

**Universidade Federal de Minas Gerais
Instituto de Geociências
Programa de Pós-Graduação em Modelagem e Análise de Sistemas Ambientais**

Luiza Cintra Fernandes

Modelagem de risco de incêndios florestais utilizando redes neurais artificiais aplicada às regiões metropolitanas

**Belo Horizonte
2019**

Luiza Cintra Fernandes

Modelagem de risco de incêndios florestais utilizando redes neurais artificiais aplicada às regiões metropolitanas

Dissertação apresentada ao programa de Pós-graduação de Análise e Modelagem de Sistemas Ambientais da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Mestre em Análise e Modelagem de Sistemas Ambientais

Orientador: Prof. Dr. Marcelo Antonio Nero

Co-orientador: Prof. Dr. Plínio da Costa Temba

Belo Horizonte
Instituto de Geociências da UFMG
2019

F363m
2019

Fernandes, Luiza Cintra.
Modelagem de risco de incêndios florestais utilizando redes neurais artificiais aplicada às regiões metropolitanas [manuscrito] / Luiza Cintra Fernandes. – 2019.
xv,163 f., enc.: il. (principalmente color.)

Orientador: Marcelo Antonio Nero.
Coorientador: Plínio da Costa Temba.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Departamento de Cartografia, 2019.
Bibliografia: f. 116-12870.
Inclui apêndices.

1. Modelagem de dados – Aspectos ambientais – Teses. 2. Incêndios florestais – Previsão – Minas Gerais – Teses. 3. Incêndios florestais – Prevenção e controle – Teses. 4. Redes neurais (Computação) – Teses. I. Nero, Marcelo Antônio. II. Temba, Plínio da Costa. III. Universidade Federal de Minas Gerais. Departamento de Cartografia. IV. Título.

CDU: 911.2:519.6(815.1)



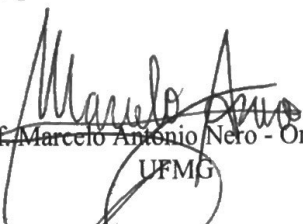
FOLHA DE APROVAÇÃO

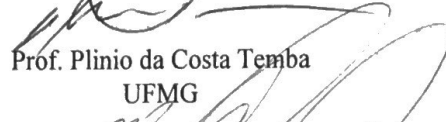
Modelagem de risco de incêndios florestais utilizando redes neurais artificiais aplicada à regiões metropolitanas

LUIZA CINTRA FERNANDES

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ANÁLISE E MODELAGEM DE SISTEMAS AMBIENTAIS, como requisito para obtenção do grau de Mestre em ANÁLISE E MODELAGEM DE SISTEMAS AMBIENTAIS, área de concentração ANÁLISE E MODELAGEM DE SISTEMAS AMBIENTAIS.

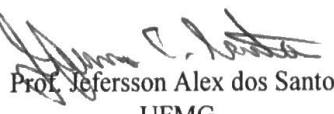
Aprovada em 05 de abril de 2019, pela banca constituída pelos membros:

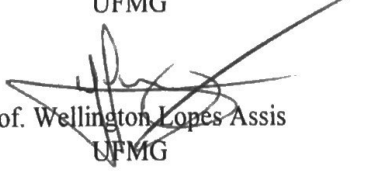

Prof. Marcelo Antonio Nero - Orientador
UFMG


Prof. Plinio da Costa Temba
UFMG


Prof. Marco Antonio Timó Emfiro
UFMG


Prof. Haroldo Fraga de Campos Velho
INPE


Prof. Jefersson Alex dos Santos
UFMG


Prof. Wellington Lopes Assis
UFMG

Belo Horizonte, 5 de abril de 2019.

AGRADECIMENTOS

Agradeço primeiramente minha família, meu pai Eleutério, minha mãe Rosângela e minha irmã Paula, que sempre foram meu porto seguro, sempre me apoiando, me incentivando e estando comigo em todos os momentos. Principalmente minha mãe, minha maior inspiração e exemplo, que me incentivou a trabalhar com Redes Neurais Artificiais, me passando seus conhecimentos e auxiliando nos meus programas. Dividiu comigo as dificuldades deste trabalho, me auxiliando da melhor forma possível.

Agradeço meus orientadores Prof. Dr. Marcelo Antonio Nero e Prof. Dr. Plínio da Costa Temba pelas orientações, boa convivência e confiança depositada em mim.

Agradeço ao programa de Pós-Graduação de Análise e Modelagem de Sistemas Ambientais e a UFMG pelas oportunidades e conhecimentos adquiridos.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de mestrado nos últimos 24 meses.

Agradeço aos membros da banca, Prof. Dr. Haroldo Fraga de Campos Velho, Prof. Dr. Wellington Lopes Assis, Prof. Dr. Marcos Antônio Timbó Elmiro e Prof. Dr. Jefersson Alex dos Santos por suas valiosas contribuições e sugestões. Agradeço principalmente ao Prof. Haroldo que disponibilizou um computador no LAC/INPE, bem como sua autorização para acesso remoto e a disponibilização do pacote MPCA para uso.

Agradeço aos meus “co-orientadores” Max, Érika, Marina e Luciana, que estavam sempre me auxiliando na medida do possível, com discussões, ajuda na confecção dos mapas, revisão do texto, entre outras coisas. Além do tempo que passamos juntos na salinha da modelagem, junto com Nina, Fernanda, Carol, Trevor e Marlon, onde dividimos os momentos de desespero, momentos de desabafo, momentos de descontração e de experiências culinárias. Gostaria de agradecer muito a todos esses amigos maravilhosos do programa onde formamos uma grande família da modelagem. Em especial a Érika que nos últimos meses “morou” comigo na salinha e aguentou todas as minhas loucuras. E a Luciana que estava sempre disponível para uma conversa e uma cervejinha para descontrair.

Agradeço a todas as amigas que eu fiz em Belo Horizonte e a todos que participaram de meu cotidiano, que de alguma forma me ajudaram a completar essa missão. A

Brenda que dividiu comigo o apartamento nesses últimos dois anos, sempre me dando apoio e incentivo. A Natália por acreditar no meu potencial, me incentivar, me ouvir e ter orgulho das minhas pequenas conquistas. A Bárbara e à Ana Luisa que me acolheram como família e me socorreram em momentos de necessidade. A Michelle, Cássia, Adriana, Cris, Babi, Aradielle e Juliana pelos momentos de descontração e alegria, que foram minha “válvula de escape” nesses últimos anos, me ajudando a manter minha saúde mental.

Agradeço a todos os meus amigos que, mesmo distantes, me apoiaram e me estimularam a dar sempre o meu melhor. Principalmente minhas amigas do Ciências sem Fronteiras, Amanda, Doca, Natalia, Barbara, Marcela, Livia e Camila e minha amiga Giovanna.

A todos que aqui não foram citados, mas que de alguma forma contribuíram diretamente ou indiretamente para o meu crescimento.

RESUMO

Incêndios Florestais representam uma das grandes preocupações relacionadas com o meio ambiente, que causam impactos na vegetação, no solo, na fauna e flora. O cálculo do risco de incêndios consiste em verificar a probabilidade de ocorrer incêndio em determinado local, sendo importante para determinar áreas com maior risco e planejar atividades para sua prevenção. Dentre as metodologias existentes para seu cálculo, as Redes Neurais Artificiais (RNA) é um método que tem obtido bons resultados. Uma RNA consiste em um conjunto de unidades de processamento simples que trabalham em paralelo, armazenando conhecimento experimental e tornando-o disponível para uso. Este conhecimento é adquirido no processo de treinamento e armazenamento dos pesos sinápticos (conexões entre os neurônios). Assim, esse trabalho teve como objetivo fazer o mapeamento das regiões de risco de incêndio na região metropolitana de Belo Horizonte (RMBH) utilizando um modelo de redes neurais artificiais com treinamento supervisionado. Esta região apresenta períodos de seca acentuados, devido aos sistemas atmosféricos atuantes, os quais deixam a região muito propícia à ocorrência de incêndios. Foram utilizadas 12 entradas na RNA: distância às áreas urbanas, distância às rodovias, declividade, orientação do terreno, cobertura e uso do solo, NDVI, umidade relativa mínima do ar, temperatura máxima do ar, radiação solar, velocidade média do vento, pressão atmosférica e precipitação total. Para treinamento da rede foram utilizados dados mensais de 2014 a 2016, sendo utilizada como saída desejada focos de incêndio dos sensores MODIS e VIIRS. A validação foi feita com dados de 2017. Quanto aos resultados, o modelo de RNA apresentou alta acurácia e coerência com dados de incêndio, na validação subestimou áreas de risco em épocas de seca. Assim foram feitos modelos complementares, sendo treinados apenas com dados de chuva e seca separados. Estes modelos apresentaram bons resultados. No modelo em épocas de seca a área de alto risco abrange praticamente a região inteira, enquanto o modelo com dados de chuva apresentou área de alto risco apenas próximo às áreas urbanas. O modelo geral, mesmo não tendo o melhor resultado, foi treinado com características das duas épocas, selecionando apenas fatores que realmente influenciam os incêndios, podendo ser utilizado como um complemento aos outros modelos. Dentre as variáveis de entrada as mais influentes no modelo foram: a distância às áreas urbanas e as variáveis meteorológicas. Assim a RNA se mostrou uma boa metodologia para este fim. O método pode ser usado como modelo de previsão e ser reaplicada a outras áreas.

Palavras-Chave: risco de incêndio, redes neurais artificiais, modelagem

ABSTRACT

Forest Fires represents one of the great concerns related with the environment, as it causes impacts in the vegetation, soil, fauna and flora. The fire risk index consists in verify the probability of fire occurrence in such place, so it is important to determine higher fire risk areas and planning activities for your prevention. From the existing methodologies for your calculation, the Artificial Neural Networks (ANN) is a methodology that have been achieving good results. The ANN it is a set of simple process units working in parallel, which stores experimental knowledge and make it available to use. This knowledge is acquired from a training process and stored in synaptic weights (connections between neurons). So, this work has the aim of mapping the fire risk areas in the Metropolitan Region of Belo Horizonte (RMBH) using an artificial neural network model with supervised training. This region presents accentuated dry seasons, due atmospheric systems activities; this season became the region propitious for fire ignition. 12 input variables were used: distance to urban areas, distance to roads, slope, aspect, land use, NDVI, minimum relative humidity, maximum air temperature, air pressure, solar radiation, wind speed and total precipitation. For network training it was used monthly data from 2014 to 2016, as output, fire active data from MODIS and VIIRS were used. A model evaluation was done with 2017 data. The RNA model results showed high accuracy and good consistency with fire data, but your validation underestimated risk areas in dry season. So complementary models were developed, training dry and rainy seasons separately. These models showed good results. In the dry model the high risk area covers almost all the region, in the other hand, the rainy model showed high risk only in locations near urban areas. The general model even dough it did not showed the best result, it was trained with characteristics of both seasons, selecting just the factors that really affects the fire, and can be used as a complementary model for the others. Among the input variables those which have more influence in the model are the distance to urban areas and meteorological variables. So, the ANN showed a good methodology for this end and can be used as a forecast model and be reapplied in other regions.

Keywords: fire risk, neural artificial networks, modeling

SUMÁRIO

Introdução	1
1. Fundamentação teórica e bases conceituais.....	3
1.1. Conceitos básicos	3
1.2. Metodologias Utilizadas no Cálculo de Risco de Incêndios	4
1.2.1. Análise Multicritério	5
1.2.1. Lógica <i>Fuzzy</i>	7
1.2.2. Regressão Logística.....	8
1.2.3. Fórmula de Monte Alegre	8
1.2.4. FIRA e FARSITE.....	9
1.2.1. Redes Neurais Artificiais	10
1.3. Agentes causadores de incêndio.....	11
1.3.1. Topografia	12
1.3.2. Uso e Ocupação do Solo	13
1.3.3. Hidrografia.....	13
1.3.4. Fatores Climáticos	13
1.4. Uso de dados de satélite para detecção de incêndios	15
1.5. Redes Neurais Artificiais (RNA).....	19
1.5.1. Modelo de um neurônio.....	20
1.5.2. Arquitetura da rede	24
1.5.3. Processo de Aprendizagem	25
1.5.4. Algoritmo <i>Backpropagation</i>	28
1.5.5. RNA autoconfigurada: Multi-Particle Collision Algorithm (MPCA).....	29
1.5.5 Algoritmo MPCA	29
2. Materiais e Métodos.....	33
2.1. Dados e materiais.....	35
2.1.1. Dados NDVI.....	35
2.1.2. Dados meteorológicos.....	35
2.1.3. Mapa de uso e cobertura do solo.....	38
2.1.4. Distância à mancha urbana e às rodovias	38
2.1.5. Dados de elevação.....	40
2.1.6. Focos de incêndio	40
2.1.7. Processamento dos dados.....	40
2.2. Construção da RNA.....	41
2.2.1. Normalização dos Dados	41
2.2.2. MPCA-NN.....	42
2.2.3. Rede Neural Artificial.....	42
2.2.4. Avaliação da RNA.....	43
2.2.5. Análise das variáveis mais características no modelo	44
2.2.5.1. Algoritmo de Garson	44
2.2.5.2. Método Perturb.....	44
2.2.5.3. Correlação	45
2.3. Validação do Modelo.....	45
3. Caracterização da Área de Estudo e Pré-análise da pesquisa	45

3.1. Informações Gerais	45
3.2. Uso e ocupação do Solo.....	52
3.2.1. Caracterização da Vegetação	52
3.3. Caracterização Climatológica RMBH	56
3.3.1. Elementos influenciadores do clima da RMBH.....	64
3.3.2. Dados meteorológicos dos anos do estudo.....	67
3.4. Análise de Incêndios na RMBH	68
4. Resultados e Discussões.....	77
4.1. Estatística Descritiva dos Dados.....	77
4.2. Resultados MPCA-NN	80
4.3. Resultados da RNA	81
4.3.1. Avaliação da Rede	83
4.4. Validação	84
4.5. Avaliação Rede Neural: Modelo época de Seca.....	88
4.6. Avaliação Rede Neural: Modelo Época de chuva.....	89
4.7. Validação dos modelos.....	91
4.8. Análise das variáveis do modelo.....	95
4.9. Análises e discussões.....	99
Conclusão.....	112
Referências	116
APÊNDICE A.....	129
APÊNDICE B.....	135
APÊNDICE C.....	136
APÊNDICE D.....	140
APÊNDICE E	145
APÊNDICE F	152

LISTA DE FIGURAS

Figura 1- Fluxograma dos dois algoritmos (adaptado de MBOW; GOÏTA; BÉNIÉ, 2004)....	10
Figura 2 - Funcionamento do Algoritmo MODIS Collection 6 (Adaptado de GIGLIO; SCHROEDER; JUSTICE, 2016).....	18
Figura 3 – Modelo de um neurônio biológico e modelo de um neurônio artificial (adaptado de ROMERO, 2017).	20
Figura 4- Modelo de um neurônio (COUTINHO; SILVA; DELGADO, 2016)	21
Figura 5 - Função degrau (NICOLAU, 2016)	22
Figura 6 - Função Linear (NICOLAU, 2016).....	22
Figura 7 - Função semi-linear (NICOLAU, 2016)	23
Figura 8- Função Logística (a) e função tangente hiperbólica (b) (NICOLAU, 2016)	23
Figura 9 - Algumas Arquiteturas de rede (ROMERO, 2017).....	25
Figura 10 - Rede Neural Perceptron Multicamadas (GOLDARAG; MOHAMMADZADEH; ARDAKANI, 2016).....	25
Figura 11 –Pseudocódigo do algoritmo MPCA (ANOCHI; CAMPOS VELHO, 2014)	30
Figura 12 – Pseudocódigo Função Exploration (ANOCHI; CAMPOS VELHO, 2014).....	31
Figura 13 – Pseudocódigo Função Scattering (ANOCHI; CAMPOS VELHO, 2014)	31
Figura 14 – Roteiro Metodológico.....	34
Figura 15 - Localização das Estações meteorológicas.....	37
Figura 16 – Mapas de distâncias: (A) Distância a Principais Rodovias (B) Distância a áreas Urbanas.	39
Figura 17 - Mapa de localização da Área de Estudo	47
Figura 18 - Densidade Populacional RMB	48
Figura 19 - Relação Área Urbana/Rural na RMBH.....	49
Figura 20 - Unidades de Conservação na RMBH.....	51
Figura 21 - Classificação Uso e ocupação do solo na RMBH para o ano de 2017. Fonte: MAPBIOMAS (2017).....	53
Figura 22 - Classes de Vegetação Fonte: IEF	54
Figura 23 - Classificação Climática da RMBH e estações do INMET para cálculo das normais climatológicas	57
Figura 24 - Climograma estações RMBH (a) Belo Horizonte, (b) Florestal e (c) Ibitité Fonte dos dados: (INMET, 2018b)	58
Figura 25 – Média da Precipitação Total das estações da RMBH da série histórica de 1987 a 2017 Fonte dos dados: ANA (2018)	59
Figura 26 – Média da Precipitação Máxima das estações da RMBH das série histórica de 1987 a 2017 Fonte dos dados: ANA (2018)	60
Figura 27 – Média da Precipitação Total Mensal das estações da RMBH das séries histórica de 1987 a 2017 Fonte dos dados: ANA (2018).....	60
Figura 28 – Precipitação Total na RMBH, média histórica de 1987-2017.....	61
Figura 29 – Balanço hídrico estações RMBH.....	62
Figura 30 – Umidade Relativa Compensada na RMBH.....	63
Figura 31 – Pressão Atmosférica na RMBH.....	63
Figura 32 – Velocidade do Vento na RMBH.....	64
Figura 33 – Altimetria na RMBH	65
Figura 34 – Circulação Geral da Atmosfera Fonte: Adaptado de (REBOITA; AMBRIZZI; PORFÍRIO DA ROCHA, 2012)	66
Figura 35 –Comparação entre as cidades da RMBH e as categorias de incêndios, segundo dados da CBMMG Fonte: (CBMMG, 2018).....	69

Figura 36 -Comparação entre as cidades da RMBH e as categorias de incêndios, segundo dados de satélite Fonte: (IBGE, 2010a; NASA, 2018; UFMG, 2014)	70
Figura 37 - Comparação entre as cidades da RMBH e as classes de uso do solo, segundo dados de satélite Fonte: (MAPBIOMAS, 2017; NASA, 2018).....	72
Figura 38-Focos de incêndio por mês na RMBH, para os anos estudados, segundo dados do CBMMG Fonte: (CBMMG, 2018).....	73
Figura 39-Focos de incêndio por mês na RMBH, para os anos estudados, segundo dados de satélite Fonte: (NASA, 2018)	73
Figura 40 - Dados de incêndio quanto ao ano Fonte: (NASA, 2018).....	74
Figura 41 - Dados de FRP (MW) durante os meses Fonte: (NASA, 2018).....	75
Figura 42 - Dados de FRP (MW) durante o ano Fonte: (NASA, 2018).....	75
Figura 43 – Comparação entre os dados dos dois satélites Fonte: (NASA, 2018).....	76
Figura 44 – Comparação entre os dados dos dois satélites quanto ao FRP Fonte: (NASA, 2018)	76
Figura 45 - Gráfico Quantil-Quantil variáveis meteorológicas	79
Figura 46 - Gráfico Quantil-Quantil para as variáveis: distancia à rodovias, declividade, distancia a urbanização, orientação do terreno e NDVI	80
Figura 47 - Representação da RNA utilizada no modelo. Imagem produzida com o pacote R <i>NeuralNetTools</i>	82
Figura 48 - Variação do Erro de treinamento	83
Figura 49 - Curvas ROC: (A) Treinamento (B) Validação (C) Generalização	84
Figura 50 - Resultados do modelo para os dados do 1º semestre de 2017. Mapas produzidos em ambiente R.	85
Figura 51 - Resultados do modelo para os dados do 2º semestre de 2017. Mapas produzidos em ambiente R	86
Figura 52 - Variação do Erro modelo seca	88
Figura 53 - Curvas ROC modelo seca: (a) Treinamento (b) Validação (c) Generalização	89
Figura 54 - Variação do Erro para modelo de chuva.....	90
Figura 55- Curvas ROC modelo chuva: (A) Treinamento (B) Validação (C) Generalização ..	91
Figura 56 - Resultados modelo chuva. Mapas produzidos em ambiente R.	92
Figura 57 - Resultados modelo seca. Mapas produzidos em ambiente R.....	94
Figura 58 - Relação de importância entre variáveis pelo método de Garson (1991) para o modelo geral.....	95
Figura 59 - Relação de importância entre variáveis pelo método de Garson (1991) para o modelo seca.....	95
Figura 60 - Relação de importância entre variáveis pelo método de Garson (1991) para o modelo chuva	96
Figura 61 - Análise da Importância de variáveis no modelo pelo método da perturbação modelo geral.....	96
Figura 62 - Análise da Importância de variáveis no modelo pelo método da perturbação modelo seca.....	97
Figura 63 - Análise da Importância de variáveis no modelo pelo método da perturbação modelo chuva	97
Figura 64 - Correlação Risco de Fogo com as variáveis modelo geral	98
Figura 65 - Correlação Risco de Fogo com as variáveis do modelo seca.....	98
Figura 66 - Correlação Risco de Fogo com as variáveis do modelo chuva.....	98
Figura 67 - Comparação de modelos para o mês de março	101
Figura 68 - Comparação de modelos para o mês de novembro.....	102
Figura 69 - Comparação de modelos para o mês de julho.....	104
Figura 70 - Comparação de modelos para o mês de novembro.....	105

LISTA DE TABELAS

Tabela 1 - Exemplo de vetorização dos dados da RNA.....	41
Tabela 2 - Relação de área queimada em Unidades de Conservação Estaduais, entre 2012 e 2016. (FEAM, 2017).....	68
Tabela 3 – Descrição dos Dados.....	78
Tabela 4 - Teste de Normalidade dos Dados.....	79
Tabela 5 - Resultados MPCA-NN.....	80
Tabela 6 – Dados de Avaliação da Rede.....	83
Tabela 7 - Percentagem de focos de incêndios classificados quanto ao grau de risco.....	84
Tabela 8 - Percentagem de Focos de incêndio em cada classe de risco.....	87
Tabela 9 - Dados de Avaliação do Modelo.....	88
Tabela 10 - Porcentagem de focos de incêndios classificados quanto ao grau de risco modelo seca.....	89
Tabela 11 - Dados de Avaliação do Modelo.....	90
Tabela 12 - Porcentagem de focos de incêndios classificados quanto ao grau de risco modelo chuva.....	91
Tabela 13 - Classificação dos focos quanto à classe de risco modelo chuva.....	91
Tabela 14 - Classificação dos focos quanto a classe de risco modelo seca.....	93

LISTA DE QUADROS

Quadro 1 - Comparação entre diversos trabalhos da literatura que aplicam a metodologia.....	6
Quadro 2 - Modificação do cálculo conforme a precipitação ocorrida no dia (NASCIMENTO; AMARAL; PONTES, 2013).....	9
Quadro 3 - Satélites que medem a localização de incêndios (Adaptado de ICHOKU; KAHN; CHIN, 2012).....	16
Quadro 4 – Parâmetros definidos pelo MPCA.....	32

LISTA DE ABREVIACOES

A	Acurcia
AHP	Analytic Hierarquy Process (Mtodo de Anlise Hierrquica)
ANA	Agncia Nacional de guas
APM	Anticiclones Polares Migratrios
ASAS	Anticiclones Subtropicais do Atlntico
ATSR	Along-Track Scanning Radiometer
AUC	rea Sob a Curva
AVHRR	Advanced Very High-Resolution Radiometer
CBMMG	Corpo de Bombeiros Militar de Minas Gerais
CPTEC	Centro de Previso de Tempo e Estudos Climticos
E	Especificidade
FAO	Organizao das Naoes Unidas para Agricultura e a Alimentao
FARSITE	Fire Area Simulator (Simulador de rea Queimada)
FIRA	Fire Risk Assessment (Avaliao do Risco de Incndio)
FMA	Frmula de Monte Alegre
FN	Falsos Negativos
FP	Falsos Positivos
FRP	Fire Radioative Power (Potncia Radioativa do Fogo)
GOES	Geostationary Operational Environmental Satellite
INMET	Instituto Nacional de Meteorologia
INPE	Instituto Nacional de Pesquisas Espaciais
IR	Infravermelho Prximo
MLP	MultiLayer Perceptron (Perceptron Muiltiplas Camadas)
MODIS	Moderate Resolution Imaging Spectroradiometer
MPCA	Multi-Particle Collision Algorithm (Algoritmo de coliso de mltiplas partculas)
MPI	Message Passing Interface
NDVI	Normalized Difference Vegetation Index (ndice de Vegetao por Diferena Normalizada)
PCA	Particle Collision Algorithm (Algoritmo de Coliso de Partculas)
PIB	Produto Interno Bruto
PMC	Peceptron Multicamadas

PVN	Predição Valores Negativos
PVV	Predição Valores Verdadeiros
RC	Relação de Consistência
REDEMETS	Rede de Meteorologia do Comando da Aeronáutica
RMBH	Região Metropolitana de Belo Horizonte
RNA	Rede Neural Artificial
ROC	Receiver Operating Characteristic (Característica de Operação do Receptor)
R	Vermelho
S	Sensibilidade
S-NPP	Suomi National Polar-orbiting Partnership
SIG	Sistema de Informação Geográfica
SRTM	Shuttle Radar Topography Mission
TPS	Thin Plate Splines
UC	Unidades de Conservação
USDA	Departamento de Agricultura dos Estados Unidos
VIIRS	Visible Infrared Imager Radiometer Suite
VN	Verdadeiros Negativos
VP	Verdadeiros Positivos
ZCIT	Zona de Convergência Intertropical

Introdução

Incêndio florestal é uma das grandes preocupações relacionada ao meio ambiente, pois devasta grandes proporções de florestas, campos e regiões agrícolas, causando alterações no meio ambiente. Guimarães et al. (2014) verificaram em seu estudo que os principais impactos dos incêndios são: aquecimento do solo, erosão, redução de nutrientes do solo, redução do teor de matéria orgânica, acidez do solo, alteração do microclima local, redução da qualidade da água, mortandade de animais, enfraquecimento da madeira, limpeza do sub-bosque, impactos às culturas agrícolas, interferência na sucessão vegetal, redução da atividade de microrganismos e renovação da pastagem, transtornos à população do entorno, impactos paisagísticos e destruição de áreas de recreação.

A queima de biomassa também libera para atmosfera gases de efeito estufa, tais como: CO₂ (dióxido de carbono), CH₄ (gás metano) e NO_x (óxidos nitrosos) responsáveis pelo aumento do efeito estufa e de problemas respiratórios (ICHOKU; KAUFMAN, 2005). Uma pluma de fumaça tem sido observada na América do Sul, com aproximadamente 4 a 5 milhões de km². Com uma concentração mais alta que 400 mg m⁻³ de partículas inaláveis, e uma concentração de ozônio de 100 ppb que é observada a grandes distâncias de incêndios florestais (LONGO et al., 2009). Na Amazônia Legal é estimado um fluxo de carbono de 0.2 a 1,2 Pg C/ano (HOUGHTON et al., 2009).

As altas temperaturas e a ocorrência de circulações das massas de ar, envolvidas na fase de chamas das queimadas, favorecem ao movimento convectivo ascendente, elevando poluentes até a troposfera e permitindo o transporte desses poluentes sobre uma extensa área da atmosfera, que se distanciam das fontes emissoras (PINTO JUNIOR; SILVA, 2012).

O cálculo do risco de incêndios consiste em verificar a probabilidade de ocorrer incêndio em determinado local, ou seja, determinar a chance de ignição (HARDY, 2005). Com essa medida há a possibilidade de ter uma avaliação crítica da ocorrência de incêndios e então fazer o monitoramento e prevenção de queimadas dessas áreas (CHUVIECO et al., 2010).

A importância em se determinar áreas com maior risco de incêndio está relacionada com o planejamento de atividades para sua prevenção. Assim, é possível determinar as estações em que ocorrem mais queimadas, permitindo queimadas controladas ou sua

prevenção, bem como o combate destas. Permite ainda, informar a população dos riscos e fazer programas de educação ambiental e, adicionalmente, em casos extremos, limitar o acesso à área (NUNES, 2006).

Há várias abordagens para se modelar o risco de incêndio: métodos estatísticos, como regressão logística, métodos quantitativos e qualitativos baseados em conhecimento de especialista, como a análise multicritério. Além de métodos de inteligência artificial, como Redes Neurais Artificiais (RNA) e Lógica Fuzzy. RNA tem sido usada como metodologia para desenvolver modelos de risco de incêndios e tem apresentados bons resultados (ALONSO-BETANZOS et al., 2003; GOLDARAG; MOHAMMADZADEH; ARDAKANI, 2016; SATIR; BERBEROGLU; DONMEZ, 2016).

Este método tenta imitar o cérebro humano, devido a sua capacidade de conhecimento adquirido, armazenado nas conexões entre os neurônios. Diferente de outros métodos, a escolha dos pesos de cada variável do modelo não é empírico, a escolha é buscada através de um processo de aprendizagem, baseado em métodos matemáticos. O que deixa o processo menos subjetivo, tais como em métodos baseados em especialistas.

Esse trabalho tem como objetivo fazer o mapeamento das regiões de risco de incêndio na região metropolitana de Belo Horizonte (RMBH) utilizando um modelo de previsão com o método de redes neurais artificiais de treinamento supervisionado. Assim, a presente pesquisa busca responder às seguintes perguntas: Qual o desempenho do método de Redes Neurais Artificiais aplicado à previsão de risco de incêndio? Qual a localização, bem como a dimensão das áreas que têm maior risco de incêndios na RMBH? Estes riscos são mais associados aos fatores humanos ou ambientais? Como o uso de dados de satélites auxiliam em estudos ambientais?

Esses resultados fornecem o delineamento de áreas prioritárias para planejamento de atividades logísticas das equipes brigadistas, buscando minimizar possíveis danos causados pelos incêndios.

1. Fundamentação teórica e bases conceituais

1.1. Conceitos básicos

Aqui serão explicados alguns conceitos básicos necessários para a melhor compreensão dos capítulos e subcapítulos seguintes.

O *fogo* é uma reação exotérmica onde o comburente reage com o combustível na presença de uma fonte de calor, liberando energia na forma de luz e calor (ALMEIDA, 2008), processo denominado de combustão. Na *combustão* ocorre um processo de oxidação rápida autossustentada, tendo como principais produtos vários gases (CO₂, HCN, CO, HCl, SO, NO_x, ...); calor; chamas e fumaça. Três elementos básicos são necessários para se começar uma combustão (chamado de triângulo do fogo) (MOTTA, 2008):

- *Combustível*: Fornece energia para a queima, representado por elementos susceptíveis a entrar em combustão como: madeira, papel, pano, estopa, entre outros.
- *Comburente*: Substância que reage quimicamente com o combustível e inicia a combustão, sendo o principal comburente o oxigênio.
- *Fonte de Calor*: Necessário para iniciar a reação, pode ser desde luz solar, faíscas de raios atmosféricos a pontas de cigarro.

O fogo descontrolado, que se dissemina no tempo e espaço, é chamado de *incêndio*. Fatores que influenciam os incêndios são: a forma geométrica do local, a superfície dos combustíveis envolvidos, a quantidade do material combustível, o local inicial do acidente, as condições do clima, entre outros (SEITO et al., 2008).

No caso dos *incêndios florestais* as propriedades dos combustíveis têm maior influência no comportamento do fogo. Os *combustíveis florestais* são classificáveis por estrato: manta morta, sub-bosque, combustíveis de transição, copa das árvores. O combustível de superfície inclui a folhada superficial, ervas e arbustos. A continuidade de um combustível descreve sua distribuição e é importante na propagação do fogo. A *carga de combustível* é a sua quantidade por unidade de área; quanto maior a quantidade maior o potencial de queima. Além disso, o teor de umidade em combustíveis florestais é importante para definir a disponibilidade de um combustível arder (FERNANDES; REGO, 2010).

Modelos são representações simplificadas de uma realidade específica, podendo ser classificados em três grandes grupos: escala ou réplica, conceitual e matemático. O primeiro consiste em uma redução análoga ao original. O segundo mostra seu sistema principal, processos e inter-relações com subsistemas. E modelos matemáticos se utilizam de sistemas de equações para sua construção, podendo ser determinísticos e estocásticos (SOARES-FILHO, 1998).

Um modelo matemático determinístico é representado por um conjunto de equações diferenciais determinísticas com termos derivados no tempo, e é feito a partir de um conjunto de condições iniciais. Já modelos estocásticos-probabilísticos utilizam variáveis aleatórias, dependentes de fatores aleatórios ou acaso (SOARES-FILHO, 1998).

A construção de modelos é conhecida como *modelagem*. Processo desenvolvido através da definição de um conjunto de hipóteses e previsões que são comparadas ao mundo real. O aceite do modelo depende da concordância entre os dados observados e os gerados pelo modelo (SOARES-FILHO, 1998).

Um processo de *simulação* consiste na modelagem da dinâmica de um sistema em um ambiente computacional, onde é reproduzida a complexidade dos mecanismos deste sistema, representada por trocas de elementos dentro dele. A simulação é vista como um instrumento de aquisição de conhecimento, a partir da integração da informação sobre sistemas dinâmicos (SOARES-FILHO et al., 2007). Por sua vez um *sistema* é um conjunto organizado de elementos que vivem em interação, segundo determinadas leis, para atingir um objetivo específico (CHRISTOFOLETTI, 1999).

Modelos probabilísticos de distribuição binomial são os utilizados quando se tem *variáveis dicotômicas*. Esse tipo de variável diz respeito sobre o acontecimento ou não de um evento, onde a variável pode apenas assumir somente dois valores 1 (sucesso) ou 0 (fracasso). O interesse desse modelo é verificar a probabilidade de sucessos em relação às variáveis observadas que podem explicar ou prever a variável resposta (POVEDA, 2006).

1.2. Metodologias Utilizadas no Cálculo de Risco de Incêndios

Métodos para o zoneamento do risco de incêndios florestais podem ser obtidos pela sobreposição de mapas temáticos, onde cada um representa um fator de risco e há uma interação ponderada entre esses. Esses fatores condicionam um risco para início e propagação

do incêndio, podendo ser compostos de fatores físicos, biológicos, socioeconômicos e climáticos. Um Sistema de Informação Geográfica (SIG) é usado como ferramenta para esse método (FERRAZ; VETTORAZZI, 2003).

Podem-se citar outros métodos para calcular o risco de incêndio. Métodos que se baseiam em fatores meteorológicos, utilizando variáveis como temperatura, umidade, precipitação e intensidade do vento. Métodos que se baseiam no risco potencial do fogo onde são considerados o estado da vegetação, tipo de combustível e o seu conteúdo de umidade. Método, como o desenvolvido pelo Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), que se baseia na estimativa da umidade da vegetação a partir de imagens de sensores orbitais (CARAPIÁ, 2006). Além disso, existem métodos estatísticos como a regressão logística; métodos quantitativos e qualitativos como a análise multicritério, métodos de inteligência computacional utilizando redes neurais artificiais e lógica fuzzy, além de fórmulas que relacionam os parâmetros meteorológicos com o risco de incêndio, como a fórmula de Monte Alegre. A seguir são detalhados alguns desses métodos e aplicações. Uma análise crítica destas metodologias pode ser encontrada em Fernandes et al. (2018).

1.2.1. Análise Multicritério

Método desenvolvido por Saaty (1977) chamado Método de Análise Hierárquica (em inglês – *Analytic Hierarquy Process - AHP*) onde é estruturada uma hierarquia de decisão para classificar os elementos de um nível mais alto para um mais baixo, depois é feita uma matriz de comparação para definir os pesos de cada variável. A matriz de comparação registra a importância percebida, relativa de cada critério com cada outro critério, permite que os tomadores de decisão obtenham um conjunto de relação em escala de pesos para os critérios individuais. Por fim, uma relação de consistência (RC) é calculada para determinar se a avaliação foi bem-sucedida ou não.

No caso do cálculo do risco de incêndios são selecionadas variáveis que influenciam a ignição do fogo, por meio de mapas temáticos como no trabalho de Jaiswal et al. (2002) ou por dados de sensoriamento remoto, como no estudo de Chuvieco e Congalton (1989). No Quadro 1 são mostradas algumas variáveis apontadas e analisadas pelos autores supramencionados.

Quadro 1 - Comparação entre diversos trabalhos da literatura que aplicam a metodologia

(Continua)

	<i>Jaiswal et al.(2002)</i>			<i>Chuvienco et al (1989)</i>			<i>Ribeiro et al. (2008)</i>		
<i>Declividade</i>	0-4%	Baixo	2	0-3%	Baixo	2	Até 15 %	Baixo	1
	5-8%	Baixo	2	3-5%	Moderado	3	16-25 %	Moderado	2
	9-12%	Baixo	2	5-10%	Moderado	4	26-35 %	Alto	3
	13-16%	Médio	1	10-15%	Alto	5	36-45%	Muito alto	4
	17-20%	Médio	1	15-35%	Muito alto	6	>=46%	Extremo	5
	21-36%	Médio	1	>35%	Muito alto	10			
	27-40%	Médio	1						
	41-44%	Alto	0						
	>44%	Alto	0						
<i>Proximidade a estradas</i>	Buffer (< 150 m de qualquer estrada vicinais ou < 50 m de qualquer autoestrada)	Alto	0	< 100m	Muito alto	8	Sob influência (< 100 m de qualquer estrada vicinais ou < 50 m de qualquer autoestrada)	Alto	1
	Fora do buffer	Baixo	1	100-200 m	Alto	7	Sem influência	Nulo	0
				200-300 m	Moderado	5			
				300-400 m	Baixo	3			
<i>Uso do Solo</i>	Pinheiro denso	Alto	0	Mistura seca	Muito alto	10	Lago/banhado/ áreas de circulação	Baixo	1
	Pinheiro médio	Alto	0	Mistura de bambu	Muito alto	9	Floresta	Moderado	2
	Pinheiro esparsos + arbusto	Médio	1	Sal e Bambu	Alto	8	Estradas	Alto	3
	Arbusto denso	Médio	1	Bambu degradado	Alto	6	Reflorestamento	Muito alto	4
	Arbusto médio	Médio	1	Sal*	Alto	6	Campo suo/agricultura	Extremo	5
	Arbusto esparsos	Baixo	2	Plantação	Moderado	4			
	Amendoieiras	Baixo	2	Mistura úmida	Moderado	3			
	Vinhedos	Baixo	2	Espaço em branco	Baixo	1			
	Laranjeiras	Baixo	2	Agricultura	Baixo	2			
				Terra seca	Baixo	1			
			Agricultura com assentamento	Baixo	1				

Quadro 1 - Comparação entre diversos trabalhos da literatura que aplicam a metodologia

(Continuação)

	<i>Jaiswal et al.(2002)</i>			<i>Chuvieco et al (1989)</i>			<i>Ribeiro et al. (2008)</i>		
Aspecto	Sudeste	Alto	0				Sudeste/ sul/sudoeste	Baixo	1
	Sudoeste	Médio	1				Leste	Moderado	2
	Norte	Baixo	2				Nordeste	Alto	3
							Noroeste/oeste	Muito alto	4
							Norte	Extremo	5
Elevação	0-3 m	Baixo	1	Até 600 m	Extremo	5			
	3-6 m	Baixo	1	600-900 m	Muito alto	4			
	398- 400 m	Alto	0	900-1200 m	Alto	3			
	401- 404 m	Alto	0	1200-1500 m	Moderado	2			
	405 - 407 m	Alto	0	Acima de 1500 m	Baixo	1			
Habitação		---		Área de assentamento	Muito alto	8			
				< 1000 m	Alto	7			
				1000-2000 m	Moderado	5			
				2000-3000 m	Baixo	2			
Hidrografia							Sob influência (< 50 m)	Alto	1
							Sem influência	Nulo	0

1.2.1. Lógica Fuzzy

A metodologia da lógica *fuzzy* ou lógica difusa, introduzida por Zadeh (1965), é um problema de imprecisão associado com conjuntos de dados onde não há uma transição definida de uma classe a outra, ou seja, eliminando os limites entre elas. O grau da associação de um objeto varia entre 0 e 1, onde quanto mais próximo de 1 maior é a associação e quanto mais próximo de 0 mais fraca é a associação. As classes contínuas dos conjuntos *fuzzy* (ou difusos) são construídas a partir de conceitos definidos *a priori* usando a experiência, ciência ou conhecimento heurístico. O conhecimento linguístico de um fenômeno complexo é usado, transformando-o em dados numéricos para processamento.

Na construção de um conjunto *fuzzy* deve-se definir uma função de pertinência, que pode ter várias formas, as funções mais comuns são: triangular, trapezoidal e em forma de sino (função normal). Além disso, é estabelecido um conjunto de regras de acordo com o conhecimento do especialista, na forma de proposições do tipo “se E1 e/ou E2 então S1”, por exemplo, uma aplicação prática seria: “Se a umidade é baixa e a precipitação é baixa então o risco é alto”.

1.2.2. Regressão Logística

A regressão logística é a mais popular modelagem matemática que relaciona as variáveis independentes com a variável dependente dicotômica. Dessa forma, a variável dependente seria o risco de fogo, que pode ser 0 (não fogo) ou 1(fogo), sendo representado pela adaptação da equação de Bernoulli apresentada na Equação 1 (GOLDARAG et al., 2016).

$$P(Y = 1) = \frac{e^{\sum(\alpha + \beta_i X_i)}}{1 + e^{\sum(\alpha + \beta_i X_i)}} \quad (1)$$

Onde Y é a variável dependente, X_i são as variáveis independentes e α e β_i , os coeficientes da equação, sendo números reais. A partir de dados históricos conhecidos de ocorrência de incêndio é possível conhecer as relações entre as variáveis a partir do cálculo dos coeficientes, que são assim substituídos na Equação 1.

1.2.3. Fórmula de Monte Alegre

Desenvolvida por SOARES (1972) a partir de sete anos (1965 a 1971) de dados meteorológicos e da ocorrência de incêndios da Fazenda Monte Alegre, no município de Telêmaco Borba-PR. Esta fórmula foi amplamente utilizada por empresas e instituições florestais para estimar o grau de risco de incêndios. Consiste em uma fórmula simples que requer apenas duas variáveis meteorológicas: umidade relativa do ar e precipitação. É um índice acumulativo que requer valores diários de umidade relativa do ar e a quantidade de precipitação. Sua fórmula básica é apresentada na Equação 2.

$$FMA = \sum_{n=1}^n \frac{100}{H} \quad (2)$$

Onde H é a umidade relativa do ar medida às 13:00 h e n é o número de dias sem chuva. Apresenta algumas restrições referentes à precipitação diária, como mostrado no Quadro 2. O perigo de incêndio do dia é indicado através de uma escala de cinco níveis: Nulo

(FMA \leq 1,0); Pequeno (1,1 \leq FMA \leq 3,0); Médio (3,1 \leq FMA \leq 8,0); Alto (8,1 \leq FMA \leq 20,0) e Muito Alto (FMA $>$ 20,0).

Quadro 2 - Modificação do cálculo conforme a precipitação ocorrida no dia (NASCIMENTO; AMARAL; PONTES, 2013)

Chuva do dia (mm)	Modificação no cálculo
$\leq 2,4$	Nenhuma
2,5 a 4,9	Abater 30% na FMA acumulada e somar (100/H) do dia
5,0 a 9,9	Abater 60% na FMA acumulada e somar (100/H) do dia
10,0 a 12,9	Abater 80% na FMA acumulada e somar (100/H) do dia
$\geq 12,9$	Interromper o cálculo e recomeçar no dia seguinte

1.2.4. FIRA e FARSITE

Mbow, Goïta, e Béné (2004) criaram um algoritmo para o cálculo do risco de incêndios (FIRA- Fire Risk Assessment). Para tanto, utilizaram índices de temperatura de brilho e umidade que foram normalizados e combinados em um algoritmo semi-empírico. Tais índices foram utilizados devido a estes refletirem a umidade do combustível e sua distribuição, aspectos importantes na propagação de incêndios em áreas de savana. Utilizando em paralelo um simulador de área de fogo (FARSITE – Fire Area Simulator) foi desenvolvido pelo departamento de agricultura dos Estados Unidos (USDA). O simulador consiste em um modelo de crescimento do fogo que projeta seu perímetro e comportamento em ambientes complexos. O modelo usa dados espaciais de SIG em conjunto com dados de clima e vento. Os algoritmos estão representados na Figura 1.

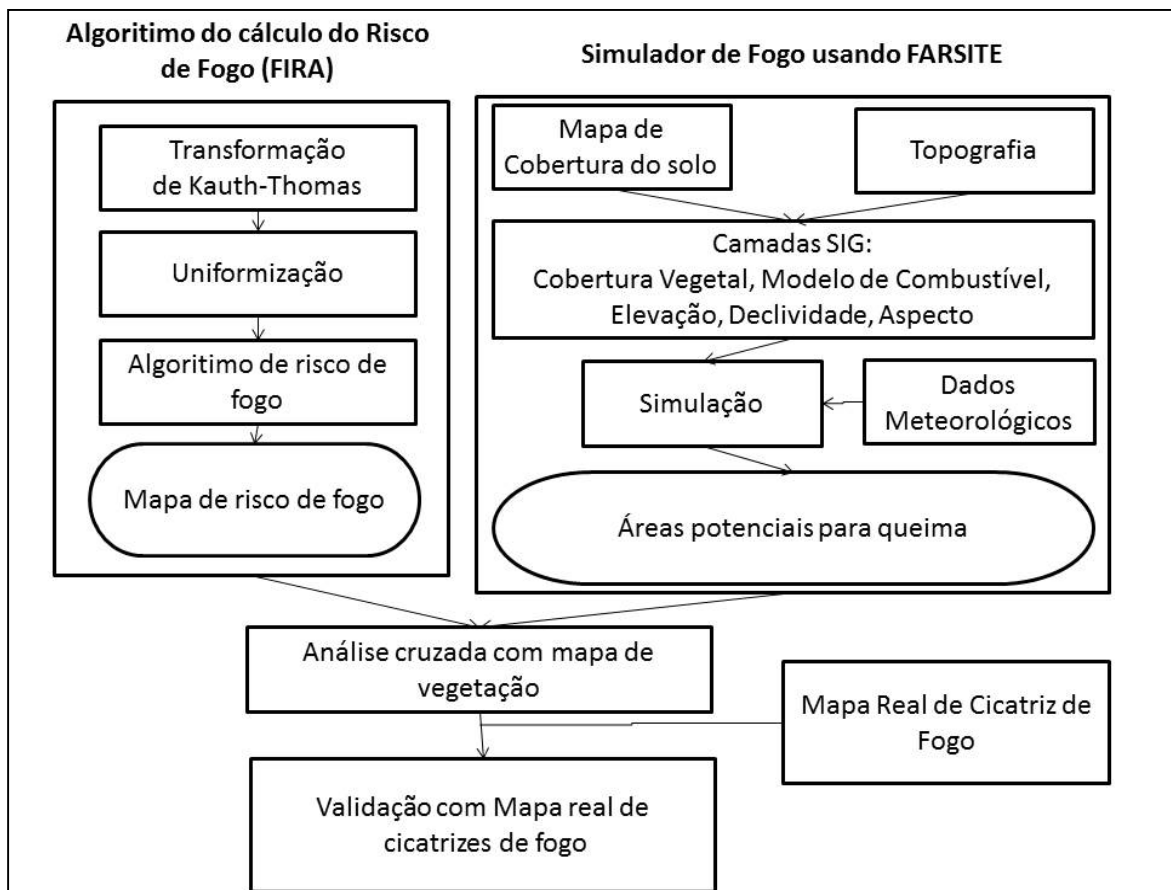


Figura 1- Fluxograma dos dois algoritmos (adaptado de MBOW; GOÏTA; BÉNIÉ, 2004)

1.2.1. Redes Neurais Artificiais

Uma rede neural artificial (RNA) é constituída de unidades de processamento simples que trabalham em paralelo, armazenando conhecimento experimental e tornando-o disponível para uso em modelagens do conhecimento adquirido. A técnica computacional procura imitar o comportamento do cérebro humano em dois aspectos: conhecimento adquirido e força de conexão entre neurônios. A aprendizagem da rede é o processo em que o conhecimento é adquirido e armazenado nos pesos sinápticos (conexões entre os neurônios) (HAYKIN, 2001). Ela difere de sistemas de especialistas, os quais oferecem ao usuário uma resposta definitiva, caso as características avaliadas corresponderem àquelas codificadas nas bases das regras. Uma RNA realiza uma análise das informações e fornece uma estimativa dependente da compatibilidade do dado com as características que foram treinadas, podendo chegar a uma acurácia de 100% de acordo com a experiência adquirida pela rede (CANNADY, 1998). Um maior detalhamento do método é feito na seção 1.5.

1.3. Agentes causadores de incêndio

Para o cálculo do risco de queimadas existem dois tipos de variáveis que podem ser consideradas: aquelas derivadas de fatores que não mudam em um curto período de tempo, chamadas de estáticas, e aquelas que têm grandes variações em um período curto de tempo, chamadas de dinâmicas. As primeiras englobam: declividade, tipos de vegetação, tipos de solo, proximidade com estradas e assentamentos. Já as variáveis dinâmicas podem-se citar variáveis meteorológicas como: temperatura, direção e intensidade do vento, umidade do ar, entre outras (LÓPEZ; SAN-MIGUEL-AYANZ; BURGAN, 2002).

A FAO (Organização das Nações Unidas para Agricultura e a Alimentação) agrupa as causas dos incêndios em oito categorias ou grupos (JUVANHOL, 2014):

- Raios: representam as descargas elétricas da atmosfera;
- Incendiários: incêndios que são provocados propositalmente por pessoas em propriedades alheias por vários motivos como retaliação, vingança, disputa pela posse de terras, urbanização e incapacidade mental entre outros;
- Queima para limpeza: originados pela negligência ou descuido do uso do fogo para a limpeza do terreno para fins florestais, agrícolas ou pecuários que escapam do controle;
- Fumantes: displicência e falta de precaução de fumantes que jogam ponta de cigarros ou fósforos acesos em locais de risco;
- Operações florestais: provocado por trabalhadores florestais que estão em atividade na floresta;
- Estrada de ferro: incêndios causados por atividades ferroviárias;
- Recreação: incêndios provocados por visitantes que utilizam a floresta como recreação;
- Diversos: causas desconhecidas ou que não são enquadradas em nenhum grupo.

Desses apenas um grupo não apresenta causa humana, os raios, mostrando assim a importância em se adicionar a influência humana nas análises. Em uma pesquisa Santos, Soares e Batista (2006) contabilizaram que 63% dos incêndios florestais ocorridos entre 1998 a 2002 foram por causa de incendiários.

Algumas formas de contabilizar essa influência é utilizar dados de densidade populacional, raios ou áreas de influência de estradas e assentamentos. Já os movimentos de atividades humanas, e veiculares nas estradas oferecem amplas oportunidades para incêndios acidentais/artificiais, pois estão sujeitas a acidentes e incêndios de veículos, pontas de cigarros em brasa jogada por motoristas e fogueiras de transeuntes, além de incêndios criminosos (JAISWAL et al., 2002).

Outras variáveis importantes são explicadas a seguir.

1.3.1. Topografia

A elevação do terreno está relacionada com o comportamento dos ventos e a propensão a incêndios. Para sua avaliação é calculada a taxa de mudança de elevação, podendo ser em graus ou porcentagem, definida como declividade do terreno (VADREVU; EATURU; BADARINATH, 2010). Os incêndios se propagam mais rápido em áreas de maior declividade e mais devagar em áreas de menor declividade. Sendo assim, a inclinação do terreno afeta tanto a taxa quanto a direção da propagação do fogo (ROTHERMEL, 1983). Muitos estudos apontam que áreas de maior declividade estão associadas a maior risco de incêndios (PEZZOPANE; NETO; VILELA, 2001).

Terrenos íngremes aumentam a taxa de propagação do fogo por causa de um sistema de pré-ignição por convecção eficiente e a ignição por ponto de contato. Mas a inclinação também influencia na supressão do fogo, pois afeta a fadiga e o material de rolamento (CHUVIECO; CONGALTON, 1989). Estudos reportam que vegetações em encostas suaves tendem a queimar mais completamente do que em locais com cursos d'água e penhascos, pois interrompe o movimento do fogo para frente (MINNINCH, 1978).

Outro fator importante é a orientação do terreno. É importante saber as faces que tem influência do calor direto do sol, pois esta afeta a taxa de secagem do material combustível. No hemisfério Sul, os raios solares incidem mais diretamente sobre faces voltadas ao Norte, sendo a face Oeste a segunda a receber mais energia, seguida da Leste, sendo que a que

menos se aquece é a Sul. Isso depende das latitudes, já que em latitudes maiores a face Sul não recebe radiação direta nem no verão (JUVANHOL, 2014).

1.3.2. Uso e Ocupação do Solo

O estudo do uso e ocupação do solo é importante para caracterização da cobertura do solo e suas respectivas localizações. Assim, é possível conhecer a distribuição e quantificação das áreas de vegetação, pastagem, agricultura, área urbana, entre outros (PRUDENTE, 2010).

Saber o tipo de vegetação é importante, pois afeta o comportamento do fogo, devido aos tipos e propriedades do combustível e condições climáticas associadas. Por exemplo, uma floresta densa e fechada proporciona um microclima mais ameno e os combustíveis possuem conteúdo de umidade maior e mais estável em comparação com floresta rala e aberta. Além disso, esses elementos interceptam a radiação solar, reduzindo a temperatura. Assim, o fogo se propaga com mais facilidade em florestas abertas. Entretanto, as florestas plantadas também apresentam condições mais favoráveis aos incêndios devido a sua homogeneidade (JUVANHOL, 2014).

1.3.3. Hidrografia

A importância em se utilizar a proximidade dos cursos d'água no risco de incêndios varia de região a região, dependendo das dimensões e do uso que se faz dos recursos. Já que pode ser um fator facilitador de incêndios devido à melhoria do acesso às áreas de vegetação, ou devido à presença de mata ciliar ao redor da hidrografia. Seguramente, nem todas as margens de rios podem significar algum tipo de risco de incêndios florestais. No entanto, para definir os riscos exatos de cada trecho de margem de rio de um município é necessário tempo e exclusividade no estudo (RIBEIRO; SOARES; BEPLER, 2012).

1.3.4. Fatores Climáticos

A ocorrência de incêndios está diretamente ligada às condições climáticas, que proporcionam a ignição, ajuda na propagação e no comportamento do fogo. Alguns parâmetros importantes que influenciam nos riscos são:

- *Temperatura:* A temperatura representa a energia interna de movimento nos átomos e moléculas, essa energia calorífica é importante na fase evaporativa da combustão (HAMADEH et al., 2017). Temperaturas mais altas fazem que os combustíveis fiquem mais suscetíveis a incêndios, já que os predispõem à ignição, além de diminuir a umidade do material. Quando a temperatura aumenta, o vapor de ar saturado aumenta, a umidade relativa diminui, afetando o teor de umidade do combustível fogo (LIU; ZHANG, 2015).
- *Umidade Relativa:* Representa a relação entre a quantidade de água existente no ar e a quantidade máxima que poderia haver na mesma temperatura e pressão (HAMADEH et al, 2017). Quanto maior a umidade relativa, mais lenta é a evaporação e maior é a absorção de umidade pelos materiais. Enquanto quando há a diminuição na absorção de umidade, aumenta a evaporação e a umidade do material combustível é reduzida. Em geral, quando a umidade está maior que 75% não ocorrem incêndios, quando está entre 55% e 75%, tem um risco baixo, com 30% a 55% o risco de incêndio é alto e quando é menor que 30% o risco é de incêndio muito alto (LIU; ZHANG, 2015).
- *Precipitação:* Afeta diretamente a quantidade de água no material combustível. Se a precipitação anual é maior que 1500 mm em uma área e bem distribuída a chance de se ter incêndios florestais é bem pequena. Se a distribuição é desigual, com períodos acentuados de seca, a chance de ocorrer incêndios nessa época é grande. No entanto, se a precipitação anual é menor que 1000 mm por ano, a chance de incêndios é maior. Entretanto, quando a precipitação mensal é maior que 100 mm o risco de incêndio é pequeno, já que tem uma boa distribuição da chuva, aumentando a quantidade de água no material combustível (LIU; ZHANG, 2015).
- *Velocidade do Vento:* É o movimento do ar em uma distância horizontal por unidade de tempo. Quando o vento envolve com o combustível pode acelerar a evaporação da água em matérias combustíveis, tornando a secagem rápida e inflamável. Ele também aumenta o fornecimento de oxigênio aumentando condições de combustão. Além de ajudar na distribuição e direção do fogo (LIU; ZHANG, 2015).

- *Pressão Atmosférica:* As condições climáticas são relacionadas com a pressão atmosférica. Em geral em lugares sob alta pressão há sol, alta temperatura, baixa umidade, deixando o clima susceptível a incêndios. Enquanto a baixa pressão ajuda na formação de nuvens e precipitação (LIU; ZHANG, 2015).

