figura 28.

Os gráficos da Figura 29 apresentam os estados do autômato celular para



a) Percentual de células “queimando” do autômato celular com perigo de incêndio dado por um sistema de aprendizado de máquina discutido na Seção 3.2.



b) Percentual de células “queimando” do autômato celular com perigo de incêndio dado por um sistema dinâmico fuzzy discutido na Seção 3.3.

Figura 28 – Séries temporais de focos de calor de 2014, perigo de incêndio florestal (dois modelos distintos) e percentual de nós “queimando” do autômato celular

diferentes instantes de tempo (em dias). As células na cor branca são consideradas no estado 0 (zero) – vazia – e representam as áreas externas ao Estado do Acre e os principais rios do Estado (estamos considerando que o fogo não atravessa os principais rios). As células em verde estão no estado 1 (um) – vegetação – e são as células passíveis de mudar para os estados 2 (dois) – queimando – na cor vermelha e 3 (três) – queimado – na cor cinza.

A simulações mostradas nos gráficos da Figura 29 foram realizadas com sistema dinâmico fuzzy gerando o perigo de incêndio florestal. Ao simular substituindo esse sistema pelo sistema de aprendizado de máquina, não houve alteração no total de células que mudaram para o estado “queimado”.

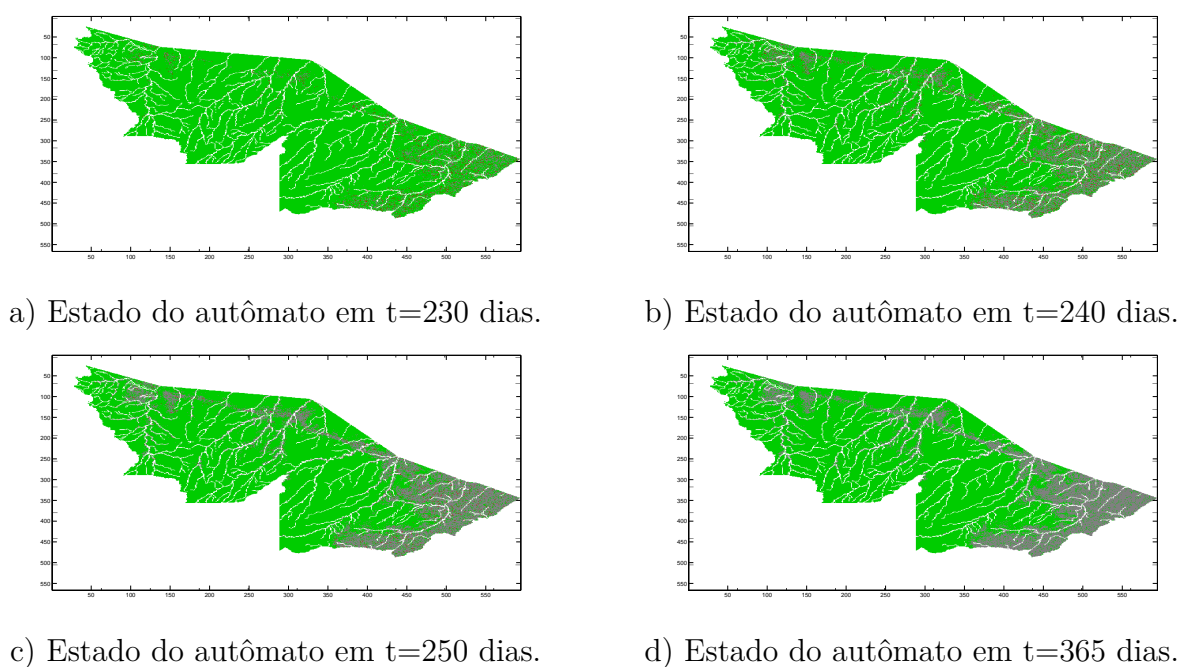


Figura 29 – Estados do autômato celular nas iterações 230, 240, 250 e 365 do ano de 2014.

Esse fato pode ser explicado pelo fato de que existem restrições para mudança de estado que dependem do risco de incêndio florestal que se mantém igual, além disso, os parâmetros do autômato e a semente de geração de números aleatórios são mantidas iguais. O que diferencia ambos os modelos é o instante no qual cada célula muda de estado, que é diferente em cada simulação. Esse fator é inerente a dinâmica temporal do alerta de incêndio florestal e não pode ser visualizado a partir do número total de células atingidas.