1.4. Uso de dados de satélite para detecção de incêndios

A luz e calor liberados na combustão modificam o espectro de energia refletida e emitida na área afetada pela queima, o que possibilita sensores de satélites orbitais a detecção de focos de incêndios, quantificação de áreas queimadas e a distribuição de nuvens de fumaça (COCHRANE, 2003). Produtos de satélites são a única fonte de informação global de ocorrência de incêndios, uma vez que fornecem uma cobertura espacial e temporal abrangente das áreas afetadas pelo fogo. Dois tipos de produtos de satélites são utilizados para mapeamento da ocorrência de incêndios: pontos de fogo ativo e áreas queimadas. Os pontos de fogo ativo são derivados de anomalias na temperatura registradas pelos sensores, enquanto áreas queimadas são detectadas pela diferença na reflexão de áreas queimadas e não-queimadas (HANTSON et al., 2013).

A detecção de incêndios por satélites começou após 1978 com o lançamento do primeiro sensor *AVHRR*, onde foram desenvolvidos algoritmos que medem a temperatura da superfície radiante em resoluções menores que a do pixel. A partir de 1990 também foram lançados outros satélites de melhor resolução como o *ATSR* e o *GOES*. Em 1999 e 2002 foram lançados os Satélites *TERRA* e *AQUA*, respectivamente, que trouxeram os sensores *MODIS* com uma resolução de 1 km. Estes representaram um grande avanço na detecção de focos de incêndio, já que combinados podem observar focos de incêndio globalmente em uma média de passagem de 4 vezes por dia, compreendendo dia e noite. Além destes, outros sensores foram lançados, como o *Visible Infrared Imager Radiometer Suite (VIIRS)* a bordo do satélite *S-NPP* em 2011 (ICHOKU; KAHN; CHIN, 2012). No Quadro 3 estão listados os principais sensores e satélites utilizados para este fim.

Quadro 3 - Satélites que medem a localização de incêndios (Adaptado de ICHOKU; KAHN; CHIN, 2012)

Sensor (Satélite)	Resolução	Cobertura	Período
OLS (DMSP)	2.7 x 2.7 km	Global	1979 - 1992
AVHRR (Tiros-N/NOAA-6_NOAA-19)	1 x 1 km	Global	1992 - presente
VIRRS (TRMM)	0,75 x 0,75 km 0,375 x 0,375 km	40° Norte – 40° Sul	1997 - presente
ASTER (TERRA)	0,03 x 0,03 km	Global	2000 - presente
HSRS (BIRD)	0,37 x 0,37 km	Global	2000 - presente
IMG (GOES)	4 x 4 km	América do Sul e do Norte	2000 - presente
SEVIRI (MSG)	3 x 3 km	África e Europa	2004 - presente
MODIS (TERRA & AQUA)	1 x 1 km	Global	2000 - presente

Uma das maiores críticas em relação ao uso de produtos de satélites é a presença de erros de omissão e comissão. Os erros de omissão estão relacionados com o tamanho e a magnitude do incêndio, que podem ser menores que os detectados pelo satélite; a presença de nuvens e fumaça; o relevo e as manchas de vegetação grandes, que camuflam os incêndios. Erros de comissão podem acontecer devido ao aumento de sinal nos canais de incêndio pelo brilho do sol, às superfícies quentes e a certos tipos de nuvens. A presença de nuvens tem sido um problema persistente em produtos de incêndio, mas Schroeder; Csiszar e Morisette (2008) investigaram esse problema detalhadamente na Amazônia brasileira e verificaram que a quantidade de focos perdidos estava entre 1,5% e 5%, em escala anual e mensal respectivamente (ICHOKU; KAHN; CHIN, 2012).

Os sensores *MODIS* e *VIIRS* utilizam o mesmo algoritmo para detecção dos focos de incêndios, o qual é baseado em um ganho duplo da temperatura de saturação (634 K) no canal do infravermelho-médio (comprimento de onda de 4 μm) e um ganho simples no canal do infravermelho termal (comprimento de onda de 11 μm) (CSISZAR et al., 2014).

O *S-NPP/VIIRS* tem uma órbita quase polar, passando pelo equador 1:30 e 13:30, tendo registros nas resoluções de 375 m e 750 m. O produto *VIIRS* 375 m possui cinco bandas. A banda I4 (comprimento de onda entre 3,55 a 3,93 μm) é a utilizada no algoritmo para detecção de incêndios, pois abrange os comprimentos de onda do pico de radiação espectral para corpos negros emitido em temperaturas entre 737 e 817 K. Sendo também possível distinguir pixels de combustão de pontos frios. É um canal de ganho único com

temperatura de saturação (temperatura máxima de brilho de saída para um canal infravermelho) de 367 K. Para complementar a banda I5 (comprimento de onda entre 10,5 a 12,4 μm) é comparada com a I4 para separar fogos ativos de pontos livres de fogo, sendo também uma banda de ganho único com uma temperatura de saturação de 380 K. As bandas I1, I2 e I3 são utilizadas para diferenciar nuvens, luz solar e corpos d'água (SCHROEDER et al., 2014).

Este produto melhorou a resolução espacial na detecção de incêndios, eliminando lacunas em baixas latitudes, aumentando a detecção de focos de incêndio menores e melhorando o monitoramento de incêndios que duram dias. Schroeder et al. (2014) fizeram uma comparação com o produto *LANDSAT-8* que possui 30 m de resolução e observaram concordância na detecção de incêndios entre os dois produtos. A área mínima em que os focos são detectados são em torno de 10 m^2 com uma média de 850 K (576,85° C) de temperatura (SCHROEDER et al., 2014). Oliva e Schroeder (2015) fizeram um estudo sobre o desempenho do produto para detecção de incêndios e verificaram que no Brasil a detecção de incêndios menores (< 10 ha) é mais baixa, o que pode ser explicado devido a ocorrência de incêndios que se alastram rápido e tem menor período de queima, sendo extinguido entre as passagens do satélite. Enquanto para incêndios maiores a omissão na detecção é bem menor, sendo quase nula.

O sensor *MODIS* por sua vez tem 1 km de resolução, mas uma frequência de passagem maior, o satélite *TERRA* passa no Equador as 10:30 e 22:30 e o satélite *AQUA* entre 1:30 e 13:30, conseguindo detectar incêndios maiores que 50 m^2 (HANTSON et al., 2013). Este produto fornece dados da localização do pixel de incêndio, o dia e hora da detecção e sua Energia Radioativa do Fogo (*FRP - Fire Radioactive Power*). O FRP representa a quantidade medida do calor radiante emitido, usado para prever a intensidade do fogo, sendo proporcional à taxa de combustão e emissões de fumaça (VADREVU; LASKO, 2018). O algoritmo desse produto como mencionado, é o mesmo utilizado pelo VIIRS e é demonstrado no fluxograma da Figura 2.

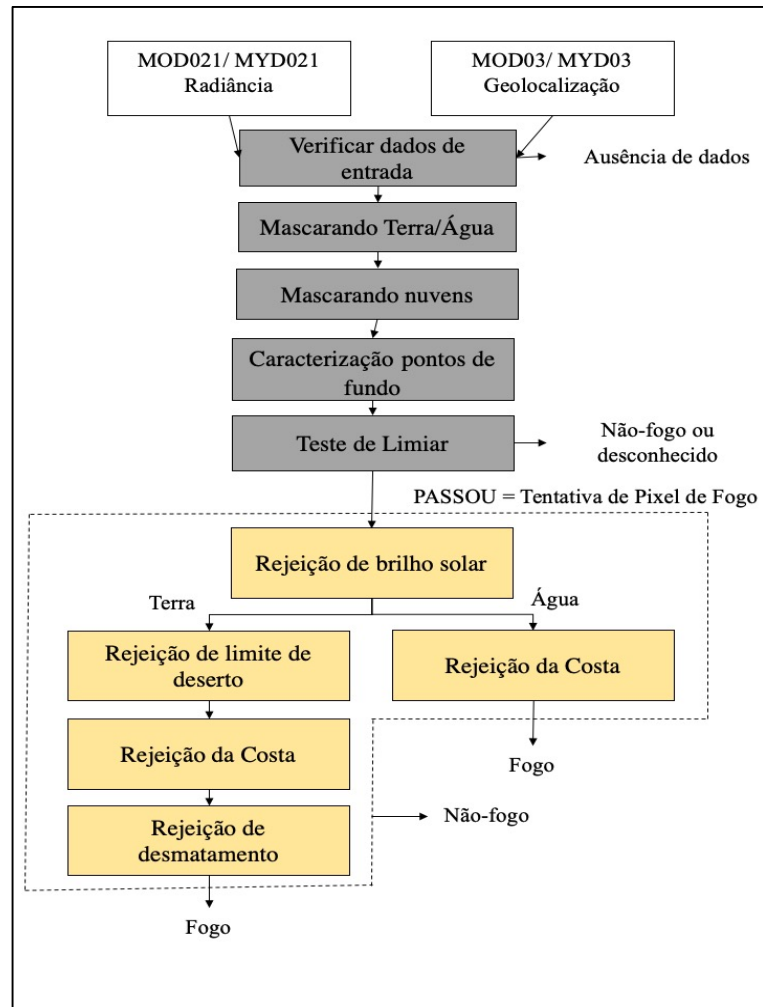


Figura 2 - Funcionamento do Algoritmo MODIS Collection 6 (Adaptado de GIGLIO; SCHROEDER; JUSTICE, 2016)

Hantson et al.(2013) fizeram um estudo com dados de área queimada do satélite de referência LANDSAT para os biomas mais importantes no planeta. Verificaram que em geral os erros de comissão eram pequenos, sendo sua maioria em áreas urbanas. Já os erros de omissão estavam relacionados com o tamanho dos incêndios, sendo que para o Brasil em áreas pequenas (<50 h) os erros de omissão chegaram em até 70%.

Chen et al. (2017) usaram dados MODIS para verificar regimes de fogo na China, e nos seus resultados concluíram que sua análise estava de acordo com a literatura, sendo que os dados utilizados foram eficientes para o método. Bedia et al. (2015) utilizaram dados do MODIS para determinar estações de incêndios em um estudo de padrões globais de incêndios. Reddy et al. (2017) utiliza dados do MODIS para o monitoramento de focos de incêndio em áreas protegidas na Índia. Goldarag; Mohammadzadeh e Ardakani (2016) utilizaram dados MODIS para produzir um modelo de risco de incêndios utilizando redes neurais artificiais.

Assim, dados de fogo ativo têm sido utilizados por muitos estudos em diversas aplicações, devido a disponibilidade e qualidade dos dados.

1.5. Redes Neurais Artificiais (RNA)

“Uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso. Ela se assemelha ao cérebro humano em dois aspectos:

O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem

Forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido”.

(HAYKIN, 2001, p.28)

Alguns benefícios da RNA são: a possibilidade de não linearidade, o fato de poder se mapear a entrada e saída através da aprendizagem supervisionada e o fato da rede se adaptar de acordo com as mudanças no meio ambiente (HAYKIN, 2001).

O cérebro humano continuamente recebe informação, a percebe e toma decisões. O sistema nervoso controla as reações rápidas do corpo. O neurônio é a célula responsável pelas atividades cerebrais. Ele é constituído por um corpo celular e por ramificações, os dendritos, e o axônio. Os impulsos elétricos são recebidos pelos dendritos, a informação é processada pelo corpo celular, onde um limiar de ação é definido. O sinal é então propagado para outro neurônio através do axônio, sendo este processo chamado de sinapse. A Figura 3 mostra a comparação entre os neurônios do corpo humano e o da rede neural artificial (ROMERO, 2017).

Na RNA os neurônios são interpretados como unidades de processamento, estes junto com suas interligações ajudam a resolver problemas complexos. O primeiro neurônio artificial, proposto por McCulloch e Pitts (1943), conhecido por MCP, é composto apenas pelas entradas, um somatório ponderado das entradas e uma função de ativação. Sendo que cada entrada x_i é multiplicada pelo peso w_i (conforme Figura 3) (ROMERO, 2017).

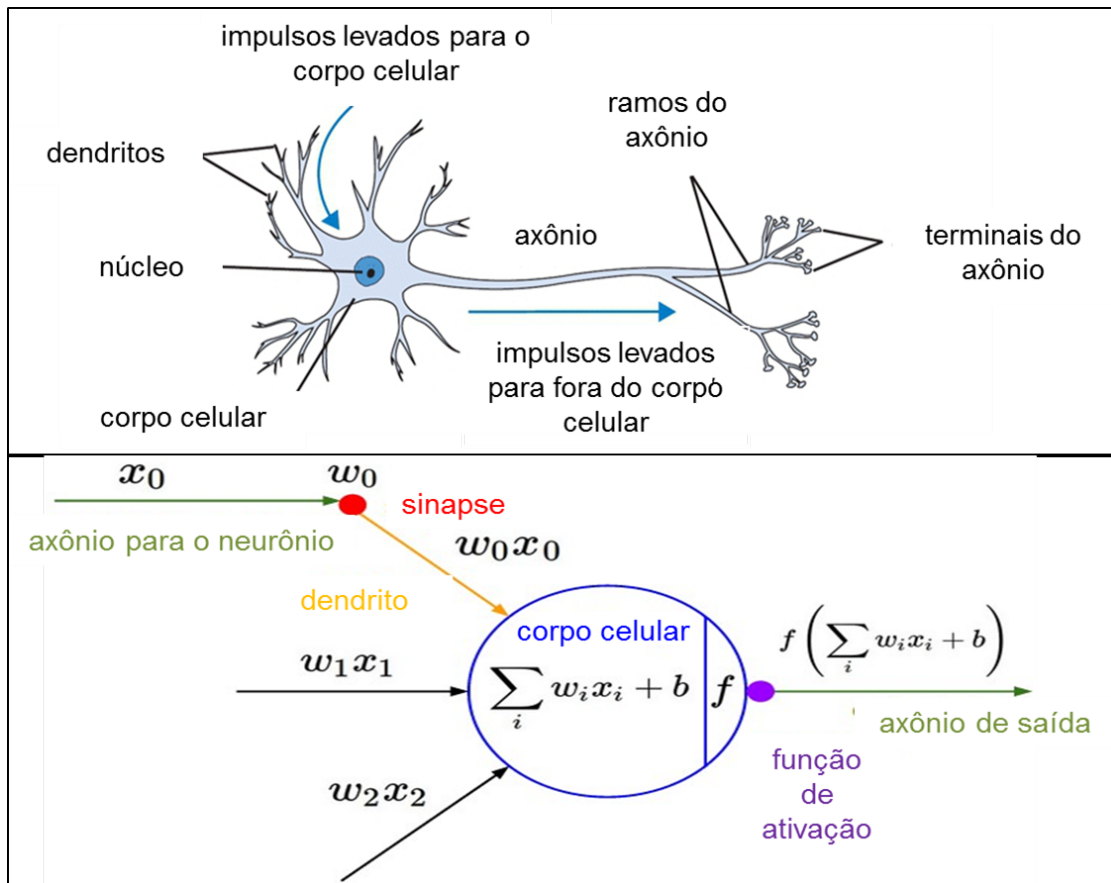


Figura 3 – Modelo de um neurônio biológico e modelo de um neurônio artificial (adaptado de ROMERO, 2017).

1.5.1. Modelo de um neurônio

Um neurônio é composto de três elementos básicos (HAYKIN, 2001):

- Conjunto de sinapses ou elos de conexões, representados pela entrada (x_i) e pelo peso sináptico (w_i)
- Uma unidade que soma os sinais de entrada, ponderados pelas respectivas sinapses do neurônio, constituindo uma combinação linear.
- Uma função de ativação que restringe a amplitude do sinal na saída do neurônio em função de seus argumentos.

Há também o bias ou viés (b_i), que é aplicado externamente com o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se for positivo ou negativo.

Um neurônio pode ser representado matematicamente pelas Equações 3 e 4. Sendo x_i os sinais de entrada, w_{ij} os pesos sinápticos do neurônio i da camada j ; u_i representa o somatório que junto com b_i denominado bias (viés) são argumentos da função de ativação $f(\cdot)$.

e y_i é o sinal de saída do neurônio. O bias permite calibrar o limiar de ativação da função (ROMERO, 2017). Uma representação pode ser observada na Figura 4.

$$u_i = \sum_{j=1}^m w_{ij}x_j \quad (3)$$

$$y_i = f(u_i + b_i) \quad (4)$$

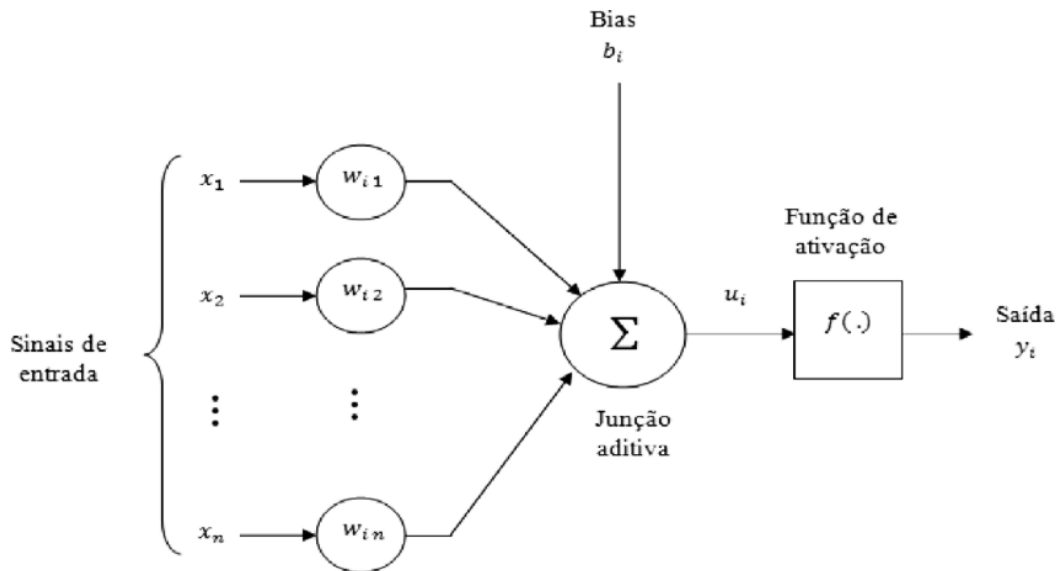


Figura 4- Modelo de um neurônio (COUTINHO; SILVA; DELGADO, 2016)

A função de ativação é importante, pois restringe a amplitude do sinal na saída do neurônio em função dos argumentos, e pode definir a linearidade ou não da saída da rede. Na literatura há vários tipos de funções encontradas, os três tipos básicos são (HAYKIN, 2001):

- *Função de Limiar (função degrau)*: neste modelo um neurônio assume o valor 1, se o campo local induzido daquele neurônio é não negativo, e 0 caso contrário. Representada na Equação 5 e Figura 5.

$$f(v) = \begin{cases} 1 & \text{se } v \geq 0 \\ 0 & \text{se } v < 0 \end{cases} \quad (5)$$

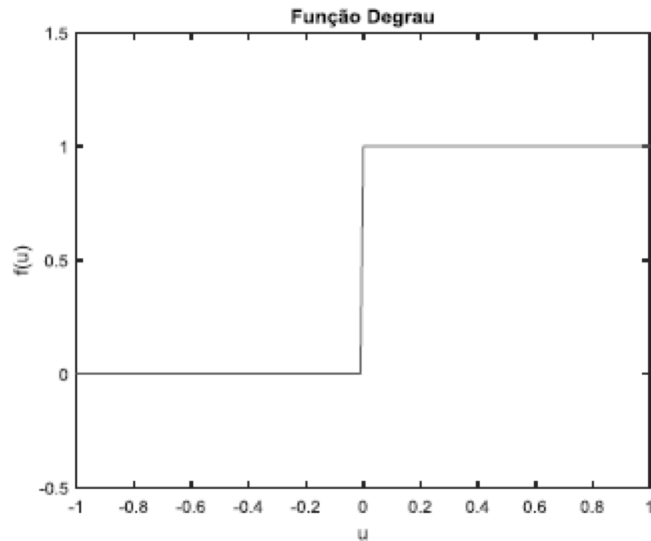


Figura 5 - Função degrau (NICOLAU, 2016)

- *Função Linear*: Representado pela Equação 6 e Figura 6.

$$f(v)=nv \tag{6}$$

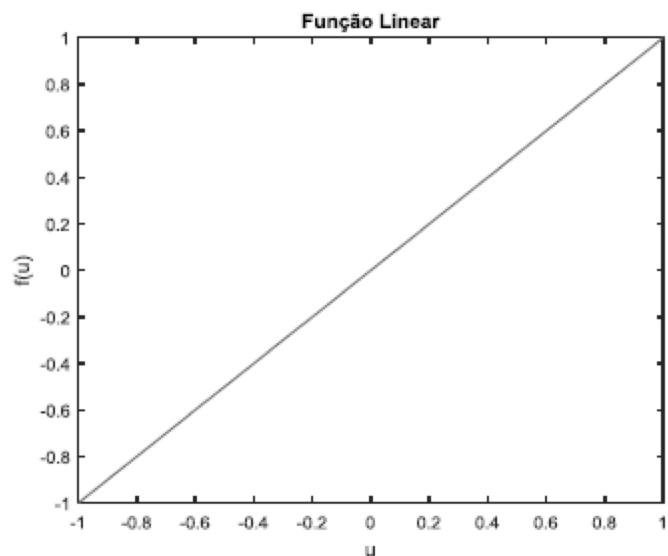


Figura 6 - Função Linear (NICOLAU, 2016)

- *Função semi-linear*: É uma aproximação de um amplificador não-linear. Onde se assume que o fator de amplificação dentro da região linear de operação é a unidade. Dado pela função da Equação 7 e representado pela Figura 7.

$$f(v) = \begin{cases} 1 & \text{se } v \geq 1 \\ pv & \text{se } 0 < pv < 1 \\ 0 & \text{se } pv \leq 0 \end{cases} \quad (7)$$

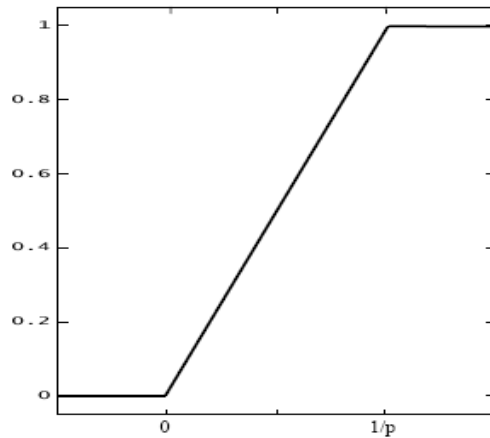


Figura 7 - Função semi-linear (NICOLAU, 2016)

- Função Sigmoide: função estritamente crescente que exibe um balanceamento adequado entre comportamento linear e não linear. Uma função muito usada é a função logística (Equação 8) e a tangente hiperbólica (Equação 9).

$$f(v) = \frac{1}{1 + \exp(-\beta v)} \quad (8)$$

$$f(v) = \frac{1 - \exp(-\beta v)}{1 + \exp(-\beta v)} \quad (9)$$

Onde β é o parâmetro de inclinação da função sigmoide, variando esse parâmetro obtêm-se funções sigmoides com diferentes inclinações. Estas funções são representadas na Figura 8.

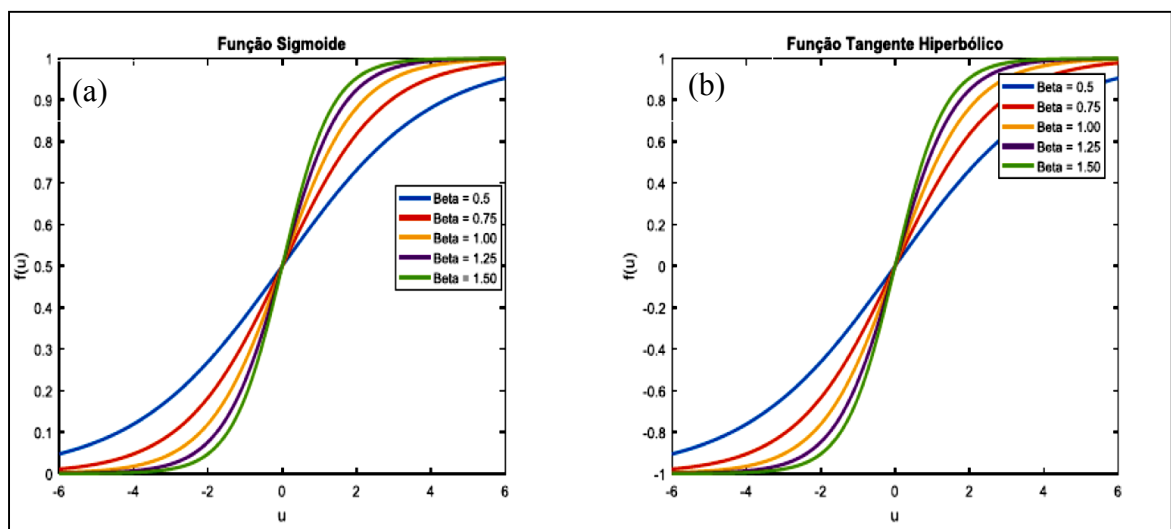


Figura 8- Função Logística (a) e função tangente hiperbólica (b) (NICOLAU, 2016)

1.5.2. Arquitetura da rede

Para utilização de uma rede neural é necessário definir sua arquitetura, ou seja, como está estruturado o número de entradas, de n camadas e de nós ou neurônios, e também o número de nós de saída. Basicamente, uma rede neural possui uma camada de entrada e uma de saída, podendo ter camadas intermediárias ou ocultas, sendo que a quantidade de neurônios depende do problema a ser aplicado. A Figura 9 mostra os três principais tipos de arquiteturas definidos em Haykin (2001):

- **Redes alimentadas diretamente com Camada Única (Figura 9a).** A forma mais simples de uma RNA, com uma camada de entrada com nós de fonte, que se projeta sobre uma camada de saída de neurônios.
- **Redes alimentadas diretamente com Múltiplas Camadas (Figura 9b).** Presença de uma ou mais camadas intermediárias chamadas de ocultas. Os neurônios ocultos interagem entre as entradas e saídas extraindo conhecimento entre as iterações, o que é necessário quando o tamanho da camada de entrada é grande. Os nós da camada de entrada fornecem os padrões de ativação (vetor de entrada) que são aplicados aos neurônios na segunda camada. Os sinais da saída da segunda camada são entrada para a terceira camada e assim por diante até a saída final ou camada de saída. A saída da última camada é a saída global, que é a resposta da RNA.
- **Redes recorrentes (Figura 9c).** Distingue-se de uma rede neural alimentada diretamente por ter pelo menos um laço de realimentação. Pode apresentar apenas uma camada, por exemplo, desde que alimente seu sinal de saída de volta para as entradas de todos os outros neurônios (como mostrado na Figura 8c). Essa estrutura tem impacto na capacidade de aprendizagem da rede e no seu desempenho. Os laços de realimentação envolvem o uso de ramos particulares compostos de elementos de atraso unitário (*time-delay*), que resulta em um comportamento dinâmico não linear, admitindo que a rede contenha unidades não lineares. Utilizadas em processos de modelagem de séries temporais, reconhecimentos de padrões, entre outras áreas (ANOCHI; CAMPOS VELHO, 2014).

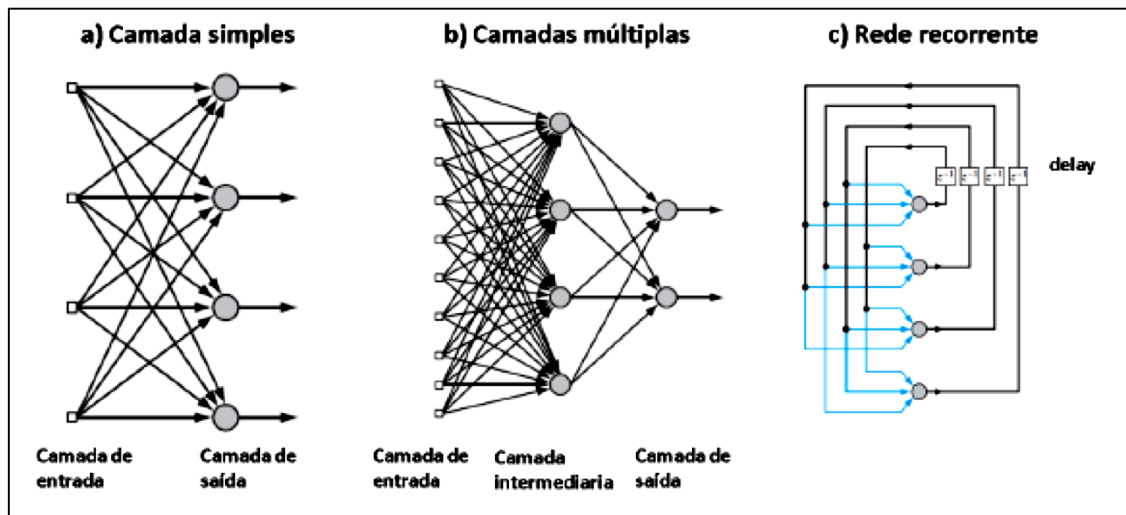


Figura 9 - Algumas Arquiteturas de rede (ROMERO, 2017)

A rede MLP (MultiLayer Perceptron) ou PMC (Perceptron Multicamadas) é a arquitetura mais empregada em aplicações para solucionar diversos problemas complexos. Na Figura 10 tem-se um exemplo de um PMC com duas camadas escondidas, sendo que cada entrada (x) é multiplicada por um peso (w) e a soma das entradas com os pesos e o bias (b) resulta em uma função de transferência ou ativação ($f(x)$), que se transforma em entrada para a camada seguinte. Nessa arquitetura cada camada envia sinais para a próxima camada, como se fossem sinapses (GOLDARAG; MOHAMMADZADEH; ARDAKANI, 2016).

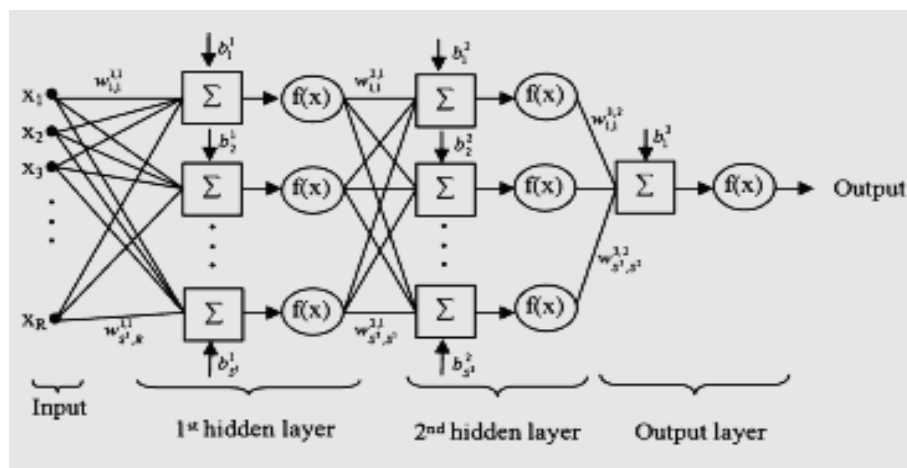


Figura 10 - Rede Neural Perceptron Multicamadas (GOLDARAG; MOHAMMADZADEH; ARDAKANI, 2016).

1.5.3. Processo de Aprendizagem

A principal característica da rede neural é sua capacidade de aprender através de exemplos e melhorar o seu desempenho através da aprendizagem (ANOCHI; CAMPOS VELHO, 2014). Na aprendizagem da rede é utilizado um processo iterativo de ajustes aplicados a seus pesos sinápticos e bias, tornando a rede mais instruída após cada interação.

Geralmente o conjunto de dados é dividido em três subconjuntos: conjunto de treinamento, validação e teste. O primeiro consiste nos padrões fornecidos à rede para os ajustes dos seus parâmetros (aprendizagem), o segundo conjunto tem a tarefa da verificação da capacidade de generalização da rede e o conjunto de teste tem a tarefa de testar o comportamento da rede neural com dados novos (ANOCHI; CAMPOS VELHO, 2014).

A aprendizagem pode ser feita de duas formas: treinamento supervisionado e o treinamento não supervisionado. No caso do não supervisionado, não é fornecida uma saída desejada (supervisor) aos padrões de entrada. Sendo estes ajustados conforme impulsos fornecidos pelo ambiente externo à rede, visando minimizar métricas pré-estabelecidas (ROMERO, 2017).

Na rede de treinamento supervisionado o treinamento é através de dados fornecidos por um especialista ou através de exemplos de entrada e saída desejada. Tendo três componentes inter-relacionados: o ambiente estacionário, quando as estatísticas do erro não mudam depois de várias iterações do aprendizado da rede; o professor, que seria a resposta desejável e o algoritmo de aprendizado, que ajusta os pesos das conexões (GOULART et al., 2006).

No treinamento supervisionado, são selecionados exemplos contendo vetores de entrada e vetores alvo contendo o que se deseja como saída. Assim, a rede compara a entrada e a saída e obtém um sinal de erro. Este é alimentado na rede para atualização dos pesos segundo o algoritmo a ser utilizado (GOULART et al., 2006).

Existem cinco regras básicas de aprendizagem, na qual os algoritmos são baseados sendo elas (HAYKIN, 2001):

- **Aprendizado por correção de erro:** fundamentada na filtragem ótima. Conhecido também como regra delta, o objetivo é minimizar a função de custo dada pela Equação 10.

$$\varepsilon(n) = \frac{1}{2} e_k^2(n) \quad (10)$$

Onde e_k é o erro dado pela Equação 11, y o sinal de saída computado pela rede, d a resposta desejada e k é o número do neurônio.

$$e_k(n) = d_k(n) - y_k(n) \quad (11)$$

O peso sináptico é atualizado pela Equação 12 e 13, sendo x o vetor de entrada e n o passo de tempo, e η a taxa de aprendizado.

$$\Delta w_{kj} = \eta e_k(n) x_j(n) \quad (12)$$

$$w_{kj}(n+1) = w_{kj}(n) + \Delta w_{kj}(n) \quad (13)$$

- **Aprendizagem baseada em memória:** ela opera memorizando explicitamente os dados de treinamento. Ela armazena todas as experiências em uma grande memória de exemplos de entrada e saída classificados como $\{(x_i, d_i)\}_{i=1}^N$, onde x representa um vetor de entrada e d uma resposta desejada correspondente. Quando um vetor de teste deseja ser classificado, o algoritmo responde procurando em uma vizinhança local a resposta correspondente. Essa aprendizagem pode utilizar regras estatísticas como a *regra do vizinho mais próximo* ou a *probabilidade de erro bayesiana* (BARRETO, 2002).
- **Aprendizagem Hebbiana:** a mais antiga e famosa regra de aprendizagem. Ela parte do princípio que quando dois neurônios estão próximos o suficiente pode haver sinapse. Se um neurônio excitar o outro, então a força da sinapse é aumentada, caso contrário se um tenta ativar o outro e outro não ativa, a sinapse é enfraquecida. A sinapse é dependente do tempo e da natureza local. Ela é interativa e correlativa. A forma mais simples de aprendizagem hebbiana é dada pela Equação 14 .

$$\Delta w_{kj}(n) = \eta y_k(n) x_j(n) \quad (14)$$

Onde η é a taxa de aprendizagem, y_k é atividade pós-sináptica e x_j a atividade pré-sináptica (BARRETO, 2002).

- **Aprendizagem competitiva:** os neurônios competem entre si para se tornarem ativos. Os neurônios inibem outros, sendo que ao fim da competição apenas um neurônio é excitado. Assim, apenas um neurônio de saída fica ativo de cada vez, diferente do anterior. Esta aprendizagem segue três elementos básicos (HAYKIN, 2001):
 1. Um conjunto de neurônios iguais, ligados por conexões sinápticas de valores aleatórios.
 2. Um valor limite que implica sobre a ativação dos neurônios
 3. Um mecanismo que permita a competição entre os neurônios, pelo direito de permanecerem excitados.

- **Aprendizagem de Boltzmann:** os neurônios constituem uma estrutura recorrente e operam de maneira binária, sendo o estado ligado representado por +1 e o estado desligado representado por -1. Isto é caracterizado por uma função de energia E (Equação 15), cujo valor é determinado pelo estado particular ocupado pelo neurônio individual da máquina.

$$E = -\frac{1}{2} \sum_j \sum_k w_{kj} x_k x_j \quad (15)$$

Onde x_j é o estado do neurônio j e w_{kj} é o peso sináptico conectado ao neurônio j ao neurônio k . Sendo ρ_{kj}^+ a correlação dos estados entre os neurônios em sua condição presa (em estado específico determinados pelo ambiente). ρ_{kj}^- representa a correlação entre os estados do neurônio j e k , que estão operando livremente. Assim a variação Δw_{kj} é dada pela Equação 16 (HAYKIN, 2001).

$$\Delta w_{kj}(n) = \eta(\rho_{kj}^+ - \rho_{kj}^-), j \neq k \quad (16)$$

1.5.4. Algoritmo *Backpropagation*

Uma das formas muito utilizadas no treinamento de redes MLP é o algoritmo conhecido como retropropagação do erro (*backpropagation*) baseado na regra de aprendizagem por correção de erro. Consiste em dois passos através de diferentes camadas de rede: um *passo à frente*, a propagação, e um *passo para trás*, a retropropagação. O erro é retropropagado da camada de saída para a camada de entrada. No passo para frente um vetor de entrada é aplicado aos nós e seu efeito se propaga e é assim produzido um conjunto de saídas com pesos fixos. E no passo para trás os pesos são corrigidos de acordo com a regra de correção de erro (HAYKIN, 2001). O erro é calculado como mostrado na Equação 17, sendo o valor instantâneo de energia do erro para o neurônio k dado pela $\frac{1}{2} e_k^2(n)$. O valor instantâneo da energia total do erro é dada pela Equação 17, onde C inclui todos os neurônios da camada de saída da rede, a energia média do erro quadrático médio dada pela Equação 18, avaliado para todos n (número de padrões) (ROMERO, 2017).

$$\mathcal{E}(n) = \frac{1}{2} \sum_{k \in C} e_k^2(n) \quad (17)$$

$$\mathcal{E}_{med} = \frac{1}{N} \sum_{n=1}^N \mathcal{E}(n) \quad (18)$$

O \mathcal{E}_{med} representa uma *função de custo* para medir o desempenho de aprendizagem, assim é necessário ajustar os parâmetros livres para minimização desse custo.

As saídas $y_k(n)$ são retro-alimentadas nas camadas de entrada para os neurônios, a somatória ponderada produz a variável $v_j(n)$, mostrado na Equação 19, conhecida como campo local induzido que é o argumento da função de ativação $\varphi(\cdot)$.

$$\mathbf{v}_j(\mathbf{n}) = \sum_{i=0}^m \mathbf{w}_{kj}(\mathbf{n}) \mathbf{y}_k(\mathbf{n}) \quad (19)$$

O algoritmo de retropropagação aplica uma correção $\Delta w_{jk}(n)$ ao peso sináptico de cada neurônio, usando o gradiente da função de custo, como na Equação 20. Devido ao uso gradiente como forma de otimização, a função de ativação da MLP deve ser contínua.

$$\Delta \mathbf{w}_{jk}(\mathbf{n}) = -\eta \frac{\partial \mathcal{E}(\mathbf{n})}{\partial \mathbf{w}_{jk}(\mathbf{n})} \quad (20)$$

Se adicionar outro parâmetro α , é possível modificar a regra Delta (correção de erro) de forma a evitar convergência para um mínimo local da função custo. Este parâmetro é denominado taxa de momento (momentum).

$$\Delta \mathbf{w}_{jk}(\mathbf{n}) = \alpha \Delta \mathbf{w}_{jk}(\mathbf{n} - 1) + \eta \delta_j(\mathbf{n}) \mathbf{y}_k(\mathbf{n}) \quad (21)$$

Sendo α definido no intervalo 0 e 1, para convergência do delta (Equação 21) e δ_j representa o gradiente local da rede, definido pela Equação 22 .

$$\delta_j(\mathbf{n}) = -\frac{\partial \mathcal{E}(\mathbf{n})}{\partial v_j(\mathbf{n})} = \mathbf{e}_j(\mathbf{n}) \varphi'_j(\mathbf{v}_j(\mathbf{n})) \quad (22)$$

1.5.5. RNA autoconfigurada: Multi-Particle Collision Algorithm (MPCA)

Um dos maiores problemas no uso de RNAs é a seleção da melhor topologia. Processo que geralmente demanda muito tempo e trabalho para encontrar os melhores parâmetros.

1.5.5 Algoritmo MPCA

Vários autores têm buscado melhorar o desempenho da rede com técnicas de otimização (TEIXEIRA et al., 2000; CARVALHO; RAMOS; CHAVES, 2011; COSTA et al., 2003a; COSTA et al., 2003b). Desses o Multiple Particle Collision Algorithm (MPCA) desenvolvido por Luz, Becceneri e Campos Velho (2008) é um algoritmo de otimização utilizado para encontrar a melhor arquitetura de uma RNA. Algoritmos de otimização têm o

objetivo de encontrar a melhor solução maximizando ou minimizando uma função, definida como função objetivo ou função de custo (ANOCHI; CAMPOS VELHO, 2014).

O MPCA foi inspirado no Particle Collision Algorithm (PCA) (SACCO; OLIVEIRA, 2005), sendo que a diferença entre eles é que ao invés de usar uma partícula como este último, apresenta várias partículas se movendo no espaço de busca. Estes algoritmos são inspirados no princípio de um reator nuclear de absorção e dispersão (ANOCHI; CAMPOS VELHO, 2014).

O algoritmo inicia com a seleção de uma solução inicial, que é modificada com uma perturbação estocástica, que produz uma solução nova. Esta é comparada com a solução anterior e pode ou não ser aceita. Desse modo se a solução é melhor que a anterior esta é absorvida, caso contrário, a partícula é dispersa no espaço. Operadores de perturbação são utilizados para a exploração de posições próximas. A melhor solução é então compartilhada por todas as partículas no processo. Este algoritmo é implementado em uma arquitetura de multiprocessador com memória distribuída usando Message Passing Interface (MPI) (ANOCHI; CAMPOS VELHO, 2014). O pseudocódigo do algoritmo está na Figura 11.

Na função de Exploração (*Exploration*) são aplicadas pequenas perturbações estocásticas usando a função *Small-Perturbation* para realizar uma pequena exploração local, a fim de verificar a existência de uma possível solução melhor em uma determinada vizinhança, o pseudocódigo é mostrado na Figura 12.

```
Generate an initial solution: Old-Config
Best-Fitness = Fitness{Old-Config}
Update Blackboard
For n = 0 to # of particles
  For n = 0 to # iterations
    Update Blackboard
    Perturbation{.}
    If Fitness{New-Config} > Fitness{Old-Config}
      If Fitness{New-Config} > Best-Fitness
        Best-Fitness = Fitness{New-Config}
      End If
      Old-Config = New-Config
      Exploration{.}
    Else
      Scattering{.}
    End If
  End For
End For
End For
```

Figura 11 –Pseudocódigo do algoritmo MPCA (ANOCHI; CAMPOS VELHO, 2014)

```

Exploration(.)
  For n = 0 of # iterations
    Small-Perturbation()
    If Fitness(New-Config) > Fitness(Old-Config)
      If Fitness(New-Config) > Best-Fitness
        Best-Fitness = Fitness(New-Config)
      End-if
      Old-Config = New-Config
    End-if
  End-if
Return

```

Figura 12 – Pseudocódigo Função Exploration (ANOCHI; CAMPOS VELHO, 2014)

A função de disseminação (*Scattering*) é uma função estocástica que tenta evitar arestas do algoritmo em regiões com locais ótimos, o pseudocódigo é mostrado na Figura 13.

```

Scattering(.)
  Pscattering = 1 - Fitness(New-Config) / Best-Fitness
  If Pscattering > Random(0,1)
    Old-Config = Random solution
  Else
    Exploration()
  End-if
Return

```

Figura 13 – Pseudocódigo Função Scattering (ANOCHI; CAMPOS VELHO, 2014)

Na aplicação em RNA o objetivo é encontrar um valor ótimo que representa a melhor arquitetura dependente de uma combinação de variáveis. Isso é obtido através da minimização de uma função custo dada pela Equação 23, proposta por Carvalho, Ramos e Chaves (2011).

$$f_{obj} = \text{penalidade} * \left(\frac{\rho_1 * E_{tre} + \rho_2 * E_{gen}}{\rho_1 + \rho_2} \right) \quad (23)$$

Onde $\rho_1 = 1$ e $\rho_2 = 0,1$, fatores e ajuste da relevância do erro de treinamento (E_{tre}) e erro de generalização (E_{gen}). O erro de treinamento reflete a capacidade de memória da rede e o erro de generalização reflete a habilidade da rede identificar os padrões no conjunto de validação. A penalidade é aplicada para se encontrar a melhor arquitetura de menor complexidade, dado pela Equação 24 (ANOCHI; CAMPOS VELHO, 2014).

$$\text{penalidade} = C_1(\varepsilon^{neurônios})^2 * C_2(\text{épocas}) + 1 \quad (24)$$

Onde $C_1 = 1$, $C_2 = 0,1$ e $\varepsilon^{neurônios}$ é o número total de neurônios.

Assim o MPCA tem como resposta: o número de camadas intermediárias, número de neurônios em cada camada, a taxa de aprendizagem, a taxa de momento e a melhor função de ativação. Os limites dos valores são mostrados no Quadro 4 (ANOCHI; CAMPOS VELHO, 2014).

Quadro 4 – Parâmetros definidos pelo MPCA

Parâmetro	Valor
Neurônios na camada escondida	1,0 ... 32
Taxa de Aprendizagem: η	0,0 ... 1,0
Constante de Momento: α	0,1 ... 0,9
Função de Ativação	Tanh Logistic Gauss

(Adaptado de ANOCHI; CAMPOS VELHO, 2014)

O MPCA é usado para gerar várias soluções de arquitetura para uma RNA. Para cada solução, a RNA é ativada, e o treinamento ocorre até o critério de parada (erro mínimo ou total de épocas), sendo esse treinamento realizado por *backpropagation*. Com os valores obtidos pela RNA, o MPCA calcula a função objetivo, atualizando os parâmetros da RNA. Esse processo é repetido até um valor ótimo da função objetivo ser encontrado (ANOCHI et al., 2016).

ANOCHI et al. (2016) compara a rede autoconfigurada com a configurada por especialistas na previsão clima de mesoescala mensal para precipitação, e mostra melhores resultados para rede configurada com o MPCA. Cintra et al., (2015, 2016, 2018) aplicaram em seus estudos essa metodologia na assimilação de dados atmosféricos para modelos atmosféricos globais. Sambatti (2017) aplicou em simulações de circulação oceânica, em equações de onda e de água rasa e conseguiu respostas eficazes.

2. Materiais e Métodos

O procedimento metodológico é apresentado na Figura 14 e será explicado nos itens a seguir. Foram utilizadas 12 variáveis para entrada no modelo. Para calibração e validação do modelo foram utilizados os dados de fogo ativo dos satélites (dados de saída). Foi feito o processamento desses dados, sendo separados os dados de 2014 a 2016 para a confecção da rede e os dados de 2017 para validação do modelo. Os primeiros foram vetorizados e separados em conjuntos para entrada no MPCA, treinamento e validação da RNA. Após validada a RNA foi feita a validação do modelo com os dados de 2017.

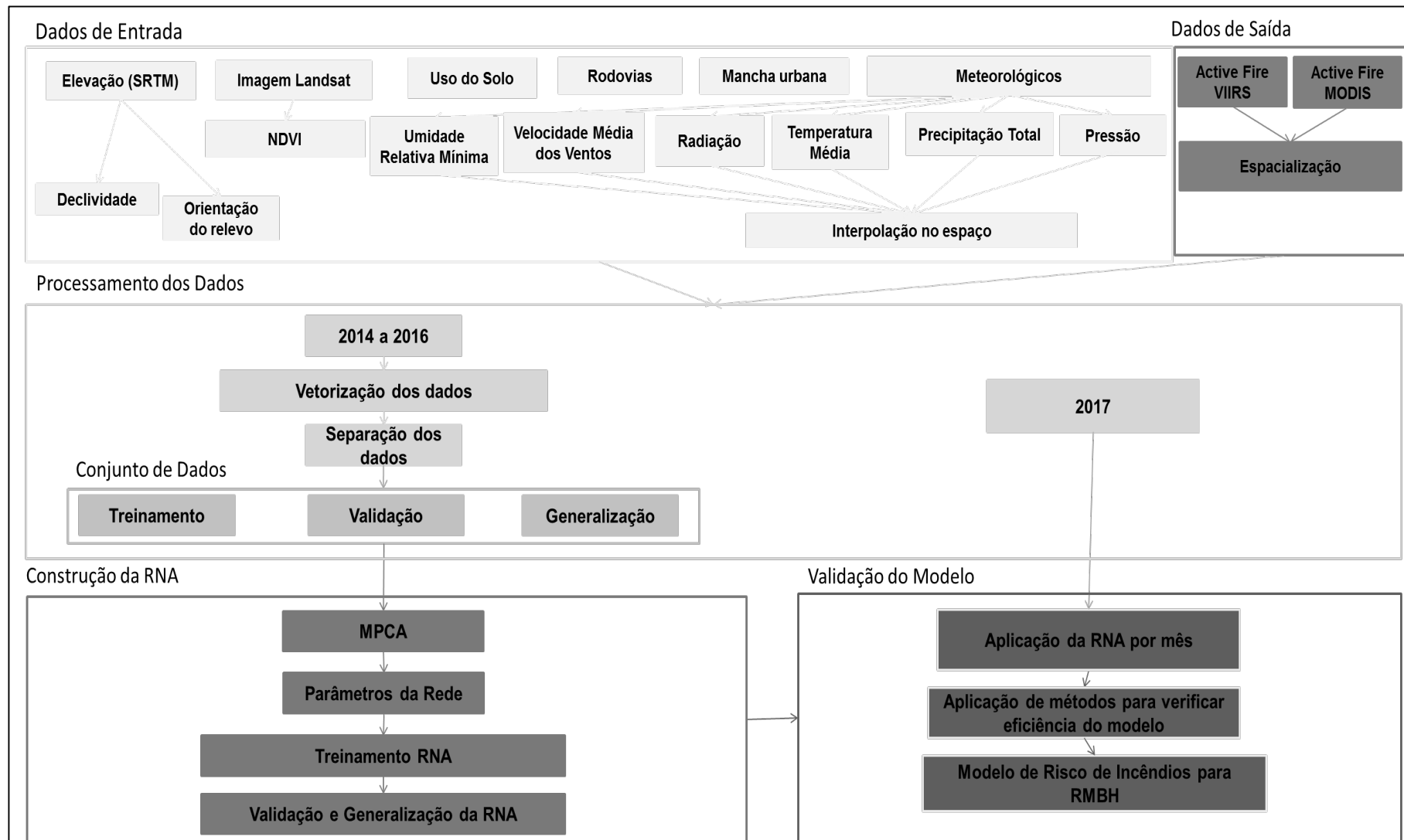


Figura 14 – Roteiro Metodológico

2.1. Dados e materiais

2.1.1. Dados NDVI

O Índice de Vegetação por Diferença Normalizada (NDVI) é utilizado para estimar a biomassa, cobertura vegetal e na detecção de mudanças de padrão de uso e cobertura da terra. O seu cálculo (Equação 25) é feito através de uma operação entre as bandas do espectro infravermelho próximo (IR) e a do espectro vermelho (R). Partindo do princípio que quanto maior a quantidade de vegetação, maior a reflexão no infravermelho e menor no vermelho, aumentando o valor da razão e realçando a vegetação (MELO; SALES; OLIVEIRA, 2011).

$$NDVI = (NIR - R) / (NIR + R) \quad (25)$$

A vegetação é caracterizada assim, por uma intensa absorção devido à clorofila na região espectral do vermelho (0,63 – 0,69 μm) e por uma intensa energia refletida na região do infravermelho próximo (0,76 – 0,90 μm) causada pela estrutura celular das folhas. Os valores de NDVI oscilam de -1 a +1, sendo que quanto mais próximo de 1, maior a cobertura vegetal, e mais próximo de -1, representam corpos d'água. As nuvens costumam refletir de forma semelhante no visível e no infravermelho próximo, assim o pixel costuma ter valor próximo de 0. Se o solo for nu ou a vegetação esparsa, costumam ter valores positivos, mas muito próximos de +1 (MELO; SALES; OLIVEIRA, 2011).

No presente trabalho calculou-se o NDVI para cada mês dos anos de 2014 a 2017, para verificar as mudanças mais sensíveis de vegetação. Para esse cálculo foram utilizadas imagens do satélite LANDSAT-8/OLI (USGS, 2017), utilizando as cenas 218/73 e 218/74 que englobam a região. Foram selecionadas as cenas de cada mês que continham a menor porcentagem de nuvens.

2.1.2. Dados meteorológicos

Os dados de precipitação foram selecionados das estações da Agência Nacional de Águas (ANA) (ANA, 2018) que continham dados recentes de precipitação e que estão presentes no entorno e dentro da área de estudo, totalizando 73 estações, de onde foi calculada a precipitação total por mês.

Os dados de temperatura, velocidade dos ventos e umidade relativa mínima foram retirados de três fontes: 6 estações automáticas (INMET, 2018b) e 6 estações convencionais

(INMET, 2018a) provenientes do Instituto Nacional de Meteorologia (INMET); e 5 estações automáticas provenientes da Rede de Meteorologia do Comando da Aeronáutica (REDEMET) (REDEMET, 2018), totalizando 17 estações na área e entorno. Destes dados de estação, foram calculadas as médias mensais para cada variável. Os dados de radiação e pressão atmosférica estão disponíveis somente nas estações do INMET (2018a). As estações estão apresentadas na Figura 15.

Os dados foram interpolados seguindo os resultados da comparação dos métodos constante no APÊNDICE A. Assim, a maioria das variáveis foram interpoladas utilizando o método da krigagem ordinária exponencial, apenas a pressão foi interpolada utilizando o método IDW (ponderação pelo inverso da distância ao quadrado) e a umidade relativa mínima foi feita utilizando o método TPS (Thin Plate Splines).

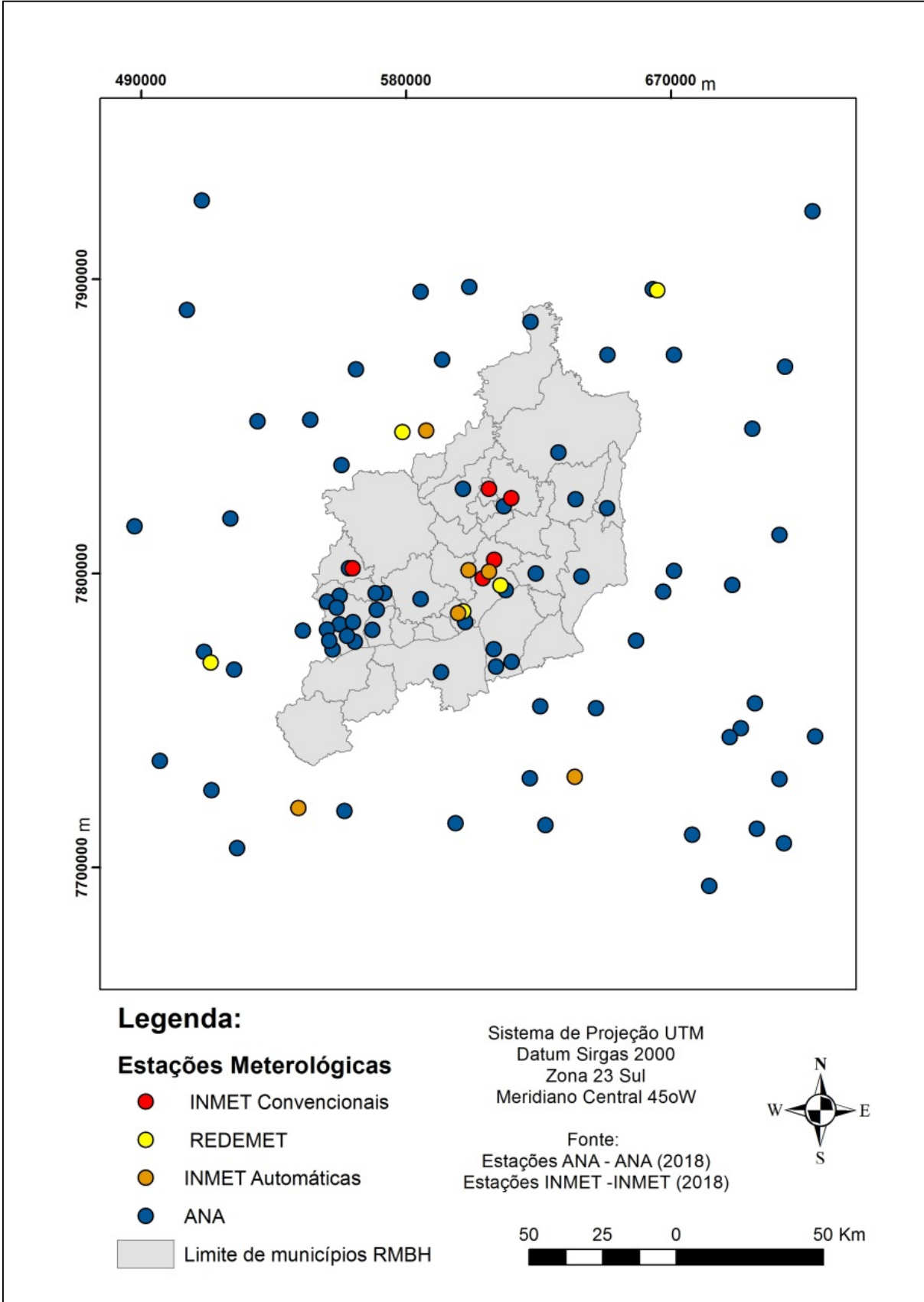


Figura 15 - Localização das Estações meteorológicas

2.1.3. Mapa de uso e cobertura do solo

O mapeamento do uso e cobertura do solo é importante para definir os tipos de uso e cobertura do solo, como: tipos de vegetações e áreas antropizadas, como áreas urbanas. O mapeamento utilizado nesse trabalho é o do Projeto de Mapeamento Anual da Cobertura e Uso do Solo do Brasil (MAPBIOMAS, 2017), mapeamento feito por uma rede de especialistas que utiliza processamento em nuvem e classificadores automatizados desenvolvidos e operados a partir da plataforma Google Earth Engine para gerar uma série histórica de mapas anuais de cobertura e uso da terra do Brasil. Os dados utilizados nesse trabalho foram os mapas de cobertura de 2014, 2015, 2016 e 2017.

As classes de uso presentes no mapeamento para a RMBH são mostradas no capítulo 2.1.1.

2.1.4. Distância à mancha urbana e às rodovias

Para tentar quantificar a influência humana nos incêndios foi utilizada distância às rodovias e às áreas urbanas. A classe de infraestrutura urbana do MAPBIOMAS (2017) foi usada como referência da mancha urbana. E para rodovias foi utilizado a base de rodovias principais, em *shapefile* (UFMG, 2014). As distâncias foram calculadas pela função Distância Euclidiana do ArcGis. Os resultados são apresentados na Figura 16.

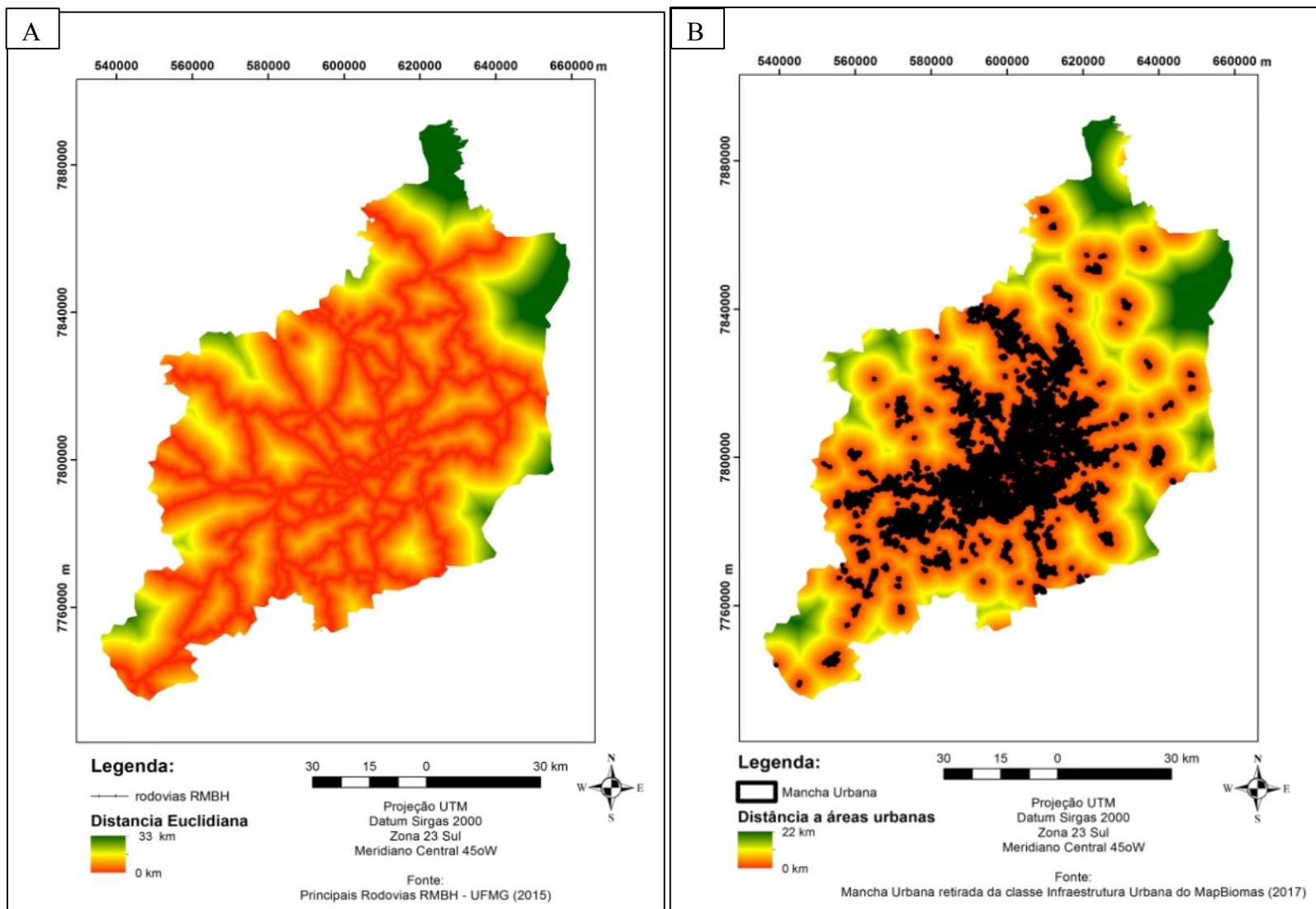


Figura 16 – Mapas de distâncias: (A) Distância a Principais Rodovias (B) Distância a áreas Urbanas.

2.1.5. Dados de elevação

Com respeito às altimetrias foram utilizados dados da Missão SRTM-Shuttle Radar Topography Mission (USGS, 2014), que possui uma resolução espacial de 30 m e foi produzido no ano 2000.

A partir destes dados foram calculadas a declividade e a orientação do relevo com a função “terrain” da biblioteca Raster do R.

2.1.6. Focos de incêndio

Foram utilizados dois produtos de Fogo Ativo: *Near real-time (NRT) Moderate Resolution Imaging Spectroradiometer (MODIS) Thermal Anomalies/Fire locations-Collection 6*, que utilizam produtos da faixa (MOD14/ MYD14), com resolução espacial de 1 km e *Near real-time (NRT) Suomi National Polar-orbiting Partnership (Suomi-NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) Active Fire detection product*, que tem resolução espacial de 375 m. Estes dados estão disponíveis em NASA (2018), e eles são fornecidos no formato *shape* ou *textfile*, contendo informações de localização dos pontos, dia e hora da detecção, confiabilidade dos dados e FRP. Para tanto, foram utilizados dados de 2014 a 2017.

Os pontos de fogo foram espacializados de acordo com suas coordenadas e com sua resolução espacial. Depois foram convertidos em *raster*, onde o valor 1 representa ocorrência de incêndios e 0 representa não ocorrência. Depois de colocados na mesma resolução espacial, foi feita a soma do *raster* dos dois produtos, para se obter um produto único de dados de focos de incêndio para cada mês.

2.1.7. Processamento dos dados

Todos os dados foram transformados em Raster e colocados na mesma resolução espacial, tamanho e projeção. Como projeção foi utilizado o sistema de coordenadas geográficas no datum WGS84, devido à maioria das variáveis apresentarem esse sistema. A grade utilizada foi a de 30 m, seguindo os dados de NDVI, uso do solo e elevação. E os dados foram recortados seguindo o limite da área disponível em UFMG (2014), ficando nos limites do seguinte quadrante: $-45,77^{\circ}$ e $-42,18^{\circ}$ de longitude e $-20,95^{\circ}$ e $-18,62^{\circ}$ de latitude. Assim no final de cada arquivo utilizou-se uma grade regular de 8404 linhas e 12948 colunas, correspondente a 388,44 x 252,12 m.

Os dados de 2014 a 2016 foram separados para construção da RNA. Para isso foram vetorizados e as células da matriz de cada variável foram associadas à sua correspondente, como no exemplo apresentado na Tabela 1.

Tabela 1 - Exemplo de vetorização dos dados da RNA

UR	Rodo	Decl	Urb	Ori	Uso	Vento	Prec	Temp	Rad	Pres	NDVI	Fogo
54.4	1.6	13.7	4.1	219.7	21	4.9	76.4	23.7	1.8	88.7	0.8	0
53.8	1.4	11.5	0.7	175.2	21	4.9	76.6	23.7	1.8	89.0	0.7	0
54.5	2.1	14.2	2.0	256.1	15	4.8	76.7	23.7	1.8	89.0	0.7	0
55.2	0.0	11.4	2.4	18.7	15	4.9	77.3	23.7	1.8	88.6	0.8	0
54.4	1.9	9.6	5.1	265.8	3	4.9	77.8	23.7	1.8	88.8	0.9	0
40.6	8.5	9.8	7.9	255.0	9	4.8	74.3	24.5	1.8	93.7	0.8	1
43.4	1.5	18.5	0.0	67.1	24	4.0	58.7	24.5	1.8	92.8	0.1	1
43.1	1.7	10.0	0.0	30.9	24	4.1	59.0	24.5	1.8	92.8	0.3	1
43.5	1.2	2.5	0.0	90.0	24	4.0	59.1	24.5	1.8	92.8	0.1	1
43.2	1.7	2.4	0.1	46.5	4	4.0	58.5	24.5	1.8	92.8	0.5	1

*UR=Umidade Relativa mínima; Rodo= Distância às rodovias; Urb=Distância à áreas urbanas; Ori=Orientação do Relevo; Vento= Vento médio; Prec=Precipitação total; Temp=Temperatura Média; Pres=Pressão

Ao todo o conjunto de dados teria: 8404 (número de linhas) x 12948 (número de colunas) x 36 (número de meses) de elementos. Um número muito grande, que demandaria muito tempo de processamento e muito provavelmente teria resultados ruins. Isso porque a quantidade total de focos no período foi de 7388. Assim, foram selecionadas apenas as células de foco de incêndio e para pontos de não ocorrência de incêndio foram criadas janelas selecionando um ponto aleatório em cada uma. Para tanto, foram testadas janelas de 1 km, 5 km, 10 km e 20 km.

Posteriormente, esses dados foram separados 50% para treinamento, 25% para validação e 25% para generalização. Para a utilização no MPCA-NN e para a construção da rede.

2.2. Construção da RNA

2.2.1. Normalização dos Dados

Para entrada da RNA os dados devem ser normalizados. Esta normalização foi feita seguindo a Equação 26, onde x_n é o valor normalizado, x é o valor não normalizado, i é o número do padrão, $\min(x)$ é o valor mínimo e $\max(x)$ é o valor máximo do conjunto inteiro de uma mesma variável. Esse procedimento é feito para cada variável de entrada separadamente.

$$x_{ni} = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (26)$$

2.2.2. MPCA-NN

O pacote MPCA-NN foi utilizado para encontrar a melhor arquitetura da RNA. Como utiliza o algoritmo *backpropagation* para arquitetura da PMC, o programa devolve a taxa de aprendizagem, taxa de momento e a função de ativação a serem utilizados. O pacote utiliza um método de busca (meta-heurística) o MPCA, o qual minimiza (ou otimiza) uma função custo onde considera o erro de treinamento e o erro de generalização da rede em função de uma penalidade que é a complexidade da RNA.

Uma máquina com multiprocessador é necessária, pois o sistema usa MPI (*Multi Parallel Interface*), um padrão para comunicação de dados em computação paralela. Este pacote foi feito em linguagem Fortran 90 e está disponível para *download* em Campos Velho et al. (2018).

Para utilização do MPCA-NN é necessário obter os três conjuntos de dados (treinamento, validação e generalização) e editar o arquivo de configuração que vem na estrutura do software; adicionando parâmetros como: número de padrões para o treinamento e validação, o número de entradas e saídas, o valor de erro alvo, o número máximo de épocas para critério de parada, dentre outras. E para executar o programa é necessário indicar o número de experimentos (para se obter vários resultados e escolher qual é a melhor solução dentre as rodadas) e o número de processadores (pois é um programa paralelo – se a máquina for de um só processador; colocar 1 no número de processadores).

Para este trabalho o software foi executado em um computador com multiprocessador disponibilizado pelo LAC/INPE através de acesso remoto. Foram utilizados 4 processadores e realizados 4 experimentos.

2.2.3. Rede Neural Artificial

A RNA construída foi do tipo *perceptron* multicamadas (PMC): com 12 variáveis de entrada (que constituem a camada de entrada da rede) e uma saída (indicando se houve incêndio ou não naquele ponto de grade). As variáveis de entrada utilizadas foram: declividade do relevo, orientação do relevo, uso e cobertura do solo, imagem NDVI, distância

às rodovias, distância à área urbana, velocidade dos ventos, radiação, temperatura, pressão, umidade relativa mínima e precipitação.

A RNA foi construída utilizando-se linguagem R (algoritmo em FERNANDES, 2019). Foram utilizados os parâmetros resultantes do MPCA-NN e a rede foi treinada com o algoritmo *backpropagation*. O resultado do treinamento são os pesos fixos às conexões rede. Na validação da rede é verificado se a rede está adaptada para solucionar o problema com valores de entrada não apresentados no treinamento. Posteriormente foi realizado um teste para ver se a rede extrapola o conhecimento adquirido no treinamento, que é a generalização da RNA ou a previsão de focos de incêndio para outros valores de entrada.

2.2.4. Avaliação da RNA

Para avaliação do desempenho da rede foi utilizada a metodologia de Weiss & Kulikowski (1991) onde é computada: a quantidade de verdadeiros positivos (VP), isto é, pontos classificados corretamente como fogo; a quantidade de verdadeiros negativos (VN), onde os pontos classificados corretamente como não fogo; quantidade de falsos positivos (FP) e negativos (FN) e pontos classificados incorretamente. Para isso foi estabelecido como valor limite de detecção 0,5. Todos os pontos acima de 0,5 foram considerados como positivo e abaixo foram considerados negativo. Estas medidas foram calculadas conforme as Equações 27, 28, 29, 30 e 31.

$$\text{Acurácia (A)} = (\text{VP} + \text{VN}) / (\text{VP} + \text{FN} + \text{FP} + \text{VN}) \quad (27)$$

$$\text{Sensibilidade (S)} = \text{VP} / (\text{VP} + \text{FN}) \quad (28)$$

$$\text{Especificidade (E)} = \text{VN} / (\text{FP} + \text{VN}) \quad (29)$$

$$\text{Predição Valores Verdadeiros (PVV)} = \text{VP} / (\text{VP} + \text{FP}) \quad (30)$$

$$\text{Predição Valores Negativos (PVN)} = \text{VN} / (\text{VN} + \text{FN}) \quad (31)$$

A acurácia é a probabilidade do modelo de fornecer resultados corretos, tanto positivos, quanto negativos. A sensibilidade (S) é verificada, se o modelo é capaz de identificar corretamente os valores positivos e a especificidade (E) é verificada, se o modelo é capaz de identificar os valores negativos. A Predição de Valores Verdadeiros (PVV) e a Predição de Valores Negativos (PVN) verificam a porcentagem de valores positivos e negativos, respectivamente, em relação à quantidade total de cada grupo. Em todos esses

parâmetros quanto mais próximo de 100%, maior é a capacidade do modelo de retornar valores corretos.

Adicionalmente, foram calculadas as curvas ROC (termo em inglês para *Receiver Operating Characteristic*), técnica muito utilizada para avaliação de problemas binários de decisão, que mostra como o número de verdadeiros positivos varia em relação aos falsos positivos. Assim, foi obtida a Área Sob a Curva - AUC (termo em inglês para *Area Under Curve*), sendo que quanto mais próximo de 1 melhor o desempenho da rede e quanto mais próximo de 0,5, pior é o desempenho (AKOBENG, 2007).

2.2.5. Análise das variáveis mais características no modelo

Existem alguns métodos desenvolvidos para tentar mensurar quanto uma variável está influenciando no modelo de RNA, algumas delas são explicadas no decorrer do texto.

2.2.5.1. Algoritmo de Garson

O algoritmo proposto por Garson (1991) relaciona os pesos de conexões da camada escondida e da camada de saída, para obtenção da importância relativa de cada variável. Utilizando o procedimento de cálculo apresentado nas Equações 32, 33, 34 e 35, para encontrar a importância relativa (R_j).

$$P_{ij} = \text{abs}(w_{ij}) \times \text{abs}(w_{io}) \quad (32)$$

$$Q_{ij} = P_{ij} / \sum_{j=1}^n P_{ij} \quad (33)$$

$$S_j = \sum_{i=1}^n Q_{ij} \quad S_j = \sum_{i=1}^n Q_{ij} \quad (34)$$

$$R_j = (S_j / \sum_{j=1}^n S_{ij}) \times 100 \quad (35)$$

Sendo i o número de neurônios na camada intermediária, j o número de entradas e o número de neurônios na camada de saída.

2.2.5.2. Método Perturb

Avalia o efeito de pequenas perturbações da variável de entrada na saída da rede. Sendo adicionada a uma variável de entrada de 10 a 50% do seu valor e é verificado o erro quadrático médio de da perturbação, isso é feito para todas as entradas. Espera-se que esse erro aumente à medida que uma maior quantidade de ruído é adicionada à variável. As

variáveis que tendem a perturbar mais o modelo geralmente são as de maior importância relativa (LIMA, 2012).

2.2.5.3. Correlação

A correlação mede a “força” ou “grau” de relacionamento entre duas variáveis, variando no intervalo de -1 a 1; quanto mais próximo de 1 e -1, maior a correlação, indicando que a variável exerce mais influência na determinação da saída da rede (LIMA, 2012).

2.3. Validação do Modelo

Serão utilizados dados de 2017 (somente para dados de entrada) aplicados à rede (PMC com os pesos fixos) e os resultados comparados às observações de focos de incêndio. Esta verificação mostrará se o modelo consegue uma boa previsão para focos de incêndio em um período para o qual o PMC (modelo) não foi treinado.

O risco de incêndio será separado em cinco classes: Risco Muito Alto (100-80%), Risco Alto (80-60%), Risco Médio (60-40%), Risco Baixo (40-20%), Risco Muito Baixo (20-0%).

Será feito para dados mensais, obtendo o total de 12 mapeamentos para 2017.

3. Caracterização da Área de Estudo e Pré-análise da pesquisa

3.1. Informações Gerais

A Região Metropolitana de Belo Horizonte, está inserida no Estado de Minas Gerais e seus acessos principais são pelas rodovias BR-040, BR-381 (Rodovia Fernão Dias), BR-356 e BR-262, como mostrado na Figura 17. A região é composta por 34 municípios, com uma extensão Territorial de 9460 Km², sendo a 3^a maior Região Metropolitana do Brasil e a 7^a maior região metropolitana da América Latina. Possui uma população de aproximadamente 5 milhões de habitantes e um PIB de 56,8 milhões, correspondente a 40% do PIB estadual (AGENCIA RMBH, 2016). A região abrange três bacias hidrográficas: Bacia do Rio das Velhas, Bacia do Rio Paraopeba e Bacia do Rio Pará, sub-bacias do Rio São Francisco.

As principais unidades geomorfológicas da RMBH são: Quadrilátero Ferrífero, Depressão do São Francisco e Serra do Espinhaço. O Quadrilátero Ferrífero é delimitado pela Serra do Curral, Serra da Moeda, Serra do Gandarela, Serra de Ouro Branco e Serra do Caraça (Figura 33), situado a sudeste da região metropolitana. É uma região clássica da geologia, ocorrendo jazidas de ferro, manganês, ouro, bauxita e pedras preciosas, como topázio e esmeralda (ROESER; ROESER, 2010). A Unidade Geomorfológica da Serra do Espinhaço inicia-se na extremidade meridional desta, nas regiões da nascente do Rio Cipó, apresentando altitudes de 1200 m, sendo predominantemente constituída por quartzitos e as formas de relevo resultantes são representadas por cristas, escarpas e vales profundos adaptados às direções tectônicas e estruturais. A Depressão San-franciscana desenvolveu-se ao longo da drenagem do Rio São Francisco, podendo ser considerada Inter-planáltica. Seu arcabouço litológico é bastante variável, ocorrendo desde rochas granito-gnáissicas até metassedimentares (REIS JUNIOR, 2008).

As cidades mais populosas da RMBH são: Belo Horizonte, Contagem, Betim, Ribeirão das Neves e Ibité (Figura 18). Belo Horizonte apresenta a maior densidade populacional da região com 72 hab/ha, seguido de Contagem com 30 hab/ha, Ibité com 22 hab/ha e Ribeirão das Neves com 19 hab/ha. É possível notar que a densidade populacional diminui quando se afasta do centro da região.

A população da RMBH é em sua maioria urbana (98%), principalmente nos grandes centros (maiores detalhes no APÊNDICE B). A área rural compreende 732.891 ha, enquanto que a área urbana é de apenas 214.273 ha, correspondente a apenas 22% da área total, como visto na Figura 19 (IBGE, 2010a).

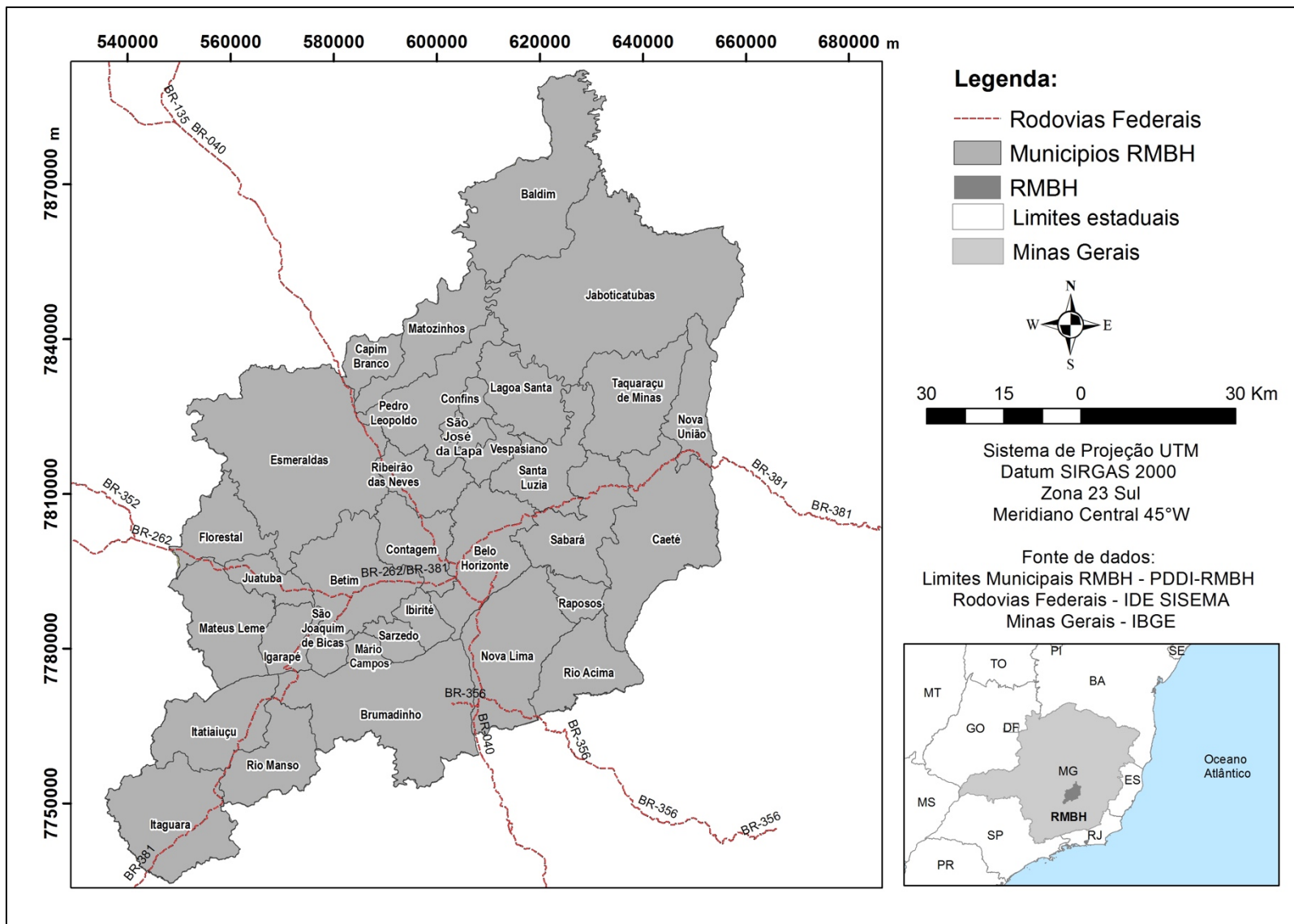


Figura 17 - Mapa de localização da Área de Estudo

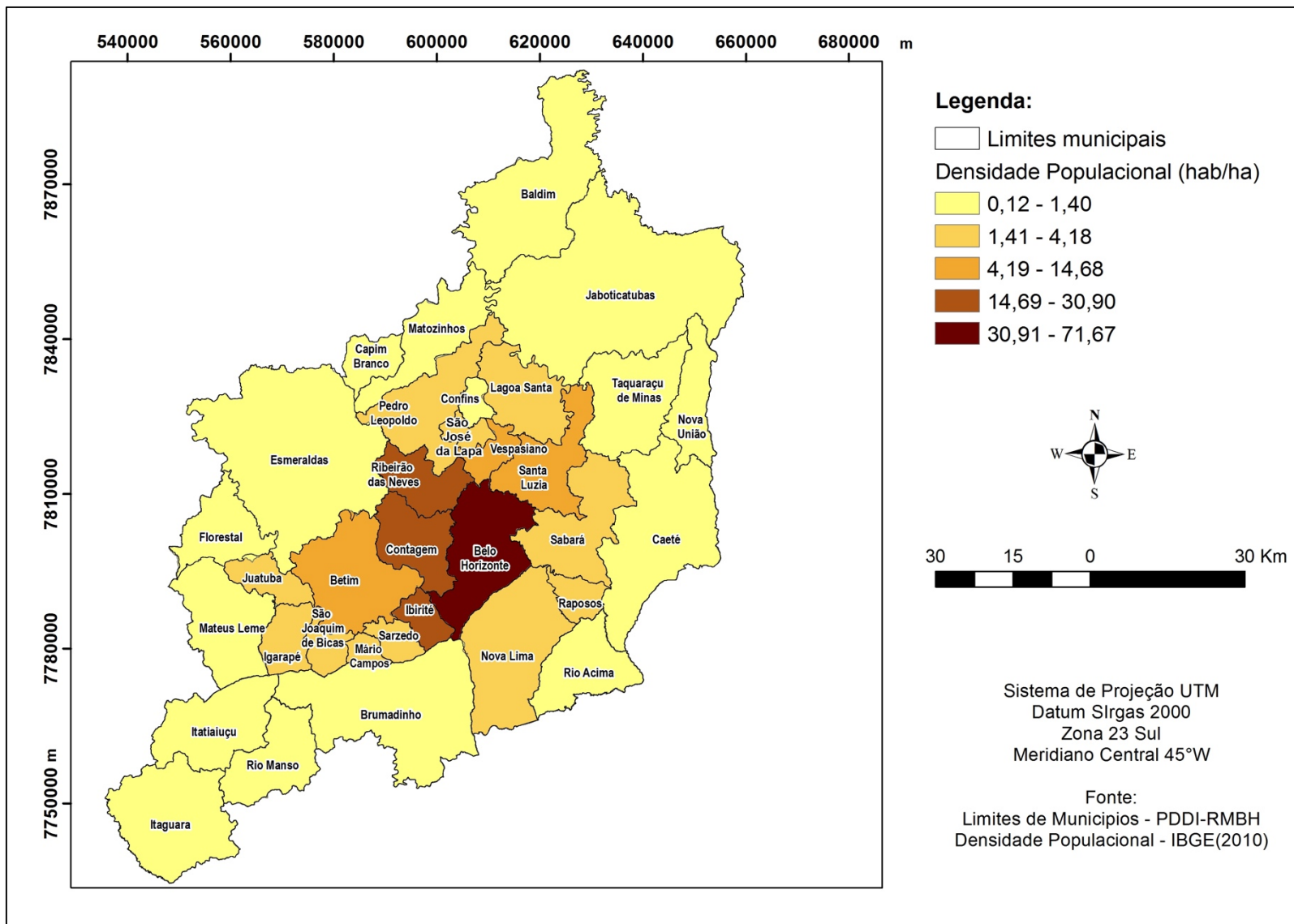


Figura 18 - Densidade Populacional RMB

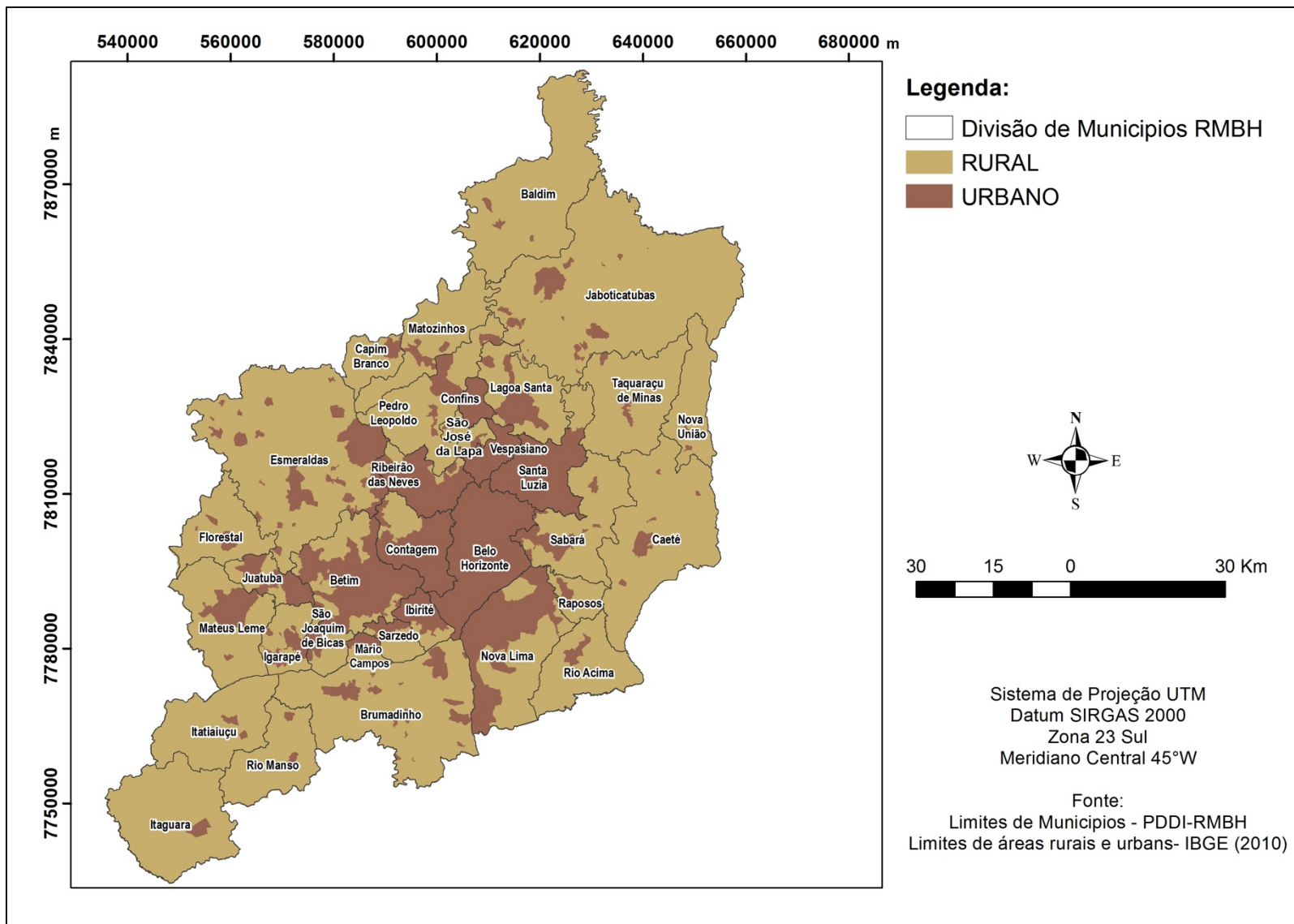


Figura 19 - Relação Área Urbana/Rural na RMBH

Na RMBH existem aproximadamente 5.000 km² de áreas protegidas, totalizando 140 unidades de conservação (Figura 20) (UFMG, 2014). Dentre essas áreas pode se destacar a Área de Proteção Ambiental ao Sul da Região Metropolitana de Belo Horizonte (APA Sul-RMBH), região que compreende os municípios de Barão de Cocais, Belo Horizonte, Brumadinho, Caeté, Catas Altas, Ibirité, Itabirito, Mário Campos, Nova Lima, Raposos, Rio Acima, Santa Bárbara e Sarzedo. Essa região abrange parte das bacias do Rio Paraopeba, Rio das Velhas e Rio Doce. Estando localizada em área coberta por Floresta Estacional Semi-decidual e pequenas áreas de Cerrado e Campo. Dentro da APA existem várias áreas consideradas de preservação permanente, em diferentes categorias de unidade de conservação. São elas: Reserva Particular do Patrimônio Natural (RPPN) da Serra do Caraça, APA Municipal do Mingu, Reserva Biológica da Mata do Jambreiro (RPPN), Parque Estadual da Baleia, Parque Estadual do Jatobá, Estação Ecológica de Fechos e Parque Estadual da Serra do Rola Moça. De 11 mananciais que abastecem a RMBH, oito estão presentes nela, sendo eles: Taboões, Rola-Moça, Bálamo, Mutuca, Fechos, Catarina, Barreiro e Cercadinho (SPÓSITO; STEHMANN, 2007).

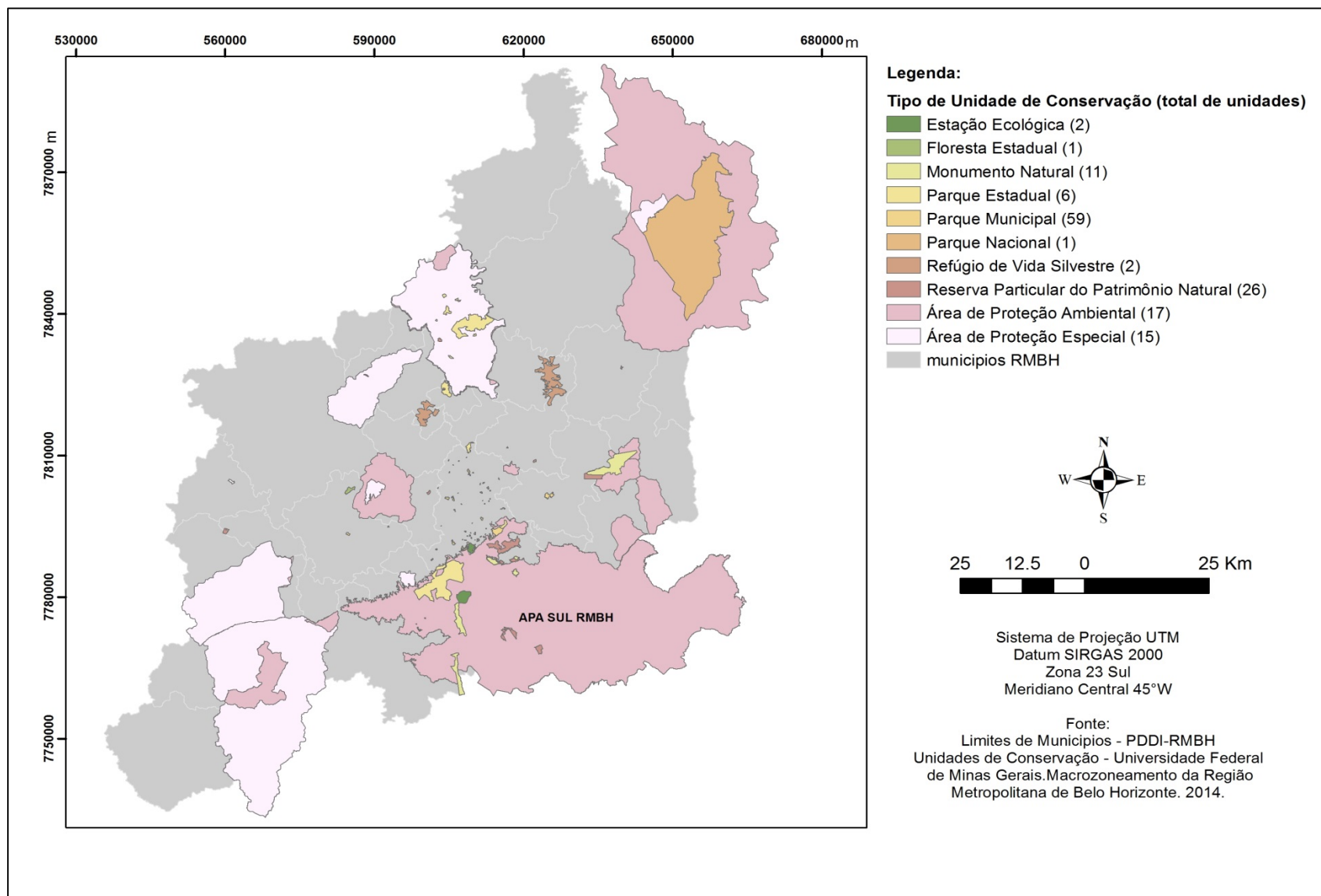


Figura 20 - Unidades de Conservação na RMBH

3.2. Uso e ocupação do Solo

As classes de uso do solo segundo o MAPBIOMAS (2017) se apresenta como na Figura 21. Segundo Alencar (2017) a classe de formação florestal, compreende as florestas densas, entre elas mata ciliar, mata seca e cerradão. Já a formação savânica, inclui as florestas abertas, cerrado denso e cerrado rupestre. E a formação rupestre representa os campos, incluindo campo limpo, sujo e campo rupestre. Pela distribuição de classes é possível notar que a maior parte da região compreende formação florestal, pastagem e mosaico de agricultura e pastagem.

3.2.1. Caracterização da Vegetação

A RMBH está presente entre dois biomas: mata atlântica e o cerrado. O primeiro se caracteriza pela predominância de fisionomias florestais ombrófilas e estacionais (semi-decíduais e decíduais). Enquanto que o cerrado é caracterizado pela predominância de fisionomias savânicas (OLIVEIRA FILHO, 2006).

A Mata Atlântica representa a segunda maior floresta pluvial tropical do continente americano, sendo um dos 25 “hotspots” mundiais de biodiversidade. Embora tenha sido em grande parte destruída, ela ainda abriga mais de 8.000 espécies endêmicas de plantas vasculares, anfíbios, répteis, aves e mamíferos (MYERS et al., 2000). Segundo último relatório da SOS MATA ATLÂNTICA/INPE (2018) apenas 15,20% da sua floresta original ainda é remanescente.

O Cerrado é uma das principais formações de savana nos trópicos, sendo o segundo maior bioma da América do Sul, ocupando 22% do território brasileiro, é o bioma que maior concentra a produção agropecuária do país (MARQUES et al., 2017).

Analisando o último inventário florestal disponibilizado pelo IEF ⁽²⁰⁰⁹⁾, a RMBH apresenta as feições vegetais apresentadas na Figura 22. Nota-se que mais de 45% da sua vegetação é composta por floresta semi-decidual montana, 25% Cerrado, 12% Campo, 7% Campo rupestre e 5% Campo Cerrado.

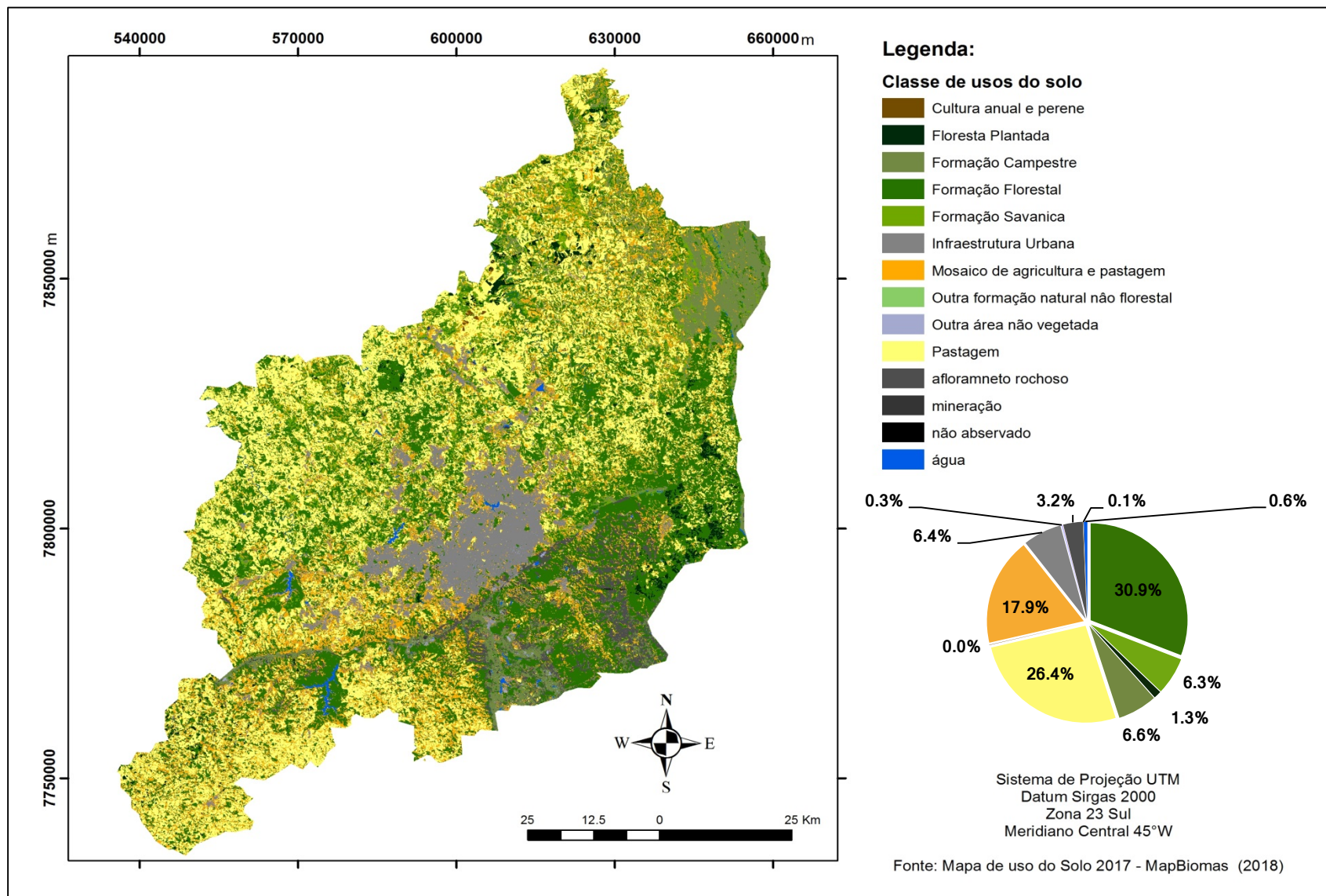


Figura 21 - Classificação Uso e ocupação do solo na RMBH para o ano de 2017.
Fonte: MAPBIOMAS (2017)

A floresta estacional semi-decidual se caracteriza por possuir dossel de 4 m a 25 m de altura, dependente do tipo de solo. Suas árvores emergentes podem chegar a 40 m e possui um sub-bosque denso. Apresentam deciduidade (perda de folhas em determinada estação) intermediária nas épocas mais secas. No Domínio da Mata Atlântica, é a tipologia predominante e no Domínio do Cerrado ocorre na forma de encraves e florestas associadas a corpos d'água permanentes ou intermitentes (OLIVEIRA FILHO, 2006).

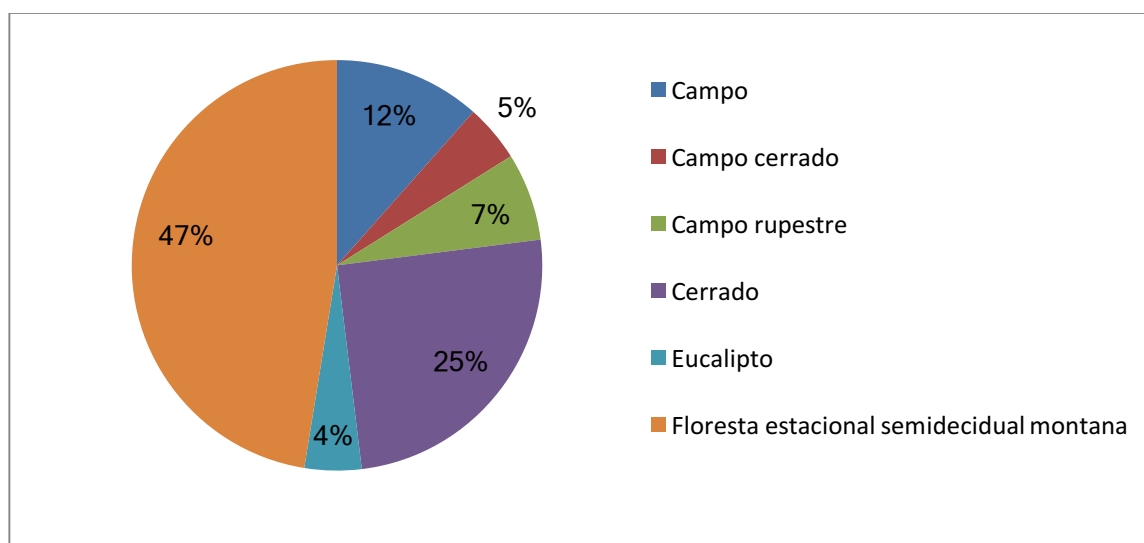


Figura 22 - Classes de Vegetação
Fonte: IEF (2009)

A vegetação de cerrado possui dois estratos, a saber, o herbáceo-subarbustivo (ou campestre) e o arbóreo-arbustivo (ou lenhoso), sendo esse último composto por árvores e arbustos tortuosos, com casca grossa. O estrato campestre apresenta densidade inversamente proporcional à cobertura do estrato lenhoso (OLIVEIRA FILHO, 2006). O Cerrado se apresenta com vários tipos de cobertura vegetal, sendo essas classificadas em:

- Campo Limpo: apresenta apenas o estrato campestre
- Campo Sujo: Apresenta 10% de cobertura lenhosa e altura média de 1,5 m
- Campo Cerrado ou Cerrado: apresenta altura média de 7 m
- Cerradão: 80% das suas fisionomias é de estrato lenhoso com altura média de até 15 m, apresentando dossel contínuo acima de 7 m de altura e sub-bosque pobre em graminóides.

Estes são presentes em climas com estações de seca e chuva bem definidos, ocorre em solos em geral bem drenados e ácidos, principalmente latossolos, cambissolos e neossolos quartzarênicos. Sendo alta a incidência de fogo na estação seca (OLIVEIRA FILHO, 2006).

Os incêndios têm influência na dinâmica e estrutura das populações de plantas, removendo serrapilheira e controlando espécies dominantes, além de atuar na hierarquia de competição em comunidades vegetais. As vegetações sob influência dos incêndios possuem adaptações que permitem sua sobrevivência ou a rápida regeneração (FIDELIS; PIVELLO, 2011).

Entre as adaptações do Cerrado aos incêndios florestais pode-se destacar: forma tortuosa de árvores e arbustos; casca grossa das árvores; gramíneas com gemas protegidas através da estrutura das folhas; órgãos subterrâneos, muitas vezes lenhosos, como xilopódios e rizomas, providos de gemas capazes de formar ramos aéreos, tanto vegetativos como reprodutivos, após a passagem do fogo. Outras estratégias incluem acúmulo de biomassa na porção subterrânea e aumento da produção de frutos e sementes após queimadas. Assim, as queimadas periódicas estimulam o rebrotamento, a ciclagem dos nutrientes, a frutificação e o estabelecimento de plântulas de diversas espécies. Além disso, a ausência do fogo geralmente promove um aumento da cobertura lenhosa, levando à perda de espécies herbáceas características e a mudanças na fisionomia da vegetação (FIDELIS; PIVELLO, 2011).

Mesmo o Cerrado sendo um ecossistema adaptado e dependente do fogo, queimadas de grande intensidade e proporções espaciais não são benéficas para a vegetação ou a fauna e podem causar perda de espécies, inclusive endêmicas, invasão por espécies exóticas, desequilíbrio nos ciclos ecológicos e mudanças nas fisionomias dos ecossistemas atingidos. Isso pode provocar a perda de nutrientes, compactação e erosão dos solos; perda da biota local devido ao acúmulo da biomassa vegetal seca. Além disso, queimadas frequentes podem afetar negativamente o estabelecimento de arbustos e árvores, e ainda liberar muitos gases responsáveis pelo efeito estufa na atmosfera (KLINK; MACHADO, 2005). As simulações realizadas por Hoffmann & Jackson (2000) mostraram que a precipitação pode ser afetada, tendo uma diminuição de 10% de chuvas. Assim, as queimadas são importantes para esses ecossistemas, mas devem ser feitas de forma programada e controlada, para manutenção de suas fisionomias e diversidade.

3.3. Caracterização Climatológica RMBH

A classificação climática da região abrange quatro zonas climáticas definidas pelo IBGE (2002) com base em Nimer (1979), como mostrado na Figura 23. As zonas de maior expressão são: Tropical Brasil Central mesotérmico brando semiúmido e Tropical Brasil Central sub-quente semiúmido. O clima Tropical Brasil Central é denominado assim devido à influência da Alta Subtropical do Atlântico Sul (ASAS). O tipo Tropical Brasil Central semiúmido se destaca por um regime anual de chuvas caracterizado por uma acentuada concentração de chuvas de instabilidades tropicais no verão, enquanto o inverno é fortemente marcado pela seca (entre 4 a 5 meses).

Quanto à classificação de sub-quente, esta compreende as terras situadas imediatamente acima de cotas altimétricas, sofrendo grande influência das altitudes. Possui pelo menos um mês com temperatura média inferior a 18°C, o mês mais frio (junho ou julho) varia de 18 a 15°C, com mínimas diárias de 6 a 10°C. Temperatura média anual inferior a 22°C, no verão apresenta temperaturas mais elevadas com média superior a 22°C (NIMER, 1979).

O clima mesotérmico brando apresenta temperaturas amenas durante o ano, devido ao relevo acidentado. Apresentando verões amenos com temperatura média inferior a 22°C e invernos mais frios com temperaturas médias inferiores a 15 °C (NIMER, 1979).

Para uma melhor análise climática dados das normais climatológicas produzidas pelo INMET (2018) foram utilizadas, elas são referentes ao período de 01/01/1981 a 31/12/2010. Foram selecionadas as três estações existentes na RMBH (Belo Horizonte, Ibirité e Florestal) para análise (Figura 23).

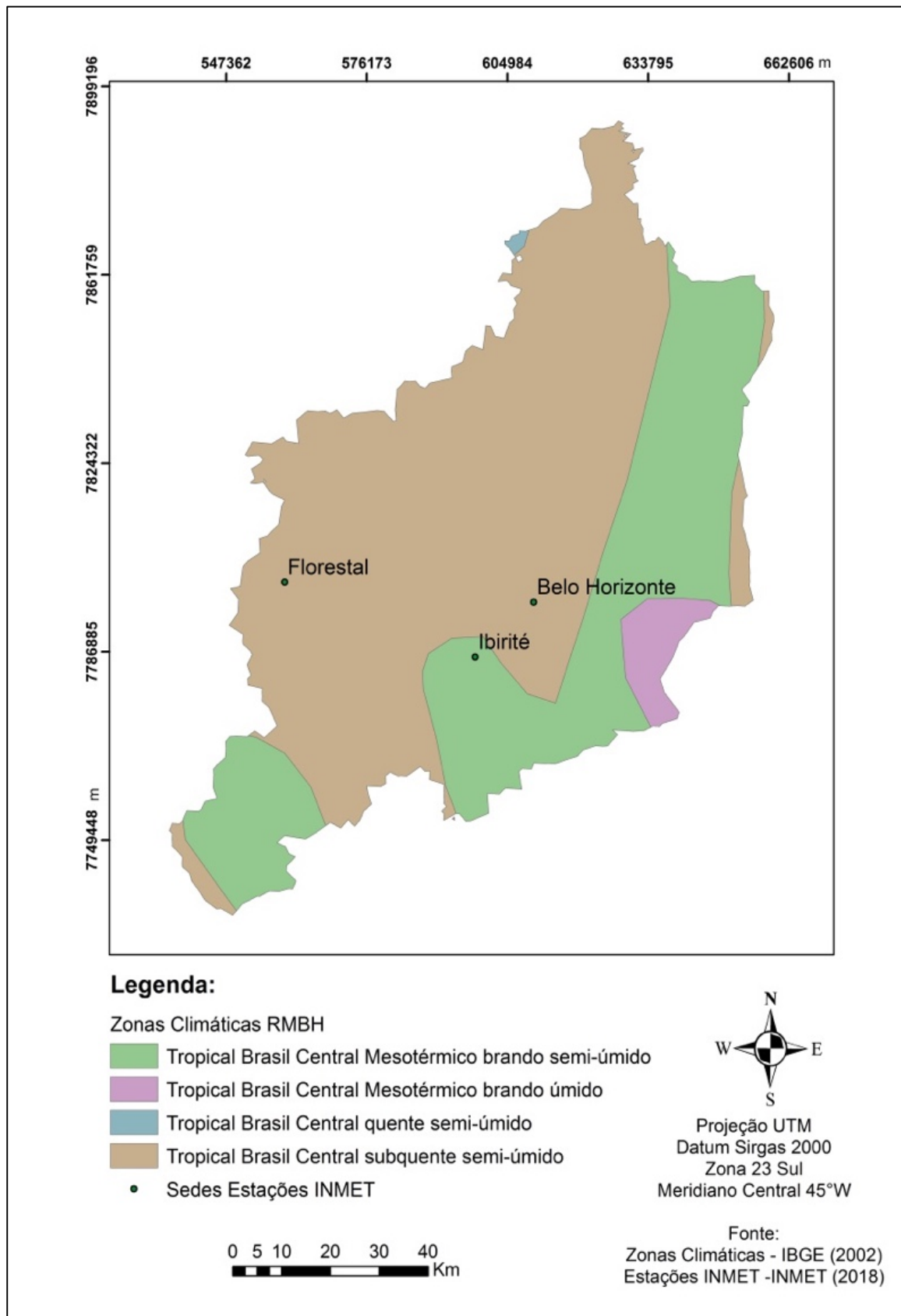


Figura 23 - Classificação Climática da RMBH e estações do INMET para cálculo das normais climatológicas

Nota-se que não há uma variação muito brusca de temperaturas médias na RMBH durante o ano (Figura 24). Em Belo Horizonte a amplitude térmica é menor tendo no inverno uma temperatura média de 20,55°C e mínima de 14,7°C; enquanto no verão apresenta média de 22,9°C e máxima de 29°C. Os municípios de Florestal e Ibirité apresentam as maiores variações entre as temperaturas do inverno e do verão, apresentando temperaturas médias mais próximas de suas temperaturas mínimas. Florestal apresenta as temperaturas mais altas das estações estudadas, apresentando no inverno 18,15°C de temperatura média e 8,5°C de temperatura mínima e 22,05°C de temperatura média e 33°C de temperatura máxima nos meses de verão. Ibirité, única região classificada como mesotérmico brando, apresenta média de 19,33°C e mínima de 9,7°C no inverno e no verão média de 22,6°C e máxima de 30,4°C. Assim, é possível notar diferenças climáticas dentro da região metropolitana.