Ao verificar a ocorrência espacial dos focos de calor em 2014 e o resultado final da simulação do autômato celular, obteve-se que 62.86% dos focos de calor que ocorreram em 2014 pertencem a células que foram atingidas pelo “incêndio”. A Figura 30 mostra os resultados das simulações. As células em vermelho representam o estado final das células queimadas durante a simulação do AC considerando os 365 dias de 2014. Já os pontos em azul representam o total de focos de calor ocorridos nesse mesmo ano. Pode-se verificar que existem muito mais células vermelhas do que azuis, isso acontece no fenômeno real, pois conforme [Pantoja e Brown \(2009\)](#) quando a floresta é apenas impactada por incêndio existe uma dificuldade maior de detecção por satélite do que quando as copas das árvores é afetada.

Nesse Capítulo 5 abordamos o modelo matemático para obtermos o alerta de incêndio florestal. Esse índice foi criado nesse trabalho para ser possível o acoplamento entre o risco e o perigo de incêndio florestal em um único índice que possui variação espaço-temporal e que consegue agregar as características importantes dos índices de entrada. Na Seção 5.1.1 mostramos a Figura 24 que apresenta seis cenários do vídeo criado a partir da evolução do alerta de incêndio no tempo e no espaço. Apresentamos na Seção

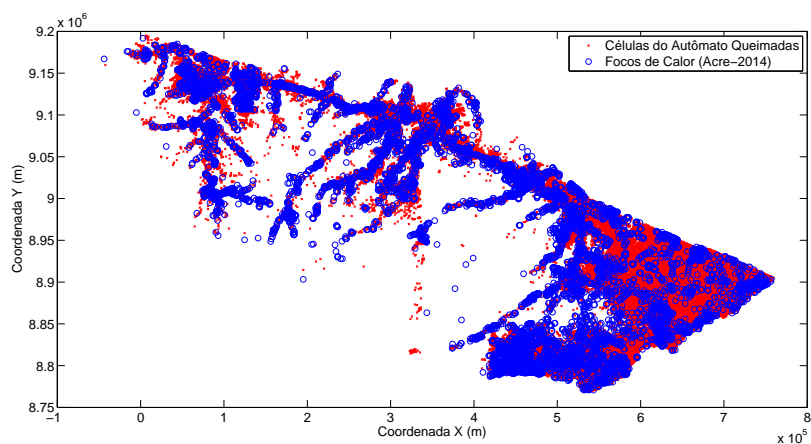


Figura 30 – Focos de calor ocorridos em 2014 e as células “queimadas” do autômato celular.

5.3 um modelo inicial de propagação de incêndio florestal baseado na metodologia de autômatos celulares. O diferencial do modelo matemático apresentado é que ele utiliza o perigo, o risco e o alerta de incêndio florestal para simular a dinâmica espaço-temporal do incêndio.

6 Considerações Finais

O trabalho desenvolvido modela três índices relacionados ao fenômeno de incêndio florestal: o perigo, o risco e o alerta de incêndio florestal todos por diferentes ferramentas matemáticas. Utilizamos esses índices para realizar propagação de incêndio florestal através da metodologia de Autômatos Celulares.

O perigo de incêndio florestal foi obtido a partir da combinação de subconjuntos fuzzy com técnicas de aprendizado supervisionado e não-supervisionado para associar fatores climáticos à ocorrência de focos de calor (Capítulo 3). Já o risco de incêndio florestal foi obtido a partir de um sistema de aprendizado computacional – kNN fuzzy – que associou características físicas do meio estudado (distâncias para estrada e curso d’água mais próximos, altitude e tipologia florestal) ao fenômeno da ocorrência de focos de calor baseado no histórico e na localização espacial da ocorrências deles (Capítulo 4).

Embora essas técnicas sejam conhecidas, o modo e o objetivo de seu uso se constituem em fator original nesse trabalho.

A estimativa do perigo de incêndio florestal para o Estado do Acre no ano de 2014, ao utilizarmos o modelo de aprendizado supervisionado e não supervisionado, foi avaliada a partir de uma classificação binária, na qual se obteve acurácia de 0.6712 e o valor da área abaixo da curva ROC foi 0.8562 (Seção 3.2.1). Esse resultado é de significativa importância na validação do modelo desenvolvido, visto que indica uma adequada assunção sobre o efetivo perigo do estado como um todo.

Ao analisar qualitativamente a série temporal de perigos fuzzy para o ano de 2014 – Figura 19 – verifica-se que esse índice permanece acima do percentual de focos de calor, exceto para o dia em que houve um máximo de focos de calor (outlier).