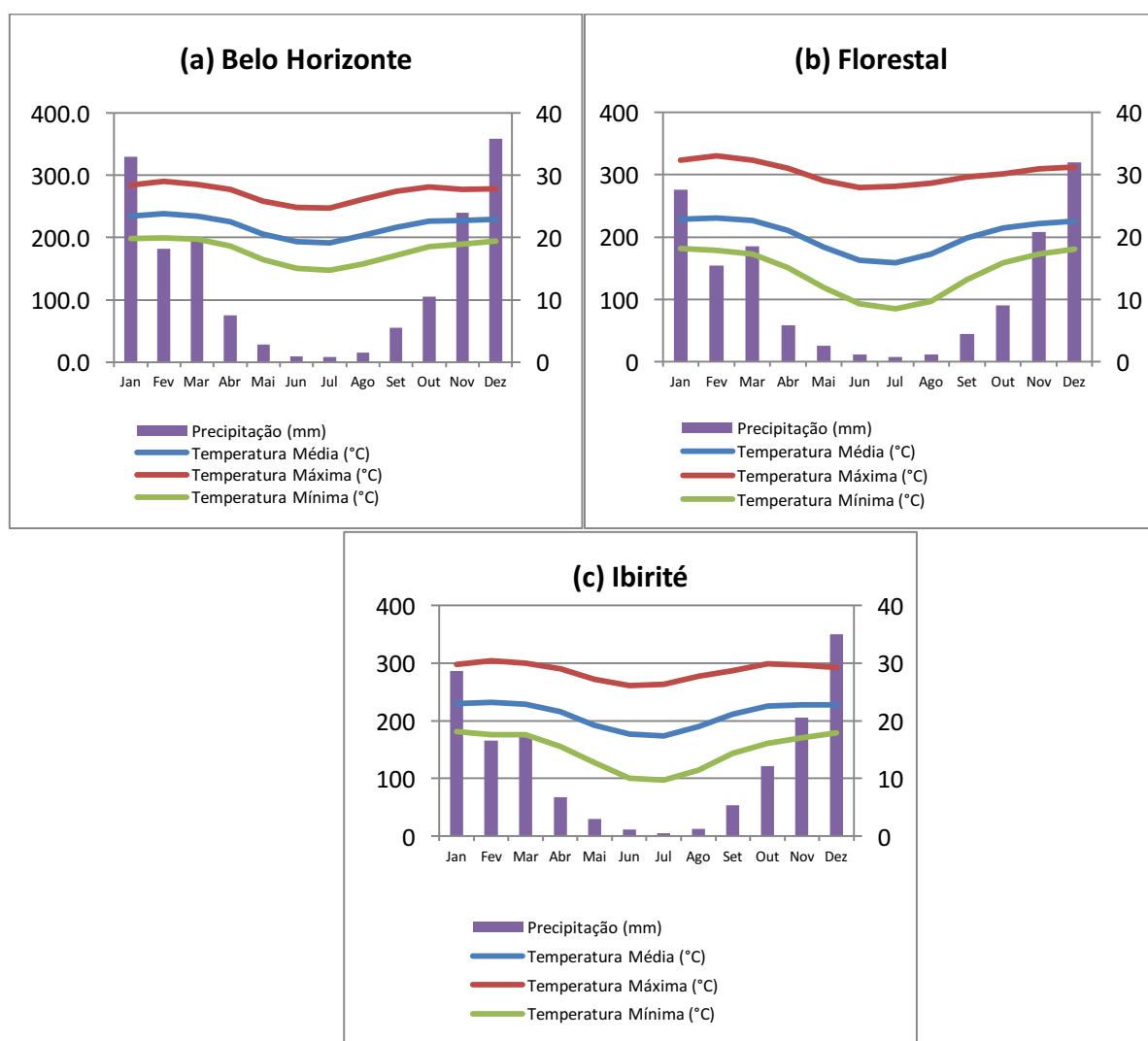


Figura 24 - Climograma estações RMBH (a) Belo Horizonte, (b) Florestal e (c) Ibirité
 Fonte dos dados: (INMET, 2018b)

A precipitação total anual na RMBH é de 1493,6 mm. Nota-se na Figura 24, uma estação de seca bem definida entre maio e setembro. Aproximadamente 80% da precipitação total anual está concentrada entre os meses de novembro e março. Entre as estações analisadas a cidade de Belo Horizonte é a que apresenta a maior quantidade de precipitação e a cidade de Florestal a que possui menor quantidade de precipitação. Por estes gráficos é confirmada as características de clima semiúmido, uma vez que os meses de inverno são meses de seca e os meses de verão, meses de chuva.

Para uma análise mais aprofundada da precipitação foram utilizados dados das estações da ANA (2018) utilizadas nesse estudo. Na Figura 25 é apresentado a precipitação total anual média de todas as estações presentes na RMBH para os anos de 1987 a 2017. Nele é possível perceber que a média dos anos estudados é de 1405,5 mm. Nota-se que os anos que apresentaram menores precipitações foram 1990, 2007, 2015 e 2017 e o ano 2014 apresentou o maior défict de precipitação dos últimos 30 anos com apenas 729,30 mm de chuva.

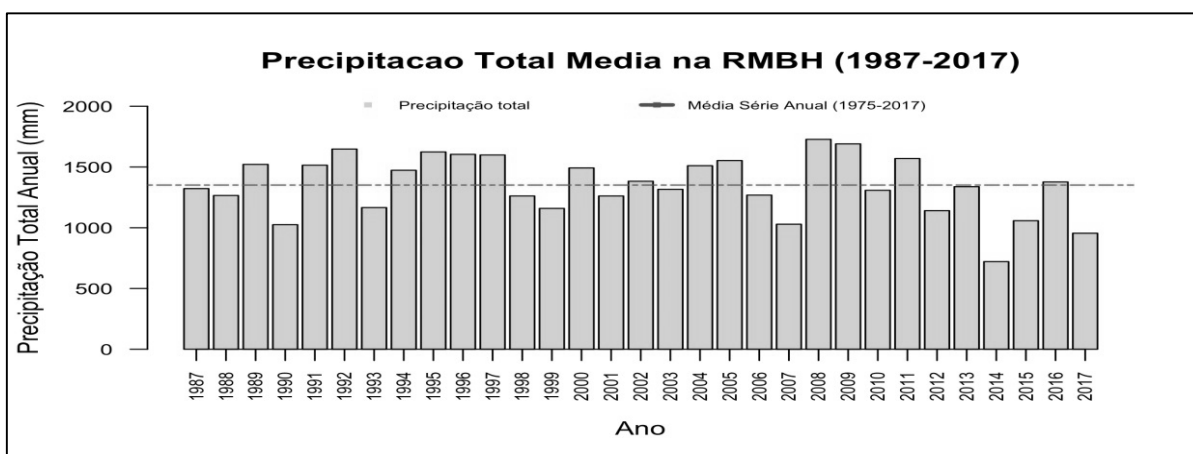


Figura 25 – Média da Precipitação Total das estações da RMBH da série histórica de 1987 a 2017
Fonte dos dados: ANA (2018)

Quanto às precipitações máximas registradas no ano, nota-se que em 2003 teve uma máxima de 800 mm (Figura 26). Os anos de 1991, 2008 e 2011 também apresentam máximas maiores de 700 mm. Com isso pode se verificar que há anos que a precipitação total anual ocorre em apenas alguns meses. Nestes casos estas foram registradas no mês de janeiro.

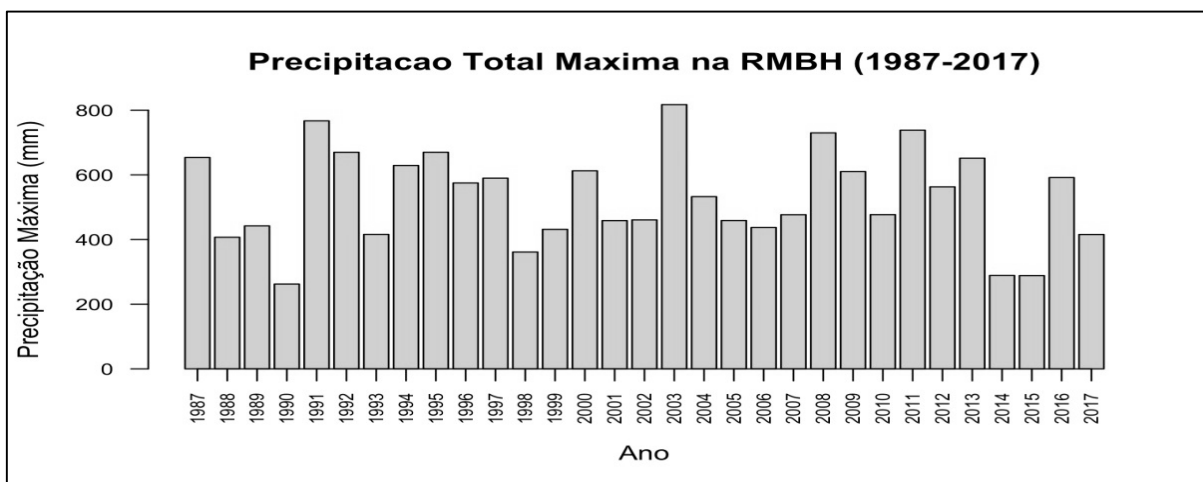


Figura 26 – Média da Precipitação Máxima das estações da RMBH das série histórica de 1987 a 2017
 Fonte dos dados: ANA (2018)

Quanto aos meses de maior precipitação (Figura 27), nota-se que o mês de dezembro é o que apresenta maior precipitação, seguido de fevereiro, janeiro e novembro. Nota-se uma pequena diferença com os dados das normais climatológicas, que pode ser devido aos dados das estações utilizadas nas análises e o período utilizado. No entanto, os meses de chuva e seca se mantiveram, seca de abril a setembro e período chuvoso de outubro a março.

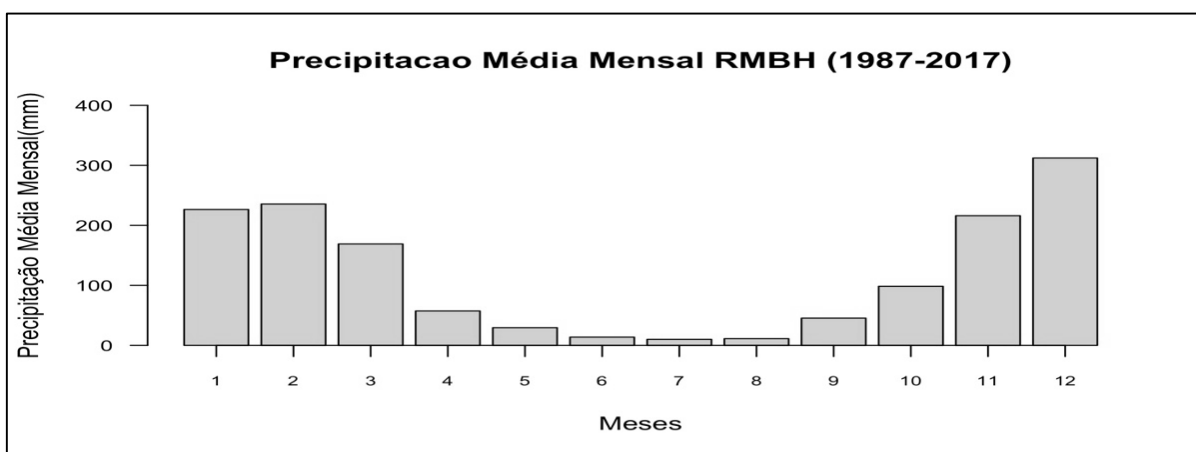


Figura 27 – Média da Precipitação Total Mensal das estações da RMBH das séries histórica de 1987 a 2017
 Fonte dos dados: ANA (2018)

A média da série histórica dos dados de precipitação total foi interpolada a partir das 73 estações utilizadas no estudo (Figura 28). Assim foi possível perceber as regiões que tem maior precipitação, estão concentradas a sudeste da região e as regiões de menor precipitação, ao norte da região. Nota-se que no sudeste da região há maiores valores de precipitações totais anuais e ao norte da região os menores. Uma análise de 10 estações meteorológicas de chuva da RMBH pode ser observada no APÊNDICE C.

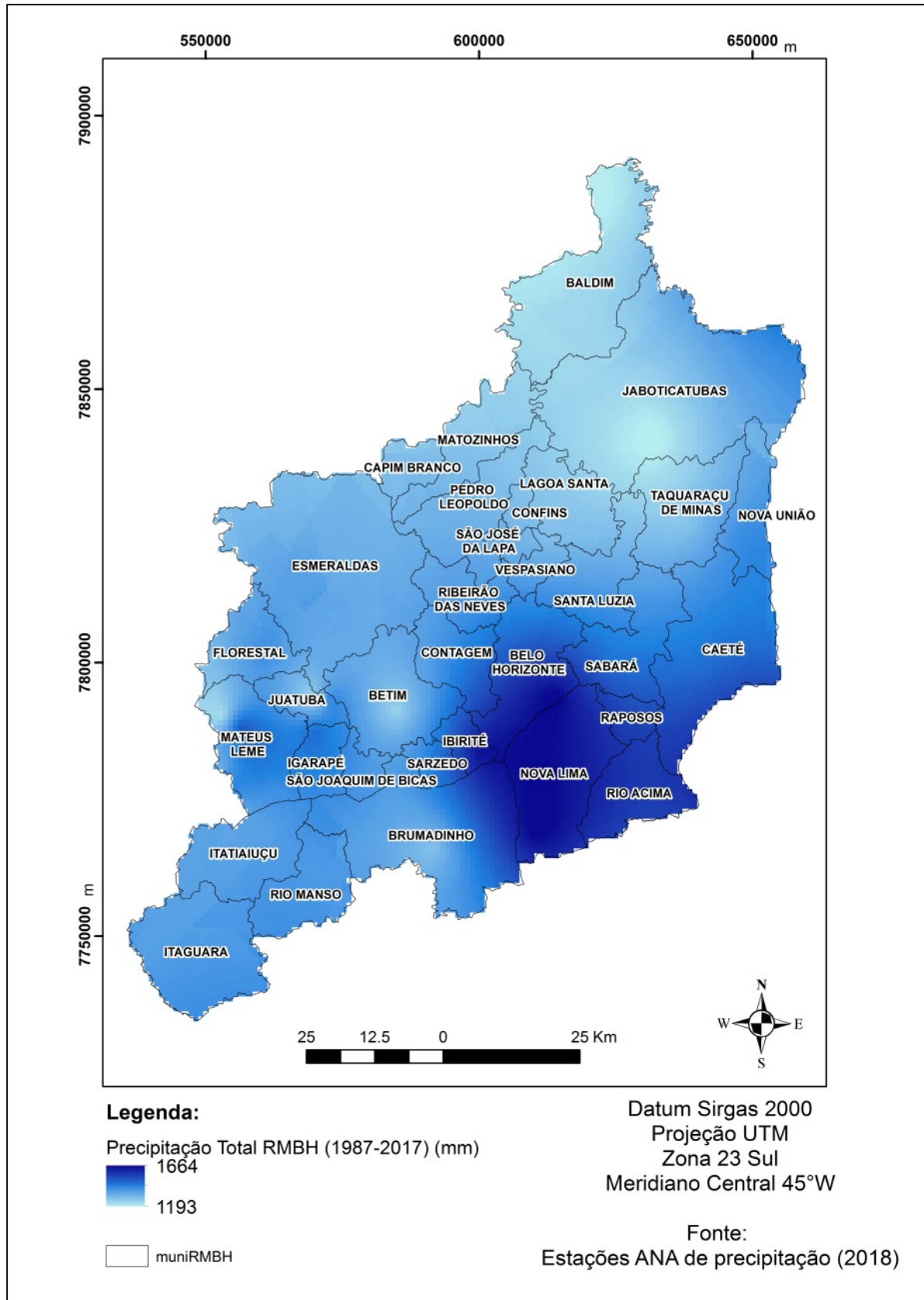


Figura 28 – Precipitação Total na RMBH, média histórica de 1987-2017

Os balanços hídricos climatológicos (Figura 29) Foram feitos com os dados das normais climatológicas das mesmas estações, nota-se que em agosto há maior deficiência hídrica no solo em todas as estações e o excedente maior é em dezembro, o que compreende as épocas de maior e menor precipitação. Há uma deficiência hídrica no solo de sete meses, de abril a outubro. Quanto à retirada de água no solo ocorre principalmente em abril e a reposição é feita principalmente no mês de novembro. Há um alto excedente de água no solo em janeiro e dezembro. Em Belo Horizonte há uma deficiência hídrica de 347,7 mm e um excedente hídrico de 498,3 mm. Já Florestal apresenta déficit e excedente com valores menores, de 244,2 mm e 248,8 mm, respectivamente. E Ibirité apresenta um déficit de 272,2 mm e um excedente de 239,2 mm.

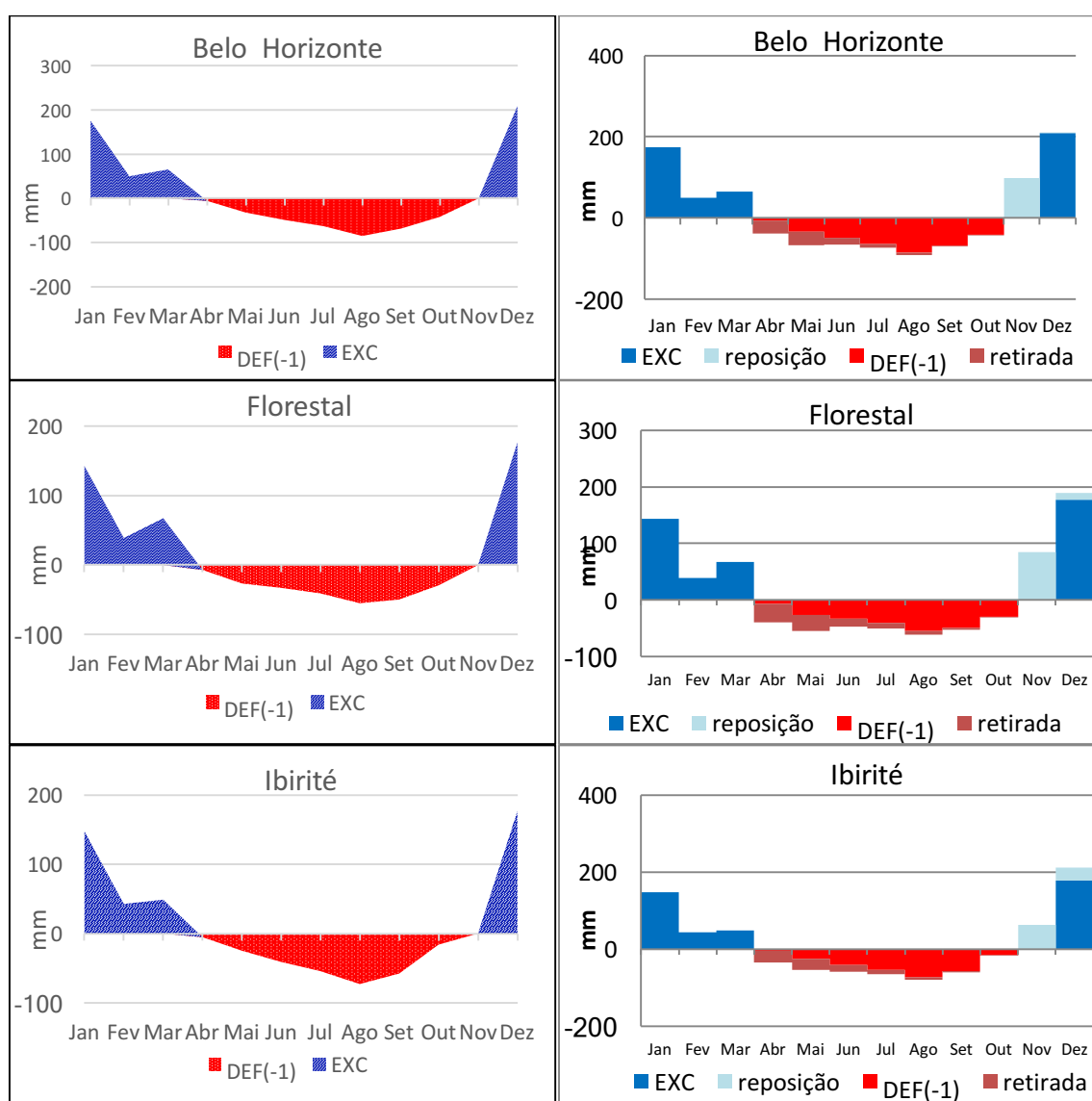


Figura 29 – Balanço hídrico estações RMBH

Fonte dos dados: INMET (2018)

Nas Figura 30 é verificar a umidade relativa da região, onde nota-se uma diminuição da umidade entre julho e outubro, chegando ao mínimo de 57%. Nota-se que Ibirité possui valores de umidade relativa maiores do que as outras regiões.

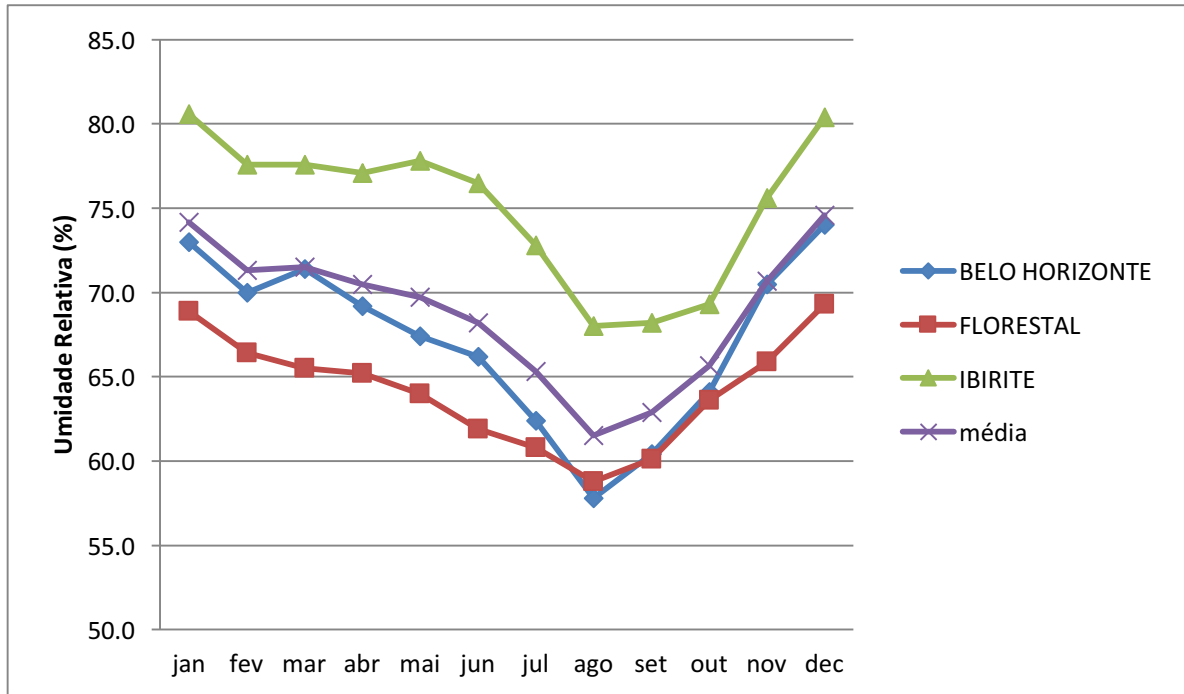


Figura 30 – Umidade Relativa Compensada na RMBH

Fonte dos dados: INMET (2018)

Na Figura 31 nota-se um aumento da pressão atmosférica de abril a outubro, uma relação com a diminuição da umidade relativa e da precipitação.

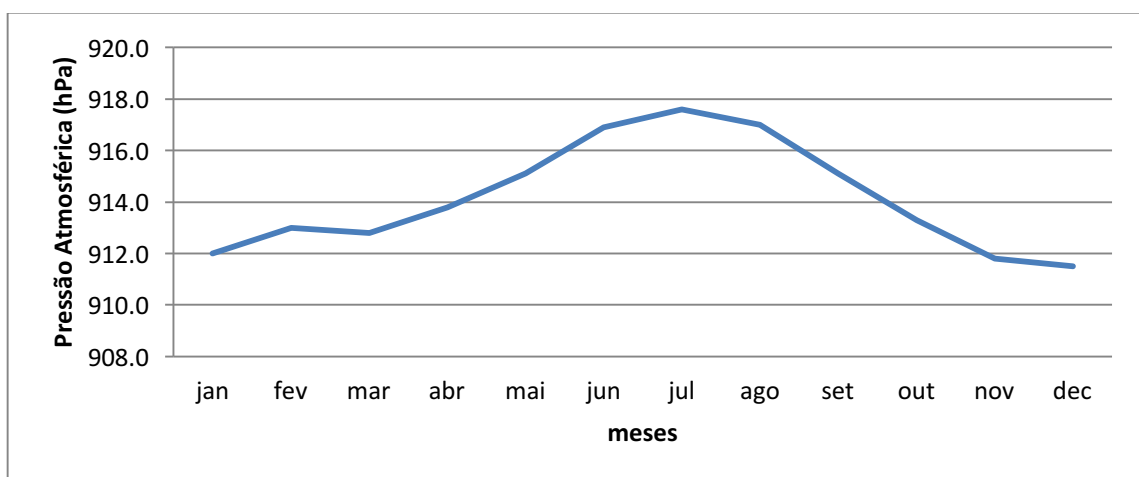


Figura 31 – Pressão Atmosférica na RMBH

Fonte dos dados: INMET (2018)

Nota-se aumento da velocidade do vento de julho a outubro, que também compreende com o período de seca (Figura 32), onde a velocidade do vento média é maior em Belo Horizonte.

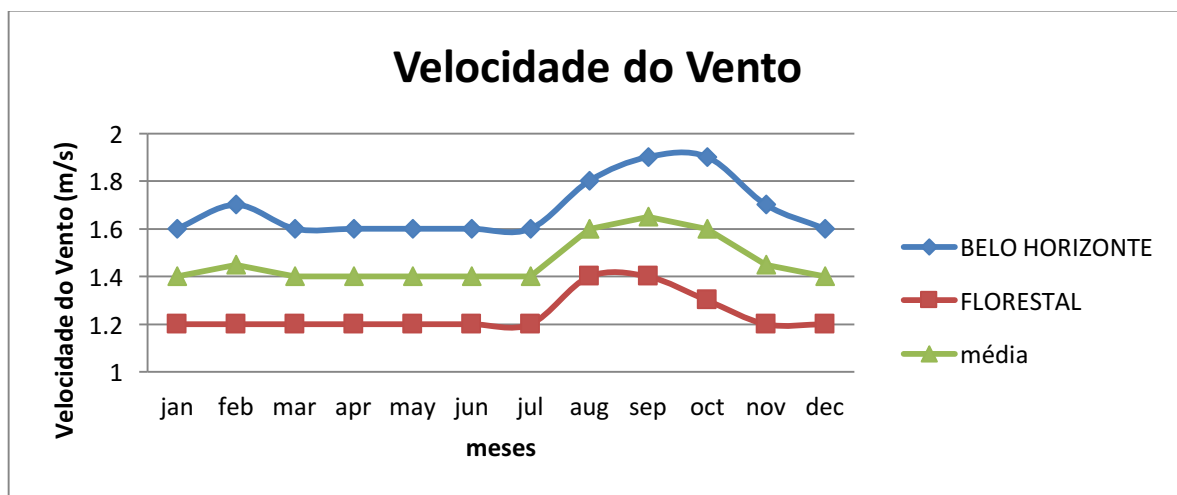


Figura 32 – Velocidade do Vento na RMBH

Fonte dos dados: INMET (2018)

3.3.1. Elementos influenciadores do clima da RMBH

Fatores que podem influenciar o clima de uma região são: continentalidade, topografia e elementos da circulação geral da atmosfera. O efeito da continentalidade faz com que a superfície continental se aqueça e se resfrie mais rapidamente do que a superfície hídrica, interferindo na circulação atmosférica (LUCAS, 2007).

O relevo atua como atenuante de temperatura, pois diminui com a altitude a uma taxa aproximada de $0,6^{\circ}\text{C}$ por 100 m (AYOADE, 2004) e influência no regime de chuvas. A Figura 33 representa a topografia da RMBH, nela podem-se perceber regiões de grandes altitudes a Leste e Norte da RMBH, que são exatamente as regiões classificadas como clima mesotérmico brando, com temperaturas mais amenas. Comparando com a Figura 28 nota-se que o local que apresenta maior precipitação é o local de maiores altitudes, na região do quadrilátero ferrífero.

Moreira (2002) analisou a espacialização de chuvas em Belo Horizonte e seu entorno, analisando o papel da Serra do Curral na climatologia da região. No seu estudo ele aponta que a Serra do Curral serve de amparo para as maiores precipitações, o que pode ser notado na Figura 28. Constata-se que o mês de setembro, mês de transição entre a estação seca e

chuvosa é caracterizada pela influência da topografia, gerando chuvas orográficas. No mês de outubro há acúmulo de chuvas ao longo da serra. Em fevereiro há uma diminuição da precipitação, assim como o acoplamento ao longo da serra. E em março há desarticulação entre a topografia mais elevada e o máximo de precipitação, já que os maiores totais se encontram transversalmente à serra.

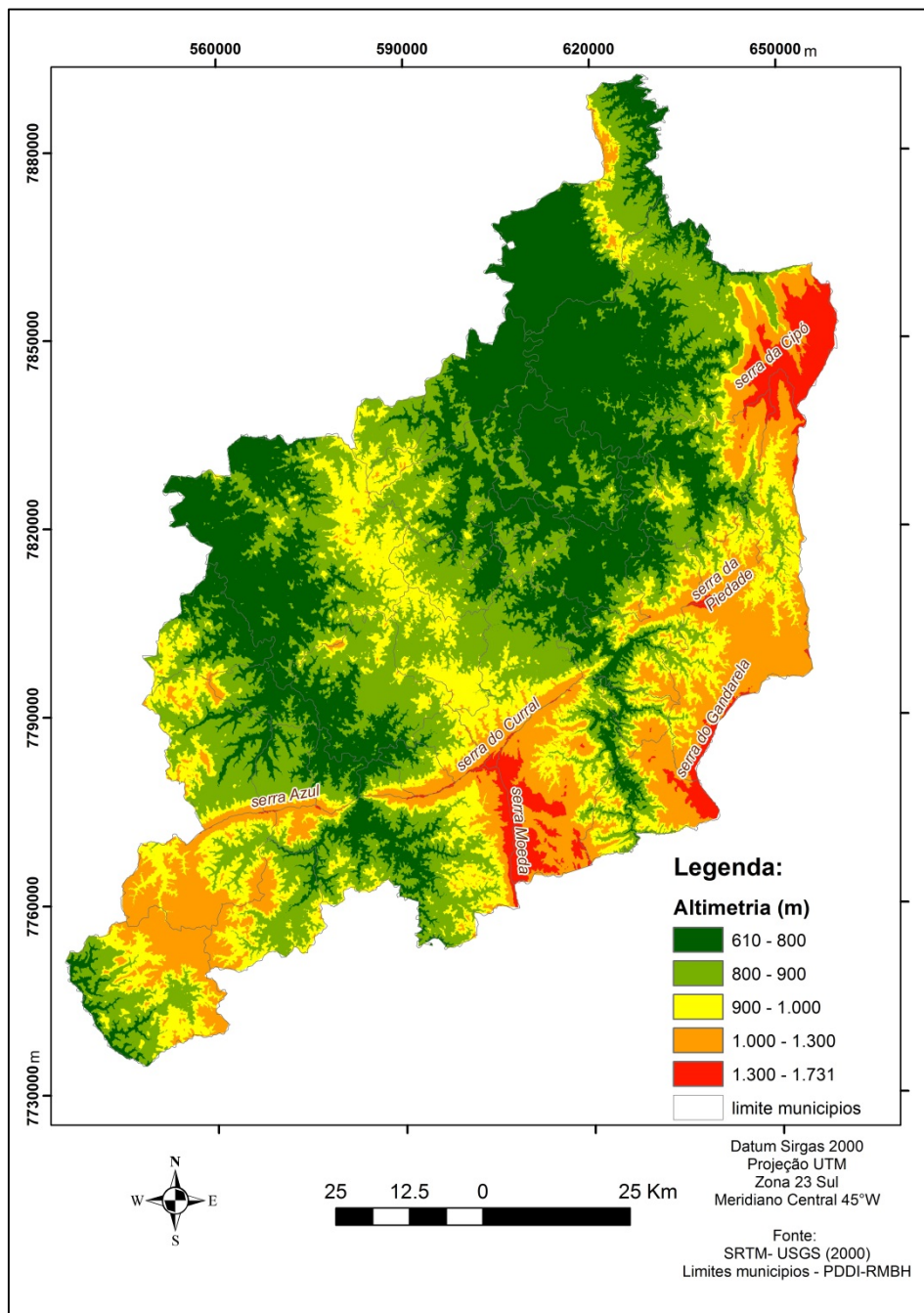


Figura 33 – Altimetria na RMBH

A circulação geral da atmosfera e a interação com sistemas atmosféricos de escala sinótica definem as características básicas do clima, representados na Figura 34. A célula de

Hadley tem como componentes principais: zonas de convergência, anticiclones subtropicais e ventos alísios (ventos úmidos), que sopram de sudeste em direção ao Equador. Na célula de Ferrel está presente duas importantes zonas de convergência extratropical, onde há o encontro do ar quente úmido com o ar frio e seco, de origem polar, originando frentes ou Sistemas Frontais (LUCAS, 2007).

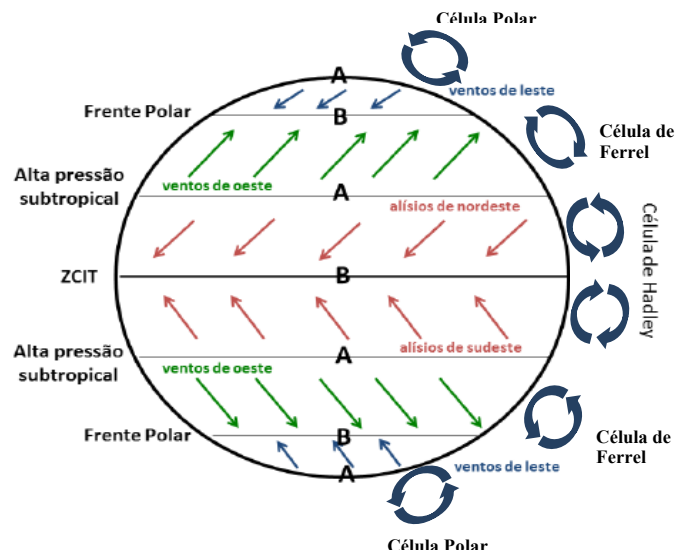


Figura 34 – Circulação Geral da Atmosfera
 Fonte: Adaptado de (REBOITA; AMBRIZZI; PORFÍRIO DA ROCHA, 2012)

A circulação geral da atmosfera origina massas de ar, porções de ar com características uniformes de temperatura e umidade. Na RMBH o verão é caracterizado pela presença da massa equatorial continental, que é formada em áreas de baixa pressão e muita umidade, assim é responsável pela precipitação na região. No inverno há maior atuação da massa tropical marítima e da massa polar marítima, que atuam através de anticiclones. Os anticiclones são centros de alta pressão atmosférica, que resultam em condições de tempo estáveis, céu limpo e baixa umidade (MUSK, 1989). Os principais anticiclones atuantes na região são os Anticiclones Subtropicais do Atlântico Sul (ASAS) e os Anticiclones Polares Migratórios (APM).

No verão, dentro da estação chuvosa, os ASAS se restringem mais ao Oceano Atlântico, sendo que sua ação na RMBH se restringe a eventos de reduções de precipitações e elevação nas temperaturas (NIMER, 1979). No inverno os ASAS se estabelecem no continente, diminuindo a umidade do ar e as precipitações, com predominância de céu limpo. Episódios extremos de baixa umidade relativa no Brasil central estão associados a esse anticiclone, que interrompe a ação de frentes frias (PADILHA; SATYAMURTY, 2004).

Os APM são sistemas que invadem periodicamente o continente sul-americano com ventos de oeste a sudoeste (NIMER, 1979). Caracteriza-se por possuir ar frio, denso e seco. Na passagem do anticiclone no inverno há uma redução da umidade relativa do ar, em conjunto com quedas de temperatura e formação de geadas nas áreas mais elevadas. No verão atua como um dos principais mecanismos para organização da convecção tropical e ocorrência de chuvas frontais (FRANCA, 2009).

Quando duas massas de ar se encontram produzem uma superfície frontal que geralmente tem precipitação associada. Podendo ser frentes frias, que quando estão próximas elevam as temperaturas, seguidas do desenvolvimento de trovoadas e precipitação. Ou frentes quentes, acarretando chuvas leves e contínuas (LUCAS, 2007).

Linhas de instabilidades tropicais são zonas de baixa pressão em forma de linhas alongadas, representando áreas de convergência próximas à superfície. A Zona de Convergência do Atlântico Sul (ZCAS) é o principal fenômeno atmosférico que provoca precipitações no verão (LUCAS, 2007). Definido por Quadro (2005) como uma banda de nebulosidade com orientação noroeste-sudeste, desde a Amazônia até o Atlântico Sul, por pelo menos quatro dias. Apresentando convergência de umidade na baixa e média troposfera alterando o regime de chuvas.

3.3.2. Dados meteorológicos dos anos do estudo

Uma análise detalhada dos dados meteorológicos usados neste trabalho é apresentada no APÊNDICE D. Dessa análise alguns pontos são importantes de destacar.

A pressão atmosférica e a radiação são inversas, sendo a primeira maior nos meses de seca e a segunda mais baixa nos meses de junho e julho. A temperatura média durante os anos ficou entre 24/25°C nos meses de verão e entre 18/19°C nos meses de inverno, enquanto as máximas se apresentaram entre 28/29°C no verão e 22°C no inverno. A umidade relativa seguiu o padrão das normais, alta no verão e baixa no inverno, sendo importante notar que a umidade relativa mínima chega a valores próximos a 40% no inverno, apresentando condições de umidade alarmantes. A velocidade do vento média é bem maior do que a apresentada nas normais, ficando próxima a 5 km/h. Com velocidade do vento máximas próximas de 10 km/h, tendo um pequeno aumento nas rajadas no período do inverno.

Quanto a precipitação observou-se que em 2014 a quantidade de precipitação mensal foi menor que 100 mm em quase todos os meses, apenas em novembro que teve um evento maior de precipitação, com 200 mm de chuva. Isso ocorreu devido a um persistente e anômalo sistema de alta pressão atmosférica, que enfraqueceu a ação de sistemas frontais, bem como a ação do ZCAS (“Infoclima”, [s.d.]). Em 2015 o período de chuva e seca também foi bem acentuado, com precipitações quase nulas nos meses de junho a agosto. O ano de 2016 apresentou precipitações altas em janeiro, novembro e dezembro, com baixas a partir de fevereiro. E 2017 apresentou precipitações bem altas a partir de novembro.

3.4. Análise de Incêndios na RMBH

A baixa umidade, principalmente no inverno, época de seca, faz com que a região esteja propícia a incêndios florestais. Na Tabela 2 é possível notar como é expressivo o tamanho de área queimada em Unidades de Conservação estaduais de Minas Gerais, que ficam dentro e no entorno de Belo Horizonte, entre os anos de 2012 a 2016 (FEAM, 2017).

Tabela 2 - Relação de área queimada em Unidades de Conservação Estaduais, entre 2012 e 2016. (FEAM, 2017)

Unidade de Conservação	Tipo	Área da UC (ha)	Área interna			Área do entorno	
			Total por UC 2012 a 2016	Média de área queimada (ha) por ano	Porcentagem da média de área atingida em relação ao tamanho da UC (%)	Total por UC 2012 a 2016	Média de área queimada (ha) por ano
Serra do Rola Moça	<i>PQ</i>	3 940.00	2168.06	433.612	11.01	2025	405
Serra das Aroeiras	<i>RVS</i>	1 035.41	271.59	90.53	8.74	348.54	116.18
Serra Verde	<i>PQ</i>	142.02	272.99	54.598	38.44	58.66	11.732
Cercadinho	<i>EE</i>	224.89	46.29	9.258	4.12	27.36	5.472
Serra do Sobrado	<i>PQ</i>	383.60	23.32	4.664	1.22	389.22	77.844
Fechos	<i>EE</i>	603.00	1.2	0.24	0.04	13.29	2.658
Baleia	<i>PQ</i>	102.00	0	0	0	26.58	13.29
APA Sul RMBH	<i>APA</i>	163 206.00	3 706.87	741.37	0.45	-	-
Vargem das Flores	<i>APA</i>	12 263.00	93.84	18.77	0.15	-	-

Uma análise de incêndios ocorridos na RMBH foi feita utilizando dados do CBMMG (Corpo de Bombeiros Militar de Minas Gerais) (CBMMG, 2018). Estes dados foram disponibilizados indicando a cidade de ocorrência, tipo de incêndio e data da ocorrência, para o período de 2014 a 2017. Desse banco de dados foram utilizadas as ocorrências classificadas na categoria de Incêndio Florestal, de acordo com Sistema Integrado de Defesa Social (s.d.). Para complementar a discussão dos incêndios foram utilizados também dados de fogo ativo dos sensores MODIS e VIIRS (NASA, 2018), para o mesmo período.

O que pode ser observado na Figura 35, segundo os dados dos bombeiros, é que a maior parte dos registros ocorreram na área urbana e lote vago, em sua maioria nas cidades de Belo Horizonte, Contagem e Betim, cidades mais povoadas da RMBH. Foram assinaladas poucas ocorrências em área rural, agricultura e unidades de conservação. Seguindo o mesmo padrão nos anos de 2015, 2016 e 2017. Em 2014 a classificação dos bombeiros era diferente, então foi só possível comparar os dados de lote vago que ficaram entre 800 e 1200 ocorrências para os quatro anos.

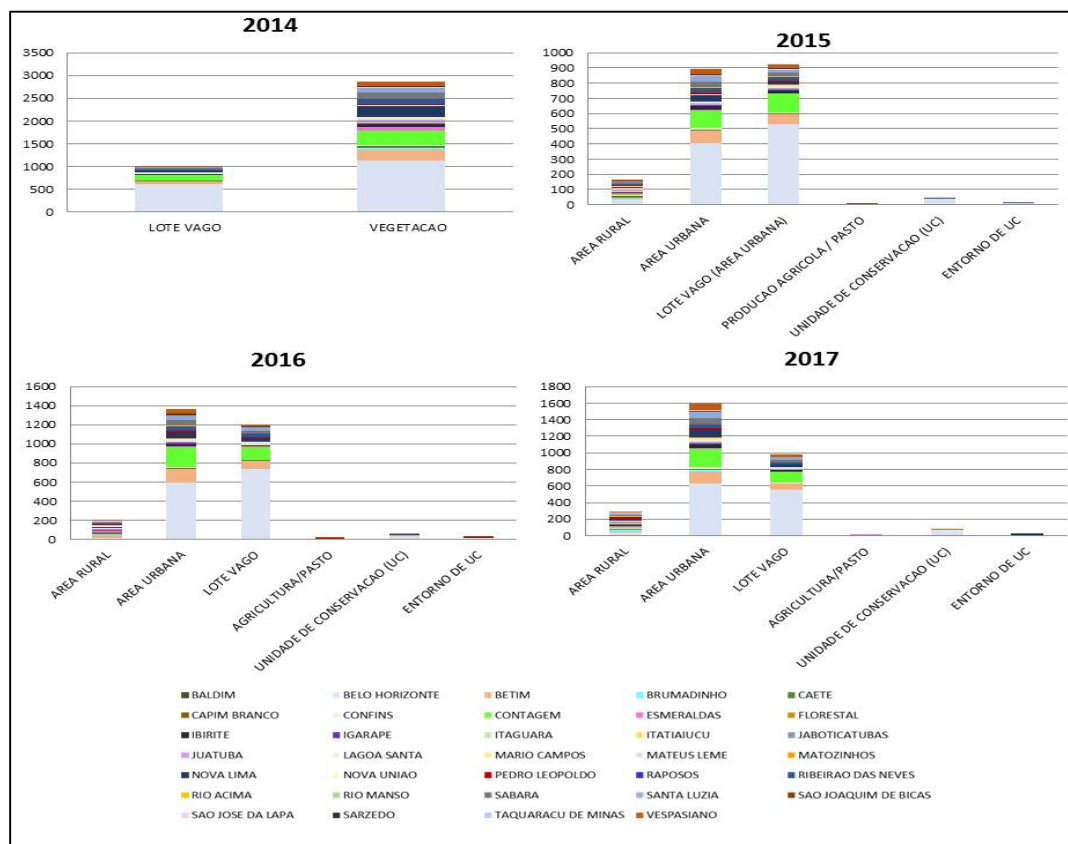


Figura 35 – Comparação entre as cidades da RMBH e as categorias de incêndios, segundo dados da CBMMG
 Fonte: (CBMMG, 2018)

Para efeitos de comparação foram verificados os focos de incêndios de satélite em áreas urbanas e rurais (IBGE, 2010a) e em unidades de conservação (UFMG, 2014). Na Figura 36 é observada uma mudança de cenário sendo que as regiões que mais ocorrem incêndios estão na área rural e em unidades de conservação. No meio rural as cidades com maiores incidências são Pedro Leopoldo e Esmeraldas, Jaboticatubas, com destaque para Brumadinho em 2017 que teve uma grande quantidade de incêndios. Já as cidades que tiveram maiores queimadas em UCs variaram muito em cada ano, com destaque para Baldim e Nova União em 2015 e Brumadinho e Pedro Leopoldo em 2017. Os incêndios em áreas urbanas foram identificados mais em Belo Horizonte e Santa Luzia.

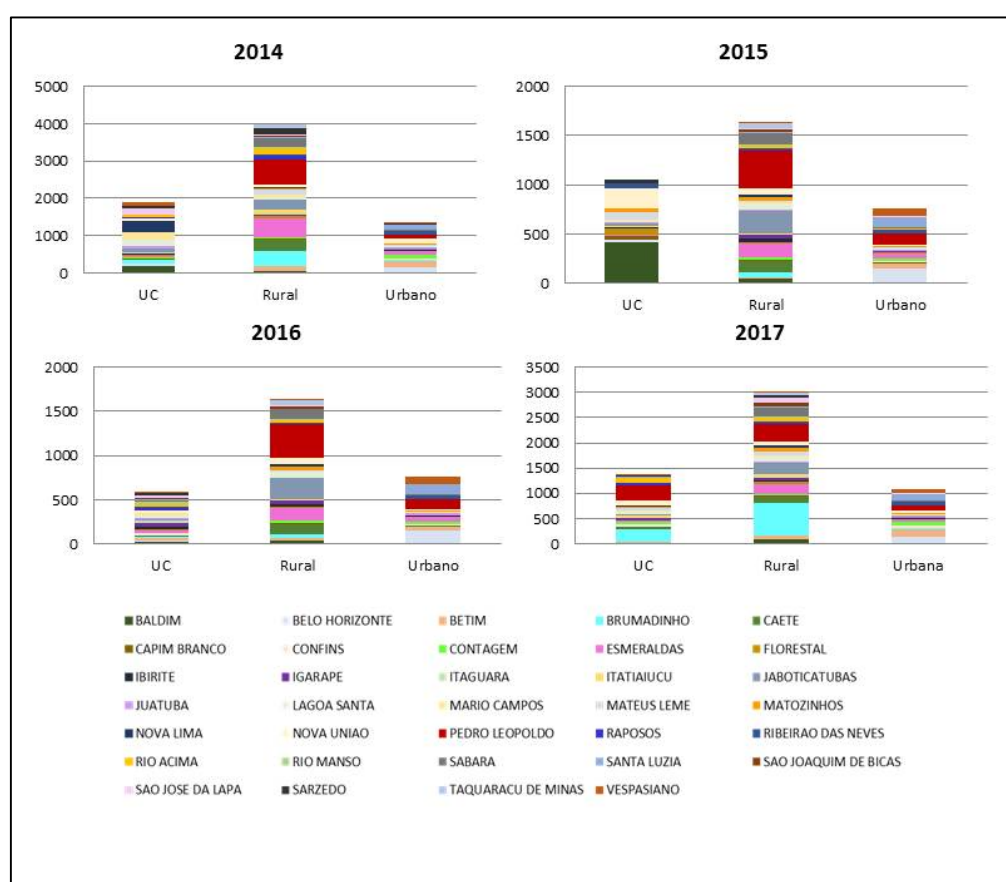


Figura 36 -Comparação entre as cidades da RMBH e as categorias de incêndios, segundo dados de satélite
Fonte: (IBGE, 2010a; NASA, 2018; UFMG, 2014)

É interessante notar que pelos dados dos bombeiros foram identificados números muito maiores de incêndios em áreas urbanas do que os identificados pelo satélite. Isso porque incêndios urbanos são rapidamente contidos e de menor magnitude, portanto mais difíceis de serem identificados pelos satélites. Por outro lado, em áreas mais afastadas, como zonas rurais e unidades de conservação, muitas vezes, a identificação dos incêndios pelos

bombeiros é dificultada, devido à falta de denúncias e, dependendo do local, dificuldade de acesso. Como os remanescentes florestais estão mais concentrados em áreas não urbanas o uso de dados de satélite foi necessário para a identificação dos incêndios.

Na Figura 37 é possível ver os focos de incêndio em cada classe de uso do solo, segundo o MAPBIOMAS (2017), onde a maior quantidade de focos está na classe de floresta, sendo que em 2014 foi o ano de maior incidência nessa classe. Em 2015 teve um destaque na classe de formação Campestre, na cidade de Jaboticatubas. Nota-se uma quantidade de focos considerável em áreas de agricultura e pastagem, principalmente em 2017 devido aos registros no município de Brumadinho. Vale ainda ressaltar que no mesmo ano, a quantidade de incêndios para classe floresta também foi maior para o município de Brumadinho, o que coincide com os dados de queimadas na área Rural e UCs.

Pelo fato da Região Metropolitana estar localizada em uma área de transição de Cerrado e Mata Atlântica, é interessante ressaltar a importância dessa análise. Já que a maior quantidade de incêndios ocorre em áreas florestais e não savânicas. Deve-se considerar ainda, que o cerrado tem um ciclo de incêndios naturais importante para o ecossistema, enquanto em áreas florestais o incêndio é prejudicial, ressaltando a importância desse estudo de risco de incêndios.

Nota-se que a Figura 38, segundo os dados do corpo de bombeiros, há ocorrência de incêndios durante o ano inteiro. Enquanto na Figura 39, segundo dados de satélite, a ocorrência é concentrada entre os meses de julho e outubro, tendo poucas ocorrências nos outros meses. Esse aumento, nesses meses, também é encontrado nos dados dos bombeiros, sendo interessante notar que em outubro a ocorrência foi maior em 2014 e 2015 nas duas bases de dados, como os meses de agosto e setembro que são os meses críticos, o que compreende a época de seca na região. O que pode explicar essas diferenças nas bases de dados é o fato de dados de satélite estarem mais relacionados com incêndios em áreas rurais, que são mais dependentes de fatores climáticos para acontecerem, enquanto que os dados dos bombeiros compreendem mais centros urbanos que tem muita incidência de incêndios causados intencionalmente, não tendo influência de fatores climáticos.

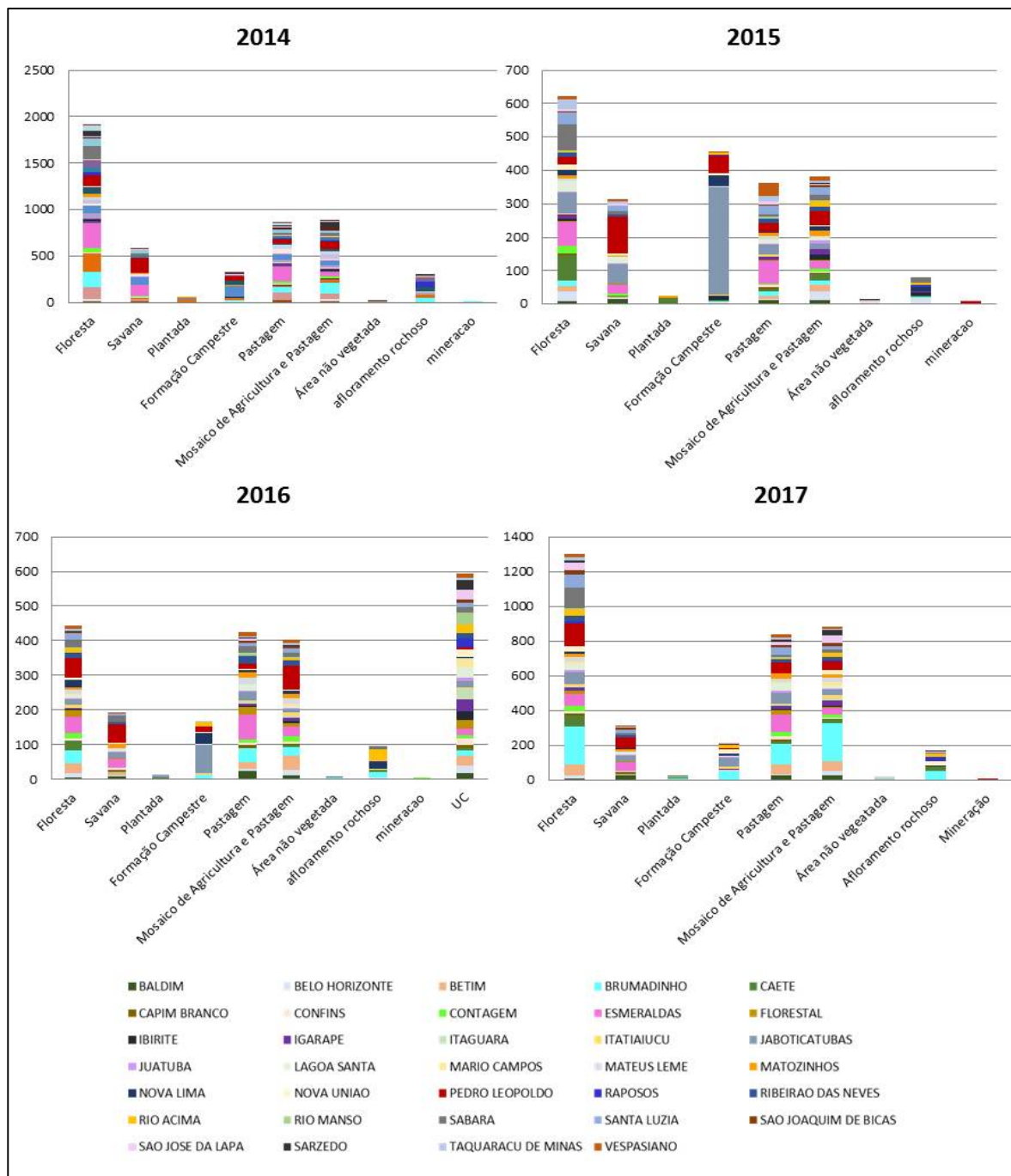


Figura 37 - Comparação entre as cidades da RMBH e as classes de uso do solo, segundo dados de satélite
 Fonte: (MAPBIOMAS, 2017; NASA, 2018)

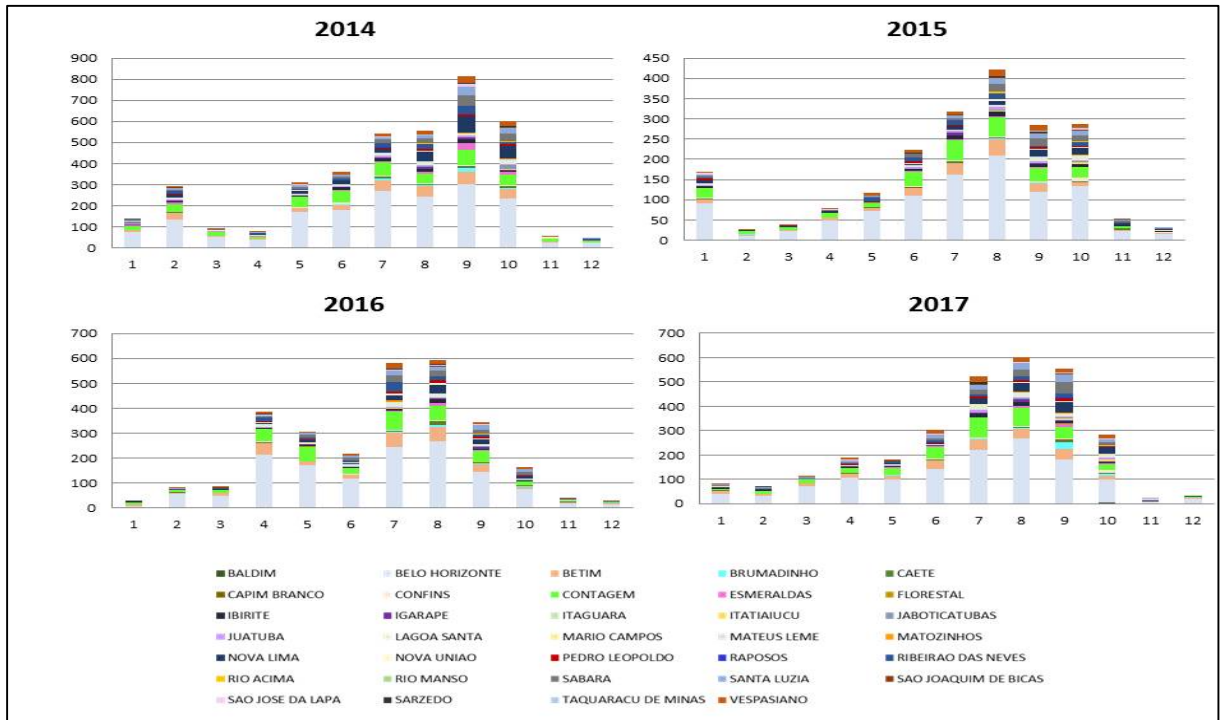


Figura 38-Focos de incêndio por mês na RMBH, para os anos estudados, segundo dados do CBMMG
 Fonte: (CBMMG, 2018)

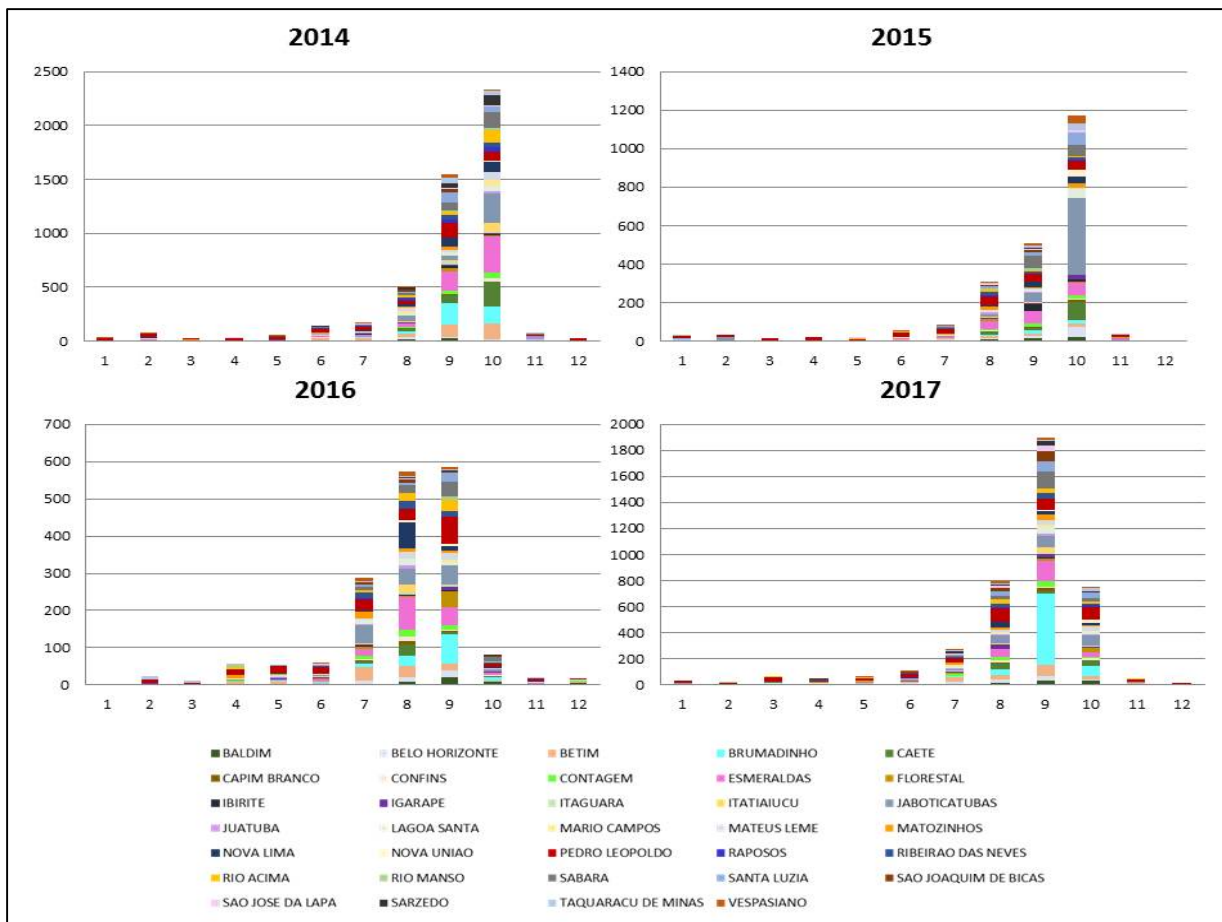


Figura 39-Focos de incêndio por mês na RMBH, para os anos estudados, segundo dados de satélite
 Fonte: (NASA, 2018)

Segundo a Figura 40, é possível observar a quantidade de incêndios por ano, onde observa-se que nas duas base de dados o ano de 2014 foi o ano que teve a maior incidência de incêndios, sendo que o ano de 2017 também se destacou por um alto número de ocorrências. Nota-se mais uma vez que os dados do CBMMG apontam para mais ocorrências em Belo Horizonte, Contagem, Betim e Nova Lima. Enquanto os dados de satélite, para esses municípios, são pouco expressivos, tendo maior incidência em Esmeraldas, Pedro Leopoldo, Jaboticatubas, Caeté e Brumadinho.