Na caracterização do **Fator de Perigo** f (Seção 3.2.1), verificou-se que os subconjuntos que estão associados a valores mais altos de fator de perigo tinham centros com valores de umidade e precipitação mais baixos (longos períodos de seca), o que concorda com a literatura da área de engenharia florestal, pois são nessas condições climáticas que ocorre o fenômeno do incêndio florestal (NEPSTAD; MOREIRA; ALENCAR, 1999).

O risco de incêndio florestal do padrão de teste x_t foi definido como a pertinência desse padrão na classe que representa a presença de foco de calor obtida pelo classificador do tipo kNN fuzzy. Ao aplicar o modelo para dados do Estado do Acre, o erro desse classificador junto ao conjunto de teste foi 0.1641 e utilizando três medidas de similaridade entre o mapa de risco de incêndio florestal predito e o conjunto de focos de calor para o ano de 2014 obteve-se $S_h = 0.8291$, $S_1 = 0.8452$ e $S_2 = 0.8427$ (Seção 4.1), o que confirma uma

alta similaridade entre o fenômeno ocorrido e o mapa risco de incêndio florestal predito, importante na verificação da utilidade do que, aqui, se propõe.

O risco e perigo de incêndio florestal foram agregados através da t-norma de Lukasiewski para formar um novo índice – o alerta de incêndio florestal – que consegue expressar as características importantes dos índices de entrada para a modelagem do fenômeno. Sendo gerado por um modelo espaço-temporal (acopla o perigo (variação temporal) e o risco (variação espacial)), o alerta de incêndio é bem mais adequado para promover o alerta antecipado de incêndio florestal e, assim, facilitar a tomada de decisão pelo poder público. Este índice é uma contribuição relevante deste trabalho para as pesquisas (Seção 5.1).

Assim, esses índices podem ser utilizados para quantificar a influência das variáveis umidade, precipitação, tipologia florestal, altitude e distâncias para estrada e curso d'água mais próximos sobre a ocorrências de focos de calor e também como ferramentas de alerta antecipado para incêndio florestal, um instrumento fundamental na adoção de políticas públicas de prevenção.

Na Seção 5.3, apresentamos simulações de propagação de incêndios florestais através da metodologia de autômatos celulares (AC). O AC apresentado nesta tese possui dinâmica temporal regulada pelo perigo de incêndio florestal (Capítulo 3) que sincroniza a mudança de estado das células do AC de acordo com a ocorrência do fenômeno real (Seção 5.3.2). Além disso, as áreas atingidas pelo “fogo” possuem risco de incêndio positivo e o incêndio se propaga das áreas de maior risco para as de menor (Seção 5.3.2). A modelagem apresentada nesta tese é uma primeira tentativa em virtude da ausência de variáveis fundamentais para a modelagem desse tipo de fenômeno tal como o vento, declividade do solo, entre outras (MCARTHUR; FORESTRY; AUSTRALIA., 1962), porém já apresenta resultados importantes que estimulam a melhoria deste tipo de abordagem.

Para todas as simulações realizadas nessa tese foram utilizados os softwares ARCGIS versão 10.5 – responsável pela geração dos dados geo-referenciados utilizados na modelagem do risco de incêndio florestal no Capítulo 4 – e MATLAB versão 7.10.0.499(R2010a) para todas as outras simulações.

Outro aspecto da originalidade apresentado nesse trabalho reside no fato dos modelos utilizarem aprendizado computacional de um conjunto representativo de dados em que a “ligação” entre as variáveis de entrada e o risco ou perigo não é definida *a priori*, mas sim, aprendida a partir dos dados disponíveis. Além disso, a quantidade de parâmetros a serem estimados é muito pequena e essas estimativas podem ser feitas a partir dos dados disponíveis, dispensando assim a necessidade de experimentos físicos/químicos e o uso de fórmulas empíricas.

Outro ponto importante a ser destacado é que apresentamos nessa tese mo-

delos matemáticos relativamente simples, pois possuem poucas variáveis de entrada em comparação com outros modelos da literatura ((BATISTA; OLIVEIRA; SOARES, 2002; SIVRIKAYA *et al.*, 2014; CHUVIECO *et al.*, 2003)), que fornecem aproximações muito boas e confiáveis para um problema extremamente complexo e que está longe de ser resolvido completamente devido a esta alta complexidade.

Propostas de Trabalhos Futuros

A primeira ideia que podemos explorar é a de aumentar o número de variáveis de entrada no modelo de risco de incêndio florestal acrescentando o NDVI (índice extraído a partir de imagens de satélite e que está relacionado à vegetação do local) e a declividade do solo. O modelo de perigo de incêndio também pode ser testado com as variáveis temperatura do ar e déficit de vapor de pressão (valor que pode ser obtido através de estações meteorológicas e pode ser relacionado à umidade presente nas florestas). Podemos acrescentar ao modelo de autômato celulares as variáveis direção e velocidade do vento e também aqui a declividade do solo.

Desejamos ainda realizar um estudo, utilizando os modelos apresentados nesta tese, para avaliar quais variáveis são mais relevantes para a estimativa do risco e perigo de incêndios florestais em certas situações específicas.

É importante também, para fins de comparação, simular a propagação de incêndios florestais a partir de três modelos matemáticos elaborados por teorias distintas: Percolação, Autômatos Celulares e por propagação de ondas elípticas. Outro trabalho futuro é o estudo da propagação de incêndio florestal em pequenas áreas que possuam informações sobre a vegetação local, quantidade de combustível e histórico de queimadas no local.

Podemos também utilizar os modelos matemáticos desenvolvidos para detectar regiões com maior risco, perigo e alerta de incêndio florestal do Estado do Acre ou de outras regiões florestais (alteradas antropicamente ou não) de interesse para fornecer informações importantes ao poder público.

Pretendemos elaborar um sistema computacional que possa realizar a quantificação dos índices elaborados nessa tese em tempo real, tanto com a finalidade de realizar o alerta de incêndio, quanto para validar as informações fornecidas pelo sistema de propagação de incêndio utilizado e disponibilizar estas informações a autoridades locais.

Referências

- ACRE, G. E. A. *Zoneamento Ecológico-Econômico do Estado do Acre, Fase II (Escala 1:250.000): Documento Síntese*. [S.l.]: Ed. Rio Branco: SEMA, 2010.
- ACRE, G. E. A. *Plano integrado de prevenção, controle e combate às queimadas e aos incêndios florestais do Estado do Acre*. [S.l.]: Ed. Rio Branco: SEMA, 2013.
- ADAMATZKY, A. *Identification of Cellular Automata*. 1. ed. [S.l.]: Taylor & Francis Ltd., 1994. ISBN 07-484-0172-5.
- ALBRIGNT, D.; MEISNER, B. N. Classification of fire simulation systems. *Fire Mgmt Notes*, v. 59, n. 2, p. 5–12, 1999.
- ALLGÖWER, B.; CARLSON, J. D.; WAGTENDONK, J. W. van. Introduction to fire danger rating and remoting sensing - will remote sensing enhance wildland fire danger rating? In: _____. *Wildland Fire Danger Estimation and Mapping : the Role of Remote Sensing Data*. River Edge, N.J.: World Scientific, 2003. cap. 1.
- ALONSO-BETANZOS, A. *et al.* A neural network approach for forestal fire risk estimation. In: *Proceedings of the 15th European Conference on Artificial Intelligence*. Amsterdam, The Netherlands, The Netherlands: IOS Press, 2002. (ECAI'02), p. 643–647. ISBN 978-1-58603-257-9.
- ANDERBERG, M. R. Book. *Cluster analysis for applications*. [S.l.]: Academic Press New York, 1973. xiii, 359 p. : p. ISBN 0120576503.
- ANDERSON, D. H. *et al.* Modelling the spread of grass fires. *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, Cambridge University Press, v. 23, n. 4, p. 451–466, 1982.
- ANDERSON, H. E. *Aids to Determining Fuel Models for Estimating Fire Behavior*. Intermountain Forest and Range Experiment Station, Ogden, UT, 1982. 22 p.
- ARTERO, A. O. *Inteligência Artificial: Teoria e prática*. São Paulo: Livraria da Física, 2009. ISBN 857861029-6.
- BARROS, L. C.; BASSANEZI, R. C.; LODWICK, W. A. *A First Course in Fuzzy Logic, Fuzzy Dynamical Systems, and Biomathematics*. 1. ed. [S.l.]: Springer-Verlag Berlin Heidelberg, 2017. v. 347. 299 p. ISBN 978-3-662-53322-2.
- BATISTA, A. C.; OLIVEIRA, D. S.; SOARES, R. V. *Zoneamento de Risco de Incêndios Florestais para o Estado do Paraná*. Curitiba-PR: FUPEF, 2002.
- BATISTA, A. C. *et al.* Análise dos impactos das mudanças climáticas sobre o risco de incêndios florestais no estado do paraná. *Scientia Forestalis*, v. 42, n. 104, p. 491–501, 2014.
- BERTONI, E. Queimadas recordes avançam sem controle. *Folha de São Paulo.*, 2015. Caderno Cotidiano. Disponível em: <<http://folha.com/no1699987>>.