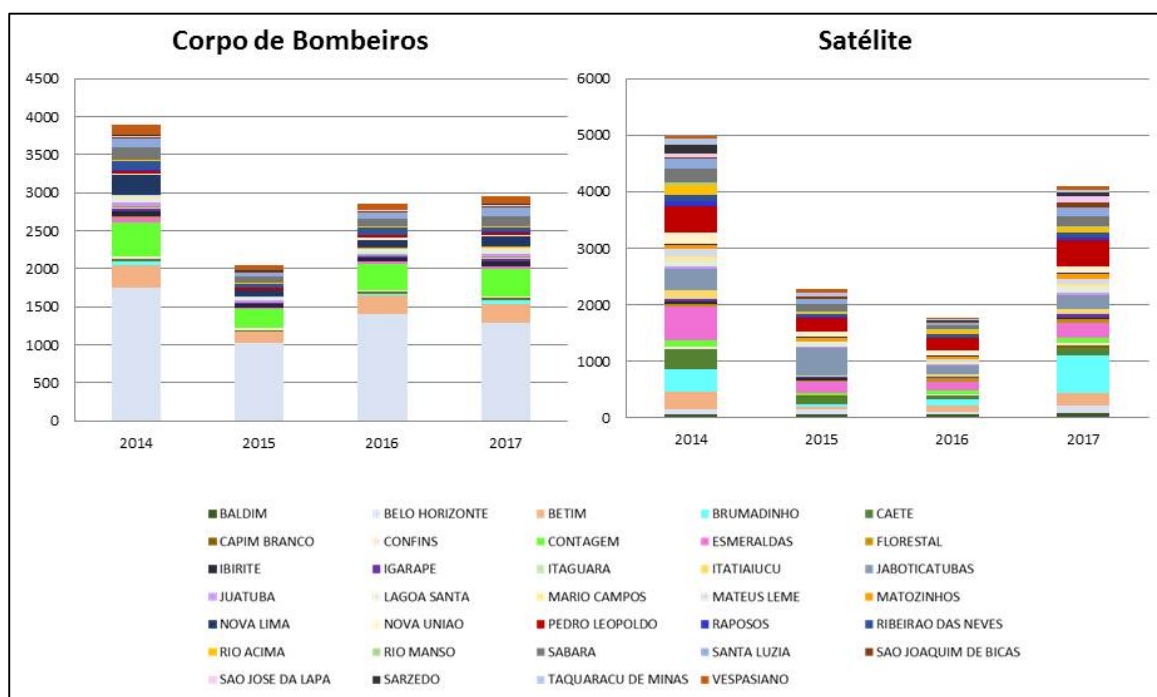


Figura 40 - Dados de incêndio quanto ao ano
Fonte: (NASA, 2018)

Conforme Capítulo 1.4, o FRP é uma medida da taxa de energia emitida pelo fogo por unidade de tempo, em megawatts. Assim, este parâmetro está relacionado com a intensidade do fogo. Na Figura 41, nota-se que os meses que têm incêndios de maior intensidade são de julho a outubro, sendo que os poucos que ocorrem nos outros meses possuem intensidade menor que 5 MW. Na Figura 42, nota-se que a maior parte dos incêndios detectados pelos satélites, tem magnitude pequena, menor que 5 MW e os focos chegam até 50 MW de intensidade. As intensidades maiores estão nos anos de maiores incidências: 2014 e 2017.

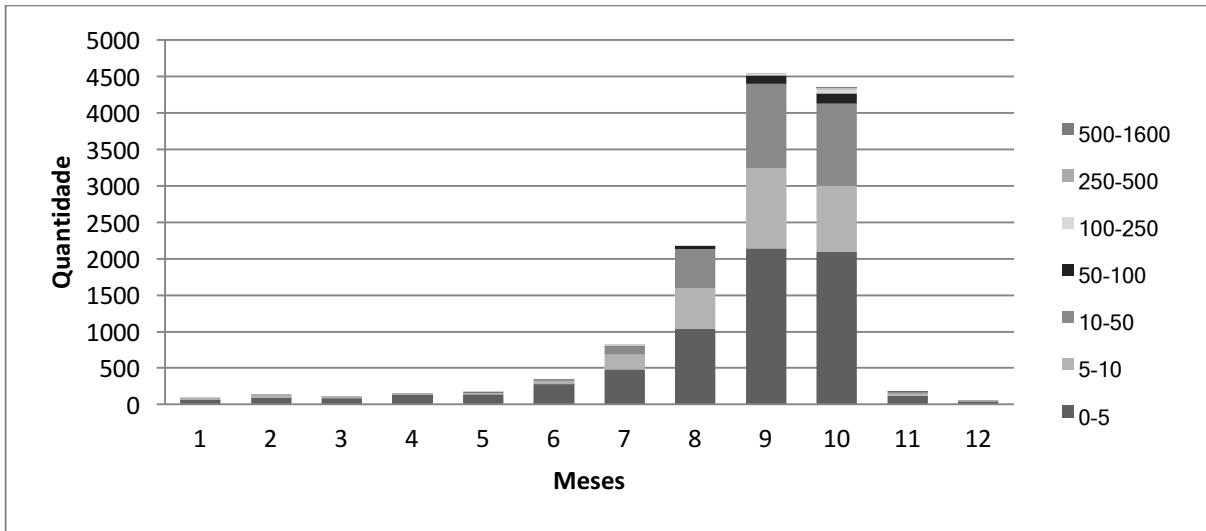


Figura 41 - Dados de FRP (MW) durante os meses
 Fonte: (NASA, 2018)

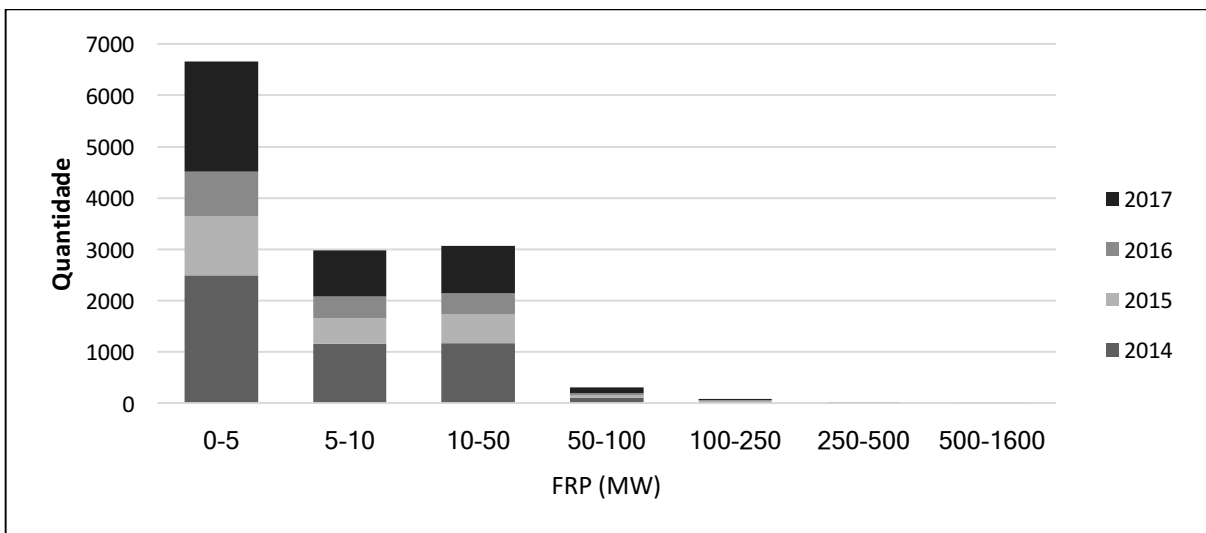


Figura 42 - Dados de FRP (MW) durante o ano
 Fonte: (NASA, 2018)

Conforme citado anteriormente os dados de dois satélites foram utilizados para se esta análise, com os sensores MODIS e VIIRS, como neste trabalho. Comparando os dados obtidos pelos dois satélites nota-se que os dados do VIIRS são os que possuem mais focos identificados, devido a sua melhor resolução espacial, como mostrado na Figura 43. Quanto à intensidade nota-se que o MODIS não detecta incêndios menores, detecta os incêndios que tem FRP entre 10-50 MW, como na Figura 44.

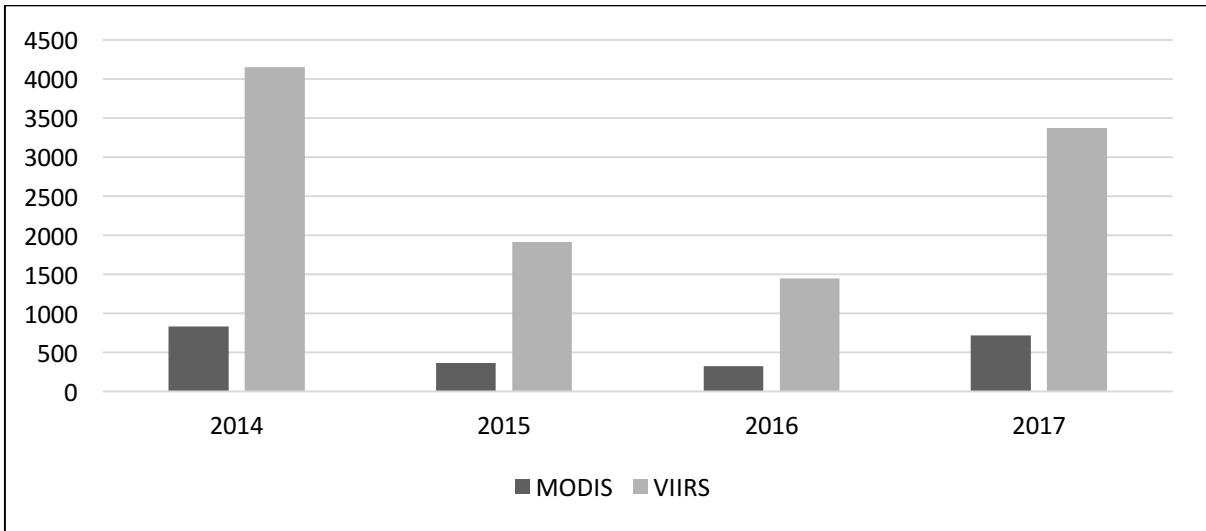


Figura 43 – Comparação entre os dados dos dois satélites
 Fonte: (NASA, 2018)

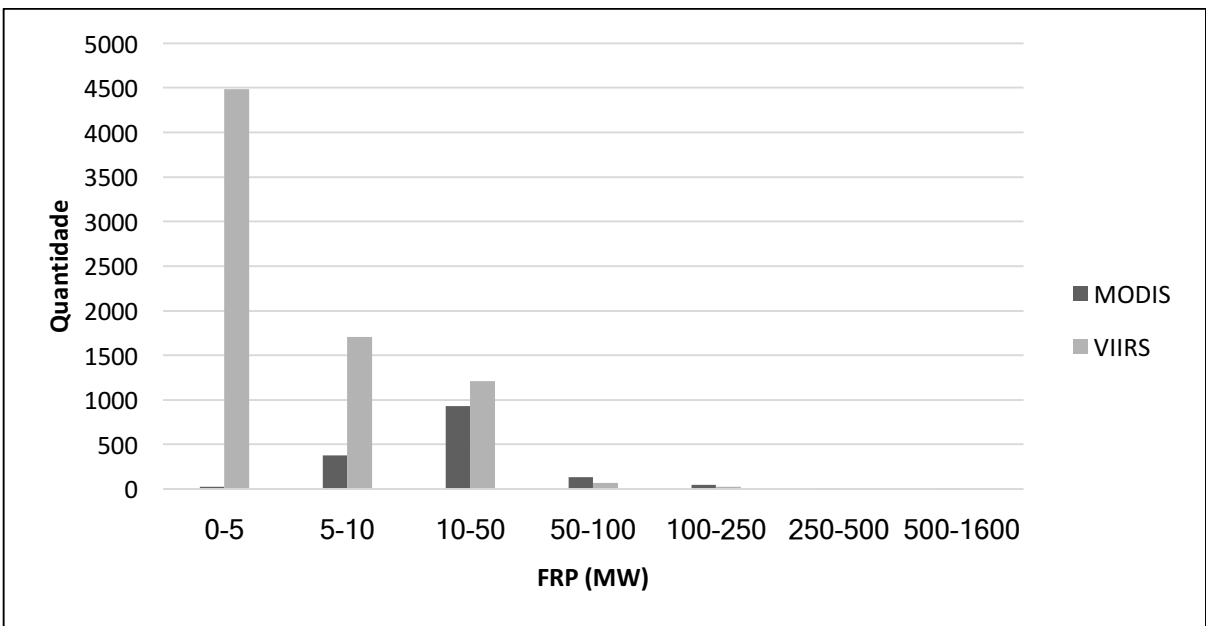


Figura 44 – Comparação entre os dados dos dois satélites quanto ao FRP
 Fonte: (NASA, 2018)

4. Resultados e Discussões

Este capítulo apresenta uma breve estatística descritiva dos dados utilizados, os resultados do MPCA, bem como a configuração final da RNA. Apresenta também os resultados finais da RNA, com seus pesos fixos e sua validação. A validação do modelo e discussões e análises em cima dos resultados.

4.1. Estatística Descritiva dos Dados

Na Tabela 3 são apresentadas as principais medidas de estatística descritiva para os dados de entrada do modelo entre os anos 2014 e 2016.

Quanto aos dados de temperatura máxima é interessante notar que a maior parte dos dados está concentrada próximo da sua média. Nota-se uma grande variação entre os dados de umidade relativa, já que apresenta um desvio de 9,25%, um valor mínimo de 22,98% e um valor máximo de 76,05%.

Quanto à velocidade do vento nota-se que não há grande variação entre os dados, ficando perto da média de 5 m/s. Contudo apresenta eventos de rajadas mais fortes, já que a máxima é de 13,6 m/s e desvio de 0,98 m/s. A precipitação, como esperado, é a que apresenta a maior variação dos dados, já que compreende meses de chuva e seca, chegando a mínimo de 0 mm a 661,59 mm de precipitação por mês.

A radiação apresenta uma média de 1,48 MJ/m² e os valores não variam muito em relação à média, tendo um desvio de 0,19 MJ/m². A pressão apresentou valores de média e mediana próximos de 92 kPa, sendo que a maior parte dos dados se encontram próximos desse valor. No entanto, apresenta valores de mínimo de 87,7 kPa e máximo de 100,25, que representam os extremos de épocas de chuva e seca, respectivamente.

Analisando os dados de distância às rodovias nota-se que a maior parte das regiões tem alguma rodovia por perto, já que a média de distância é de 3,54 km e a mediana 2,07 km. O mesmo ocorre com áreas urbanas em que a sua média é de 3,38 km de distância, sendo que pelos quartis a maioria dos dados está até 5 km de áreas urbanas. O maior valor encontrado para esta variável foi de 20,92 km, o que explica o alto desvio de 3,57 km.

Tabela 3 – Descrição dos Dados

Variáveis de Entrada	Média	Desvio Padrão	Mediana	Mínimo	Máximo	1º Quartil	3º Quartil
Temperatura Máxima (°C)	27,09	2,57	27,16	18,01	35,20	25,42	28,87
Umidade Relativa Mínima (%)	49,64	9,11	49,20	22,98	76,05	42,95	56,22
Velocidade do vento média (m/s)	5,02	0,98	5,00	0,83	13,60	4,56	5,52
Precipitação total (mm)	84,48	96,03	58,81	0,00	661,59	8,18	116,22
Radiação (MJ/m ²)	1,48	0,19	1,49	1,02	1,83	1,33	1,63
Pressão (kPa)	92,28	1,58	92,17	87,70	100,25	91,26	93,20
NDVI	0,34	0,21	0,30	-0,52	1,00	0,18	0,46
Distância a Rodovias (km)	3,54	4,24	2,07	0,00	32,05	0,88	4,47
Distância a áreas urbanas(km)	3,36	3,57	2,28	0,00	20,92	0,61	4,94
Declividade (%)	10,32	6,46	9,13	0,00	64,41	5,41	14,03
Orientação (graus)	179.63	107.16	181.12	0.00	360.00	82.95	274.93

A média da declividade é de 10,13%, sendo que a maioria das declividades são menores que 14%. No entanto, em locais de encostas de declividade mais elevadas, atinge-se 64,41%. Quanto à orientação, a média e mediana são próximos de 180°, de forma que a média das regiões estão voltadas para o Sul.

O teste de Anderson-Darling (ANDERSON; DARLING, 1954) foi aplicado com um nível de significância $\alpha=5\%$ para verificar-se se os dados têm uma distribuição normal. Como

o valor-p encontrado foi inferior a 0,05, há evidências para afirmar que estes não seguem uma distribuição normal. Os resultados do teste são apresentados na Tabela 4.

Tabela 4 - Teste de Normalidade dos Dados

Variáveis	Anderson-Darling	p-value
Temperatura máxima (°C)	25,806	<2,2e-16
Umidade Relativa Mínima(%)	34,096	<2,2e-16
Distância a Rodovias (km)	3250,9	<2,2e-16
Declividade (%)	556,04	<2,2e-16
Distância a áreas urbanas(km)	1864,6	<2,2e-16
Orientação (graus)	661,38	<2,2e-16
Velocidade do vento média(m/s)	412,28	<2,2e-16
Precipitação total (mm)	2242,3	<2,2e-16
Radiação (MJ/m ²)	269,78	<2,2e-16
Pressão (kPa)	95,947	<2,2e-16
NDVI	607,58	<2,2e-16

Além do teste, foram plotados os gráficos Quantil-Quantil para verificar a normalidade dos dados (Figuras 45 e 46), em que uma distribuição normal teórica é confrontada com os dados. Caso a distribuição fosse normal, os valores iriam se encontrar na linha vermelha pontilhada, o que não ocorre na prática em nenhum dos dados.

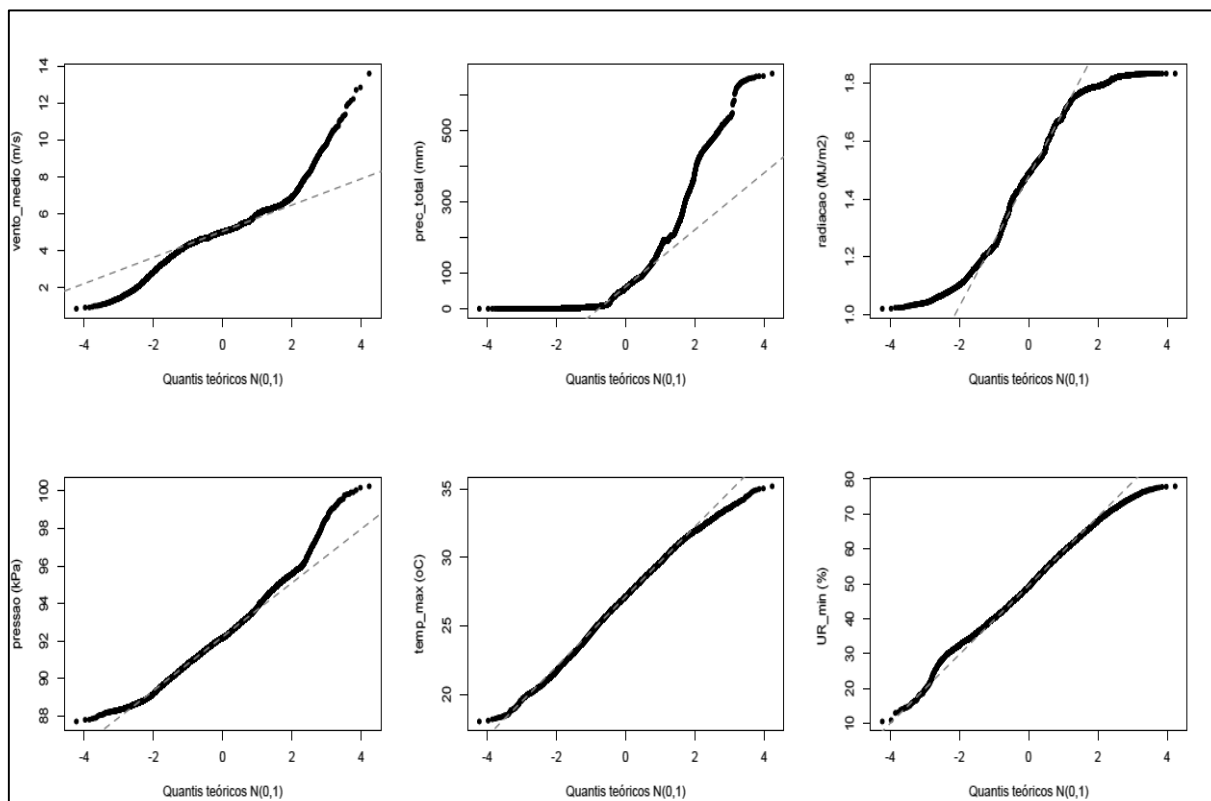


Figura 45 - Gráfico Quantil-Quantil variáveis meteorológicas

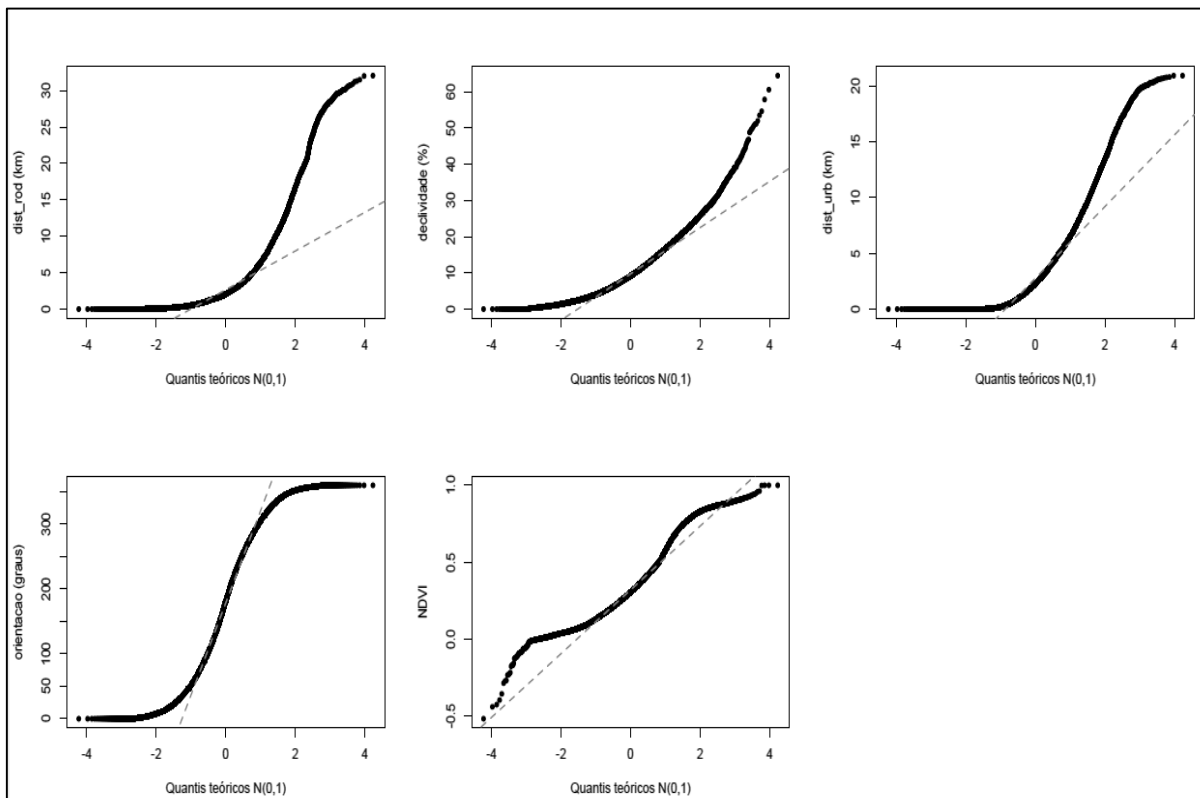


Figura 46 - Gráfico Quantil-Quantil para as variáveis: distancia à rodovias, declividade, distancia a urbanização, orientação do terreno e NDVI

A não normalidade dos dados justifica o uso de RNA, pois é uma técnica que ganha importância para ser usada com dados que se afastam da normal.

4.2. Resultados MPCA-NN

Cada experimento realizado no MPCA-NN resulta em uma configuração de rede diferente. Assim, cada um desses resultados é apresentado na Tabela 5.

Tabela 5 - Resultados MPCA-NN

Parâmetros	Experimento 1	Experimento 2	Experimento 3	Experimento 4
Taxa de Aprendizagem: η	0,0674	0,1185	0,6215	0,1177
Constante de Momento: α	0,6304	0,3811	0,8513	0,8799
Neurônios	6	15	6	13
Função de Ativação	Tanh	Tanh	Tanh	Tanh

*Tanh=tangente hiperbolica

É interessante notar que apesar da maior parte da literatura (ALONSO-BETANZOS et al., 2003; GOLDARAG; MOHAMMADZADEH; ARDAKANI, 2016; SATIR; BERBEROGLU; DONMEZ, 2016) utilizar a função logística como função de ativação para

este tipo de problema, o MPCA retornou como melhor função a ser usada a tangente hiperbólica. Foram realizados testes com as duas funções na RNA onde verificou-se que a função tangente hiperbólica apresentou melhores resultados.

Todos os resultados dos experimentos do MPCA-NN foram testados e o que resultou em melhores resultados foi o Experimento 1.

Notou-se que quando se aumentava o número de neurônios, como no experimento 2 e 4, o erro quadrático médio da rede aumentava muito e a rede não convergia. Outro fator determinante no desempenho da rede foi a taxa de aprendizagem (η), que quando muito alta, como no Experimento 3, incrementava bastante o erro da rede.

Além desses experimentos também foram testadas outras arquiteturas de rede, aplicando-se, por exemplo, o aumento de neurônios para 7 e 8, e a variação da taxa de aprendizagem e constante de momento. Constatou-se em todos os casos o incremento do erro. Assim, o melhor número de neurônios, considerando-se os resultados, foram 6 neurônios, com taxas de aprendizagem mais baixas.

4.3. Resultados da RNA

A rede neural final ficou igual à representada na Figura 47, com 12 neurônios na camada de entrada, 6 neurônios da camada intermediária e 1 neurônio na camada de saída. A função de ativação utilizada nas duas camadas foi a tangente hiperbólica (Equação 9).

As entradas foram escolhidas seguindo as referências na literatura quanto aos fatores de influência nos incêndios. Preferiu-se utilizar os dados de temperatura máxima ao invés da temperatura média, já que é sabido que temperaturas mais altas influenciam na ocorrência de incêndios, bem como a umidade relativa mínima que apresentou valores baixos (próximos de 30-40 %). Assim, quando foram utilizados estes últimos valores ao invés dos parâmetros médios os resultados melhoraram.

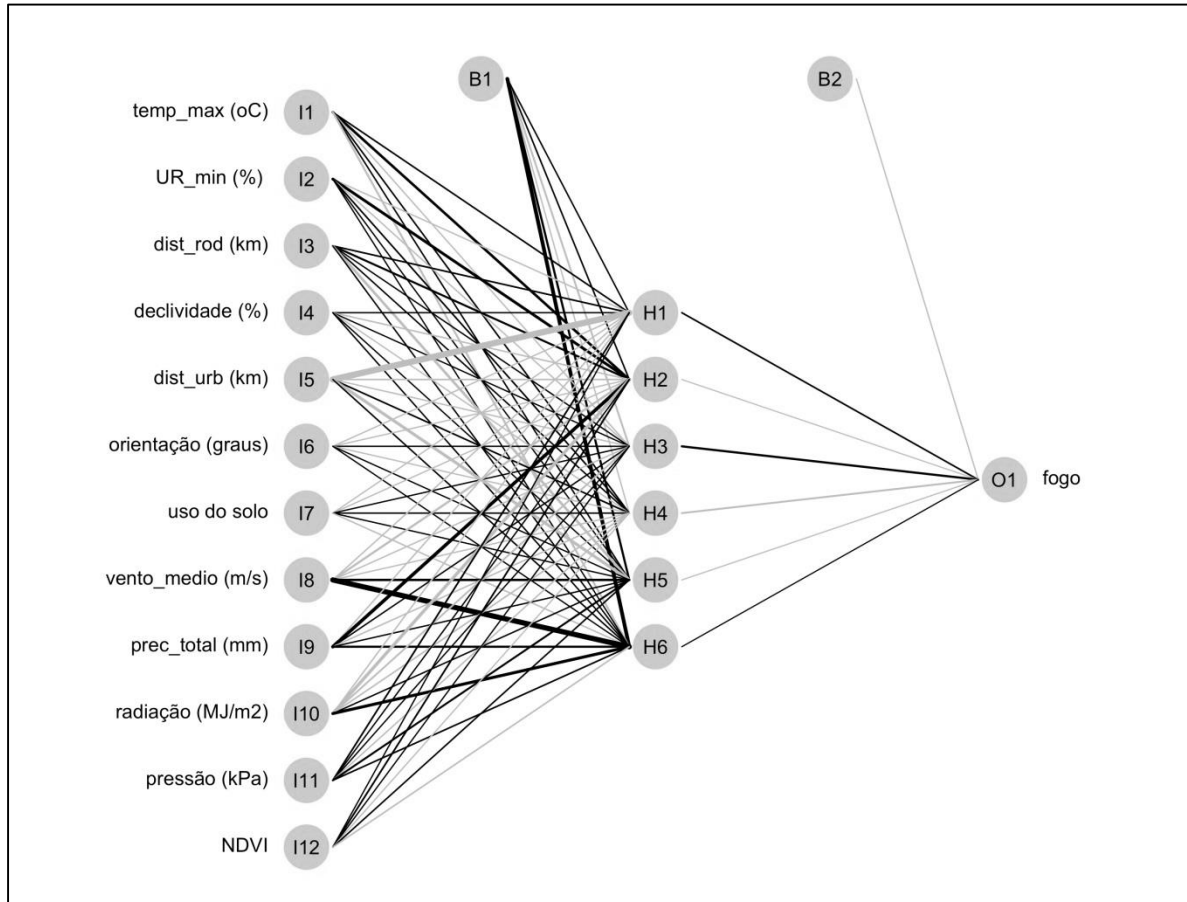


Figura 47 - Representação da RNA utilizada no modelo. Imagem produzida com o pacote R *NeuralNetTools*.

Na representação das conexões (Figura 47) quanto mais espessas as linhas, maiores os valores dos pesos. Nota-se que os dados de Distâncias às áreas urbanas e a velocidade do vento média apresentaram maiores valores de peso.

Como mencionado no item 2.1.7, foram selecionados pontos aleatórios de não ocorrência de incêndio através de janelas. Notou-se que quando essas janelas eram muito pequenas, como as janelas de 1 km e 5 km, o número de pontos de ausência de incêndio era consideravelmente maior que os pontos de incêndio e conseqüentemente, a rede não conseguia reconhecer os pontos de risco alto, apenas o risco baixo. Já quando se diminuía muito a janela, como a janela de 20 km, o contrário acontecia, a rede apenas reconhecia os pontos de risco mais alto. A janela de 10 km foi a que obteve melhores resultados, pois conseguiu reconhecer tanto o risco mais alto, quanto o risco mais baixo.

A rede convergiu pelo número máximo de épocas (2000 épocas), com um erro de 0,063. A variação do erro é mostrada na Figura 48. O tempo de treinamento foi de 3,47 horas.

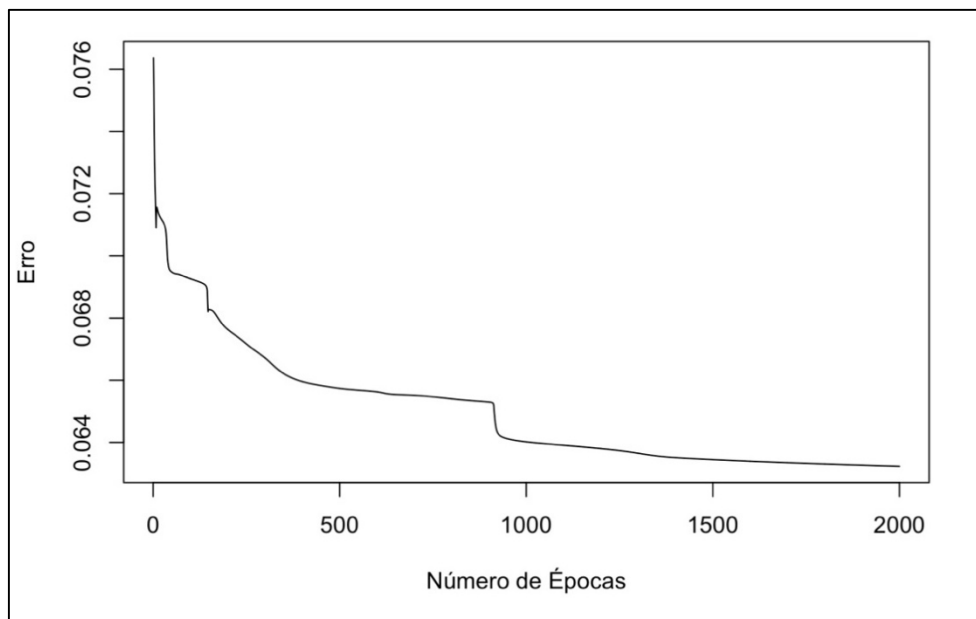


Figura 48 - Variação do Erro de treinamento

4.3.1. Avaliação da Rede

Na Tabela 6, são apresentados os parâmetros de avaliação da rede. A acurácia é a probabilidade do modelo de fornecer resultados corretos, tanto positivos, quanto negativos. A acurácia da RNA no presente trabalho atingiu aproximadamente 80%. A Sensibilidade e a Predição Valores Verdadeiros estão relacionadas com a capacidade da rede de prever os valores positivos dessas medidas, os quais apresentaram valores mais baixos do que o esperado, ficando entre 50-60%. A Especificidade e a Predição Valores Negativos medem a capacidade do modelo de prever valores negativos. Como esses resultados próximos de 90%, justifica-se o valor da acurácia.

Tabela 6 – Dados de Avaliação da Rede

	Treinamento	Validação	Generalização
Acurácia (A)	84%	80%	78%
Sensibilidade (S)	64%	54%	53%
Especificidade (E)	89%	91%	90%
Predição Valores Verdadeiros (PVV)	64%	69%	71%
Predição Valores Negativos (PVN)	90%	83%	81%
AUC	0,87	0,85	0,85

Já a medida do AUC confronta a taxa de falsos positivos com a taxa de verdadeiros positivos, sendo que quanto mais próximo de 1 melhor seu resultado. Os valores de AUC ficaram próximos de 0,85, indicando um bom desempenho do modelo. As curvas ROC são apresentadas na Figura 49.

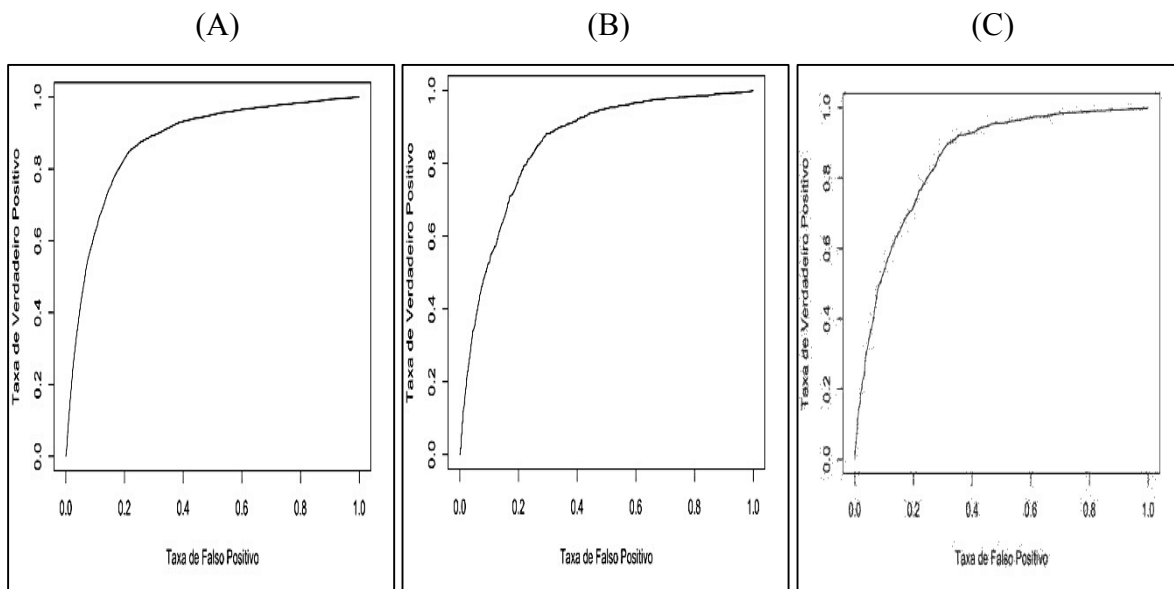


Figura 49 - Curvas ROC: (A) Treinamento (B) Validação (C) Generalização

Apesar dos valores de Sensibilidade e PPV não estarem muito altos, nota-se que mais de 60% dos pontos foram classificados nas classes muito alto e alto, e menos de 20% foi classificado como baixo e muito baixo, indicando bons resultados da rede.

Tabela 7 - Percentagem de focos de incêndios classificados quanto ao grau de risco

	Treinamento	Validação	Generalização
Muito Alto (MA)	11,7%	17,4%	16,7%
Alto (A)	40,3%	41,7%	46,1%
Médio (M)	27,7%	22,4%	19,7%
Baixo (B)	8,4%	7,4%	8,7%
Muito Baixo (MB)	11,9%	11,0%	8,8%

4.4. Validação

Depois de finalizada, a RNA foi utilizada com os dados mensais de 2017. Os resultados são apresentados nas Figuras 50 e 51.

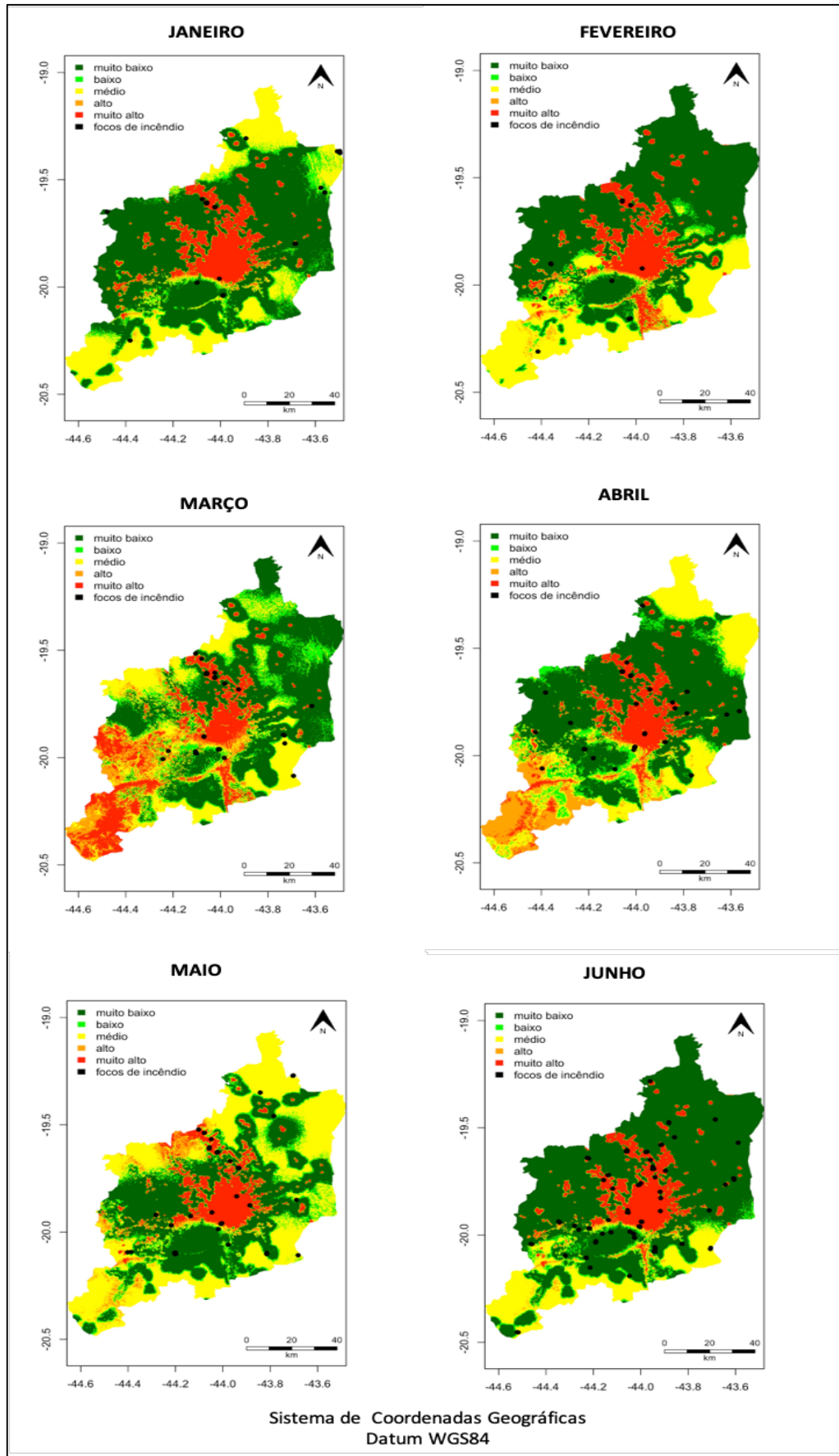


Figura 50 - Resultados do modelo para os dados do 1º semestre de 2017. Mapas produzidos em ambiente R.

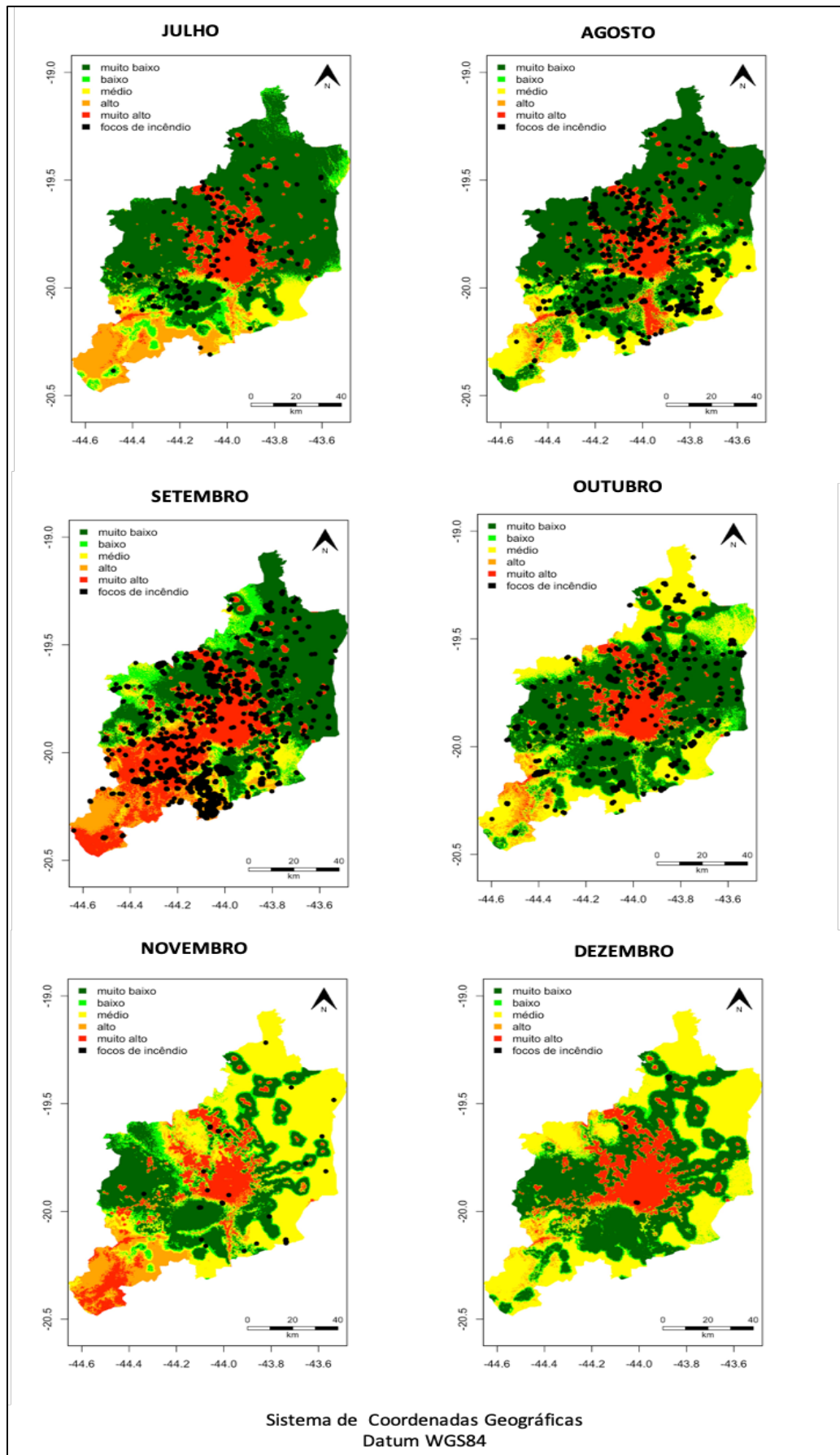


Figura 51 - Resultados do modelo para os dados do 2º semestre de 2017. Mapas produzidos em ambiente R

A partir do resultado apresentado é possível verificar que, em todos os meses, regiões próximas às áreas urbanas possuem um risco muito alto de incêndio e isso é coerente com a presença dos focos de incêndio na região e com os dados apresentados no Item 3.4. As regiões mais afastadas de áreas urbanas (áreas essencialmente rurais) e mais próximas ao limite da RMBH, também apresentaram riscos altos e médios de incêndio, o que pode estar associado a queimas utilizadas na agricultura, ou apenas com os próprios fatores meteorológicos associados ao tipo de vegetação na região. A classificação “risco muito alto” foi identificada dentro das áreas urbanas em todos os meses; no entanto, nos meses de setembro, março e outubro, essa classificação foi verificada também em outras áreas. As variações observadas entre os meses não estão relacionadas com o fator precipitação, pois nas duas estações (seca e chuva) foram observados locais de risco muito alto.

Foi verificado a incidência dos focos de incêndio em relação às áreas classificadas, mostrado na Tabela 8, onde é possível perceber que nos meses de janeiro a junho e dezembro os focos incidiram em sua maioria em áreas classificadas como risco muito alto. Já nos meses de julho a outubro a maior parte dos focos de incêndio foi classificada como risco muito baixo. Desta forma, o modelo não funcionou muito bem para os meses de seca. Quanto aos meses de chuva, apesar da classificação estar correta, a quantidade de áreas identificadas como risco alto é maior do que o esperado para essa época. Devido a esses resultados, foi testada uma rede separada para cada estação para verificar o comportamento da RNA.

Tabela 8 - Percentagem de Focos de incêndio em cada classe de risco

Mês	Muito Alto	Alto	Médio	Baixo	Muito Baixo
Janeiro	58%	3%	18%	0%	21%
Fevereiro	56%	0%	13%	0%	31%
Março	58%	2%	18%	7%	15%
Abril	42%	12%	9%	12%	26%
Mai	61%	2%	5%	11%	21%
Junho	52%	1%	5%	4%	38%
Julho	29%	5%	5%	10%	51%
Agosto	21%	5%	13%	5%	56%
Setembro	27%	23%	9%	12%	29%
Outubro	11%	4%	17%	9%	59%
Novembro	36%	9%	34%	0%	21%
Dezembro	69%	0%	0%	0%	31%

4.5. Avaliação Rede Neural: Modelo época de Seca

Foram selecionados como meses de seca os meses de déficit hídrico (maio a outubro), que também compreendem os meses de maior incidência de focos de incêndio. O treinamento da rede levou 1,89 horas, com um erro médio quadrático de 0,071 (Figura 52).

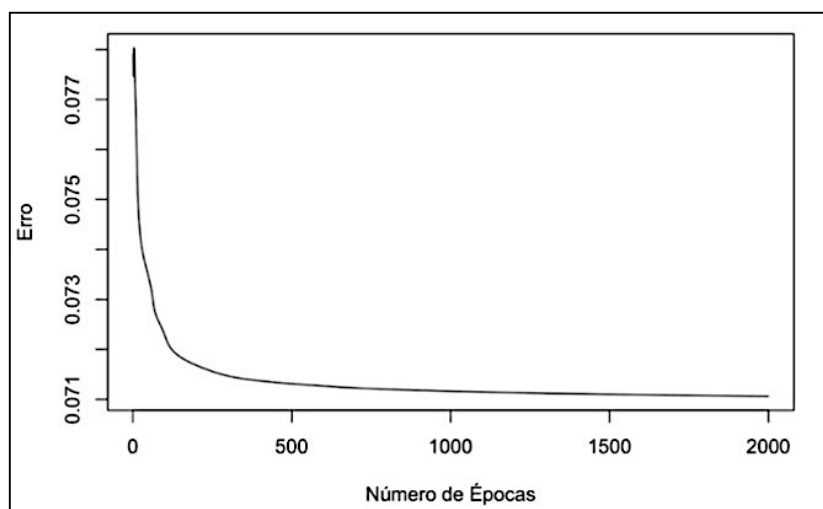


Figura 52 - Variação do Erro modelo seca

Na Tabela 9 é possível notar que, quando usado o modelo separado para a época seca, a acurácia diminuiu, mas continua satisfatória. A sensibilidade diminuiu, enquanto a especificidade aumentou; isso pode ser explicado devido ao número de predição de valores positivos que aumentou e a predição de valores negativos que diminuiu. Para esta rede a janela de 10 km foi mantida para a seleção de pontos de não ocorrência de incêndio. Como a maior parte dos focos de incêndio ocorreu nesse período, o número de focos em relação a pontos de não ocorrência de incêndio foi maior que o modelo geral, o que pode ter causado esses resultados. Já o parâmetro de AUC não variou muito, ficando próximo de 0,85, indicando um bom modelo, sendo as curvas ROC apresentadas na Figura 53.

Tabela 9 - Dados de Avaliação do Modelo

	Treinamento	Validação	Generalização
Acurácia (A)	78%	77%	76%
Sensibilidade (S)	54%	50%	50%
Especificidade (E)	92%	91%	92%
Predição Valores Verdadeiros (PVV)	78%	75%	79%
Predição Valores Negativos (PVN)	78%	77%	75%
AUC	0,86	0,84	0,84

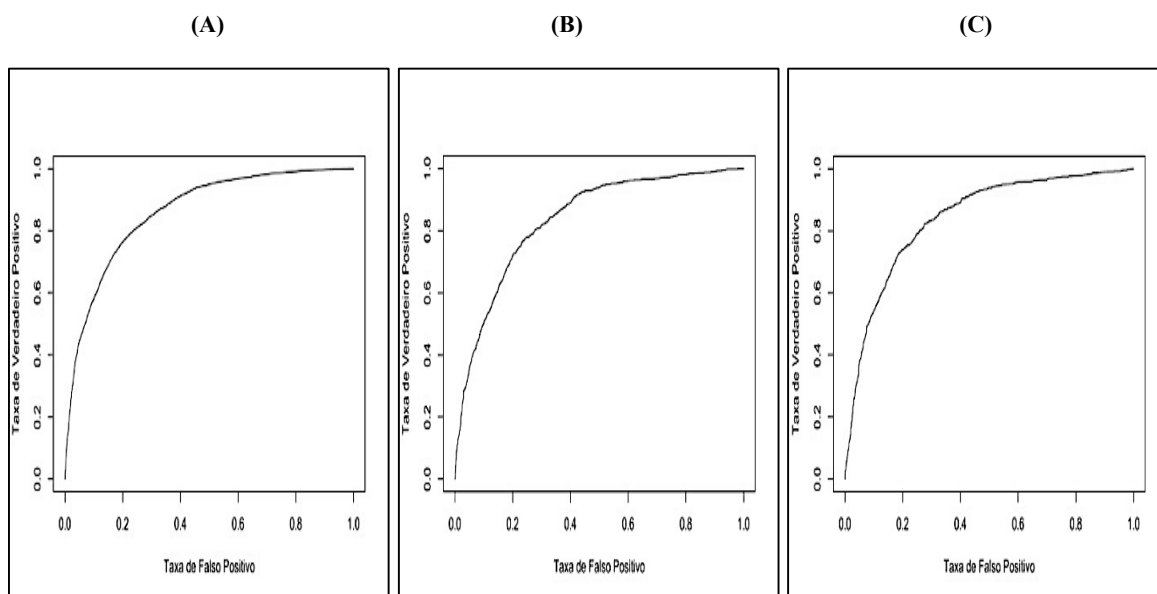


Figura 53 - Curvas ROC modelo seca: (a) Treinamento (b) Validação (c) Generalização

Na Tabela 10 pode ser visto que pelo menos 60% dos focos ocorreram em áreas classificadas com risco alto ou muito alto, tendo em média 6% classificados no risco baixo, indicando a boa capacidade da rede em prever o risco de incêndio.