- BEZDEK, J. C. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Norwell, MA, USA: Kluwer Academic Publishers, 1981. ISBN 0306406713.
- BISHOP, C. M. *Neural Networks for Pattern Recognition*. New York, NY, USA: Oxford University Press, Inc., 1995. ISBN 0198538642.
- BRADLEY, A. P. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, v. 30, n. 7, p. 1145 – 1159, 1997. ISSN 0031-3203.
- BRAGA, A. C. S. *Curvas ROC: aspectos funcionais e aplicações*. Tese (Tese de Doutorado) — Universidade do Minho, 2000.
- BRASIL. *Instituto Brasileiro de Geografia e Estatística*. 2013. Disponível em: <<http://cod.ibge.gov.br/MOG>>.
- BRASIL, M. A. P. A. *Instituto Nacional de Meteorologia*. 2011. Disponível em: <http://www.inmet.gov.br>. Acesso em: 20/06/2015.
- BROWN, A. A.; DAVIS, K. P. *Forest Fire: Control and use*. 2. ed. New York: McGraw Hill, 1973. 686 p.
- CASTANHO, M. J. P. *et al.* Fuzzy receiver operating characteristic curve: An option to evaluate diagnostic tests. *IEEE Transactions on Information Technology in Biomedicine*, v. 11, p. 244–250, 2007.
- CECCONELLO, M. S. *Modelagem alternativa para dinâmica populacional: Sistemas dinâmicos fuzzy*. Campinas: IMECC- UNICAMP, 2006. (Dissertação de Mestrado).
- CHERKASSKY, V.; MULIER, F. *Learning From Data: Concepts, Theory, and Methods*. New York, USA: Wiley, 1998.
- CHIU, S. L. *Extracting Fuzzy Rules from Data for Function Approximation and Pattern Classification*. In: Dubois D. and Prade H. and Yager R. New York, NY, USA: John Wiley & Sons, 1997.
- CHUVIECO, E. *et al.* Design of an empirical index to estimate fuel moisture content from noaa-avhrr images in forest fire danger studies. *International Journal of Remote Sensing*, v. 24, p. 1621–1637, abr. 2003.
- CHUVIECO, E.; ALLGÖWER, B.; SALAS, J. Integration of physical and human factors in fire danger assessment. In: _____. *Wildland Fire Danger Estimation and Mapping : the Role of Remote Sensing Data*. River Edge, N.J.: World Scientific, 2003. cap. 7.
- CHUVIECO, E.; KASISCHKE, E. S. Remote sensing information for fire management and fire effects assessment. *Journal of Geophysical Research: Biogeosciences*, v. 112, n. G1, 2007. ISSN 2156-2202. G01S90.
- COELHO, G. P. *Geração, Seleção e Combinação de Componentes para Ensembles de Redes Neurais Aplicadas a Problemas de Classificação*. Dissertação (Dissertação de Mestrado) — Universidade Estadual de Campinas - Campinas-SP, 2006.
- COUNTRYMAN, C. M. Rating fire danger by the multiple basic index system. *Journal of Forestry*, v. 64, n. 8, p. 531–536, 1966.

- DA SILVA, I. D. B.; BARROS, L. C.; MEYER, J. F. C. A. Modelagem da dinâmica de incêndios florestais em rio branco-ac: uma abordagem baseada em autômato celular acoplado a sistema fuzzy. *Biomatemática*, v. 25, p. 131–144, 2015.
- DA SILVA, I. D. B. *et al.* Estimativa de risco e perigo de incêndios florestais utilizando subconjuntos fuzzy, k-nn fuzzy e subtractive clustering. In: _____. *Recentes Avanços em Sistemas Fuzzy*. São Carlos: Sociedade Brasileira de Matemática Aplicada e Computacional, 2016. v. 1, p. 398–409.
- DA SILVA, I. D. B.; PONTES JR., A. C. F. Elaboração de um fator de risco de incêndios florestais utilizando lógica fuzzy. *Biomatemática*, v. 21, p. 113–128, 2011.
- DAVIS, J.; GOADRICH, M. The relationship between precision-recall and roc curves. In: *Proceedings of the International Conference on Machine Learning*. Pittsburgh, PA: [s.n.], 2006. p. 233–240.
- DE BAETS, B.; DE MEYER, H. Transitivity-preserving fuzzification schemes for cardinality-based similarity measures. *European Journal of Operational Research*, v. 160, n. 3, p. 726 – 740, 2005. ISSN 0377-2217. Decision Analysis and Artificial Intelligence.
- DE BAETS, B.; DE MEYER, H.; NAESSENS, H. A class of rational cardinality-based similarity measures. *Journal of Computational and Applied Mathematics*, v. 132, p. 51–69, 2001.
- DE MAESSCHALCK, R.; JOUAN-RIMBAUD, D.; MASSART, D. L. The Mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems*, Elsevier, v. 50, n. 1, p. 1–18, 2000.
- DEEMING, J. E.; BURGAN, R. E.; COHEN, J. D. *Aids to Determining Fuel Models for Estimating Fire Behavior*. Intermountain Forest and Range Experiment Station, Ogden, UT, 1977. 63 p.
- DUDA, R. O. *et al.* *Pattern Classification, 2nd Ed.* 2001.
- DUDANI, S. A. The distance-weighted k-nearest-neighbor rule. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6, n. 4, p. 325–327, April 1976. ISSN 0018-9472.
- EVERITT, B.; LANDAU, S.; LEESE, M. *Cluster Analysis*. London, UK: Arnold, 2001.
- FAO, F. *Wildland fire management terminology. Terminologie de la lutte contre les incendies de forêt. Terminologia del control de incendios en tierras incultas*. Roma, 1986. 257 p.
- FIORUCCI, P.; GAETANI, F.; MINCIARDI, R. Development and application of a system for dynamic wildfire risk assessment in Italy. *Environmental Modelling & Software*, v. 23, n. 6, p. 690 – 702, 2008. ISSN 1364-8152.
- FLANNIGAN, M.; WOTTON, B. Chapter 10 - climate, weather, and area burned. In: JOHNSON, E. A.; MIYANISHI, K. (Ed.). *Forest Fires*. San Diego: Academic Press, 2001. p. 351 – 373. ISBN 978-0-12-386660-8.
- FUKUNAGA, K.; HOSTETLER, L. D. k-nearest-neighbor bayes risk estimation. *IEEE Trans. Inform. Theory*, v. 21, p. 285 – 293, 1975.

- GALVÃO, M. *Governo do Acre valida seu plano de combate às queimadas e incêndios florestais*. 2013. Disponível em: http://www.agencia.ac.gov.br/wp-content/uploads/2013/06/2013_maiou_thumbnails_Foto01.jpg. Acesso em: 19/05/2017.
- GOLDAMMER, J. G.; BUCHHOLZ, G.; RESCH, F. The contribution of the global fire monitoring center (gfmc) for early warning and management of wildfires. In: _____. *Zschau, Jochen and Küppers, Andreas. Early Warning Systems for Natural Disaster Reduction*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. p. 585–592. ISBN 978-3-642-55903-7.
- GOLDARAG, Y. J.; MOHAMMADZADEH, A.; ARDAKANI, A. S. Fire risk assessment using neural network and logistic regression. *J Indian Soc Remote Sens.*, v. 44, n. 6, p. 885–894, 2016.
- GOMIDE, F.; PEDRYCZ, W. *An Introduction to Fuzzy Sets-Analysis and Design*. Massachusetts: Massachusetts Institute of Technology, 1998.
- GROUP, F. C. F. D. *Development and Structure of the Canadian Forest Fire Behavior Prediction System*. Forestry CA, Ottawa, Ontario, 1992.
- HAJIAN-TILAKI, K. Receiver operating characteristic (roc) curve analysis for medical diagnostic test evaluation. *Caspian Journal of Internal Medicine*, v. 4, n. 2, p. 627–635, 2013.
- HANLEY, J. A.; MCNEIL, B. J. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology*, v. 143, n. 1, p. 29–36, 1982.
- HANSEN, P.; JAUMARD, B. Cluster analysis and mathematical programming. *Math. Program.*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, v. 79, n. 1-3, p. 191–215, 1997.
- HAYKIN, S. *Redes Neurais: Princípios e Práticas*. Porto Alegre-RS: Ed. Bookman, 2001.
- INC, G. *Google Earth Pro 7.1.8.3036 (32-bit)*. 2017. Disponível em: www.google.com.br/intl/pt-BR/earth/.
- INPE, D. P. I. *Divisão de Processamento de Imagens*. 2011. Disponível em: <https://prodwww-queimadas.dgi.inpe.br/bdqueimadas/>. Acesso em: 20/03/2015.
- JAIN, A. K.; DUBES, R. C. *Algorithms for Clustering Data*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1988. ISBN 0-13-022278-X.
- KELLER, J. M.; GRAY, M. R.; GIVENS JR., J. A. A fuzzy k-nearest neighbor algorithm. *IEEE Transact. on Syst., Man, and Cybern.*, v. 15, n. 4, p. 580–585, 1985.
- KOHONEN, T. The self-organizing map. In: *Proceedings of IEEE 78*. [S.l.: s.n.], 1990. p. 1464–1480.
- KWON, O.; LEE, J. Text categorization based on k-nearest neighbor approach for web site classification. *Information Processing & Management*, v. 39, n. 1, p. 25 – 44, 2003. ISSN 0306-4573.
- LAUREANO, E. E. *θ -FAMs: Mamórias Associativas Fuzzy Baseadas em Funções θ* . Tese (Doutorado) — Universidade Estadual de Campinas - Unicamp, 2014.