Tabela 10 - Porcentagem de focos de incêndios classificados quanto ao grau de risco modelo seca

	Treinamento	Validação	Generalização
Muito Alto (MA)	40,62%	40,35%	40,89%
Alto (A)	20,67%	18,68%	22,90%
Médio (M)	32,99%	34,25%	29,26%
Baixo (B)	5,72%	6,60%	6,83%
Muito Baixo (MB)	0,00%	0,12%	0,12%

4.6. Avaliação Rede Neural: Modelo Época de chuva

Os meses de chuva considerados foram janeiro, fevereiro, março, abril, novembro e dezembro. O treinamento da rede levou 1,21 horas, convergiu com o número máximo de épocas e obteve um erro médio quadrático de 0,019 (Figura 54).

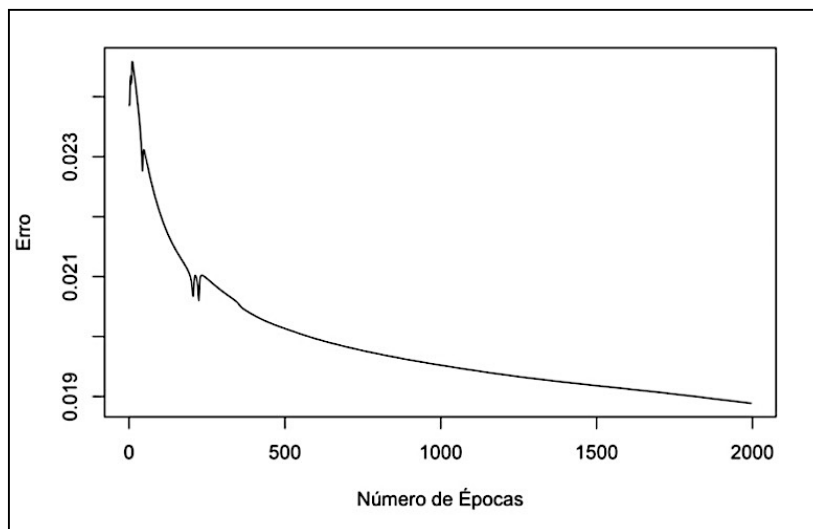


Figura 54 - Variação do Erro para modelo de chuva

Na Tabela 11 é possível notar um aumento da acurácia, da especificidade e do PVN, pois os valores de não ocorrência de incêndio foram muito bem reconhecidos. Por outro lado, a sensibilidade e a predição de valores verdadeiros diminuíram, ou seja, a rede não teve uma resposta satisfatória para os pontos de ocorrência de incêndio. Tal fato pode estar relacionado ao número de pontos de ocorrência de incêndio ser muito pequeno em relação aos pontos de não ocorrência de incêndio, e realmente nessa época ocorrem poucos focos. No entanto a RNA aplicada para esse modelo conseguiu prever muitos pontos, da forma que seu desempenho não foi prejudicado. O AUC foi de 0,84, mas no conjunto de validação este valor reduziu para 0,78. As curvas ROC se encontram na Figura 55.

Tabela 11 - Dados de Avaliação do Modelo

	Treinamento	Validação	Generalização
Acurácia (A)	95%	94%	95%
Sensibilidade (S)	46%	44%	51%
Especificidade (E)	97%	96%	98%
Predição Valores Verdadeiros (PVV)	47%	35%	49%
Predição Valores Negativos (PVN)	97%	98%	98%
AUC	0,84	0,78	0,83

Na Tabela 12, os focos estão distribuídos em todas as classes de risco, tendo porcentagem alta tanto nas classes de risco mais alto quanto nas classes de risco mais baixo.

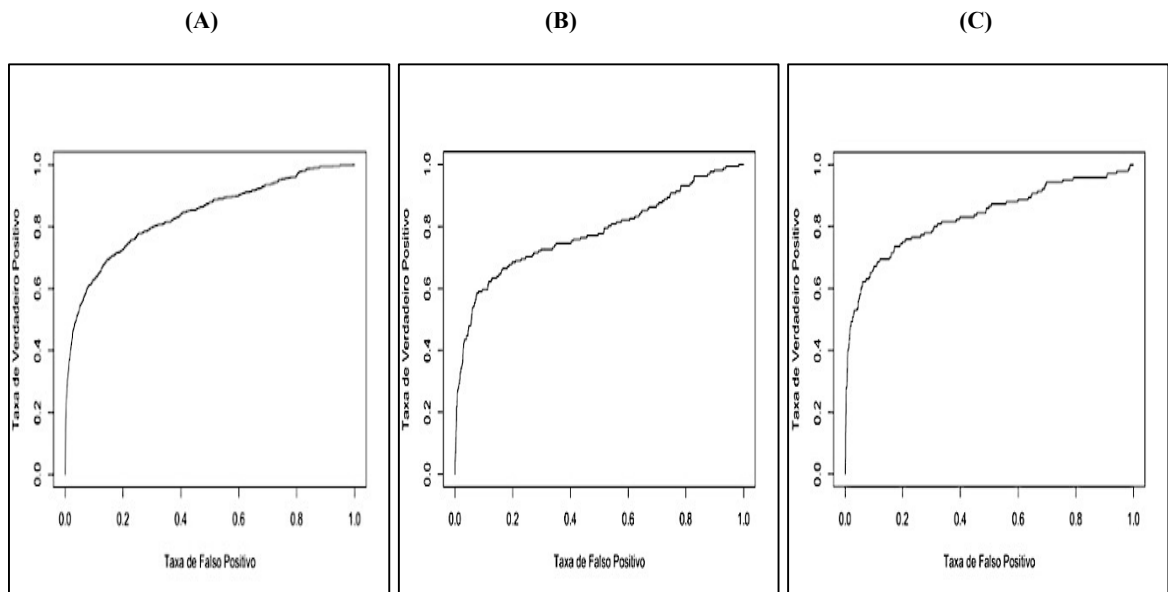


Figura 55- Curvas ROC modelo chuva: (A) Treinamento (B) Validação (C) Generalização

Tabela 12 - Porcentagem de focos de incêndios classificados quanto ao grau de risco modelo chuva

	Treinamento	Validação	Generalização
Muito Alto (MA)	32%	23%	34%
Alto (A)	9%	9%	10%
Médio (M)	10%	10%	9%
Baixo (B)	17%	18%	16%
Muito Baixo (MB)	32%	40%	31%

4.7. Validação dos modelos

Na Tabela 13 é possível notar que a maior parte dos focos ocorreu em áreas classificadas como muito alta. Apenas o mês de novembro teve uma discrepância, isso devido ao fato do mês de novembro ainda estar próximo aos meses de seca. Pela Figura 56 observa-se que as áreas classificadas com risco alto são exatamente as áreas próximas a centros urbanos.

Tabela 13 - Classificação dos focos quanto à classe de risco modelo chuva

Mês	Muito Alto	Alto	Médio	Baixo	Muito Baixo
Janeiro	61%	0%	0%	0%	39%
Fevereiro	56%	0%	0%	0%	44%
Março	71%	7%	0%	0%	22%
Abril	70%	0%	5%	0%	26%
Novembro	40%	6%	0%	0%	53%
Dezembro	69%	0%	0%	0%	31%

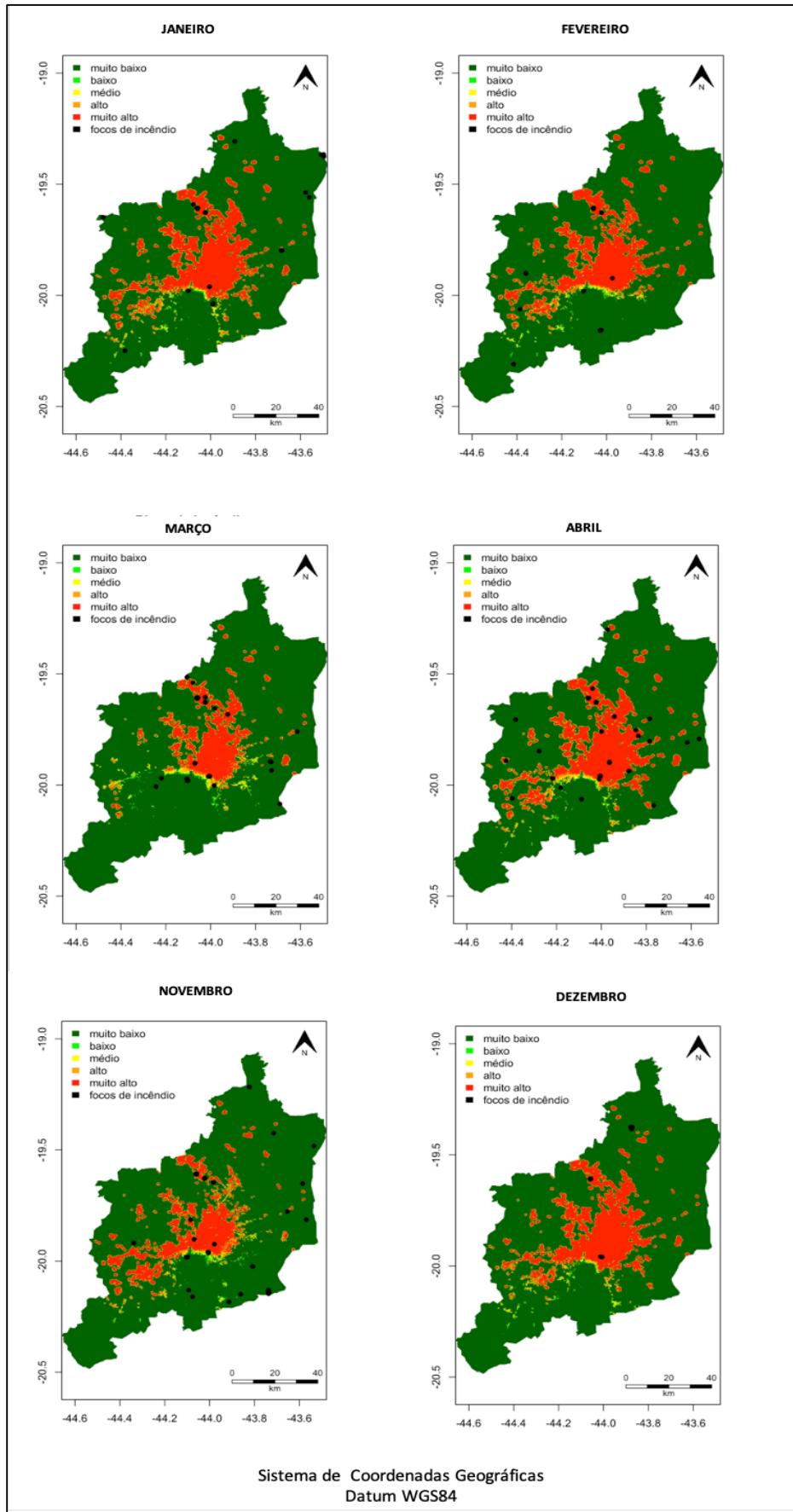


Figura 56 - Resultados modelo chuva. Mapas produzidos em ambiente R.

Na Tabela 14, pode ser observado que quase todos os focos de incêndio são previstos em áreas de risco alto ou muito alto, e a porcentagem em áreas de risco baixo é muito pequena. Na Figura 57, observa-se que o modelo muda bastante dependendo do mês, assim há outras variáveis influenciando no modelo além da área urbana. Mas também é possível notar que grande parte da área é classificada como risco mais alto em alguns meses.

Tabela 14 - Classificação dos focos quanto a classe de risco modelo seca

Mês	Muito Alto	Alto	Médio	Baixo	Muito Baixo
Maio	64%	14%	21%	0%	0%
Junho	64%	14%	21%	0%	0%
Julho	91%	3%	6%	0%	0%
Agosto	72%	20%	5%	1%	1%
Setembro	52%	22%	25%	1%	0%
Outubro	40%	41%	15%	4%	2%

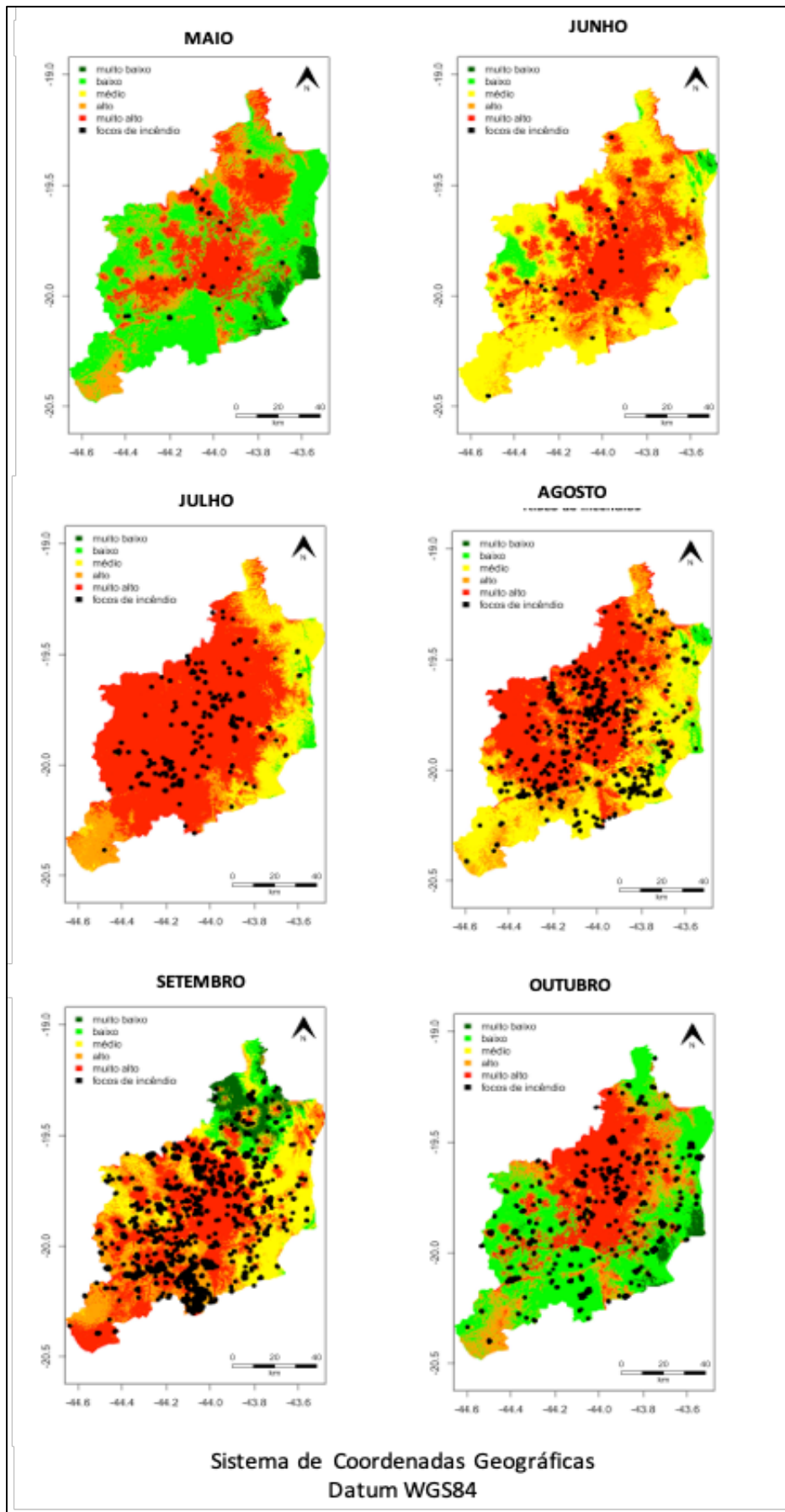


Figura 57 - Resultados modelo seca. Mapas produzidos em ambiente R.

4.8. Análise das variáveis do modelo

Os resultados dos testes para avaliar quanto cada variável influencia no modelo é apresentado a seguir. As Figuras 58, 59 e 60 apresentam o resultado do método de Garson (1991) que relaciona os pesos de cada variável. O que pode ser notado é que para os três modelos a variável de maior importância é a distância às áreas urbanas. Depois desta, as variáveis de destaque são as meteorológicas (radiação, vento médio e umidade relativa). No modelo de seca, a pressão também apresentou valor alto. Nota-se que as variáveis estáticas como declividade, orientação do relevo, uso do solo e NDVI foram as que tiveram a menor importância no modelo.

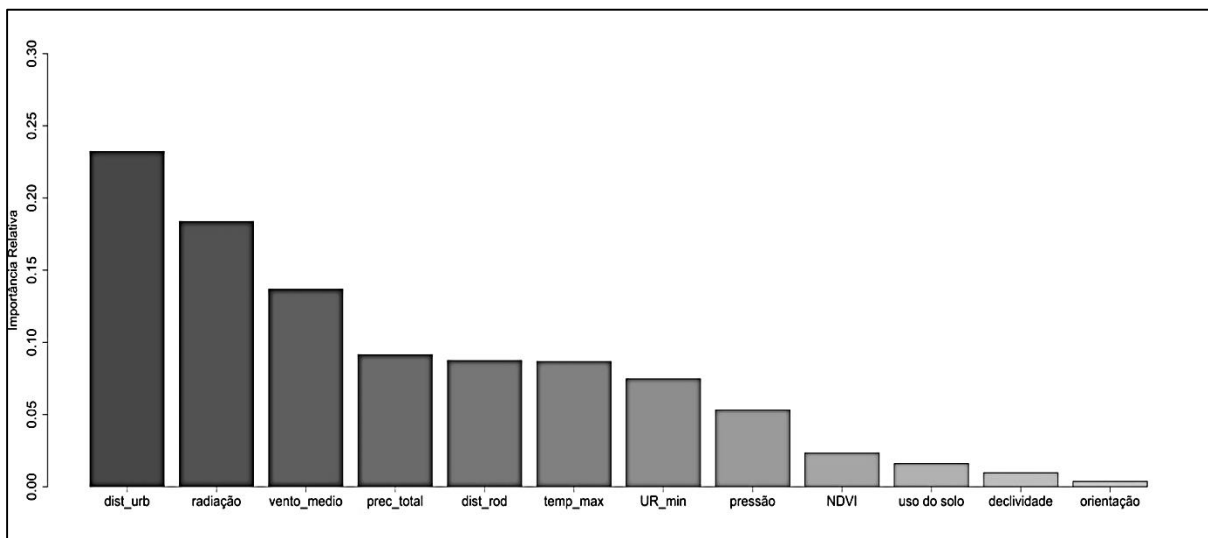


Figura 58 - Relação de importância entre variáveis pelo método de Garson (1991) para o modelo geral

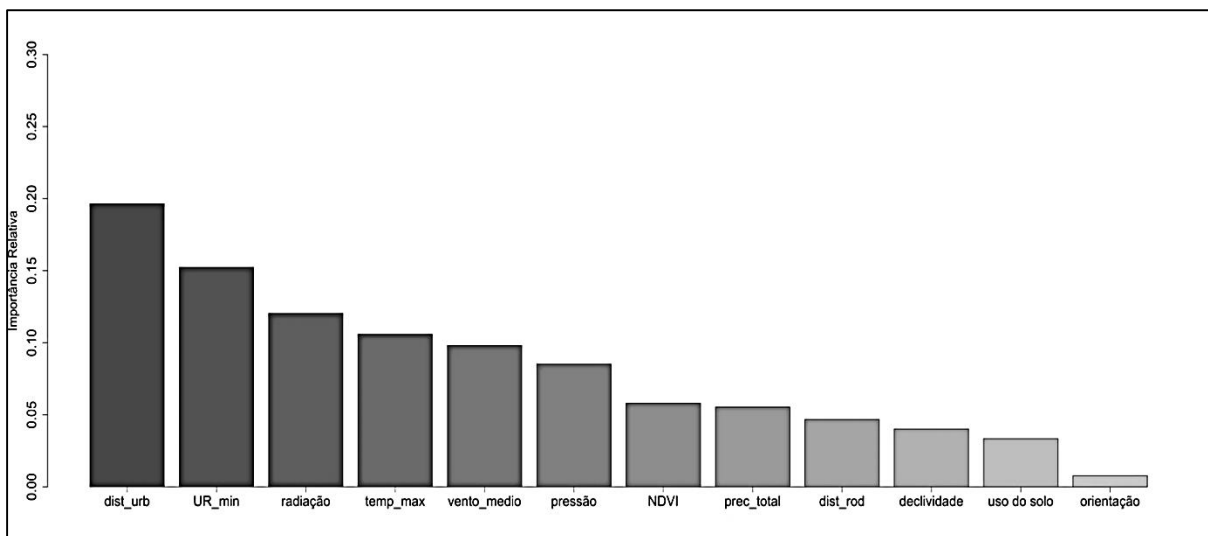


Figura 59 - Relação de importância entre variáveis pelo método de Garson (1991) para o modelo seca

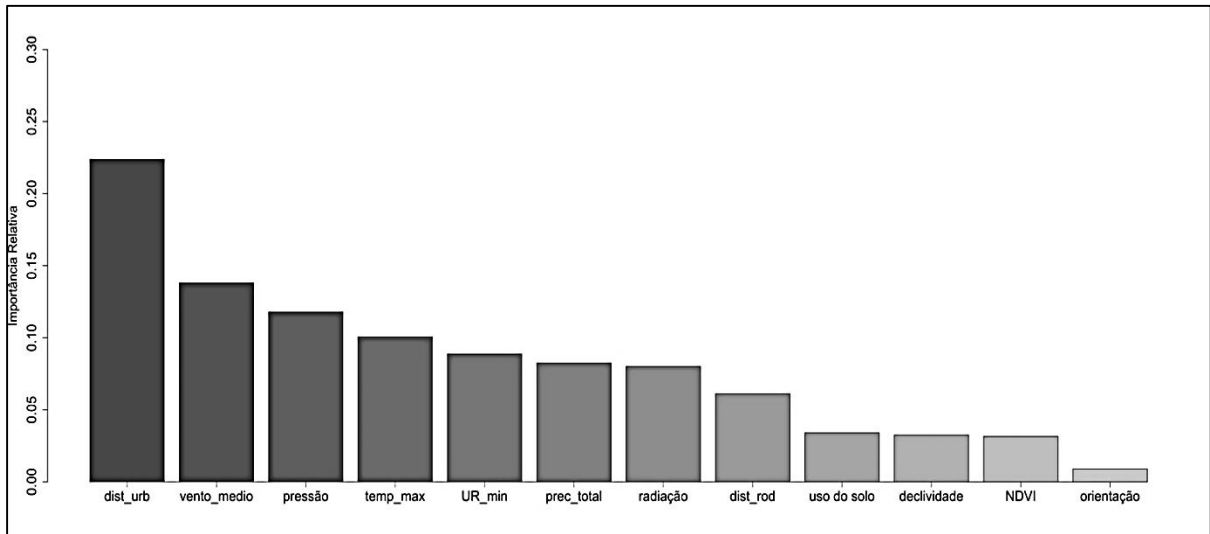


Figura 60 - Relação de importância entre variáveis pelo método de Garson (1991) para o modelo chuva

As Figuras 61, 62 e 63 mostram a resposta dos modelos às perturbações nas variáveis, onde é possível notar que a radiação é a que mais modifica a resposta do modelo geral e do modelo de seca. A temperatura é outro parâmetro que apresenta valor alto em todos os modelos e a pressão altera os modelos de chuva e seca.

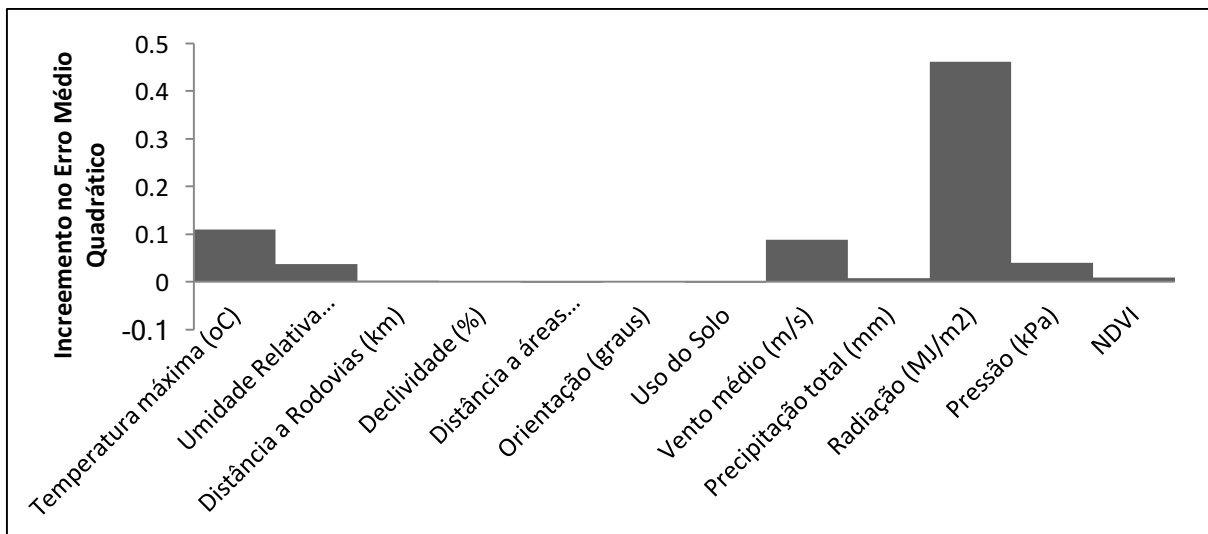


Figura 61 - Análise da Importância de variáveis no modelo pelo método da perturbação modelo geral

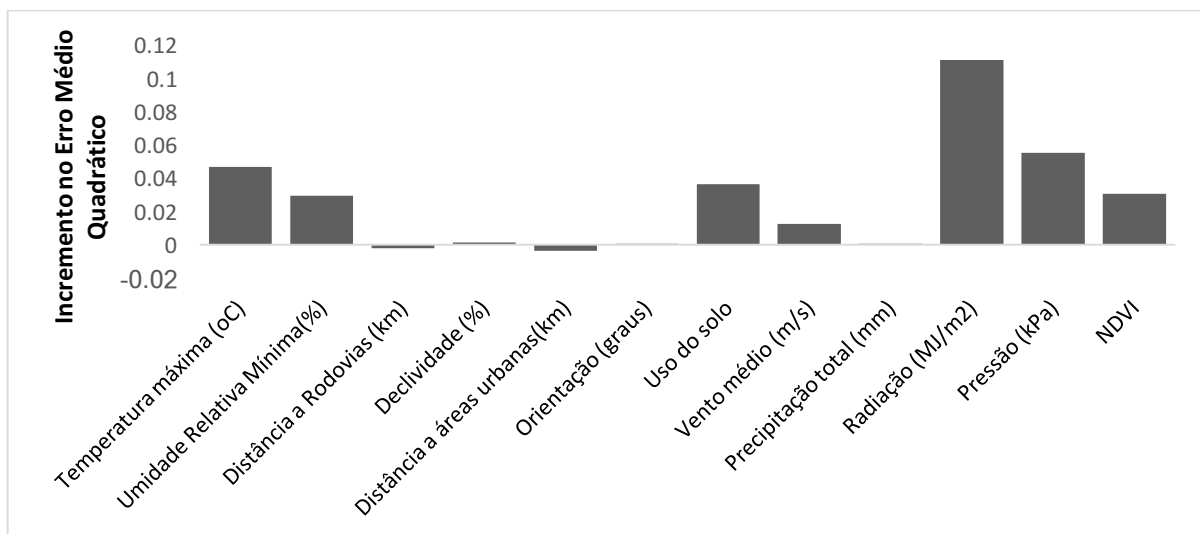


Figura 62 - Análise da Importância de variáveis no modelo pelo método da perturbação modelo seca

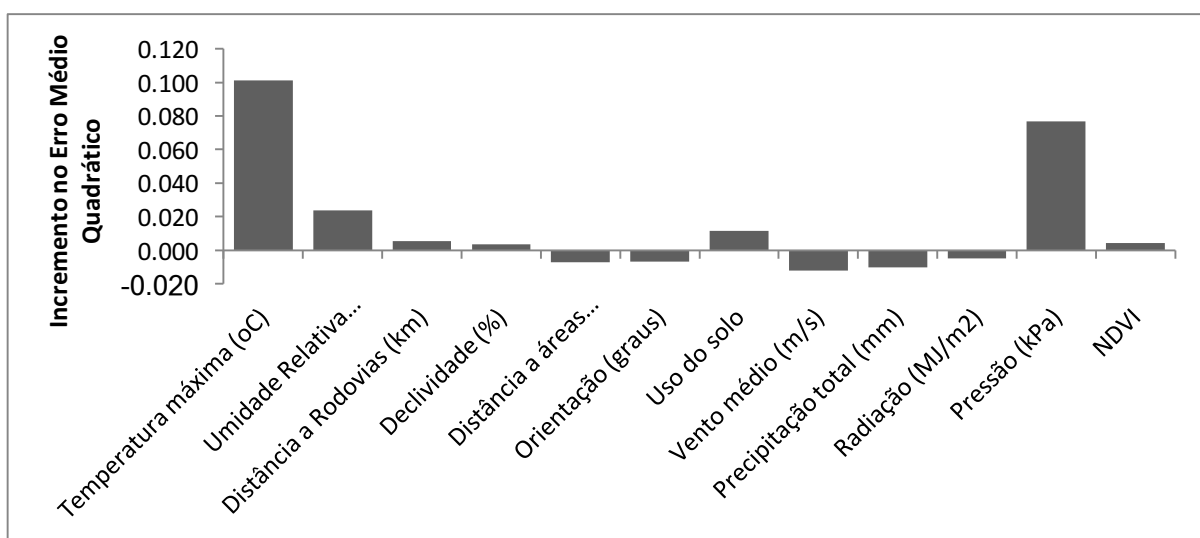


Figura 63 - Análise da Importância de variáveis no modelo pelo método da perturbação modelo chuva

Também foi feita uma análise de correlação de Spearman entre as variáveis e o risco de incêndio. Nas Figuras 64, 65 e 66, é possível notar, para o modelo geral e o modelo de seca, uma alta correlação entre a umidade relativa e a radiação. As variáveis precipitação, vento médio e distância a áreas urbanas também apresentam uma boa correlação.

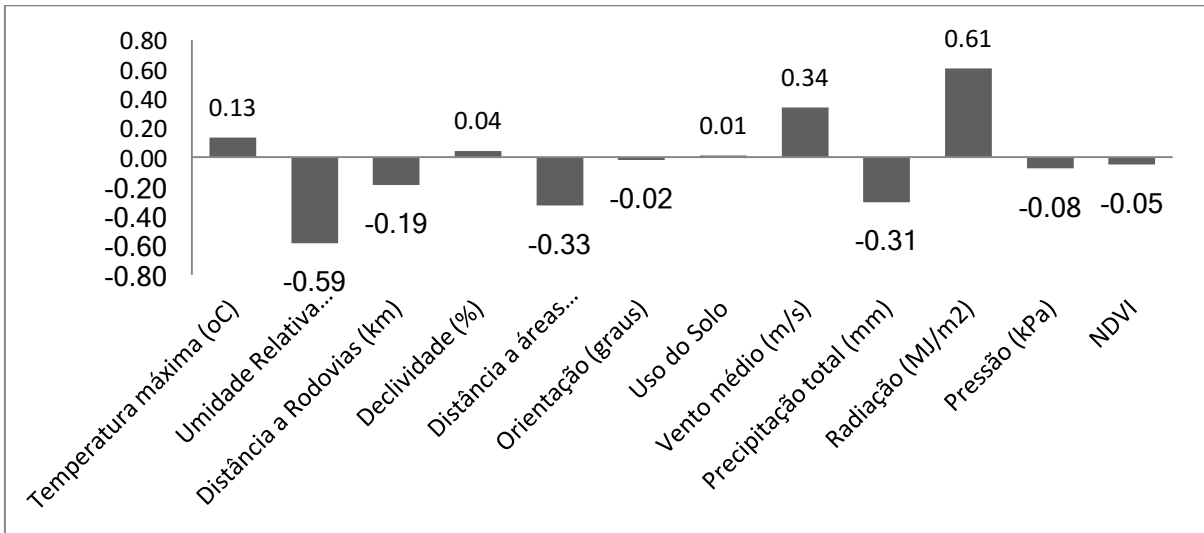


Figura 64 - Correlação Risco de Fogo com as variáveis modelo geral

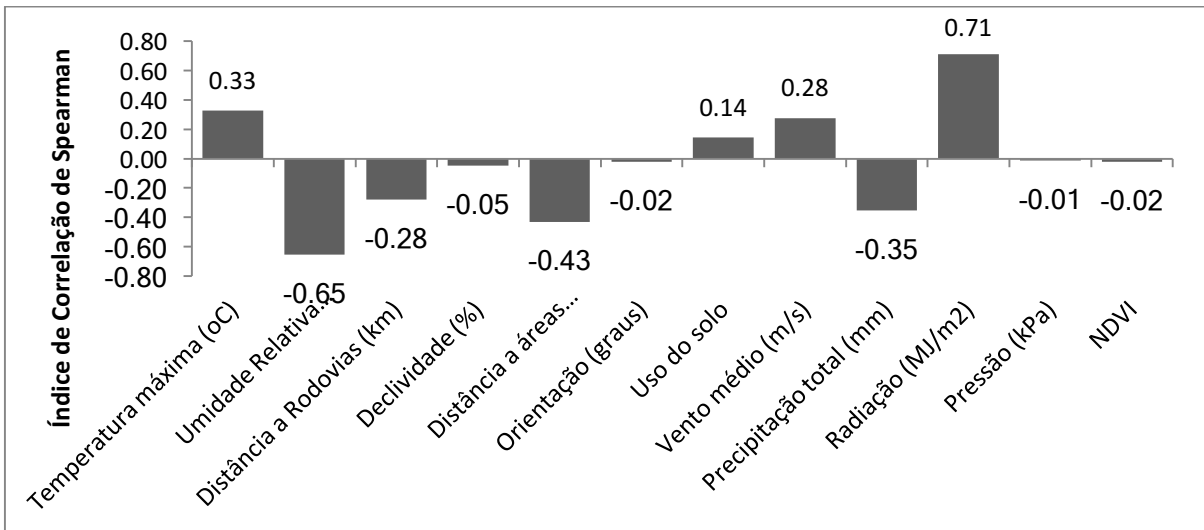


Figura 65 - Correlação Risco de Fogo com as variáveis do modelo seca

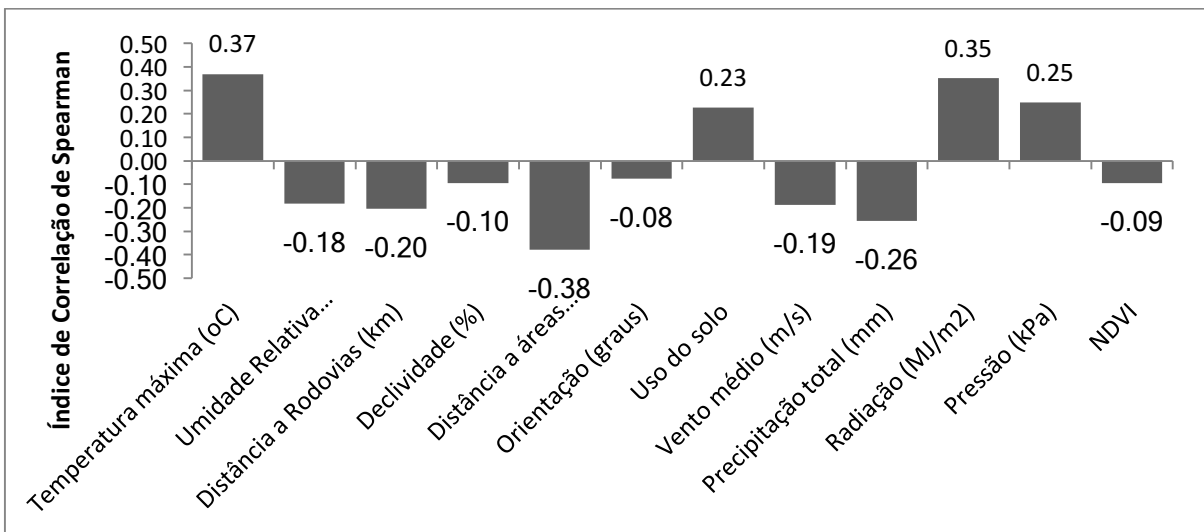


Figura 66 - Correlação Risco de Fogo com as variáveis do modelo chuva

Com essa análise foi possível perceber que as variáveis que mais influenciam o modelo são as variáveis de “distância às áreas urbanas” e as “meteorológicas”, principalmente umidade relativa mínima, radiação solar, temperatura máxima e pressão. Contudo, as variáveis como declividade e orientação do relevo contribuem muito pouco. Variáveis como NDVI e uso do solo apresentaram uma contribuição abaixo do esperado.

4.9. Análises e discussões

O esperado era que a Rede Neural pudesse identificar sozinha as sazonalidades, não sendo necessária a construção de modelos separados. Porém a ocorrência de incêndios pode estar ligada a muitos fatores que não foram possíveis de mensurar. Um fator que pode ter contribuído para a não diferenciação de sazonalidades pode ter sido a existência de focos em épocas de chuva fora da área em que se esperaria incêndio nesta época, já que a rede acabou associando os incêndios às características das duas épocas. Assim, foram escolhidos alguns meses do estudo para análise detalhada.

Na Figura 67 são apresentados os modelos para o mês de março, bem como as variáveis meteorológicas associadas. O modelo geral do mês de março, apesar de ter muitas áreas de risco baixo, também apresenta muitas áreas de risco médio. Esse resultado pode ser devido à combinação de valores mais altos de radiação, aliados a valores baixos de umidade relativa e valores mais altos da velocidade do vento. A precipitação também varia muito na região, apresentando valores baixos em alguns locais. É interessante notar que as áreas de risco mais alto, além de estarem influenciadas pelas áreas urbanas, é também perceptível, a relação com a umidade mínima na porção Sul da região.

Já o modelo de época de chuva para o mês de março apenas indicou risco alto nas áreas urbanas. Neste local já era esperado de se encontrar riscos mais altos, principalmente porque, como visto no Item 3.4, é o local que ocorrem mais incêndios nesse período. Comprovando que algumas características meteorológicas, principalmente umidade relativa baixa, podem influenciar às queimadas, e essa característica é contemplada no modelo geral. Mesmo com a pequena incidência de focos de incêndio fora de áreas urbanas, o risco deles é previsto no modelo geral.

Na Figura 68 são apresentados os modelos para o mês de novembro, que mesmo sendo um mês em época de chuva apresentou áreas de risco muito alto. Observa-se que isso é

devido à baixa umidade apresentada no sul da região, aliado a uma radiação alta e rajadas de vento mais altas. Nota-se ainda que em áreas classificadas com risco médio a umidade relativa estava com valores associados a risco mais altos, mas as outras variáveis, como a precipitação e o vento estavam com valores associados a risco mais baixos, o que diminui o fator de risco. Mas é interessante notar que nessa região ocorreram muitos focos de incêndio no mês, que não foram contemplados com o modelo apenas da época de chuva, que mais uma vez apresentou como alto risco apenas as áreas urbanas.

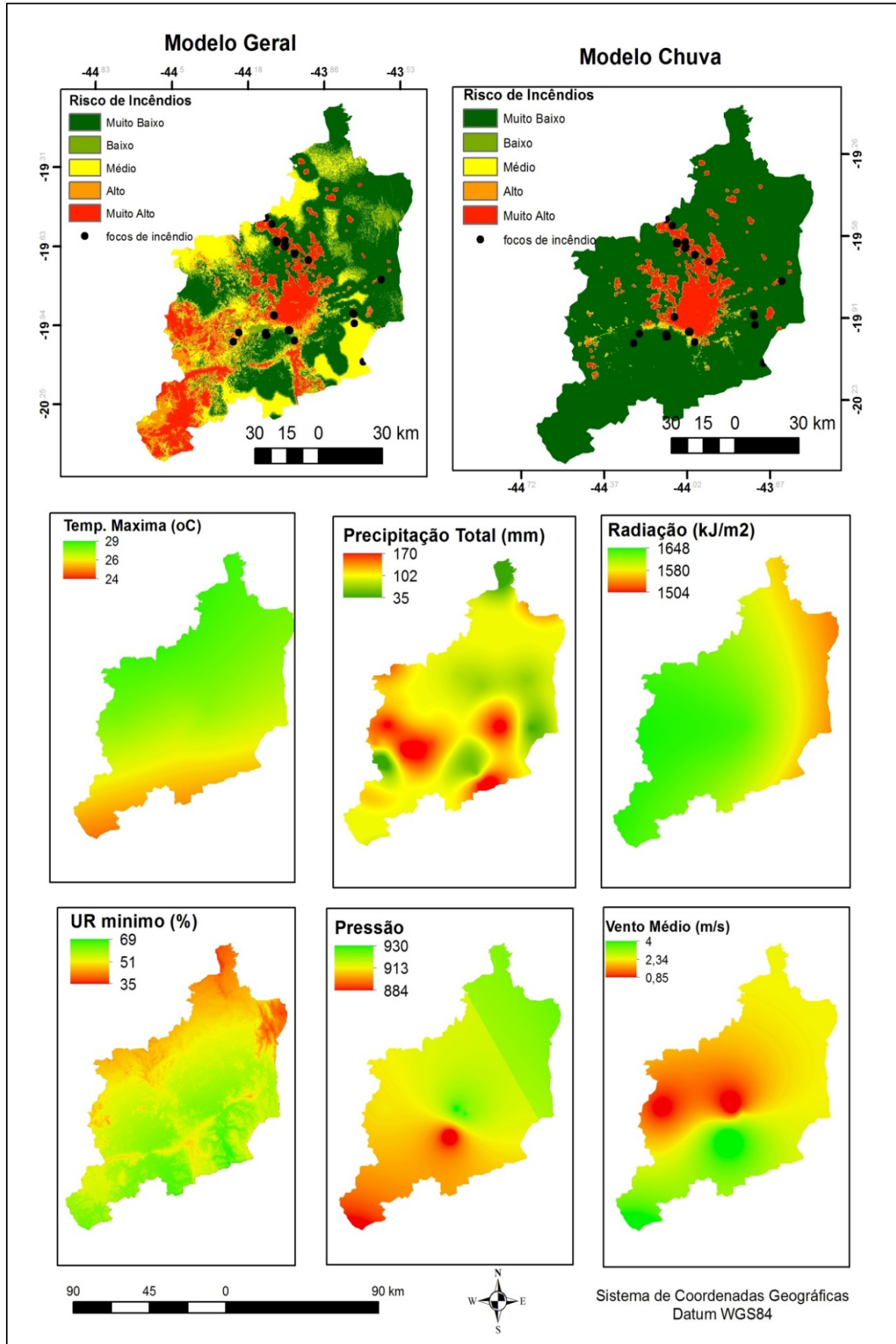


Figura 67 - Comparação de modelos para o mês de março

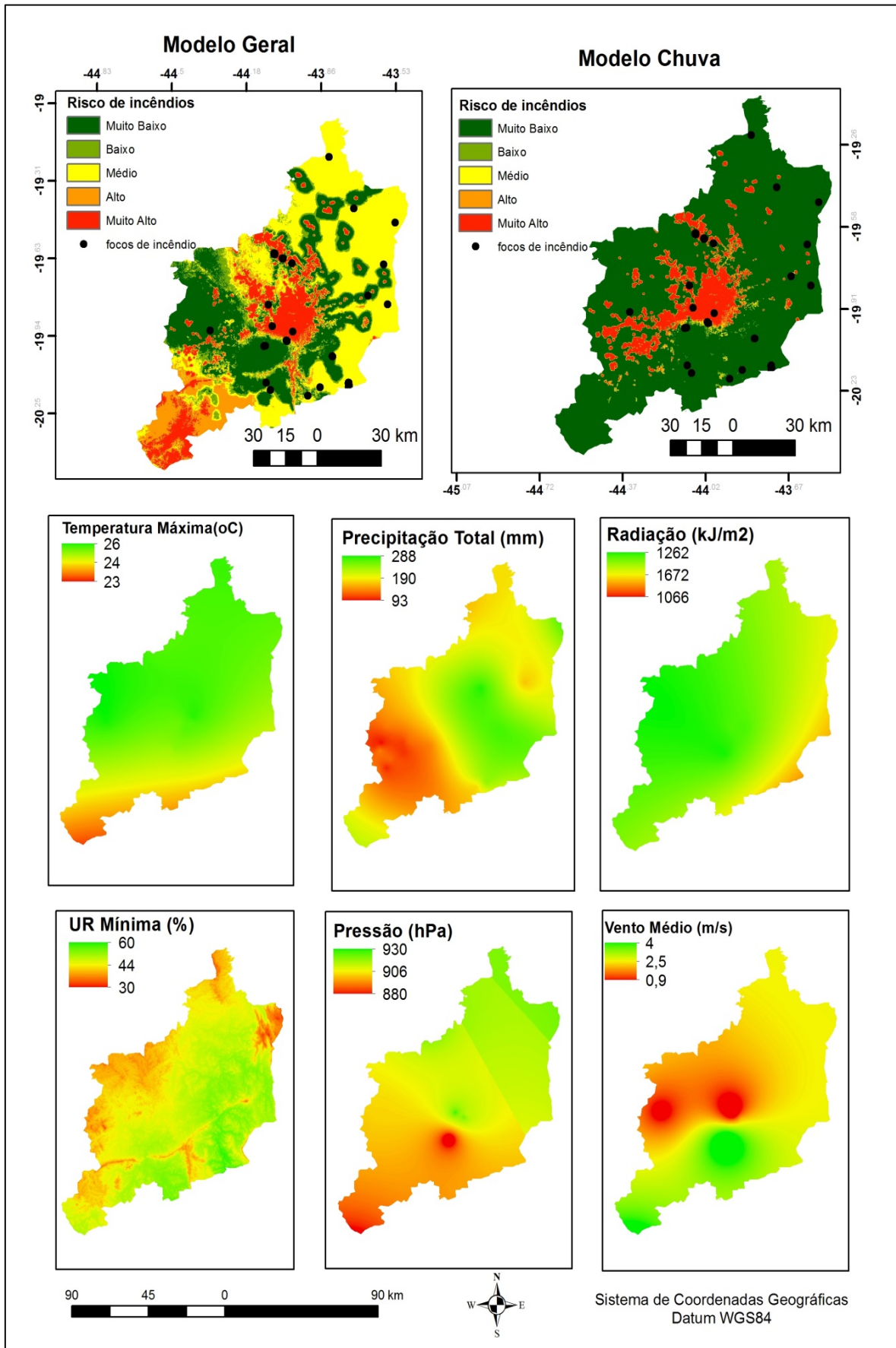


Figura 68 - Comparação de modelos para o mês de novembro

Na Figura 69 mostra o modelo de época de seca, onde quase a totalidade da área apresenta risco alto e isto está muito associado com os níveis muito baixos de precipitação no mês. Interessante notar que no local que há maior risco de incêndio no modelo geral (região sul) é classificado apenas como risco alto. Isso pode estar associado à temperatura e a pressão com valores menores nesse local. O modelo geral está mais associado com os dados de umidade relativa que são mais baixos na região sul, bem como rajadas de ventos mais intensas.

Na Figura 70, relacionada ao mês de setembro, nota-se que no modelo geral o risco de incêndio está concentrado na região sul, o que pode ser relacionado com os dados de temperatura, vento médio, pressão e radiação nessa região, que estão com valores mais extremos. Já o modelo de época de seca apresenta alto risco em grande parte da região, o que pode ter grande influência da radiação solar e da umidade relativa.

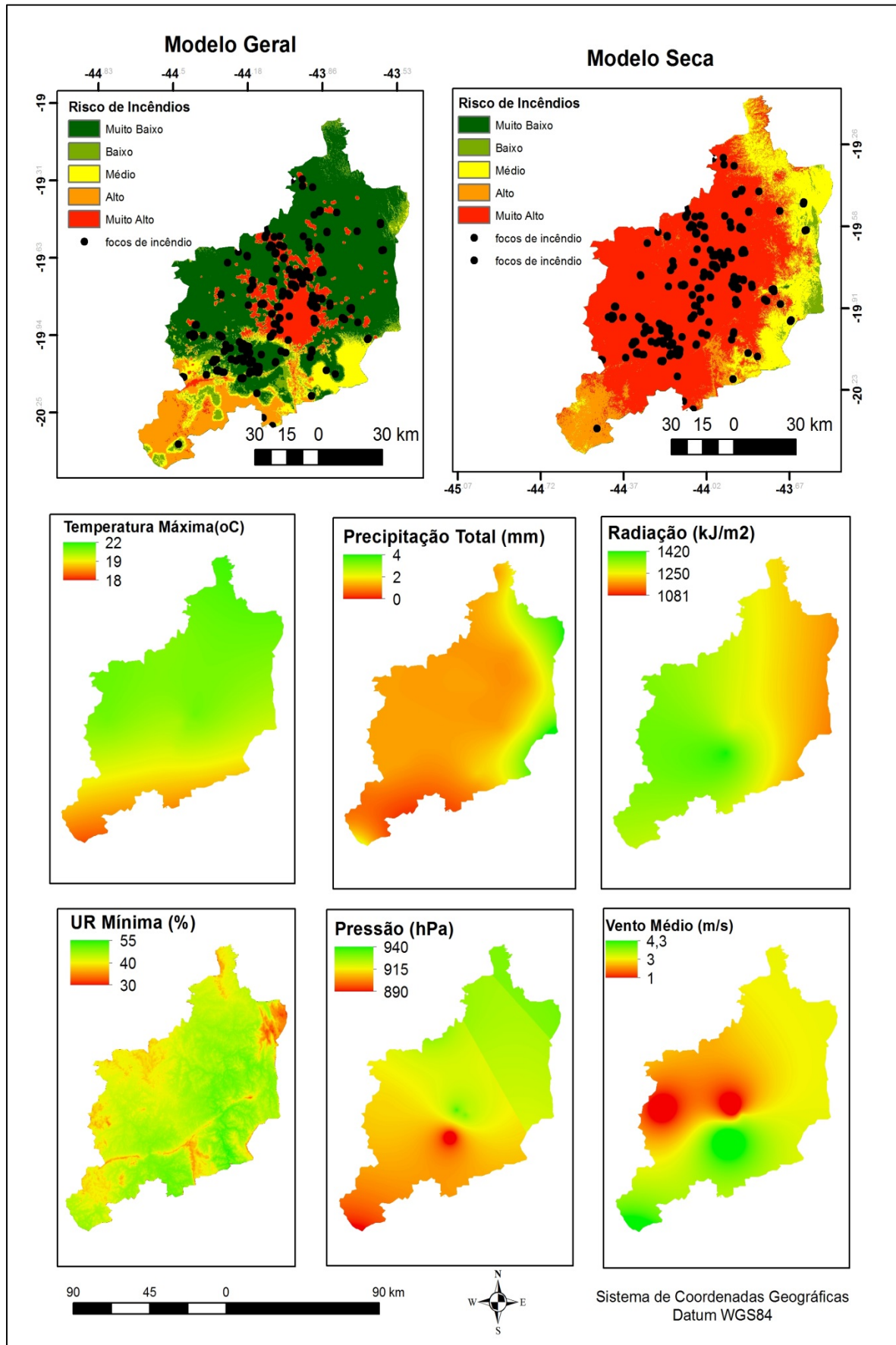


Figura 69 - Comparação de modelos para o mês de julho

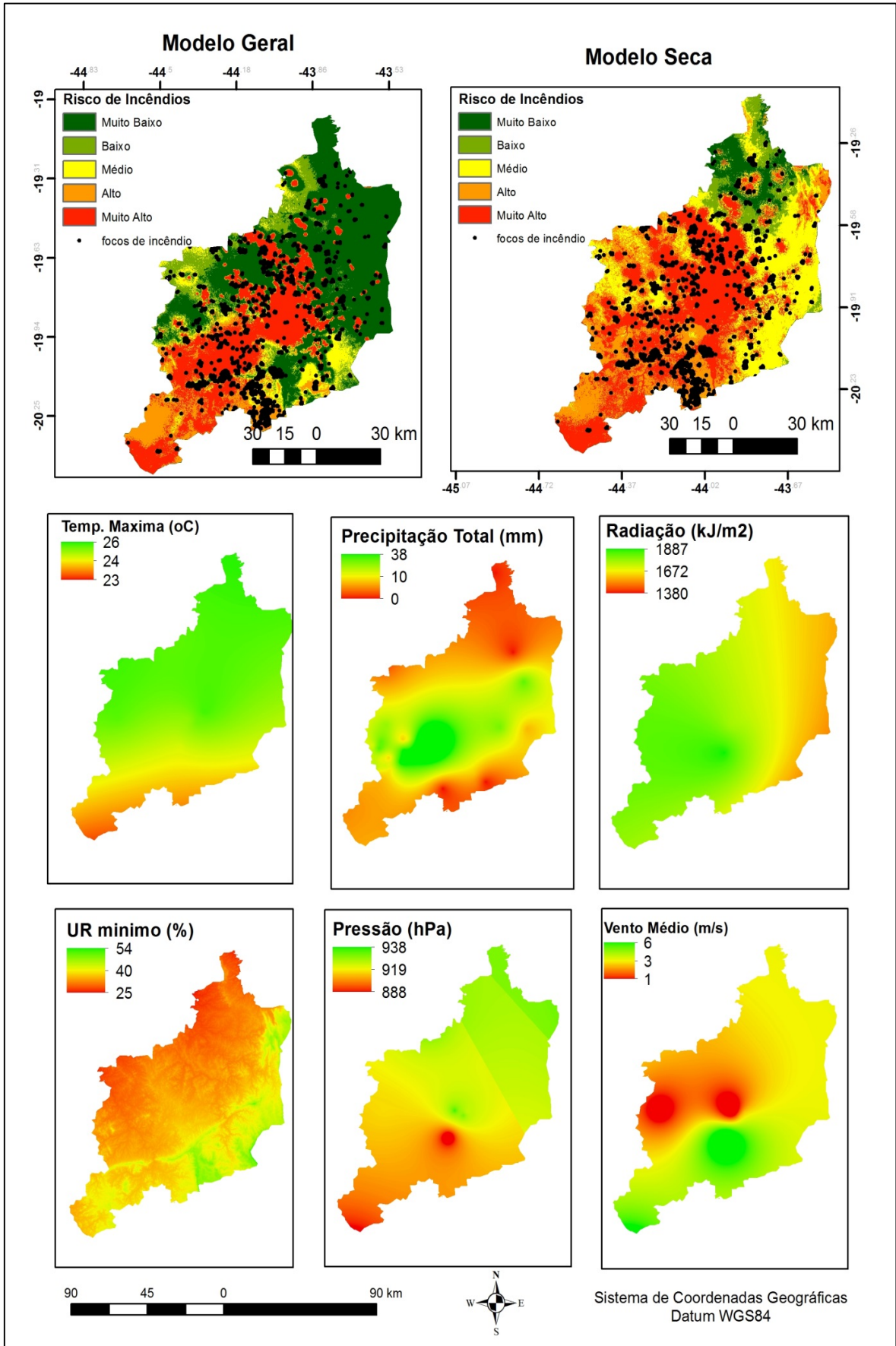


Figura 70 - Comparação de modelos para o mês de novembro

Assim, o modelo geral conseguiu captar áreas que tem perigo de incêndio em épocas em que se acredita que o perigo é baixo, enquanto o modelo de época de chuva captou apenas as áreas em que já é conhecido o risco alto. E para épocas de seca o modelo geral pode não ter tido grande coerência com os dados de incêndio, mas selecionou áreas onde a concentração de focos é maior. Dessa forma um modelo complementa o outro: em épocas de seca, como o mês de setembro, a área de risco abrange praticamente a região inteira, dificultando a seleção de locais a serem monitorados, e isso é conseguido com o modelo geral.

O modelo de redes neurais artificiais mostrou-se uma boa metodologia para esse problema, pois conseguiu captar particularidades que outras metodologias, baseadas em especialistas, não conseguiria captar.

A metodologia mais utilizada na literatura para cálculo de risco de incêndios é o da análise multicritério, baseada apenas no conhecimento dos especialistas para dar pesos às variáveis. O modelador coloca o peso na variável de acordo com seu conhecimento e sua área de estudo. Os resultados geralmente são subjetivos e não apresentam boa validação, como no de Chuvieco et al. (1989) que comparou áreas de pixels de queimadas e as áreas de risco encontradas, sendo que o resultado não foi muito satisfatório, já que apenas 22,72% das áreas foram classificadas como de alto risco. Jaiswal et al. (2002) aplicou em seu trabalho a metodologia e atestou que áreas de maior risco de incêndio, coincidiram com áreas próximas a assentamentos e rodovias. Deve-se ressaltar, ainda, que os métodos utilizados por Chuvieco et al. (1989) e Ribeiro et al. (2008) foram aplicados por Oliveira et al. (2017) para o Município de João Pessoa-PB, considerando as sazonalidades (estações do ano) e o resultado não foi satisfatório.

Outra metodologia utilizada em modelos de risco de incêndio é a lógica *Fuzzy*. Este método foi aplicado por Oliveira et al. (2017) com variáveis estáticas como declividade, distância de estradas, vegetação, entre outros e obteve um resultado satisfatório, comparando seu modelo com dados de ocorrência de incêndio fornecidos pela Brigada de Incêndios do Município de João Pessoa-PB, conseguiu explicar que o modelo em 85% em um dos períodos, considerando as sazonalidades (estações de seca e de chuvas). Silva e Pontes (2011) também fizeram uma abordagem *fuzzy* utilizando variáveis dinâmicas, como umidade relativa e precipitação. Ao compararem os índices de focos de calor registrados pelo INPE para validação do modelo os resultados foram positivos. Dessa maneira, conseguiram atribuir altos valores de fator de risco, exatamente no período em que ocorrem os focos de incêndio e até

mesmo quando existe um número pequeno de focos isolados, ou seja, o modelo atingiu a previsão desse comportamento com valores não nulos. A desvantagem dessa metodologia como aplicada pelos autores é que as classes contínuas dos conjuntos *fuzzy* são construídas a partir de conceitos definidos *a priori* usando a experiência, ciência ou conhecimento heurístico.

A regressão logística é mais uma técnica muito utilizada na literatura para predição e análise de ocorrência de incêndio, alguns mais recentes são Hoyo, Isabel e Vega (2011); Guo et al. (2016); Hernandez-leal et al.(2006); Chuvieco et al.(2010) e Goldarag, Mohammadzadeh e Ardakani (2016). Alguns trabalhos não tiveram resultados muito acurados, por exemplo o trabalho de Hoyo, Isabel e Vega (2011). Neste trabalho o método de validação por AUC, obteve o valor de 0,709 para o modelo utilizando pontos de ignição como variável independente e de 0,672 para o modelo utilizando densidade de pontos. Já Guo et al. (2016) utilizaram em seu modelo variáveis como topografia, vegetação, infraestrutura, dados socioeconômicos e meteorológicos; onde a variável dependente consistia nos dados de focos de calor obtidos de dados de satélite. Neste último trabalho o valor de AUC foi de 0,843, indicando que o modelo teve um bom resultado compatível com a expectativa do modelo.

Goldarag, Mohammadzadeh e Ardakani (2016) fizeram uma comparação entre os o método da regressão logística e das Redes Neurais Artificiais e obtiveram um resultado muito melhor com a segunda metodologia. O modelo de regressão logística apresentou uma acurácia de apenas 50,84%, com um R^2 de 0,533, ou seja, o método não foi bem explicado. Com as Redes Neurais Artificiais, o resultado foi de 95% de acurácia; além disso, no cruzamento com os dados de incêndio, os resultados apontaram que todos os locais com predição de ocorrência de incêndio foram atingidos no cruzamento com os dados de ocorrências reais.

Foram encontrados poucos trabalhos na literatura que aplicam as redes neurais artificiais para a análise do risco de incêndios, os encontrados estão mencionados nesse trabalho. Goldarag, Mohammadzadeh e Ardakani (2016) fizeram uma abordagem similar à utilizada nesse trabalho, apresentando variáveis estáticas como declividade, orientação do relevo, altitude, uso do solo, NDVI, distâncias à rodovias e à áreas urbanas e variáveis meteorológicas como temperatura do ar, umidade, precipitação, vento e duração do sol, somando 12 variáveis de entrada da rede e utilizando como dados de saída para treinamento dados do sensor MODIS. Neste trabalho os autores fizeram um experimento variando os neurônios de 5 até 65, mas a acurácia máxima foi obtida com 20 neurônios. O algoritmo

usado para treinamento foi o algoritmo de otimização *Levenberg-Marquardt* e sua função de ativação foi a função logística.

Alonso-Betanzos et al.(2003) também fizeram um estudo utilizando RNA para previsão de riscos de ocorrência de incêndios. O experimento utilizou apenas dados meteorológicos, empregando também o algoritmo *Levenberg-Marquardt* para treinamento. Os autores realizaram vários testes com as melhores topologias de redes neurais artificiais, obtendo o melhor resultado com 9 neurônios e atingindo uma acurácia de 80%, na validação com dados históricos de ocorrência de incêndio.

Maeda et al.(2009) elaboraram um estudo em três municípios do norte do Mato Grosso, na região Amazônica, onde utilizou RNA do tipo *feedforward* (alimentação direta), treinadas com o algoritmo *backpropagation* (algoritmo de retropropagação do erro), tendo como entrada imagens NDVI, adquiridas pelo sensor MODIS durante cinco períodos de 2005, além de amostras de florestas queimadas e não queimadas para treinamento. Os resultados obtidos apontaram um erro médio quadrático de 0,07.

Outro trabalho que aplica o método da RNA, é o de Satir, Berberoglu e Donmez (2016), onde os autores mapeiam a probabilidade de ocorrer incêndios em florestas mediterrâneas. Como dados de entrada, foram utilizados dados históricos de incêndios, dados climáticos, distância às estradas e assentamentos, modelo digital de elevação e porcentagem de cobertura de árvores. Em seus resultados, foi obtido um resultado positivo, com um ROC de 0,83, indicando que essa metodologia é adequada para o cálculo de riscos de incêndio. Este último trabalho também realizou o treinamento com o algoritmo *backpropagation*.

A autora deste trabalho, previamente utilizou RNA para prever o risco de incêndio na cidade de Belo Horizonte (FERNANDES et al., 2019), utilizando como variáveis de entrada dados meteorológicos, uso do solo, declividade e orientação do relevo. Neste trabalho foi selecionado apenas os meses de agosto, setembro e outubro, por ser o período com maior incidência de focos de incêndio. Verificou-se que a maior parte da área foi classificada com risco de incêndios alto.

Os trabalhos de Fernandes et al. (2019), de Goldarag, Mohammadzadeh e Ardakani (2016) e de Alonso-Betanzos et al. (2003) utilizaram como algoritmo de treinamento o algoritmo *Levenberg-Marquardt*. Diferente do *backpropagation*, que

possui um algoritmo de descida mais íngreme, o treinamento por *Levenberg-Marquardt* é uma aproximação ao método de Newton e é usado para acelerar a convergência do algoritmo de aprendizado *backpropagation*. Este não foi utilizado por não estar implementado no MPCA-NN.

O que pode ser percebido é que as RNA geralmente possuem boa acurácia quando aplicada na resolução desse problema. Uma das dificuldades encontradas foi o fato da rede não conseguir reconhecer os dois tipos de dados, locais de ocorrência e não ocorrência de incêndio. A estratégia adotada de selecionar *pixels* por janelas foi inspirada em Goldarag, Mohammadzadeh e Ardakani (2016), mas outros autores como Alonso-Betanzos et al.(2003) selecionaram igualmente o número de pontos de ocorrência e não ocorrência de incêndios, que foi testado esse método também; porém a RNA reconhecia praticamente todos os pontos como pontos de grande probabilidade de incêndios.

Outro fator observado é que se tinham muitos pontos de ocorrência de incêndio próximos e muitos pontos de não ocorrência também próximos, no conjunto de padrões, a rede também não conseguiu uma boa acurácia na generalização. Por esse motivo os dados do conjunto foram embaralhados aleatoriamente antes de começar o treinamento.

Um ponto negativo da RNA é a escolha da melhor arquitetura e parâmetros de treinamento, que é baseado na experiência e testes, mas isso foi facilmente resolvido com o algoritmo do MPCA-NN, que auxiliou no processo e diminuiu o trabalho e o tempo de teste.

A avaliação do modelo foi outro problema enfrentado. A avaliação das RNA geralmente é feita quando o erro quadrático médio fica bem próximo de zero em todos os conjuntos (treinamento, teste e validação); no entanto, isso não aconteceu. Como a saída desejada da rede é sempre 0 ou 1, e a saída da rede é sempre entre 0 e 1, o erro se acumula, resultando em valores mais altos. Assim foram utilizados métodos de avaliação da classificação correta da saída da RNA com os dados de incêndio classificados.

A opção de se utilizar um algoritmo próprio faz com que o modelador tenha maior controle sobre seu modelo. Quando se sabe o que está ocorrendo dentro do modelo é possível fazer pequenas modificações e benfeitorias que ajudam a melhorar seu desempenho, não depende apenas das modificações nas variáveis de entrada. Em Fernandes et al. (2019) foi utilizado o *Neural Network Toolbox* disponível no *Matlab*, nesse toolbox é apenas necessário

selecionar os dados de entrada e o número de camadas escondidas. Apesar da facilidade em se usar essa ferramenta, os resultados não foram satisfatórios. Com o desenvolvimento da RNA própria foi possível fazer testes de modificações de taxa de aprendizado e taxa de momento, funções de ativação e número de neurônios, o que possibilitou o melhor resultado da rede e entendimento do modelamento da RNA

A escolha das variáveis certas também é outro fator que influencia nos modelos de incêndio. Em Hernandez-Leal et al. (2006) foram inicialmente utilizadas apenas variáveis estáticas e este concluiu que havia necessidade de adicionar variáveis meteorológicas em seu modelo. Pelos trabalhos apresentados aqui, nota-se que quando adicionadas variáveis meteorológicas há uma tendência de melhora do modelo. A FMA, metodologia de risco de incêndio, que envolve apenas variáveis meteorológicas, apresenta bons resultados que podem ser vistos em Soares (1998) e Nascimento, Amaral e Pontes (2013), chegando à correspondência de 91% entre os dados de incêndio e o risco alto. O modelo de Alonso-Betanzos et al.(2003) também utiliza apenas variáveis meteorológicas. Como é conhecido que outros fatores influenciam os incêndios, outras variáveis foram aplicadas no modelo de RNA.

A RNA apresentada aqui também sofre muita influência das variáveis meteorológicas, que é o que confere as diferenças mensais entre os resultados do modelo. Mas é importante salientar que a adição da variável “distância às áreas urbanas” foi essencial. As variáveis de uso do solo e NDVI, que se esperava maior representação da vegetação, não contribuíram muito. A variável “uso do solo” foi a única categórica utilizada, pode ser uma explicação para sua baixa influência, seria necessário utilizar outros meios de quantificar essa variável. Satir, Berberoglu e Donmez (2016) utilizam, por exemplo, porcentagem de cobertura de árvores como medida. Ou possibilidade é a utilização de parâmetros que indiquem a presença ou não de vegetação.

O fato de distância á áreas urbanas ter tido grande influência no modelo mostra que esta variável foi eficiente em tentar mensurar o fator humano na ocorrência dos incêndios. Porém, como esse fator é aleatório, também prejudica esse tipo de modelo, sendo que qualquer pessoa pode iniciar um incêndio em qualquer local, e é difícil de ser mensurado.

Outra dificuldade encontrada na construção da metodologia foi a disponibilidade de dados. Os dados meteorológicos disponíveis foram os dados das estações do INMET, que são

poucas em relação ao tamanho da área, esse fato pode ter prejudicado o modelo, superestimando ou subestimando valores para algum local.

A disponibilidade de dados de incêndio foi mais uma dificuldade, já que não foi possível conseguir dados de ocorrência efetiva dos incêndios. Estes foram solicitados junto ao Corpo de Bombeiros e não puderam ser disponibilizados por questões de segurança. Foi feita uma tentativa de pedido desses dados junto ao órgão responsável pelas florestas em Minas Gerais, o Instituto Estadual de Florestas de Minas Gerais (IEF) e este também não possuía esse tipo de dados observados e computados.

Assim, a alternativa foi utilizar dados de focos de incêndio de satélite. Apesar de apresentarem problemas de omissão e superestimação, são dados que se utilizam de algoritmos de recuperação que estão cada vez mais sendo aprimorados e apresentando bons resultados, como mostrado no Item 1.4. Na análise feita entre os dados estatísticos de satélite e do Corpo de Bombeiros nota-se a importância dos satélites em detectar incêndios em áreas de acesso dificultado, como por exemplo, unidades de conservação e áreas rurais, que não são verificados pelos bombeiros.

Uma vantagem da RNA é que ela própria reconhece a relação entre as variáveis com a ocorrência de incêndios, não sendo necessário o conhecimento de especialista na seleção de pesos. Uma vez treinada, ela pode ser utilizada como um modelo de previsão e ser reaplicada para a região, apenas adicionando-se as variáveis para qualquer época desejada. Isso pode auxiliar brigadas de incêndio que estarão preparadas para atuar nas regiões, podendo servir como modelo de prevenção.

Outra aplicação deste modelo, também apresentado no trabalho de Alonso-Betanzos et al. (2003) e Bonazountas et al. (2007), é que a partir do risco calculado foi feito um sistema acoplado a um modelo de propagação de fogo e assim foram preparados planos de ação e planos emergenciais. Também foram verificadas as áreas com maior dificuldade de controle, classificando as regiões que devem ter maior alerta. O modelo de RNA pode ser facilmente implementado e também atuar como ferramenta de preservação ambiental, pois se conhecendo o local da queima, pode-se analisar o impacto causado por essa e auxiliar na importância na preservação. Resultados desse tipo de trabalho poderiam ser utilizados em programas de educação ambiental.

Conclusão

Para a obtenção de melhores resultados foram feitos três modelos, um geral, um da época de seca e um da época de chuva. Todos os modelos apresentaram boa acurácia e coerência com os pontos de ocorrência de incêndios. Já na validação, o modelo geral não conseguiu reconhecer muito bem as áreas de risco de incêndio na época de seca.

Quanto ao modelo da época de seca este obteve boa coerência de dados de incêndio, pois mais de 50% dos pontos se apresentaram em áreas de alto risco em quase todos os meses, mas por outro lado a maior parte da área estava também classificada com valores mais altos de risco. Esse resultado seria interessante para alertar sobre o risco de incêndios nesse período, mas não auxilia muito bem as equipes brigadistas em termos de localização com boa precisão.

No modelo da época de chuva, apenas as regiões próximas às áreas urbanas foram classificadas como de risco muito alto e o restante como de risco baixo, mesmo apresentando boa coerência com os dados de focos de incêndio, é interessante notar que há focos externos a essas áreas. Um número maior de focos de incêndio nessa época seria interessante para melhorar o treinamento da RNA regiões, e isso só seria possível se aumentasse o período de coleta de dados.

O modelo geral, mesmo não tendo o melhor resultado, foi treinado com características das duas épocas, podendo ter selecionado apenas fatores que realmente influenciam os incêndios. Assim, seria interessante na utilização como informação adicional, pois auxiliaria na seleção das áreas de maior risco de incêndios na época de seca e para alertar sobre possíveis riscos em outras áreas em épocas de chuva.

É importante ressaltar que dentre as variáveis de entrada as que mais influenciaram o modelo foram: a distância às áreas urbanas e as variáveis meteorológicas. A primeira não é alterada nos anos estudados, sendo que em todos os modelos às áreas mais próximas receberam risco muito alto. Já as variáveis meteorológicas como se alteram todos os dias, podem apresentar eventos atípicos, como a influência dos sistemas atmosféricos atuantes na região. Assim, mesmo que seja conhecido que não há muitos focos de incêndio em março e novembro, meses geralmente chuvosos, nota-se que o modelo se baseou em dados de umidade relativa mais baixas para apresentar o resultado de risco alto na região. Portanto, o

modelo reconheceu as características meteorológicas que mais influenciam os incêndios e demonstra isso em seu resultado, importante para que as variáveis de entrada do modelo estejam coerentes com a realidade.

Quanto às variáveis de entrada, estas são produtos da interpolação de dados de estações meteorológicas, pois na região de estudo estas são escassas. Dessa maneira, o produto pode não ter conseguido extrapolar os dados meteorológicos para a região inteira. Para pesquisas futuras seria necessário o aprimoramento desses modelos ou busca de outras fontes de dados meteorológicos.

O sensoriamento remoto pode ser uma forma de melhorar os modelos de interpolação, pois existem satélites meteorológicos que fazem a medição indireta de parâmetros meteorológicos, por exemplo dados de temperatura que são derivados de radiâncias. Os produtos dos satélites são utilizados em modelos de previsão meteorológica. Este tipo de dado é interessante, pois possui informações de todo o globo, depende da resolução espacial; as estações (medidas diretas) possuem informações apenas de alguns pontos.

É importante destacar aqui que os dados de sensoriamento remoto foram essenciais na construção deste trabalho, uma vez que os focos de incêndios foram retirados de produtos do MODIS e do VIIRS. Não foi possível a obtenção de dados de focos de incêndios espacializados de uma outra fonte mais precisa (a exemplo de endereços codificados da Brigada de Incêndio). Apesar disso, estudos demonstram que esses dados têm melhorado bastante e tem grande coerência com a realidade, a única desvantagem é que só conseguem verificar a ocorrência dos maiores incêndios e que são registrados na hora da passagem do satélite. Assim, alguns dados de focos de incêndio por imagens de satélite não foram contemplados nesse estudo.

É importante ressaltar que este tipo de modelagem de risco de incêndios tem duas grandes dificuldades intrínsecas. Uma delas é a influência humana na ocorrência de incêndios, por ser uma ação aleatória, é difícil de ser prevista, se os incêndios fossem dependentes apenas de fatores meteorológicos os resultados seriam bem melhores. Devem ser buscadas melhores formas de prever esta ação humana.

Outra grande dificuldade é que os incêndios se tratam de variáveis não contínuas, apresentando uma não linearidade essencial. Sendo difícil de encontrar uma aproximação

matemática para o problema. Uma forma que poderia ser utilizada para melhorar o modelo seria a utilização de uma classificação hierárquica, utilizando quatro redes neurais uma para cada classe de risco.

Outra forma de melhorar o modelo seria fazer um estudo prévio da correlação das variáveis do modelo com a ocorrência de incêndios, aonde poderia ser utilizada, por exemplo, a medida da entropia de Shannon.

O desenvolvimento do modelo de Redes Neurais Artificiais é trabalhoso devido a seu processo de treinamento, onde devem ser parametrizados a melhor taxa de aprendizagem, a taxa de momento, o número de neurônios e a função de ativação. Além disso, necessita de um conjunto de padrões que represente bem o problema. Parte desse trabalho foi minimizado quando se utiliza o *software* MPCA-NN para definição dos parâmetros.

Apesar das dificuldades encontradas, após o entendimento e familiarização com a RNA, as possibilidades de utilização deste método são quase inesgotáveis, em praticamente todos os domínios do saber. De fato, as redes neurais artificiais têm permitido novas fronteiras na resolução de vários problemas, neste caso em especial, da análise do risco de incêndio, a RNA revelou resultados muito interessantes.

Pelo fato da RNA ser uma metodologia baseada em aprendizagem, ela pode ser reaplicada para qualquer região, sendo apenas necessário ter os dados específicos para o treinamento. Isso se faz necessário, pois poucos são os estudos de risco de incêndio voltados para regiões próximas às áreas urbanas, em especial considerando ainda a Região Sudeste do Brasil, onde devido aos períodos de estiagem no inverno, a incidência de incêndios é bem alta.

Os Mapas de Risco de Incêndio podem ser usados como recurso de atenuação de desastres dessa categoria, já que predizem cenários vulneráveis quanto ao risco, diagnosticando com precisão locais onde devem-se tomar medidas cabíveis para prever tragédias, como locais para estabelecer postos de vigia e bases de equipes brigadistas.

Estes mapas podem ser usados como ferramenta de alerta à população, já que a maior parte dos incêndios ocorrem em áreas urbanas. Esses alertas podem ser feitos com a criação de políticas informativas para a população com dados de cuidados e os prejuízos que podem ocorrer. Além disso, com o conhecimento das áreas de maior perigo, pode-se fazer o

monitoramento constante dessas áreas e criar estratégias para acesso às mesmas, de modo a proporcionar a melhoria no atendimento das ocorrências.

Como trabalho futuro, espera-se colocar esse modelo em um sistema no qual ele seja alimentado com dados de previsão meteorológicas vindos do CPTEC, de modo a permitir a predição dos riscos com até sete dias de antecedência. Isso seria colocado em uma plataforma de possível divulgação desses dados, que poderiam auxiliar tanto o poder público com a criação de políticas públicas, bem como às empresas privadas que dependem de recursos florestais.

Referências

- AGENCIA RMBH. **RMBH e Colar Metropolitano**, 2016. Disponível em: <<http://www.agenciarmbh.mg.gov.br/rmbh-e-colar-metropolitano/>>. Acesso em: 20 out. 2016
- AKOBENG, A. K. Understanding diagnostic tests 3: receiver operating characteristic curves. **Acta Paediatrica**, v. 96, n. 5, p. 644–647, maio 2007.
- ALENCAR, A. Apresentação do Seminário de Lançamento da Coleção 2: Cerrado. 2017.
- ALMEIDA, R. M. **Modelo de propagação do fogo em incêndios de superfície baseado no método nodal**. [s.l.] Dissertação (mestrado em Computação Aplicada)-INPE, São José dos Campos, 2008.
- ALONSO-BETANZOS, A. et al. An intelligent system for forest fire risk prediction and fire fighting management in Galicia. **Expert Systems with Applications**, v. 25, n. 4, p. 545–554, nov. 2003.
- ALVAREZ, O. et al. Comparison of elevation and remote sensing derived products as auxiliary data for climate surface interpolation. **International Journal of Climatology**, v. 34, n. 7, p. 2258–2268, jun. 2014.
- ANA, (AGÊNCIA NACIONAL DE ÁGUAS). **Séries Históricas de Estações**. Disponível em: <http://www.snirh.gov.br/hidroweb/publico/medicoes_historicas_abas.jsf>. Acesso em: 1 jun. 2018.
- ANDERSON, T. W.; DARLING, D. A. A Test of Goodness of Fit. **Journal of the American Statistical Association**, v. 49, n. 268, p. 765–769, dez. 1954.
- ANOCHI, J. et al. **New learning strategy for supervised neural network: MPCA meta-heuristic approach**. Anais do 11. Congresso Brasileiro de Inteligência Computacional. **Anais...SBIC**, 11 mar. 2016Disponível em: <<https://www.researchgate.net/publication/256729765>>. Acesso em: 11 fev. 2019
- ANOCHI, J. A.; CAMPOS VELHO, H. F. DE. **Optimization of feedforward neural network by Multiple Particle Collision Algorithm**. 2014 IEEE Symposium on Foundations of Computational Intelligence (FOCI). **Anais...Orlando, FL: IEEE**, dez. 2014Disponível em:

<<http://ieeexplore.ieee.org/document/7007817/>>. Acesso em: 11 fev. 2019

AYOADE, J. O. **Introdução à climatologia para os trópicos**. 10 ed. ed. Rio de Janeiro: Bertrand Brasil, 2004.

BARRETO, J. M. **Indrodução às Redes Neurais Artificiais**. Florianopolis - SC: Laboratório de Conexionismo e Ciências Cognitivas UFSC -Departamento de Informática e de Estatística, 2002. Disponível em: <<http://www.inf.ufsc.br/~j.barreto/tutoriais/Survey.pdf>>.

BEDIA, J. et al. Global patterns in the sensitivity of burned area to fire-weather: Implications for climate change. **Agricultural and Forest Meteorology**, v. 214–215, p. 369–379, 2015.

BOER, E. P. J.; DE BEURS, K. M.; HARTKAMP, A. D. Kriging and thin plate splines for mapping climate variables. **ITC Journal**, v. 3, n. 2, p. 146–154, 2001.

BONAZOUNTAS, M. et al. A decision support system for managing forest fire casualties. **Journal of Environmental Management**, v. 84, n. 4, p. 412–418, 2007.

BURROUGH, P. A.; MCDONNELL, R. A. **Principles of Geographical Information Systems**. Oxford: Oxford University Press, 1998.

CAMPOS VELHO, H. F. et al. **MPCA-ANN**, 2018. Disponível em: <<https://bitbucket.org/reynierhdez/mpcaann/overview>>

CANNADY, J. D. Artificial neural networks for misuse detection. **Proceedings of the 21st National information systems security conference**, p. 368–381, 1998.

CARAPIÁ, V. R. **Predição do índice de risco de incêndios e modelagem computacional do comportamento do avanço da frente do fogo no Parque Nacional da Floresta da Tijuca**. Rio de Janeiro: Tese (Doutorado em Engenharia Civil) - Universidade Federal do Rio de Janeiro, 2006.

CARVALHO, A. R.; RAMOS, F. M.; CHAVES, A. A. Metaheuristics for the feedforward artificial neural network (ANN) architecture optimization problem. **Neural Computing and Applications**, v. 20, n. 8, p. 1273–1284, 9 nov. 2011.

CBMMG, C. DE B. M. DE M. G. **Dados conseguidos junto ao órgão responsável através do portal da transparência**, 2018.

CHEN, D. et al. Mapping fire regimes in China using MODIS active fire and burned area data. **Applied Geography**, v. 85, p. 14–26, 2017.

CHRISTOFOLETTI, A. **Modelagem de Sistemas Ambientais**. [s.l.] Edgard Blücher, 1999.

CHUVIECO, E. et al. Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. **Ecological Modelling**, v. 221, n. 1, p. 46–58, 2010.

CHUVIECO, E.; CONGALTON, R. G. Application of remote sensing and geographic information systems to forest fire hazard mapping. **Remote Sensing of Environment**, v. 29, n. 2, p. 147–159, ago. 1989.

CINTRA, R. et al. **Data assimilation by artificial neural networks for the global FSU atmospheric model: Surface pressure**. 2015 Latin America Congress on Computational Intelligence (LA-CCI). **Anais...IEEE**, out. 2015Disponível em: <<http://ieeexplore.ieee.org/document/7435937/>>. Acesso em: 11 fev. 2019

CINTRA, R. et al. Moisture Data Assimilation to the Global FSU Atmospheric Model. **Conference of Computational Interdisciplinary Science**, 2016.

CINTRA, R.; COCKE, S.; CAMPOS VELHO, H. F. **Data assimilation by neural networks with ensemble prediction**. ICVRAM ISUMA UNCERTAINTIES conference. **Anais...2018Disponível em:** <<http://icvramisuma2018.org/cd/web/PDF/ICVRAMISUMA2018-0095.PDF>>. Acesso em: 11 fev. 2019

COCHRANE, M. A. Fire science for rainforests. **Nature**, v. 421, n. 6926, p. 913–919, 27 fev. 2003.

COSTA, M. A. et al. Improving neural networks generalization with new constructive and pruning methods. **Journal of Intelligent & Fuzzy Systems**, v. 13, n. 2–4, p. 75–83, 2003a.

COSTA, M. A. et al. Training neural networks with a multi-objective sliding mode control algorithm. **Neurocomputing**, v. 51, p. 467–473, abr. 2003b.

COUTINHO, E.; SILVA, R. M.; DELGADO, A. R. S. Using Computational Intelligence Technique for the Meteorological Data Prediction. **Revista Brasileira de Meteorologia**, v.

31, n. June, p. 24–36, 2016.

CSISZAR, I. et al. Active fires from the Suomi NPP Visible Infrared Imaging Radiometer Suite: Product status and first evaluation results. **Journal of Geophysical Research: Atmospheres**, v. 119, n. 2, p. 803–816, 27 jan. 2014.

FEAM, (FUNDAÇÃO ESTADUAL DO MEIO AMBIENTE). **Dados conseguidos junto ao órgão responsável**, 2017.

FERNANDES, L. C. et al. Modelagem de risco de incêndios florestais: uma visão geral. **Sodebras**, v. 13, n. 147, p. 79–84, 2018.

FERNANDES, L. C. **Código Backpropagation**, 2019. Disponível em: <https://github.com/luizacintraf/backproplu/blob/master/rna_.r>

FERNANDES, L. C. et al. Fire Risk Modeling Using Artificial Neural Networks. In: **EngOpt 2018 Proceedings of the 6th International Conference on Engineering Optimization**. Cham: Springer International Publishing, 2019. p. 938–948.

FERNANDES, P.; REGO, F. Ecologia do fogo e gestão de áreas ardidas. In: MOREIRA, F. et al. (Eds.). **Ecologia do fogo e gestão de áreas ardidas**. [s.l.] ISA Press, 2010. p. 13–20.

FERRAZ, S. F. D. B.; VETTORAZZI, C. A. Identificação de áreas para recomposição florestal com base em princípios de ecologia de paisagem. **Revista Árvore**, v. 27, n. 4, p. 575–583, 2003.

FIDELIS, A.; PIVELLO, V. R. Deve-se usar o fogo como instrumento de manejo no Cerrado e Campos Sulinos? **Biodiversidade Brasileira**, v. 1, n. 2, p. 12–25, 2011.

FRANCA, R. R. DA. Anticiclones e umidade relativa do ar: um estudo sobre o clima de Belo Horizonte. p. 109, 2009.

GARSON, G. D. Interpreting neural-network connection weights. **AI Expert**, v. 6, n. 4, p. 46–51, 1991.

GIGLIO, L.; SCHROEDER, W.; JUSTICE, C. O. The collection 6 MODIS active fire detection algorithm and fire products. **Remote Sensing of Environment**, v. 178, p. 31–41, 1 jun. 2016.

- GOLDARAG, Y. J.; MOHAMMADZADEH, A.; ARDAKANI, A. S. Fire Risk Assessment Using Neural Network and Logistic Regression. **Journal of the Indian Society of Remote Sensing**, v. 44, n. 6, p. 885–894, 2016.
- GOULART, D. A. et al. Redes neurais artificiais aplicadas no estudo de questionário de varredura para conjuntivite alérgica em escolares. **Arquivos Brasileiros de Oftalmologia**, v. 69, n. 5, p. 707–713, 2006.
- GUIMARÃES, P. P. et al. Análise dos impactos ambientais de um incêndio florestal. **Agrarian Academy**, v. 1, n. 1, p. 38, 2014.
- GUO, F. et al. Wildfire ignition in the forests of southeast China: Identifying drivers and spatial distribution to predict wildfire likelihood. **Applied Geography**, v. 66, p. 12–21, 2016.
- HAMADEH, N. et al. Using correlative data analysis to develop weather index that estimates the risk of forest fires in Lebanon & Mediterranean: Assessment versus prevalent meteorological indices. **Case Studies in Fire Safety**, v. 7, p. 8–22, 2017.
- HANTSON, S. et al. Strengths and weaknesses of MODIS hotspots to characterize global fire occurrence. **Remote Sensing of Environment**, v. 131, p. 152–159, 15 abr. 2013.
- HARDY, C. C. Wildland fire hazard and risk: Problems, definitions, and context. **Forest Ecology and Management**, v. 211, n. 1–2, p. 73–82, 6 jun. 2005.
- HAYKIN, S. **Redes Neurais: Princípios e prática**. 2a. ed. Porto Alegre: Bookman, 2001.
- HERNANDEZ-LEAL, P. A.; ARBELO, M.; GONZALEZ-CALVO, A. Fire risk assessment using satellite data. **Advances in Space Research**, v. 37, n. 4, p. 741–746, 2006.
- HIJMANS, R. J.; GHOSH, A. **Spatial Data Analysis with R**, 2019. Disponível em: <<https://rspatial.org/analysis/analysis.pdf>>
- HOFFMANN, W. A.; JACKSON, R. B. Vegetation–Climate Feedbacks in the Conversion of Tropical Savanna to Grassland. **Journal of Climate**, v. 13, n. 9, p. 1593–1602, maio 2000.
- HOUGHTON, R. A. et al. O Balanço do Carbono Regional. **Amazonia and Global Change Geophysical Monograph Series**, v. 186, 2009.

HOYO, L. V. DEL; ISABEL, M. P. M.; VEGA, F. J. M. Logistic regression models for human-caused wildfire risk estimation: analysing the effect of the spatial accuracy in fire occurrence data. **European Journal of Forest Research**, v. 130, n. 6, p. 983–996, 10 nov. 2011.

IBGE, (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA). **Clima do Brasil 1: 500.000**, 2002. Disponível em: <http://dados.gov.br/dataset/cren_climadobrasil_5000/resource/360f02ac-f5e9-45ee-a6d4-b7ae1ae84def>

IBGE, (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA). **Setores Censitários-MG**, 2010a. Disponível em: <ftp://geoftp.ibge.gov.br/organizacao_do_territorio/malhas_territoriais/malhas_de_setores_censitarios__divisooes_intramunicipais/censo_2010/setores_censitarios_shp/mg/>

IBGE, (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA). **Censo Demográfico**, 2010b. Disponível em: <<https://sidra.ibge.gov.br/>>. Acesso em: 13 fev. 2019

ICHOKU, C.; KAHN, R.; CHIN, M. Satellite contributions to the quantitative characterization of biomass burning for climate modeling. **Atmospheric Research**, v. 111, p. 1–28, 1 jul. 2012.

ICHOKU, C.; KAUFMAN, Y. J. A method to derive smoke emission rates from MODIS fire radiative energy measurements. **IEEE Transactions on Geoscience and Remote Sensing**, v. 43, n. 11, p. 2636–2649, nov. 2005.

IEF, (INSTITUTO ESTADUAL DE FLORESTAS). **Inventário Florestal de Minas Gerais**. Disponível em: <<http://idesisema.meioambiente.mg.gov.br/>>. Acesso em: 1 nov. 2018.

Infoclima. Disponível em: <<http://infoclima.cptec.inpe.br/>>. Acesso em: 29 abr. 2019.

INMET, (INSTITUTO NACIONAL DE METEOROLOGIA). **Dados de estações automáticas disponibilizados via email pela agência**, 2017.

INMET, (INSTITUTO NACIONAL DE METEOROLOGIA). Banco de Dados Meteorológicos para Ensino e Pesquisa. 2018a.

INMET, (INSTITUTO NACIONAL DE METEOROLOGIA). **Normais Climatológicas do**

Brasil. Disponível em:
<<http://www.inmet.gov.br/portal/index.php?r=clima/normaisClimatologicas>>. Acesso em: 10 jan. 2018b.

JAISWAL, R. K. et al. Forest fire risk zone mapping from satellite imagery and GIS. **International Journal of Applied Earth Observation and Geoinformation**, v. 4, n. 1, p. 1–10, ago. 2002.

JUVANHOL, R. S. **Modelagem da vulnerabilidade à ocorrência e propagação de incêndios florestais**. [s.l.] Universidade Federal do Espírito Santo, 2014.

KAMALI, M. I. ET AL. The determination of reference evapotranspiration for spatial distribution mapping using geostatistics. **Water Resources Management**, v. 29, n. 11, p. 3929–3940, 2015.

KLINK, C. A.; MACHADO, R. B. A conservação do Cerrado brasileiro. **MEGADIVERSIDADE**, v. 1, n. 1, p. 9, ago. 2005.

LIMA, Á. H. N. DE. **Seleção de características de dados utilizando redes neurais artificiais**. Lavras: Monografia (graduação em Ciência da Computação) - Universidade Federal de Lavras, 2012.

LIU, D.; ZHANG, Y. Research of regional forest fire prediction method based on multivariate linear regression. **International Journal of Smart Home**, v. 9, n. 1, p. 13–22, 2015.

LONGO, K. M. et al. Biomass burning in Amazonia: Emissions, long-range transport of smoke and its regional and remote impacts. In: **Amazonia and Global Change**. [s.l: s.n.]. p. 207–232.

LÓPEZ, A. S.; SAN-MIGUEL-AYANZ, J.; BURGAN, R. E. **Integration of satellite sensor data, fuel type maps and meteorological observations for evaluation of forest fire risk at the pan-European scale****International Journal of Remote Sensing**, 2002.

LUCAS, T. DE P. B. **Chuvas persistentes e ação da zona de convergência do atlântico sul na região metropolitana de belo horizonte**. Belo Horizonte: Dissertação (mestrado em Geografia) -Universidade Federal de Minas Gerais, 2007.

LUZ, E. F. P. DA; BECCENERI, J. C.; CAMPOS VELHO, H. F. DE. A new multi-particle

collision algorithm for optimization in a high performance environment. **Journal of Computational Interdisciplinary Sciences**, v. 18, n. 1, p. 3–10, 2008.

MAEDA, E. E. et al. Predicting forest fire in the Brazilian Amazon using MODIS imagery and artificial neural networks. **International Journal of Applied Earth Observation and Geoinformation**, v. 11, n. 4, p. 265–272, 2009.

MAPBIOMAS. **Projeto MapBiomass – Coleção [3] da Série Anual de Mapas de Cobertura e Uso de Solo do Brasil**. Disponível em: <<http://mapbiomas.org>>. Acesso em: 1 jun. 2018.

MARQUES, C. B. et al. **Análise de correspondência entre mapas de vegetação florestal do bioma Cerrado**. XVIII Simpósio Brasileiro de Sensoriamento Remoto - SBSR. **Anais...2017** Disponível em: <<http://urlib.net/8JMKD3MGP6W34M/3PSM4HJ>>

MBOW, C.; GOÏTA, K.; BÉNIÉ, G. B. Spectral indices and fire behavior simulation for fire risk assessment in savanna ecosystems. **Remote Sensing of Environment**, v. 91, n. 1, p. 1–13, 2004.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, v. 5, n. 4, p. 115–133, dez. 1943.

MELO, E. T.; SALES, M. C. L.; OLIVEIRA, J. G. B. DE. Aplicação do índice de vegetação por diferença normalizada (ndvi) para análise da degradação ambiental da microbacia hidrográfica do riacho dos cavalos, CRATEÚS-CE. **Raega - O Espaço Geográfico em Análise**, v. 23, n. 23, p. 520–533, 3 nov. 2011.

MINNINCH, R. A. **The geography of fire and conifer forest in the Eastern Transverse Ranges, California**. [s.l.] UCLA, 1978.

MOREIRA, J. L. B. **Estudo da distribuição espacial das chuvas em Belo Horizonte e seu entorno**. Belo Horizonte: Instituto de Geociências da Universidade Federal de Minas Gerais, Belo Horizonte, 2002.

MOTTA, D. S. **Identificação dos fatores que influenciam no comportamento do fogo em incêndios florestais**. Seropédica: Monografia (Engenharia Florestal) - Universidade Federal Rural do Rio de Janeiro, 2008.

- MUSK, L. F. **Weather systems**. Cambridge: Cambridge University Press, 1989.
- MYERS, N. et al. Biodiversity hotspots for conservation priorities. **Nature**, v. 403, p. 853, 24 fev. 2000.
- NASA. **Active Fire Data**. Disponível em: <<https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms/active-fire-data>>. Acesso em: 1 jun. 2018.
- NASCIMENTO, J. C. DO.; AMARAL, R. S. DO.; PONTES, A. N. Determinação do grau de perigo de incêndios por meio da fórmula de monte alegre em municípios do estado do Pará, Brasil. **Revista Sodebras**, v. 8, n. 94, p. 24–28, 2013.
- NICOLAU, L. **Aula 3 -Funções de Ativação e Gradiente Descendente**CEAR UFPB, 2016. Disponível em: <<http://www.cear.ufpb.br/juan/wp-content/uploads/2016/08/Aula-3c-Funções-de-Ativação-e-Gradiente-Descendente.pdf>>. Acesso em: 5 jan. 2018
- NIMER, E. **Climatologia do Brasil**Rio de Janeiro, 1979.
- NUNES, R. S. FMA + - um novo índice de perigo de incêndios florestais para o estado do Paraná - Brasil. **Floresta**, v. 36, n. 1, p. 75–91, 2006.
- OLIVA, P.; SCHROEDER, W. Assessment of VIIRS 375m active fire detection product for direct burned area mapping. **Remote Sensing of Environment**, v. 160, p. 144–155, abr. 2015.
- OLIVEIRA, A. L. S. DE et al. Comparação e validação da modelagem espacial de riscos de incêndios considerando diferentes métodos de predição. **Boletim de Ciências Geodésicas**, v. 23, n. 4, p. 556–577, dez. 2017.
- OLIVEIRA FILHO, A. T. DE. Definição e delimitação de domínios e subdomínios das paisagens naturais do estado de Minas Gerais. In: SCOLFORO, J. R.; CARVALHO, L. M. T. DE (Eds.). **Mapeamento e inventário da flora nativa e dos reflorestamentos de Minas Gerais**. 2a. ed. ed. Lavras: UFLA, 2006. p. 21–35.
- PADILHA, C. K.; SATYAMURTY, P. **Um caso de estagnação de massa de ar quente e seco sobre a região central do Brasil**. XIII Congresso Brasileiro de Meteorologia. **Anais...**Fortaleza: 2004

PEZZOPANE, J. E. M.; NETO, S. N. DE O.; VILELA, M. DE F. Risco de incêndios em função da característica do clima, relevo e cobertura do solo. **Floresta e Ambiente**, v. 8, n. 1, p. 161–166, 2001.

PINTO JUNIOR, S. C.; SILVA, C. A. DA. A dinâmica climática do Mato Grosso do Sul e as queimadas do ano de 2009: uma análise têmporo-espaical a partir das imagens do NOAA-15. **Revista ACTA Geográfica**, n. 1970, p. 223–237, 2012.

POVEDA, L. M. R. **Estimação em populações finitas assistida por modelos para variáveis dicotômicas**. Recife: Universidade Federal de Pernambuco, 2006.

PRUDENTE, T. D. **Geotecnologias Aplicadas ao mapeamento de risco de incêndio florestal no parque nacional da chapada dos veadeiros e área de entorno**. Uberlândia: Universidade Federal de Uberlândia, 2010.

QUADRO, M. F. L. DE. Estudo de episódios de zonas de convergência do Atlântico Sul (ZCAS) sobre a América do Sul. **Revista Brasileira de Geofísica**, v. 17, n. 2–3, p. 210–210, 2005.

REBOITA, M. S.; AMBRIZZI, T.; PORFÍRIO DA ROCHA, R. Entendendo o Tempo e o Clima na América do Sul O Sol como Fonte de Energia. **Terrae Didactica**, v. 8, n. 1, p. 34–50, 2012.

REDDY, C. S. et al. Monitoring of fire incidences in vegetation types and protected areas of India: Implications on carbon emissions. **Journal of Earth System Science**, v. 126, n. 1, 2017.

REDEMET, (REDE DE METEOROLOGIA DO COMANDO DA AERONÁUTICA). **Dados disponibilizados pelo Website**, 2018. Disponível em: <<http://clima.icea.gov.br>>

REIS JUNIOR, W. DOS. **Caracterização das unidades geotécnicas da porção leste da Região Metropolitana de Belo Horizonte, MG**. Belo Horizonte: Dissertação(Mestrado em Geologia) - Universidade Federal de Minas Gerais, 2008.

RIBEIRO, L. et al. Zoneamento De Riscos De Incêndios Florestais Para a Fazenda Experimental Do Canguiri, Pinhais (PR). **Floresta**, v. 38, n. 3, p. 561–572, 2008.

RIBEIRO, L.; SOARES, R. V.; BEPLER, M. Mapeamento do risco de incêndios florestais

no município de Novo Mundo, Mato Grosso, Brasil. **Cerne**, v. 18, n. 1, p. 117–126, 2012.

ROESER, H. M. P.; ROESER, P. A. O Quadrilátero Ferrífero-MG, Brasil: aspectos sobre sua história, seus recursos minerais e problemas ambientais relacionados. **Geonomos 18(1)**: v. 18, n. 1, p. 33–37, 2010.

ROMERO, L. F. S. **Classificação de previsibilidade do modelo atmosférico global do CPTEC utilizando Breeding e Inteligência Computacional**. São José dos Campos: Dissertação (Mestrado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, 2017.

ROTHERMEL, R. C. How to Predict the Spread and Intensity of Forest and Range Fires. **USDA For. Serv. Gen. Tech. Rep.**, p. 161, 1 jan. 1983.

SAATY, T. L. A scaling method for priorities in hierarchical structures. **Journal of Mathematical Psychology**, v. 15, n. 3, p. 234–281, 1977.

SACCO, W. F.; OLIVEIRA, C. R. E. DE. A new stochastic optimization algorithm based on a particle collision metaheuristic. **Proceedings of the 6th World Congress of Structural and Multidisciplinary Optimization, WCSMO, Rio de Janeiro**, n. June, 2005.

SAMBATTI, S. B. M. **Rede neural auto-configurada para assimilação de dados usando fpga para a circulação oceânica**. São José dos Campos: Tese (Doutorado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, 2017.

SANTOS, J. F.; SOARES, R. V.; BATISTA, A. C. Perfil dos incêndios florestais no Brasil em áreas protegidas no período de 1998 a 2002. **Floresta**, v. 36, n. 1, p. 93–100, 2006.

SATIR, O.; BERBEROGLU, S.; DONMEZ, C. Mapping regional forest fire probability using artificial neural network model in a Mediterranean forest ecosystem. **Geomatics, Natural Hazards and Risk**, v. 7, n. 5, p. 1645–1658, 2 set. 2016.

SCHROEDER, W. et al. The New VIIRS 375m active fire detection data product: Algorithm description and initial assessment. **Remote Sensing of Environment**, v. 143, p. 85–96, 5 mar. 2014.

SCHROEDER, W.; CSISZAR, I.; MORISETTE, J. Quantifying the impact of cloud obscuration on remote sensing of active fires in the Brazilian Amazon. **Remote Sensing of**

Environment, v. 112, n. 2, p. 456–470, 15 fev. 2008.

SEITO, A. I. et al. **A segurança contra incêndio no Brasil**. São Paulo: Projeto Editora, 2008.

SILVA, I. D. B. DA; PONTES, A. C. F. J.; PONTES JR., A. C. F. Elaboração de um Fator de Risco de Incêndios Florestais utilizando Lógica Fuzzy. **Biomatemática**, v. 21, p. 113–128, 2011.

SISTEMA INTEGRADO DE DEFESA SOCIAL. **Diretriz integrada de ações e operações**. Disponível em: <<https://diao.sids.mg.gov.br/ii-tipicas-bm/o-00000-referente-a-explosao-e-incendio/214-incendio-florestal>>. Acesso em: 1 fev. 2019.

SOARES-FILHO, B. S. Análise De Paisagem : Fragmentação E Mudanças. **Ecologia**, p. 75, 1998.

SOARES-FILHO, B. S. et al. Modelagem de dinâmica de paisagem: concepção e potencial de aplicação de modelos de simulação baseados em autômato celular. **Megadiversidade**, v. 3, n. 1–2, p. 74–86, 2007.

SOARES, V. R. Desempenho da “fórmula de monte alegre” índice brasileiro de perigo de incêndios florestais. **Cerne**, v. 4, p. 87–99, 1998.

SOARES, R. V. **Determinação de um índice de perigo de incêndio para a região centro paranaense, Brasil**. Turrialba, Costa Rica: Tese (Magister Scientiae) - Instituto Interamericano de Ciências Agrícolas da OEA, 1972.

SOS MATA ATLÂNTICA/INPE. **Atlas dos Remanescentes Florestais da Mata Atlântica - 2008**. São Paulo: [s.n.]. Disponível em: <<http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Atlas+dos+remanescentes+florestais+da+Mata+Atlantica#1>>.

SOUZA, G. S. DE et al. KRIGAGEM ORDINÁRIA E INVERSO DO QUADRADO DA DISTÂNCIA APLICADOS NA ESPACIALIZAÇÃO DE ATRIBUTOS QUÍMICOS DE UM ARGISSOLO. **Scientia Agraria**, v. 11, n. 1, p. 073, 28 fev. 2010.

SPÓSITO, T. C.; STEHMANN, J. R. Heterogeneidade florística e estrutural de remanescentes florestais da Área de Proteção Ambiental ao Sul da Região Metropolitana de

Belo Horizonte (APA Sul-RMBH), Minas Gerais, Brasil. **Acta Botanica Brasilica**, v. 20, n. 2, p. 347–362, 2007.

TEIXEIRA, R. DE A. et al. Improving generalization of MLPs with multi-objective optimization. **Neurocomputing**, v. 35, n. 1–4, p. 189–194, 1 nov. 2000.

UFMG, (UNIVERSIDADE FEDERAL DE MINAS GERAIS). **Macrozoneamento da Região Metropolitana de Belo Horizonte**, 2014. Disponível em: <<http://www.rmbh.org.br/central-cartog.php?formato=shp>>. Acesso em: 5 jun. 2017

USGS, (U.S. GEOLOGICAL SURVEY). **SRTM data**, 2014. Disponível em: <earthexplorer.usgs.gov>

USGS, (U.S. GEOLOGICAL SURVEY). **Dados LANDSAT-8**, 2017. Disponível em: <<https://earthexplorer.usgs.gov/>>

VADREVU, K.; LASKO, K. Intercomparison of MODIS AQUA and VIIRS I-Band Fires and Emissions in an Agricultural Landscape—Implications for Air Pollution Research. **Remote Sensing**, v. 10, n. 7, p. 978, 21 jun. 2018.

VADREVU, K. P.; EATURU, A.; BADARINATH, K. V. S. Fire risk evaluation using multicriteria analysis—a case study. **Environmental Monitoring and Assessment**, v. 166, n. 1–4, p. 223–239, 27 jul. 2010.

WEISS, S. M.; KULIKOWSKI, C. A. **Computer systems that learn: classification and prediction methods from statistics, neural nets, machine learning, and expert systems**. San Mateo, CA: M. Kaufmann Publishers, 1991.

ZADEH, L. A. Fuzzy sets. **Information and Control**, v. 8, n. 3, p. 338–353, 1965.

APÊNDICE A

Análise de Interpoladores

Como dados meteorológicos são escassamente distribuídos e pontuais é necessária a criação superfícies contínuas com informações meteorológicas espacialmente distribuídas, isso é possível através de métodos de interpolação (ALVAREZ et al., 2014). A Interpolação é definida como a predição dos valores de uma variável primária em pontos dentro da mesma região de locais amostrados (BURROUGH; MCDONNELL, 1998).

Existem dois grupos principais de interpolação: os determinísticos e os geoestatísticos. Os métodos determinísticos criam superfícies de dados medidos baseados na extensão de similaridade (inverso da distância ao quadrado) ou no grau de suavização (funções de base radial), métodos que podem forçar resultados a passar pelos valores dos dados ou não. Já os métodos geoestatísticos quantificam a correlação espacial entre todos os pontos medidos, levando em conta a configuração espacial dos pontos em torno do local de previsão, utilizando propriedades estatísticas dos dados medidos para produzir mapas em Raster (KAMALI, 2015).

Neste estudo foram comparados três métodos de interpolação. O método de interpolação do Thin Plate Spline (TPS) consiste em ajustar uma superfície para os dados com algum erro associado a cada ponto. Neste método cada estação é omitida da estimação e o erro médio é encontrado. Isso é repetido para um intervalo de valores de parâmetros suavizados, então o valor que minimiza o erro é base para se ter uma suavização ótima (BOER; DE BEURS; HARTKAMP, 2001).

A Krigagem, método geoestatístico, se utiliza da dependência espacial entre amostras vizinhas para estimação dos valores, e isso é feito com o ajuste de um semivariograma. No caso da krigagem ordinária a variância está relacionada com a distribuição espacial das amostras, ou seja, depende apenas da distância e não dos valores dos atributos. No método do inverso do quadrado da distância (IDW), é calculado o inverso do quadrado da distância euclidiana entre o ponto conhecido e o estimado (SOUZA et al., 2010).

E isso foi feito utilizando-se o pacote “gstat” e “fields” do R. Foi utilizado o algoritmo disponibilizado por Hijmans e Ghosh (2019), onde são utilizados para teste um modelo ótimo para o IDW, através de um processo de otimização, a Krigagem Ordinária

Exponencial e o modelo do Thin Plate Spline simples. Para validação do modelo de interpolação, é feita uma validação cruzada estratificada em duas amostras e tem-se a construção de um modelo médio ponderado, onde os pesos são feitos de acordo com o Raiz do Erro Médio Quadrático (RMSE) de cada modelo.

A Tabela 1 apresenta os valores do RMSE para cada variável de acordo com cada interpolador. Não foi possível fazer a validação do TPS com os dados de Radiação devido ao número de estações. Os melhores modelos são os que apresentaram os menores valores de RMSE. Nota-se que para a maioria dos resultados, a krigagem foi o melhor modelo, apenas a pressão apresentou como melhor o IDW e a Umidade Relativa como o TPS.

Tabela 1 - Resultados do RMSE para modelos de interpolação

	IDW	Krigagem	TPS
Temperatura	1,64	1,62	1,74
Vento	4,32	3,91	5,78
Radiação	144,23	111,19	-
Pressão	19,97	21,46	46,86
Precipitação	41,59	36,72	41,48
Umidade Relativa	6,09	6,18	5,67

Das Figuras A.1 a A.6 são apresentados os resultados das interpolações e o modelo médio. O modelo médio no caso da radiação (Figura A.3) não foi possível de ser construído por causa da falta de validação do modelo TPS.

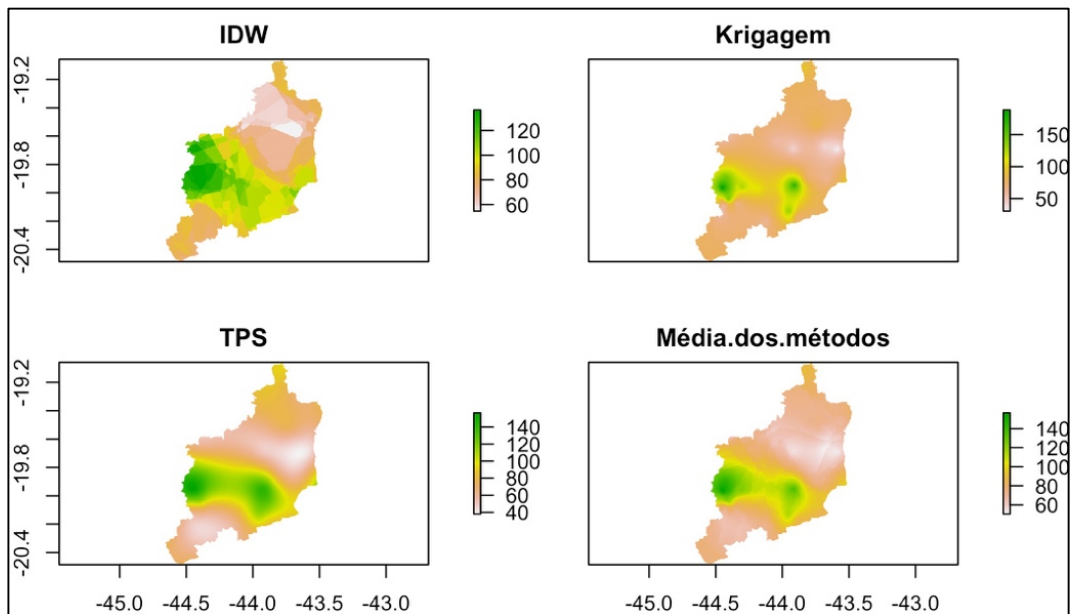


Figura A.1- Interpolação Modelos para Precipitação

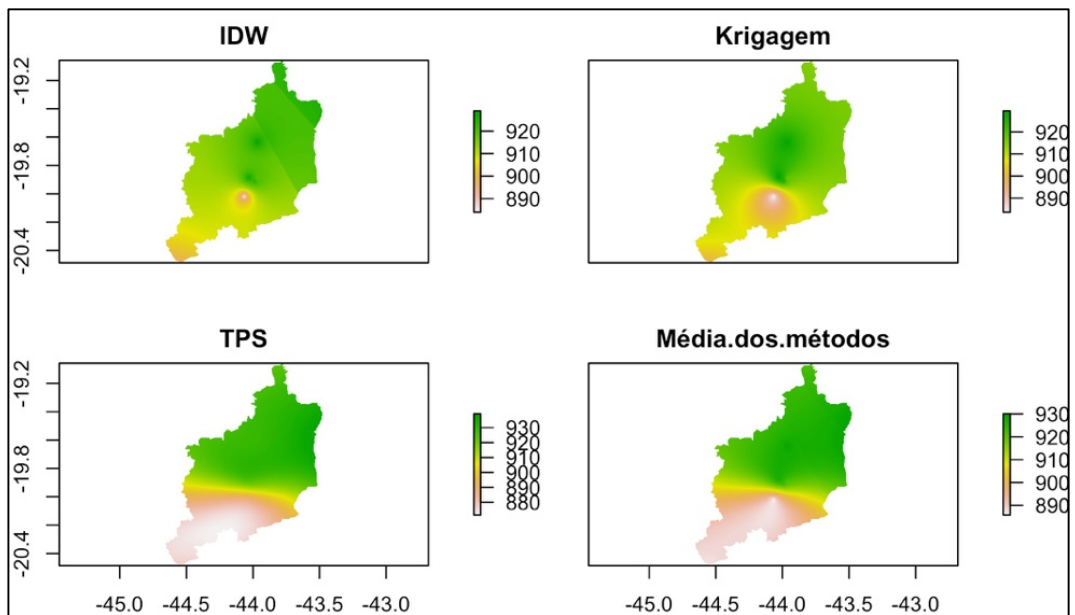


Figura A.2 - Interpolação Modelos para Pressão

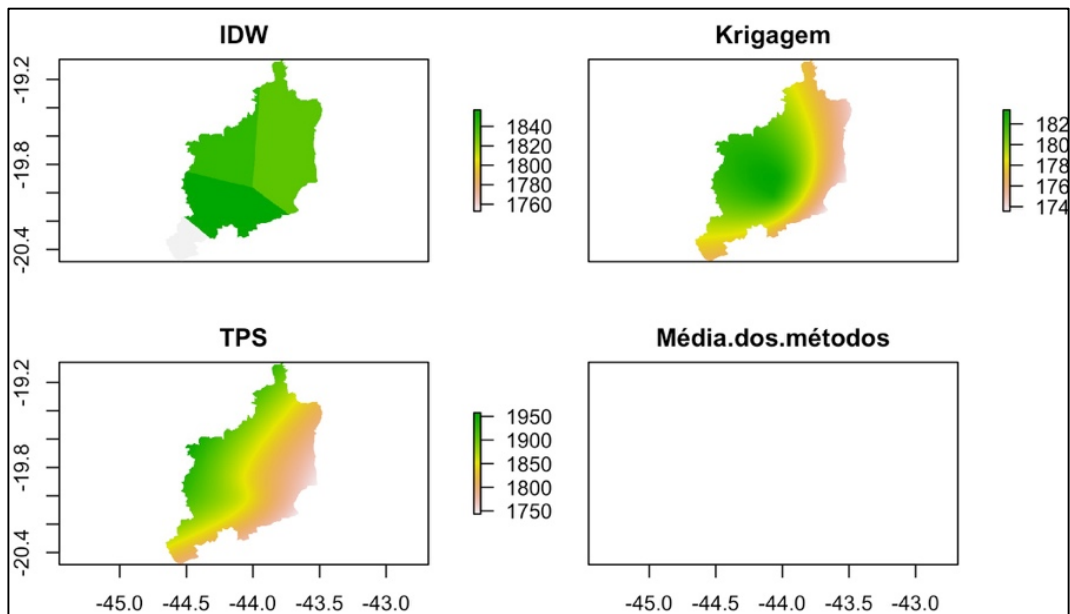


Figura A.3- Interpolação Modelos para Radiação.