- LOURENÇO, L. Uma fórmula simples para o cálculo do risco meteorológico de incêndio na floresta de Portugal. In: *Proceedings do II Congresso Florestal Nacional*. [S.l.: s.n.], 1990. p. 743–754.
- LOUZADA, V. H. P.; FERREIRA JR., W. C. Incêndios florestais em autômatos celulares, simples e grandes queimadas. *Biomatemática*, v. 18, p. 15–26, 2008.
- MASSAD, E. *et al. Fuzzy Logic in Action: Applications in Epidemiology and Beyond*. Berlin, Heidelberg: Springer-Verlag, 2008. v. 232. (Studies in Fuzziness and Soft Computing, v. 232).
- MCARTHUR, A. G.; FORESTRY; AUSTRALIA., T. B. Book; Book/Illustrated. *Control burning in eucalypt forests*. [S.l.]: Canberra : Forestry and Timber Bureau, 1962.
- MCARTHUR, A. G.; FORESTRY, A.; BUREAU, T. Book; Book/Illustrated. *Weather and grassland fire behaviour*. [S.l.]: Canberra : Forestry and Timber Bureau, 1966. Prepared for the Country Fire Authority and Victorian Rural Fire Brigades Association Group Officers Study Period, Geelong, 13th-15th August, 1965.
- METZ, C. E. Statistical analysis of roc data in evaluating diagnostic performance. In: _____. *D. E. Herbert and R. H. Myers editors. Multiple Regression Analysis: Applications in the Health Sciences, number 13*. New York, NY: American Institute of Physics, 1986. p. 365–384.
- MIRANDA, E. E. *Brasil em Relevô. Campinas: Embrapa Monitoramento por Satélite, 2005*. 2014. Disponível em: <<http://www.relevobr.cnpm.embrapa.br>>.
- NATURALEZA, I. C. *Manual de operaciones contra incendios forestales*. Madrid, Spain, 1993. 65 p.
- NEPSTAD, D. C.; MOREIRA, A. G.; ALENCAR, A. A. *Floresta em Chamas: Origens, impactos e prevenção de fogo na Amazônia*. Brasília-DF, Brasil: Editado pelos Autores, 1999.
- NOBLE, R.; BARY, G. A. V.; GILL, A. M. McArthur's fire-danger meters expressed as equations. *Australian Journal of Ecology*, n. 5, p. 201–203, 1980.
- NUNES, J. R. S.; SOARES, R. V.; BATISTA, A. C. Ajuste da fórmula de monte alegre alterada (fma+) para o estado do Paraná. *Floresta*, v. 37, p. 1–14, 2007.
- NUNES, J. R. S.; SOARES, R. V.; BATISTA, A. C. *Manual de prevenção e combate a incêndios florestais 2ª edição*. Curitiba: Editado pelos autores, 2008.
- PALMIERI, S.; INGHILESI, R.; SIANI, A. M. Un indice meteorologico di rischio per incendi boschivi. In: *Proceedings from Seminar on Fighting Forest Fires*. [S.l.: s.n.], 1993.
- PANTOJA, N. V.; BROWN, I. F. Estimativas de áreas afetadas pelo fogo no leste do Acre associadas à seca de 2005. In: *Anais do XIII Simpósio Brasileiro de Sensoriamento Remoto (SBSR)*. Natal, Brazil: [s.n.], 2009. p. 6029–6036.
- PASTOR, E. *et al. Mathematical models and calculation systems for the study of wildland fire behaviour*. *Progress in Energy and Combustion Science*, v. 29, n. 2, p. 139 – 153, 2003. ISSN 0360-1285.

- PEDRYCZ, W.; GOMIDE, F. *Fuzzy Systems Engineering: Toward human-centric computing*. New Jersey: John Wiley & Sons, Inc., 2007. ISBN 9780470168967.
- PEIXOTO, M. S. *Sistemas dinâmicos e controladores fuzzy: Um estudo da dispersão da Morte Súbita dos Citros em São Paulo*. Campinas: IMECC - UNICAMP, 2005. (Tese de Doutorado).
- PENMAN, H. Natural evaporation from open waters, bare soil and grass. In: *Proc. Roy. Soc. London*. [S.l.: s.n.], 1948. (A 193), p. 120–145.
- PREISLER, H. K. *et al.* Probability based models for estimation of wildfire risk. *International Journal of Wildland Fire*, v. 13, n. 2, p. 133–142, 2004.
- ROTHERMEL, R. C. *A Mathematical Model for Predicting Fire Spread in Wildland Fuels*. Intermountain Forest and Range Experiment Station, Ogden, UT, 1972. 40 p.
- SAN-MIGUEL-AYANZ, J. *et al.* Current methods to assess fire danger potential. In: _____. *Wildland Fire Danger Estimation and Mapping*. [S.l.]: World Scientific, 2011. cap. 2, p. 21–61.
- SEMOLINI, R. *Support Vector Machines, Inferência Transdutiva e o Problema de Classificação*. Dissertação (Dissertação de Mestrado) — UNICAMP, Campinas-SP, 2002.
- SHAW, I. S.; SIMÕES, M. G. *Controle e Modelagem Fuzzy*. São Paulo: Edgard Blucher Ltda, 1999. ISBN 8521202482.
- SILVA, J. D. M. *Análise de Estabilidade de Sistemas Dinâmicos P-fuzzy com Aplicações em Biomatemática*. Tese (Doutorado) — Universidade Estadual de Campinas - UNICAMP, Campinas-SP, 2005.
- SIVRIKAYA, F. *et al.* Evaluation of forest fire risk with gis. *Polish Journal Environmenta Studies*, v. 23, n. 1, p. 187–194, 2014.
- SOARES, R. V.; BATISTA, A. C. *Incêndios Florestais controle, efeitos e uso do fogo*. Curitiba: Editado pelos autores, 2007.
- SOL, B. Estimation du risque météorologique d'incendies de forêts dans le sud-est de la france. *Rev. For. Fr.*, v. 42, p. 296–363, 1990.
- SWETS, J. A. Roc analysis applied to the evaluation of medical imaging techniques. *Invest Radiol*, v. 14, n. 2, p. 109–121, 1979.
- TAVARES, H. R. *et al.* Modelagem da dinâmica de incendios florestais em rio branco-ac: uma abordagem baseada em autômato celular acoplado a sistema fuzzy. *Biomatemática*, v. 25, p. 131–144, 2015.
- VALLE, M. E.; SOUZA, A. C. On the recall capability of recurrent exponential fuzzy associative memories based on similarity measures. *Mathware & Soft Computing Magazine*, v. 22, n. 1, p. 33–39, 2015.
- VAN WAGNER, C. E. *Development and structure of the Canadian Forest Fire Weather Index System*. Place Vincent Massey, 3rd Floor, Ottawa, Ontario, 1987.

WHITE, G. M.; FONG, P. J. k-nearest-neighbor decision rule performance in a speech recognition system. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-5, n. 3, p. 389–389, May 1975. ISSN 0018-9472.

WOLFRAM, S. *Cellular Automata and Complexity: collected papers*. 1. ed. [S.l.]: Addison-Wesley, 1994. ISBN 0-201-62716-7.

WU, Y.; IANAKIEV, K.; GOVINDARAJU, V. Improved k-nearest neighbor classification. *Pattern Recognition*, v. 35, n. 10, p. 2311 – 2318, 2002. ISSN 0031-3203.

XU, R.; WUNSCH, D. Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, v. 16, n. 3, p. 645–678, May 2005. ISSN 1045-9227.

YAGER, R. R.; FILEV, D. P. Generation of fuzzy rules by mountain clustering. *J. Intell. Fuzzy Syst.*, IOS Press, Amsterdam, The Netherlands, The Netherlands, v. 2, n. 3, p. 209–219, maio 1994. ISSN 1064-1246.