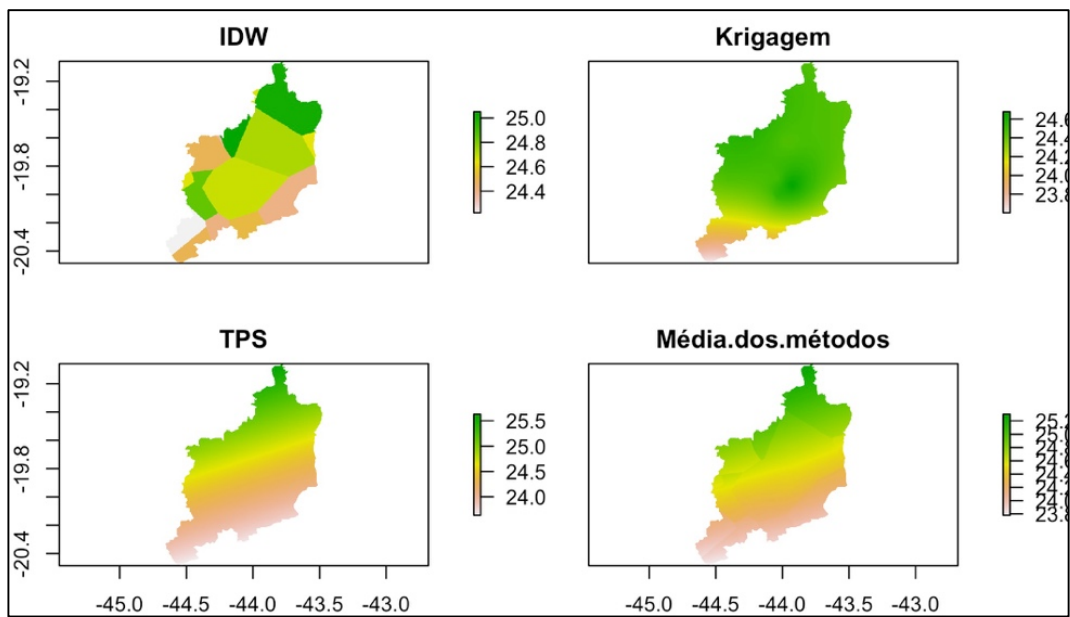


Figura A.4 - Interpolação Modelos para Temperatura

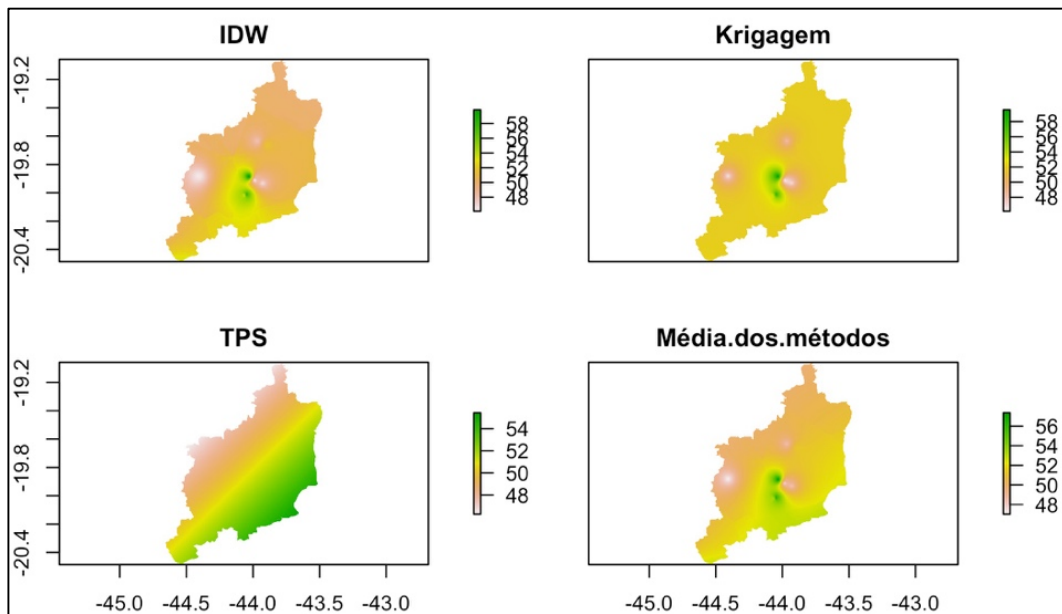


Figura A.5 - Interpolação Modelos para Umidade Relativa

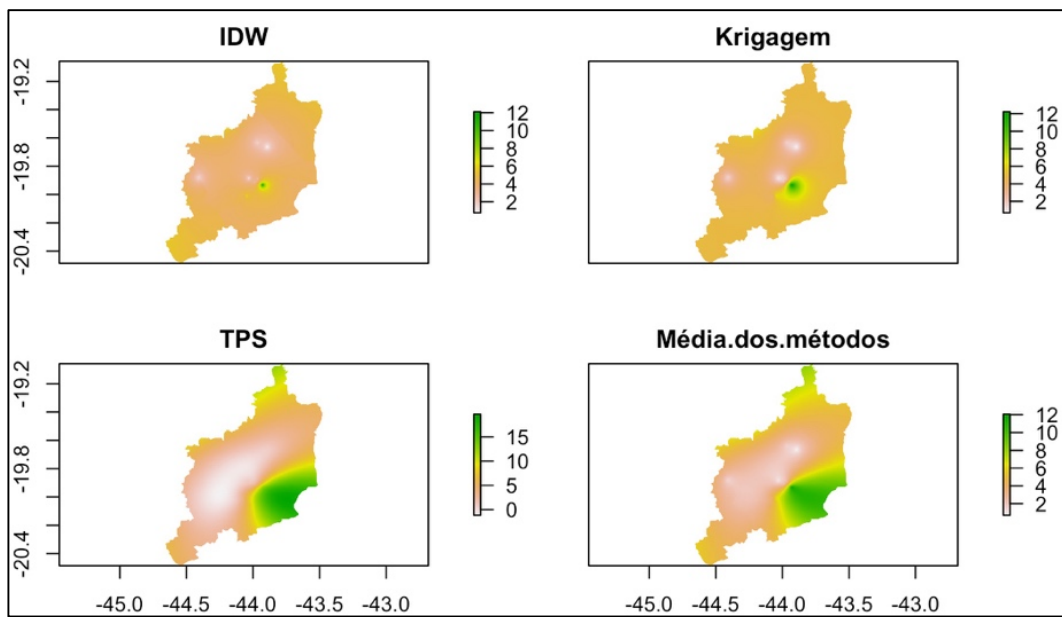


Figura A.6 - Interpolação Modelos para Vento

REFERÊNCIAS

ALVAREZ, O. et al. Comparison of elevation and remote sensing derived products as auxiliary data for climate surface interpolation. **International Journal of Climatology**, v. 34, n. 7, p. 2258–2268, jun. 2014.

BOER, E. P. J.; BEURS, K. M. DE; HARTKAMP, A. D. Kriging and thin plate splines for mapping climate variables. **ITC Journal**, v. 3, n. 2, p. 146–154, 2001.

BURROUGH, P. A.; MCDONNELL, R. A. **Principles of Geographical Information Systems**. Oxford: Oxford University Press, 1998.

HIJMANS, R. J.; GHOSH, A. **Spatial Data Analysis with R**, 2019. Disponível em: <<https://rspatial.org/analysis/analysis.pdf>>

KAMALI, M. I. ET AL. The determination of reference evapotranspiration for spatial distribution mapping using geostatistics. **Water Resources Management**, v. 29, n. 11, p. 3929–3940, 2015.

SOUZA, G. S. DE et al. Krigagem ordinária e inverso do quadrado da distância aplicados na espacialização de atributos químicos de um argissolo. **Scientia Agraria**, v. 11, n. 1, p. 073, 28 fev. 2010.

APÊNDICE B

Relação entre população Rural e Urbana na RMBH

Município	População Total	Porcentagem de População	
		Urbana	Rural
Balduim (MG)	7913	64%	36%
Belo Horizonte (MG)	2375151	100%	
Betim (MG)	378089	99%	1%
Brumadinho (MG)	33973	84%	16%
Caeté (MG)	40750	87%	13%
Capim Branco (MG)	8881	91%	9%
Confins (MG)	5936	100%	
Contagem (MG)	603442	100%	0%
Esmeraldas (MG)	60271	93%	7%
Florestal (MG)	6600	83%	17%
Ibirité (MG)	158954	100%	0%
Igarapé (MG)	34851	94%	6%
Itaguara (MG)	12372	77%	23%
Itatiaiuçu (MG)	9928	63%	37%
Jaboticatubas (MG)	17134	63%	37%
Nova União (MG)	5555	52%	48%
Juatuba (MG)	22202	98%	2%
Lagoa Santa (MG)	52520	93%	7%
Mário Campos (MG)	13192	94%	6%
Mateus Leme (MG)	27856	89%	11%
Matozinhos (MG)	33955	91%	9%
Nova Lima (MG)	80998	98%	2%
Pedro Leopoldo (MG)	58740	85%	15%
Raposos (MG)	15342	95%	5%
Ribeirão das Neves (MG)	296317	99%	1%
Rio Acima (MG)	9090	87%	13%
Rio Manso (MG)	5276	53%	47%
Sabará (MG)	126269	97%	3%
Santa Luzia (MG)	202942	100%	0%
São Joaquim de Bicas (MG)	25537	73%	27%
São José da Lapa (MG)	19799	58%	42%
Sarzedo (MG)	25814	99%	1%
Taquaraçu de Minas (MG)	3794	46%	54%
Vespasiano (MG)	104527	100%	-
Total	4883970	98%	2%

Fonte: IBGE (2010)

APÊNDICE C

Análise detalhada de 10 estações de precipitação na RMBH

Verificando a Figura C.1 é possível notar que em Baldim e Jaboticatubas, cidades mais ao Norte da RMBH, a precipitação média é mais baixa, aproximadamente 1150 mm. Ressalta-se o ano de 1991 em Jaboticatubas que teve uma precipitação menor que 500 mm. Em Taquaraçu de Minas e Vespasiano, a precipitação média registrada foi de 1250 mm aproximadamente. Os municípios de Caeté e Sabará, situadas à Nordeste da RMBH apresentaram maior precipitação média, aproximadamente 1450 mm. O município de Brumadinho, a sudeste da região, também apresentou precipitações médias menores, 1250 mm, sendo que em 2011 apresentou um evento ainda menor, precipitação com 886 mm. Por outro lado, Betim, Florestal e Mateus Leme, municípios a sudoeste da RMBH, apresentaram precipitações médias próximas de 1400 mm. Em particular, no município de Mateus Leme ocorreu precipitação mais baixa, 918 mm. Com essa análise é possível perceber que a precipitação varia na região, podendo ter eventos isolados em cada lugar.

Nota-se que não há muita diferença entre os meses apresentados na região, todos seguem o mesmo padrão (Figura B.1). Quanto à precipitação máxima (Figura B.2), nota-se que as máximas registradas estão nos municípios de Caeté e Sabará em 2003, sendo Caeté o município que apresentou o maior número de altas precipitações máximas durante os anos. Deve-se destacar também, o município de Florestal, que apresentou uma alta precipitação máxima em 1991.

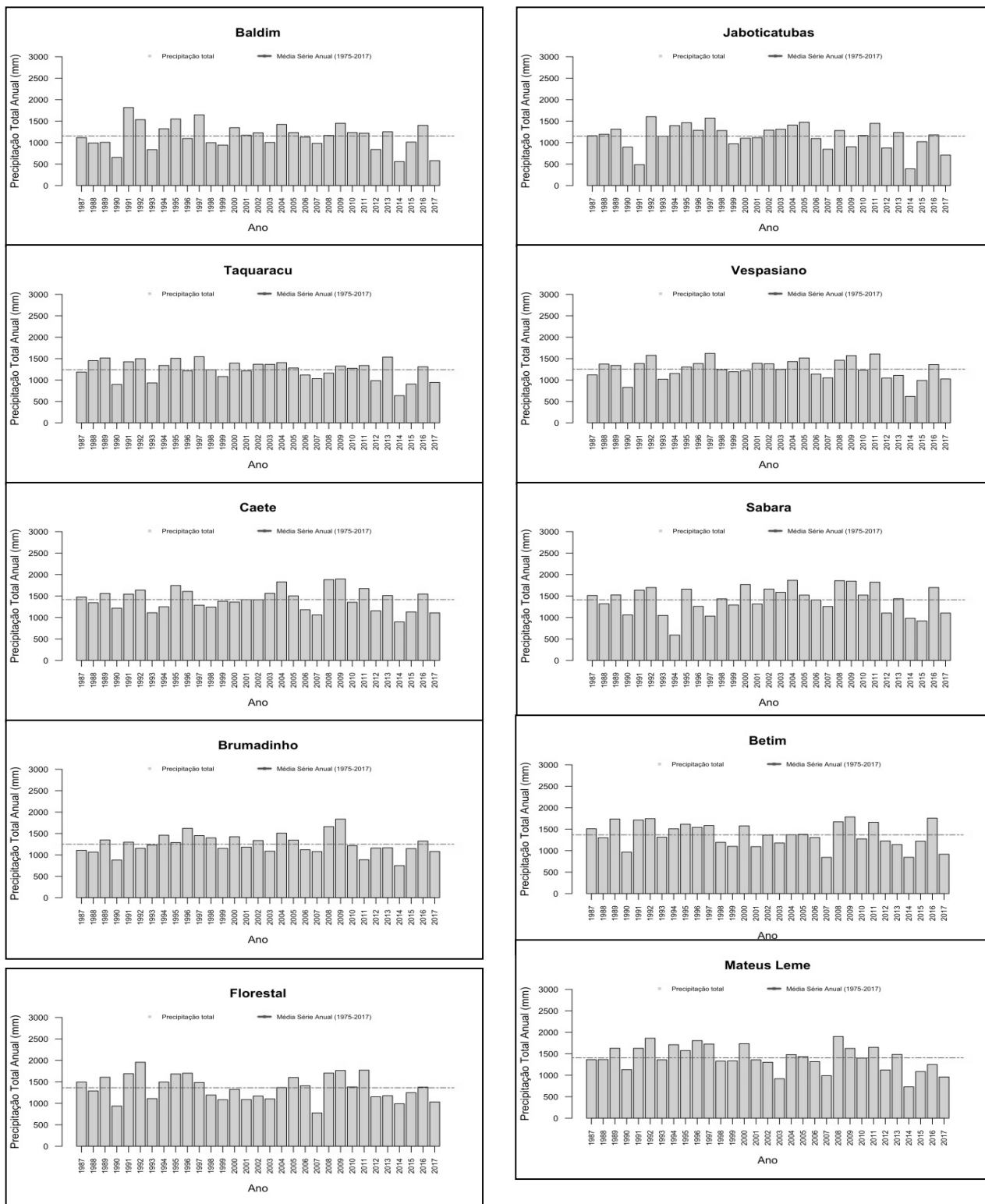


Figura C.1 – Precipitação Total Anual nas estações da RMBH das séries histórica de 1987 a 2017
Fonte dos dados: ANA (2018)

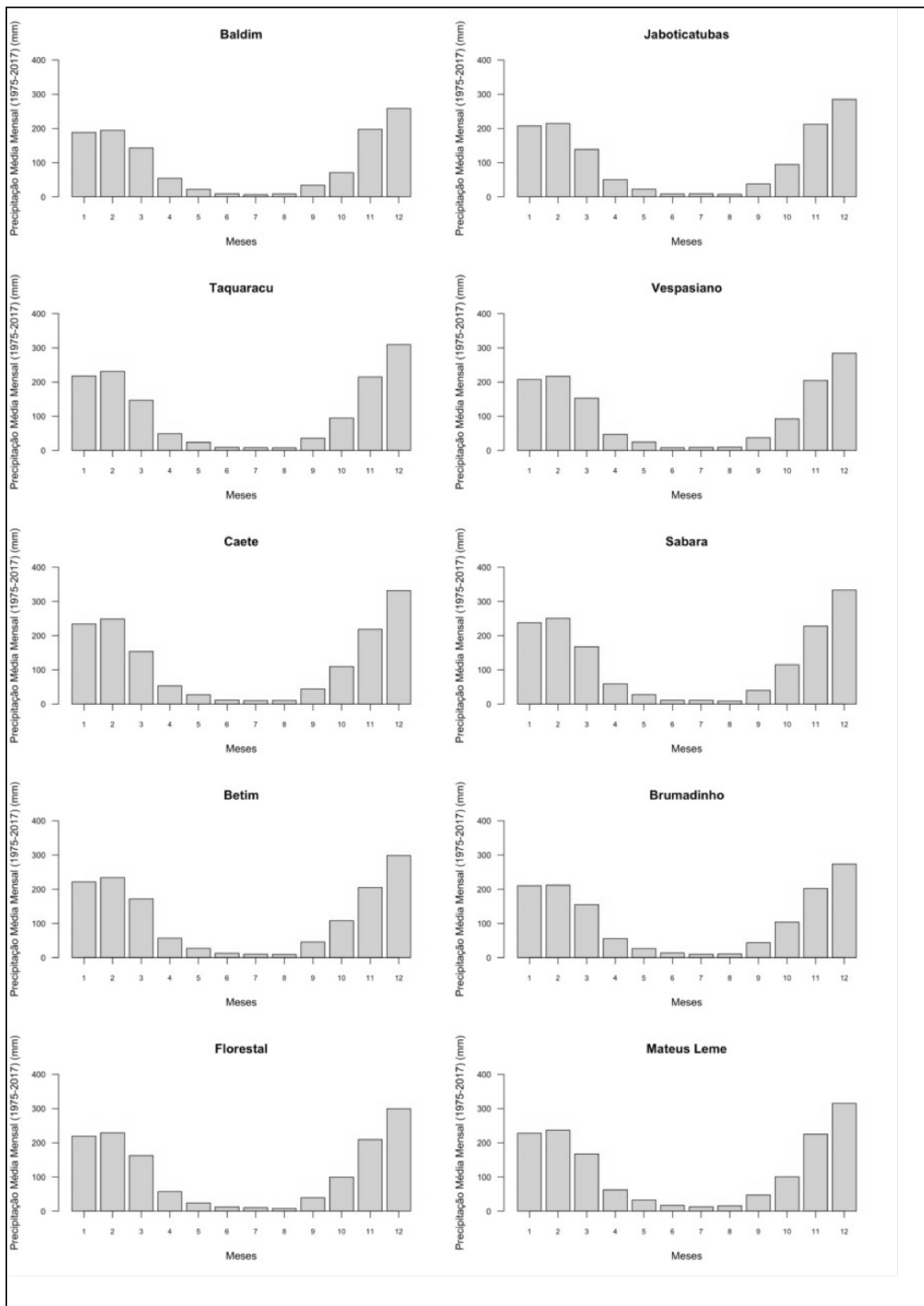


Figura C.2 – Precipitação Total Mensal nas estações da RMBH das séries histórica de 1987 a 2017 Fonte dos dados: ANA (2018)

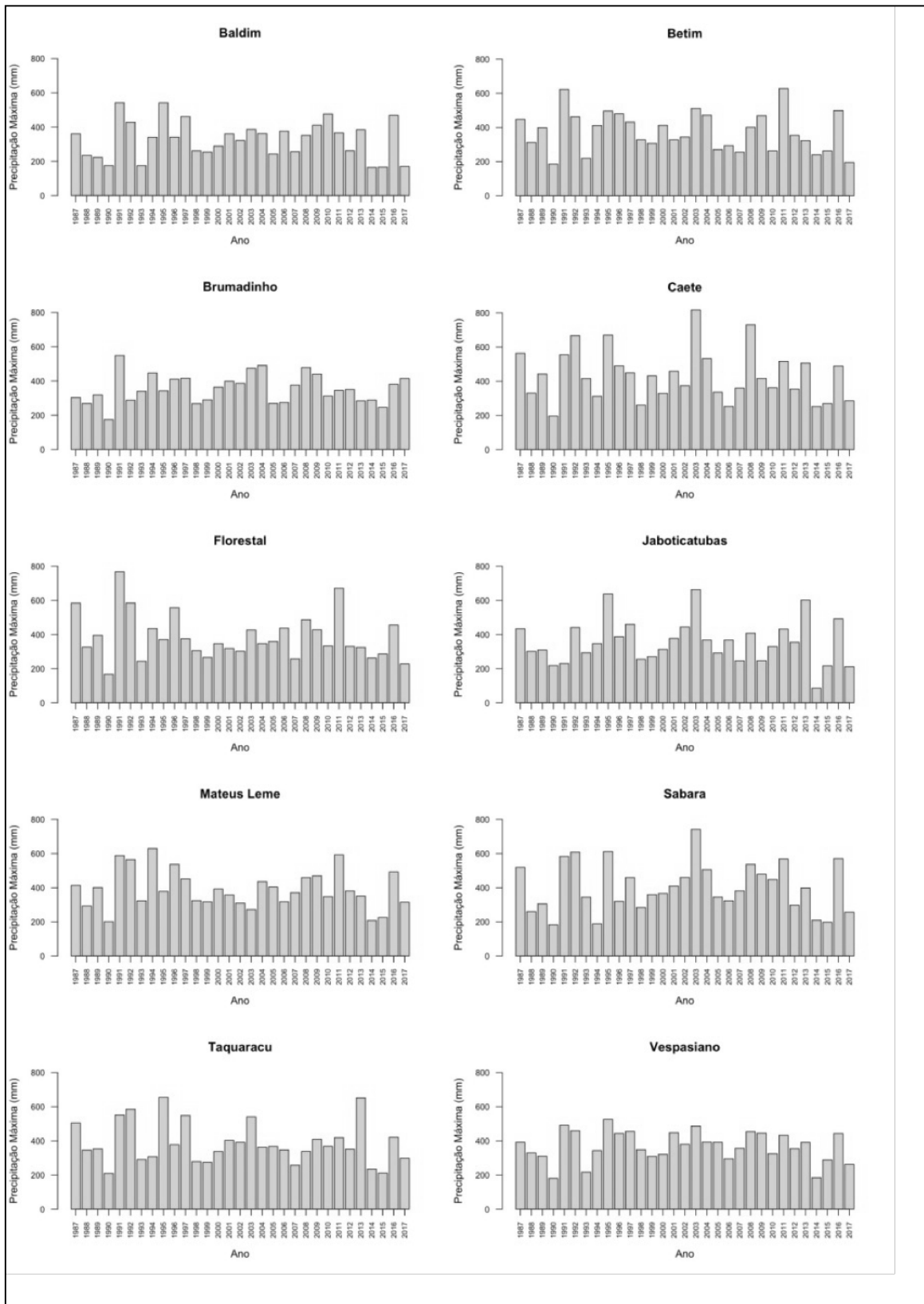


Figura C.3 – Precipitação Máxima Registrada nas estações da RMBH das séries histórica de 1987 a 2017
 Fonte dos dados: ANA (2018)

APÊNDICE D

Análise dados meteorológicos da região de estudo

Os dados meteorológicos utilizados para compor a metodologia foram analisados. É possível perceber que os dados de pressão atmosférica seguem os padrões normais, onde a pressão aumenta nos meses de inverno e diminui nos meses de verão, como apresentado na Figura D.1

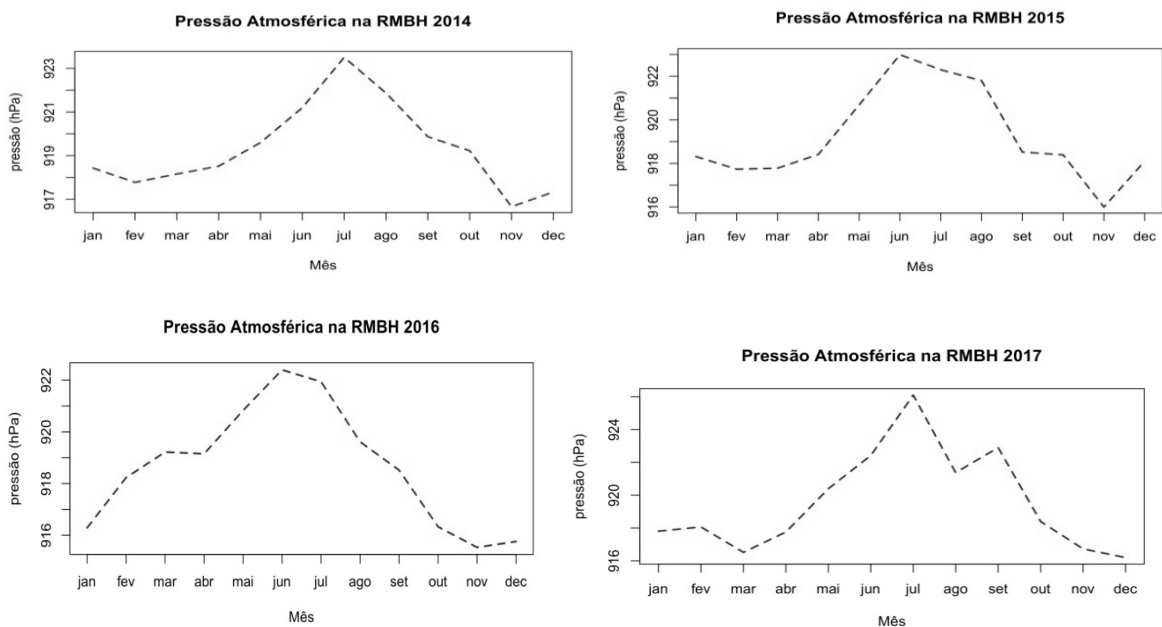


Figura D.1 - Pressão Atmosférica na RMBH
Fonte: (INMET, 2017)

Quanto à radiação (apontado na Figura D.2) apresenta os menores valores entre junho e julho, mesmo em novembro também os valores de radiação são menores, principalmente em 2017.

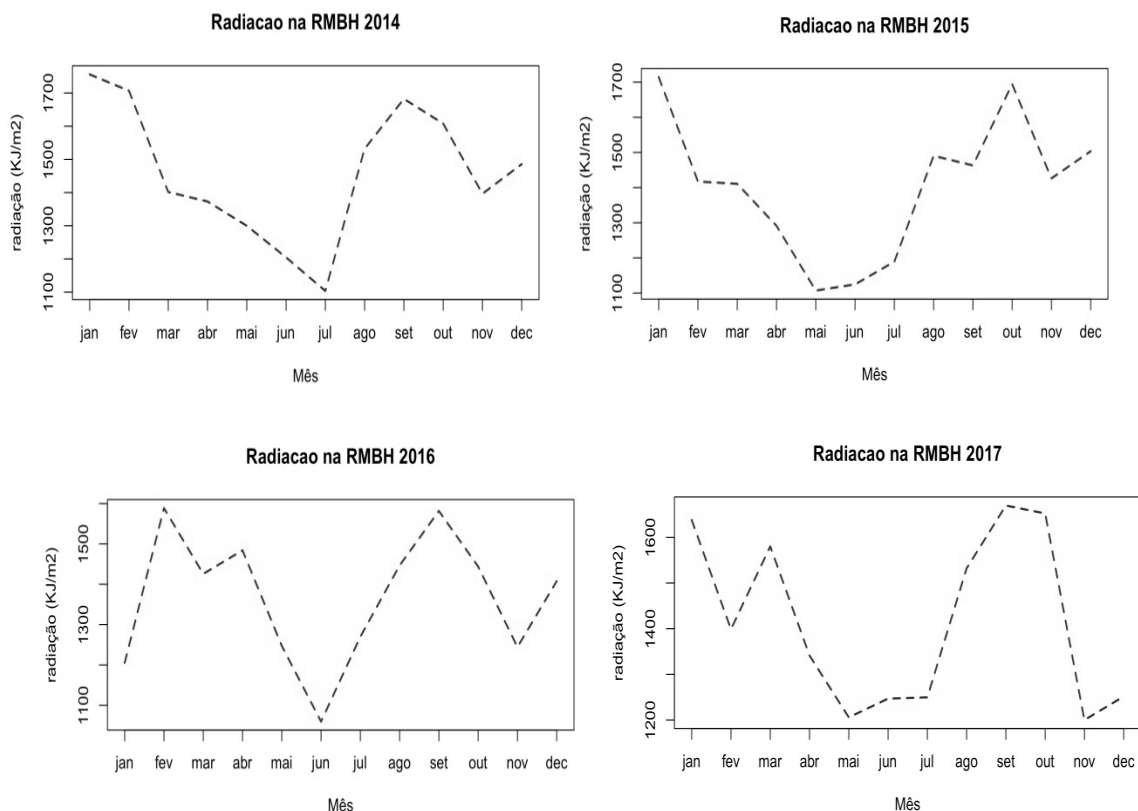


Figura D.2 - Radiação na RMBH

Fonte: (INMET, 2017)

Nota-se na Figura D.3, que a temperatura média durante os anos 2014 a 2017, ficou entre 24/25°C nos meses de verão e entre 18/19°C nos meses de inverno. Com destaque para 2017 que apresentou 16°C em julho de 2017, e temperatura mínima de 13°C. As temperaturas mínimas ficaram em 15°C no inverno e 20°C no verão. E as temperaturas máximas entre 28/29°C no verão e 22°C no inverno.

A umidade relativa do ar média apresenta valores mais altos de novembro a março, próximos a 70/75%, como pode ser observado na Figura D.4. Enquanto nos meses de agosto a outubro apresentam valores mais baixos, próximos a 50%. Nota-se que não há muita variação na distribuição da umidade entre os anos estudados. É importante notar que a umidade relativa mínima chega a valores próximos a 40% no inverno, condições de umidade alarmantes.

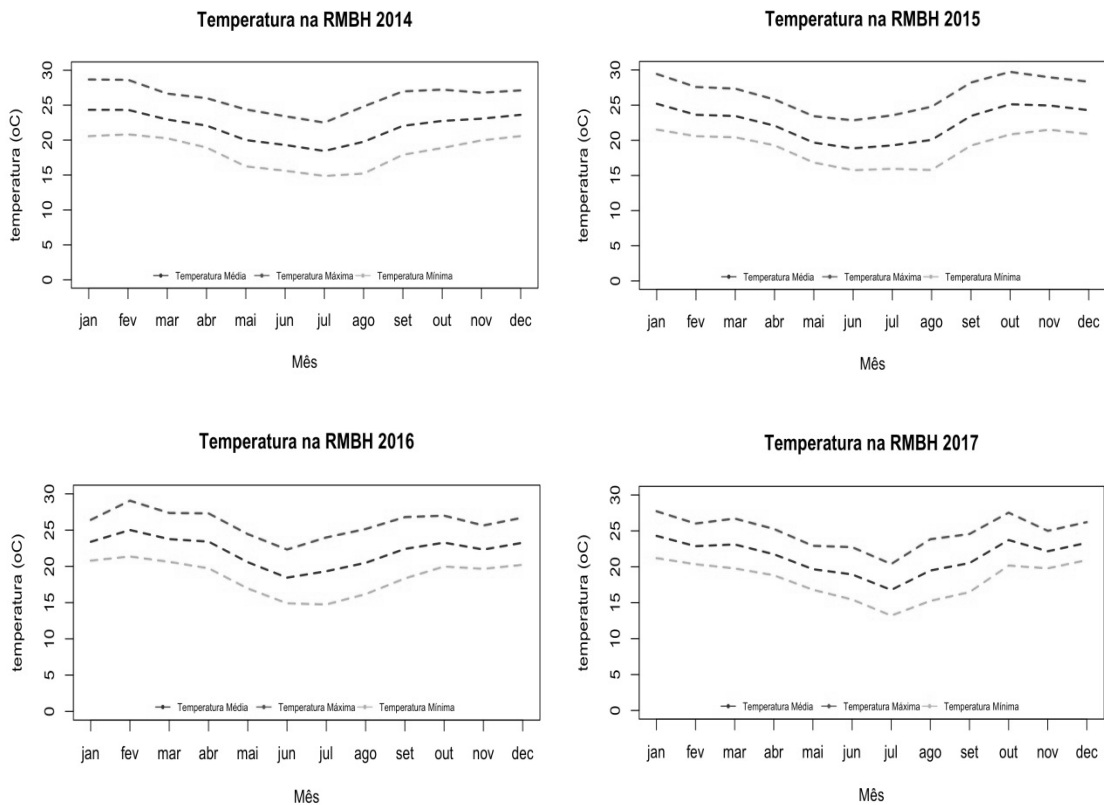


Figura D.3 - Temperatura na RMBH

Fonte: (INMET, 2017)

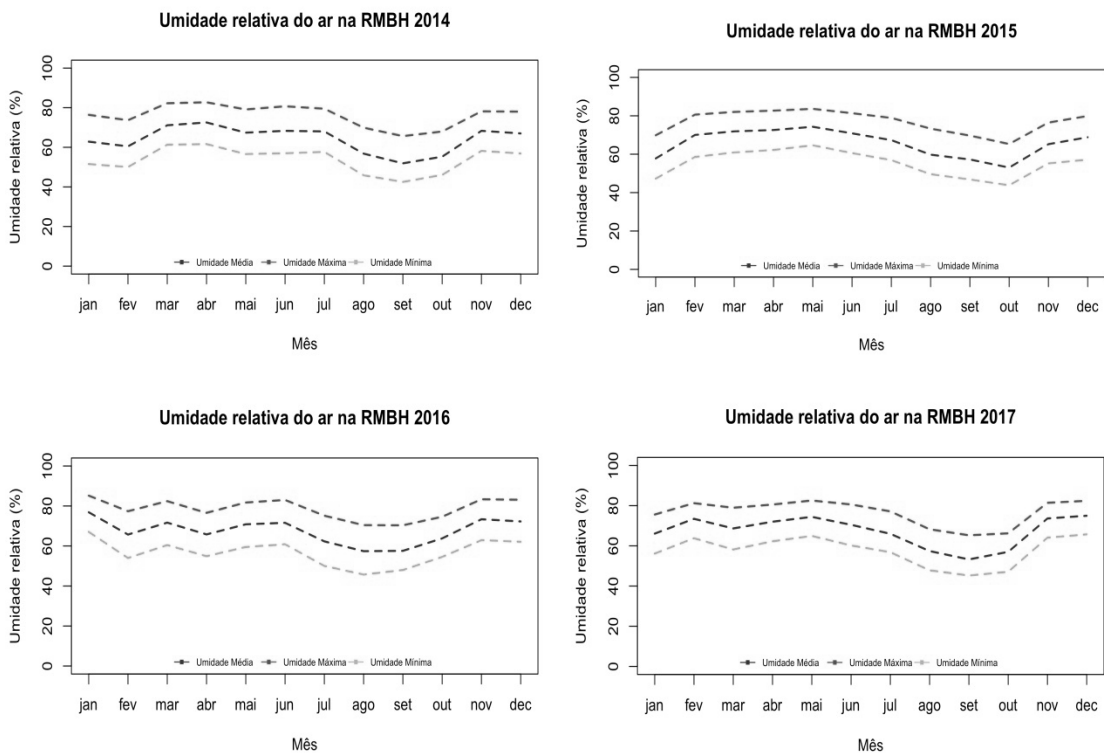


Figura D.4 – Umidade Relativa do Ar na RMBH

Fonte: (INMET, 2017)

A velocidade média do vento varia pouco durante os anos, ficando próxima a 5 km/h e a velocidade do vento máximas próximas de 10 km/h. Verificando um pequeno aumento nas rajadas no período do inverno, apontado na Figura D.5

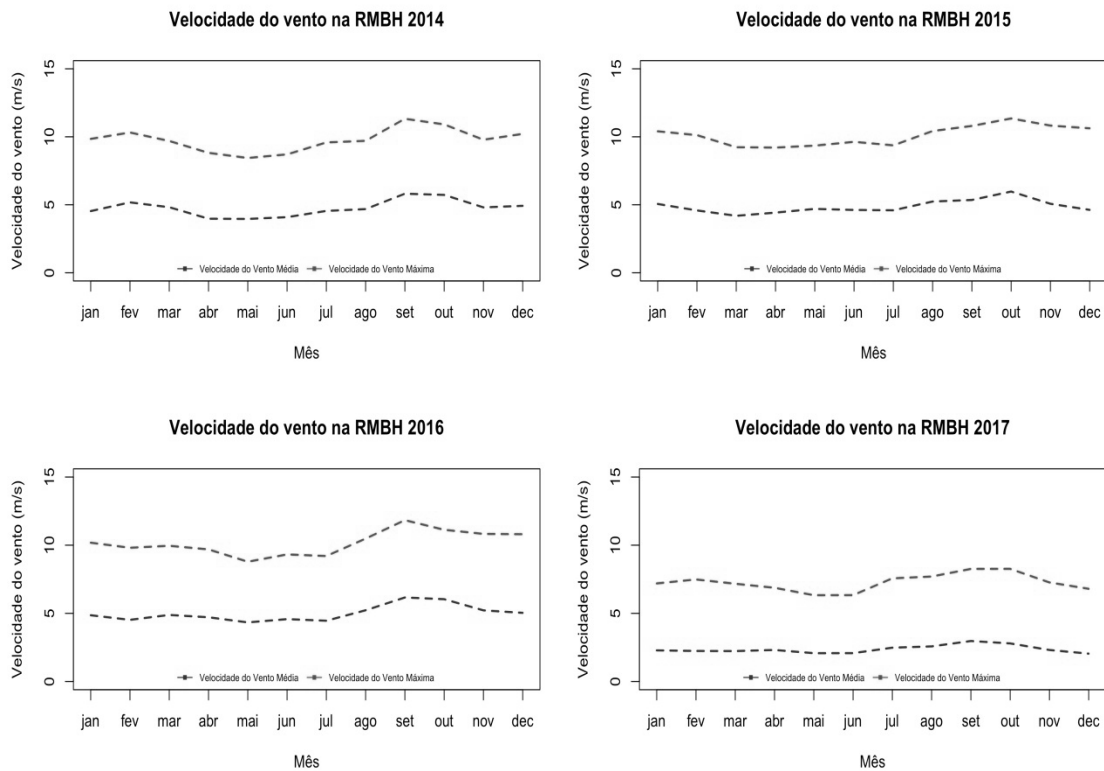


Figura D.5 – Velocidade do Vento na RMBH

Fonte: (INMET, 2017)

Nota-se que em 2014, a quantidade de precipitação mensal foi menor que 100 mm em quase todos os meses, apenas em novembro teve um evento de maior de precipitação, com 200 mm de chuva. Em 2015 o período de chuva e seca também foi bem acentuado, com precipitações quase nulas nos meses de junho a agosto. O ano de 2016 apresentou precipitações altas em janeiro, novembro e dezembro, com baixas a partir de fevereiro. E 2017 apresentou precipitações bem altas a partir de novembro, como mostrado na Figura D.6.

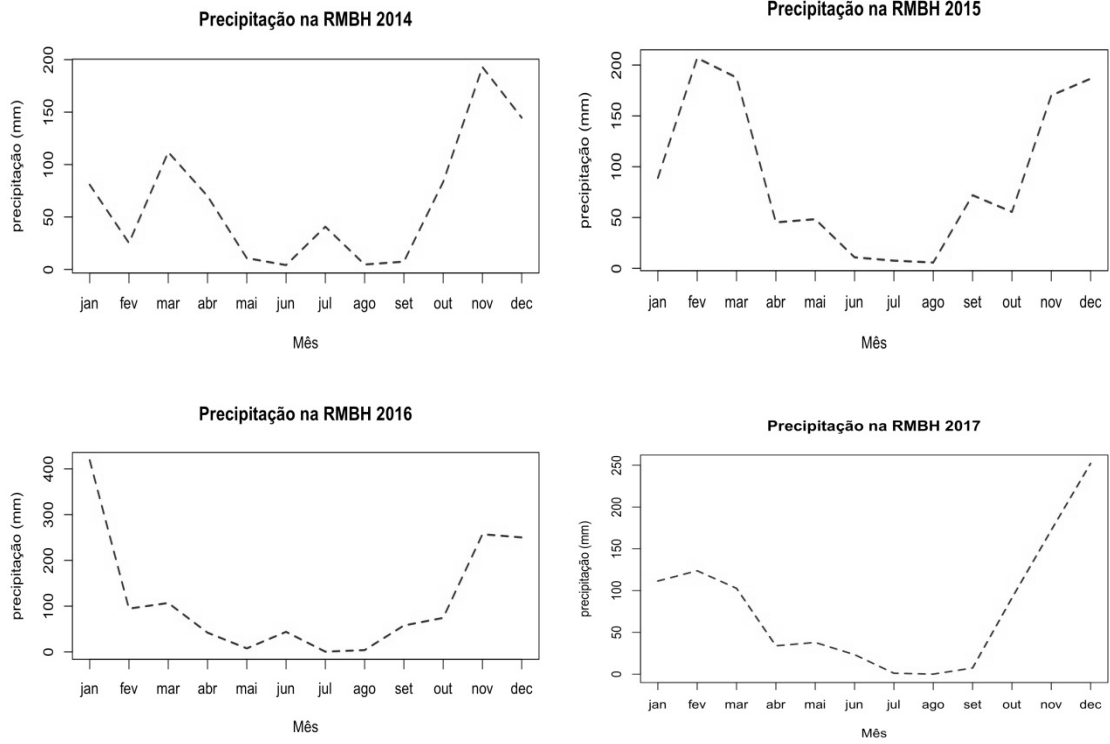


Figura D.6 – Precipitação Total na RMBH

Fonte: (ANA, 2018)

APÊNDICE E

Artigo:

MODELAGEM DE RISCO DE INCÊNDIOS FLORESTAIS: UMA VISÃO GERAL

MODELAGEM DE RISCO DE INCÊNDIOS FLORESTAIS: UMA VISÃO GERAL

FOREST FIRE RISK MODELING: A OVERVIEW

LUIZA CINTRA FERNANDES¹; MARCELO ANTONIO NERO¹; PLÍNIO TEMBA¹; BRITALDO SILVEIRA SOARES FILHO¹

1 – INSTITUTO DE GEOCIÊNCIAS/ UNIVERSIDADE FEDERAL DE MINAS GERAIS

luzacintraf@gmail.com, marcelo-nero@ufmg.br, temba.mobile@gmail.com, britaldo@csr.ufmg.br

Resumo – A avaliação do risco de incêndios é importante para estabelecer áreas prioritárias de risco e assim ter um melhor planejamento para prevenção e combate dos mesmos. Assim este artigo tem o objetivo de comparar diferentes abordagens para modelar o risco de incêndios. Foram discutidas cinco abordagens, a saber: Análise Multicritério, Lógica Fuzzy, Regressão logística, Redes Neurais e Fórmula de Monte Alegre. O método que aplica as Redes Neurais tem melhor desempenho em relação às outras, mas sua execução depende de conhecimentos computacionais e tempo para testar a melhor estrutura da rede. Então a escolha da melhor metodologia depende do objetivo do estudo e do tempo que se tempo disponível para a modelagem.

Palavras-chave: modelagem de risco de incêndios, revisão, métodos.

Abstract -The fire risk assessment is important to establish the priority risk zones to have a better planning of prevention and firefighting activities. So this paper has the aim to compare different approaches to modeling the fire risk. It was discussed four approaches the Multicriteria Analysis, Fuzzy Logic, Logistic Regression, Neural Networks and Monte Alegre Formula. The neural networks have a better performance, but your execution is more complex because depends in computational knowledge, and time to test the better structure for the network. So the choose of the methodology to be use depends of the aim of the work and the time to spend in the modeling.

Keywords: fire risk modeling, review, methods.

I. INTRODUÇÃO

Incêndios florestais são um dos maiores fatores de alteração do meio ambiente, tendo impactos globais, afetando áreas florestais com importante papel na emissão de gases de efeito estufa. O fogo tem sido usado historicamente para gestão do uso da terra e em muitos ecossistemas bem adaptados a ciclos de queimadas, mas mudanças recentes no clima e em fatores sociais transformam regimes tradicionais do fogo, aumentando seus efeitos negativos na vegetação, solo e valores humanos. Além dos efeitos globais, há efeitos locais, associados com a frequência e intensidade dos incêndios, o que implica em

degradação do solo, perda de vidas, de biodiversidade e de infraestruturas (CHUVIECO et al, 2010).

A avaliação do risco de queimadas é a parte crítica da prevenção de incêndios, já que as ações preventivas demandam recursos e ferramentas objetivas para monitorar quando e onde este pode ocorrer ou aonde terá mais efeitos negativos (CHUVIECO et al, 2010). Estabelecidas as regiões com maiores riscos é possível se ter um melhor planejamento em atividades de prevenção de incêndios florestais, como definição de estação de queimadas, permissão de queimadas controladas em períodos e lugares menos perigosos, previsão do comportamento do fogo para adotar técnicas mais efetivas de combate e informar a população quanto aos riscos. Em casos extremos, limitar o acesso em áreas de risco e proibir o uso de fogo próximo a florestas ou outras formas de vegetação (NUNES, 2006), além de verificar os lugares com prioridade no atendimento podem ser consideradas medidas preventivas efetivas.

Para o cálculo do risco de queimadas existem dois tipos de variáveis que podem ser consideradas: aquelas derivadas de fatores que não mudam em um curto período de tempo, chamadas de estáticas, e aquelas que mudam em períodos de tempo mais curtos, chamadas de dinâmicas. Os primeiros englobam declividade, tipos de vegetação, tipos de solo, proximidade com estradas e assentamentos. Já as variáveis dinâmicas podem-se citar variáveis meteorológicas, como temperatura, direção e intensidade do vento, umidade do ar, entre outras (AYANZ et al., 2003).

Na modelagem de risco de incêndios florestais existem muitas abordagens, como métodos estatísticos como a regressão logística; métodos quantitativos e qualitativos como a análise multicritério, métodos de inteligência artificial como redes neurais e lógica fuzzy, além de fórmulas que relacionam os parâmetros meteorológicos com o risco de incêndio, como a fórmula de Monte Alegre. Nesse sentido este artigo traz uma revisão e comparação destes métodos.

II. METODOLOGIAS ESTUDADAS

2.1. Análise Multicritério

Método desenvolvido por SAATY, 1977, chamado Analytic Hierarqu Process (AHP), aonde é estruturada uma hierarquia de decisão classificando os elementos de um nível mais alto para um mais baixo, depois é feita uma matriz de comparação para definir os pesos de cada variável. A matriz de comparação registra a importância percebida relativa de cada critério com cada outro critério e permite que os tomadores de decisão obtenham um conjunto de relação em escala de pesos para os critérios individuais. Por fim, uma relação de consistência (RC) é calculada para determinar se a avaliação foi bem-sucedida ou não.

No caso do cálculo do risco de incêndios são selecionadas variáveis que influenciam a ignição do fogo, por meio de mapas temáticos, como utilizado por Jaiswal et al 2002 ou dados de sensoriamento remoto, como utilizado por Chuvieco, 1989.

Na tabela 1 são mostradas algumas variáveis apontadas por esses autores analisados. Dessa maneira, é possível notar que cada autor coloca o peso na variável de acordo com seu conhecimento e sua área de estudo, assim como a definição de quais variáveis utilizar e as classes de risco de cada um (tabela 2). Para validação do modelo de Jaiswal et al (2002), esses autores atestaram que as áreas de maior risco de incêndio resultantes coincidiram com áreas próximas a assentamentos e rodovias, considerando o modelo válido, a partir de conhecimentos prévios que nessas regiões ocorrem mais queimadas. Já Chuvieco et al (1989) realizou a comparação com imagens de satélite, onde foi realizada a análise de coerência entre as áreas de pixels de queimadas, com as áreas de risco, sendo que o resultado não foi muito satisfatório. Isso foi constatado pois não houve coerência dos dados reais em para o que foi obtido na pesquisa, onde: 22,72 % das áreas foram classificadas como de alto risco, 15,28% como de médio risco e 3,74% como de baixo risco. Com isso aqueles autores concluíram que o importante do método é conhecer as áreas com maior risco, e que necessitando de estudos mais aprofundados para aprimorar a metodologia.

2.2. Lógica Fuzzy

A lógica fuzzy, introduzida por Zadeh (1965), é um problema de imprecisão associado com conjuntos aonde não há uma transição definida de uma classe a outra, ou seja eliminando os limites entre elas. O grau de uma associação de um objeto varia de 0 a 1. Onde o valor 1 representa uma associação completa e quanto mais próximo de 0 mais fraca é esta associação. As classes contínuas dos conjuntos fuzzy são construídas a partir de conceitos definidos a priori usando a experiência, ciência ou conhecimento heurístico. É usado o conhecimento linguístico de um fenômeno complexo transformando-o em dados numéricos para processamento. Esta técnica tem ampla aplicação em várias outras linhas de pesquisa, tais como exposto em Steiner et al (2013), os quais aplicaram em tomadas de decisões de investimento em ações e em Saragioto & Pereira (2012) que utilizaram para controlar a velocidade de uma linha de montagem.

Tabela 1 – Comparação entre análises de multicritério feita por diferentes autores

	Jaiswal et al 2002			Chuvieco et al 1989			Ribeiro et al 2007			
Decidivade (%)	0-4	B	2	0-3	B	2	até 15	B	1	
	5-8	B	2	3-5	M	3	16-25	M	2	
	9-12	B	2	5-10	M	4	26-35	A	3	
	13-16	M	1	10-15	A	5	36-45	M	4	
	17-20	M	1	15-35	M	6	>=46	A	5	
	21-36	M	1	>35	M	1		E		
	27-40	M	1		A	0				
	41-44	A	0							
>44	A	0								
Proximidade a estradas	< 50	A	0	< 100	M	8	< 50	A	1	
	>50	B	1	100-200	A	7	>50	N	0	
				200-300	M	5				
				300-400	B	3				
Aspecto	SE	A	0				SE/S/SW	B	1	
	SW	M	1				E	M	2	
	N	B	2					NW	A	3
								NW/W	M	4
								N	E	5
Elevação (m)	0-3	B	1	até 600	E	5				
	3-6	B	1	600-900	M	4				
	398-400	A	0	900-1200 m	A	3				
	401-404	A	0	1200-1500	M	2				
	405 - 407	A	0	acima de 1500	B	1				
Habituação (m)				0	M	8				
				< 1000	A	7				
				1000-2000	M	5				
				2000-3000	B	2				

*Legenda: E=Extremo, MA=Muito Alto, A=Alto, M=Moderado, B=Baixo, N=Nulo

Na construção de um conjunto fuzzy deve-se definir uma função de pertinência, que pode ter várias formas, sendo as mais comuns: triangular, trapezoidal e em forma de sino (função normal). Além disso é estabelecido um conjunto de regras de acordo com o conhecimento do especialista, na forma de proposições do tipo “se E1 e/ou E2 então S1”, por exemplo “Se umidade é baixa e precipitação é baixa então o risco é alto”.

Tabela 1 – Atribuição da classe de risco de incêndio de acordo com vários autores.

Jaiswal et al 2002	Chuvieco et al 1989	Ribeiro et al 2007
162-98	MA	0-30
90	A	31-60
75	M	61-160
66-54	B	161-190
		191-230
		231-255

*Legenda: E=Extremo, MA=Muito Alto, A=Alto, M=Moderado, B=Baixo, N=Nulo

Este método é aplicado por Ferreira et al (2001) e Oliveira & Nero (2013) com variáveis estáticas como declividade, distância a estradas, vegetação entre outros. Oliveira & Nero (2013) obtiveram um resultado satisfatório, comparando seu modelo com dados de ocorrência de incêndio fornecidos pela Brigada de Incêndios do Município de João Pessoa-PB (capital do Estado da Paraíba, Brasil), conseguindo explicar seu modelo em 85% em um dos períodos considerando as sazonalidades (estações de seca e de chuvas). Já Ferreira não validou seu modelo com dados de incêndio, mas indicou que os lugares que resultaram um risco mais alto estão de acordo com o esperado.

Silva & Pontes (2011) também fizeram uma abordagem fuzzy utilizando variáveis dinâmicas, como umidade relativa e precipitação. Comparando índices de focos de calor registrados pelo INPE para validação de seu modelo, tendo respostas positivas. Dessa maneira, conseguiram atribuir altos valores de Fator de Risco exatamente no período em que ocorrem os focos de incêndio e até mesmo quando existe um número pequeno de focos isolados, o modelo atingiu a previsão desse comportamento com valores não nulos.

2.3. Regressão Logística

A regressão logística é a mais popular modelagem matemática que relaciona as variáveis independentes com a variável dependente dicotômica. Assim, a variável dependente seria o risco de fogo, que pode ser 0 (não fogo) ou 1(fogo), sendo representado pela adaptação da equação de Bernoulli apresentada na equação 1 (GOLDARAG, 2016).

$$P(Y = 1) = \frac{e^{\sum(\alpha + \beta_i X_i)}}{1 + e^{\sum(\alpha + \beta_i X_i)}} \quad (1)$$

Onde Y é a variável dependente, Xi são as variáveis independentes e α e β_i números reais. A partir de dados históricos conhecidos de ocorrência de incêndio é possível conhecer as relações entre as variáveis a partir do cálculo dos coeficientes, que são assim substituídos na equação 1.

É uma técnica muito utilizada pelos autores para predição e análise de ocorrência de incêndio, alguns mais recentes são Del Hoyo, Isabel & Vega (2011); GUO et al (2016); Hernandez-Leal et al (2006); Chuvieco et al (2010) e Goldarag (2016).

Del Hoyo, Isabel & Vega (2011) utilizaram dados socioeconômicos, tais como uso da terra, rodovias, áreas protegidas e recreacionais, como variáveis independentes, aplicando as variáveis dependentes de duas formas. No primeiro foram utilizados os pontos de ignição e no segundo modelo a densidade dos pontos. Finalmente, os autores realizaram a validação com dados independentes dos utilizados na construção do modelo, aplicando o método de

validação por meio da área sob a curva (AUC), obtendo-se o valor de 0,709 para o primeiro modelo e para o segundo 0,672. Já GUO et al (2016) utilizaram em seu modelo variáveis como topografia, vegetação, infraestrutura, dados socioeconômicos e meteorológicos e como variável dependente dados de focos de calor obtidos de dados de satélite. Neste último caso o valor de AUC foi de 0,843, indicando que o modelo teve um bom resultado.

Hernandez-Leal et al (2006) utilizou variáveis dinâmicas, como declividade, altitude, insolação, proximidade a rodovias e vegetação em seu modelo, sendo utilizada a área queimada, contabilizada a partir de imagens NDVI, como dados históricos para o cálculo dos coeficientes. Em seus resultados percebeu-se que há uma sinergia entre os valores de risco de fogo e o NDVI, que indica indiretamente o estresse hídrico nas plantas. Os dados não foram validados com dados históricos, mas como foram utilizadas apenas variáveis estáticas, os autores concluíram da necessidade de se utilizar variáveis meteorológicas para o aprimoramento do modelo proposto, sugerindo a adoção da metodologia para o conhecimento prévio dos lugares de maior risco de fogo.

Goldarag (2016) por sua vez aplicou a regressão logística utilizando-se de variáveis estáticas e meteorológicas, como, temperatura do ar, umidade, precipitação, vento e duração do sol. Assim, foram utilizados dados históricos de ocorrência de incêndio obtidos do sensor MODIS. Os resultados apresentaram uma acurácia de apenas 50,84%, com um R² de 0,533, ou seja, o modelo não estava muito bem explicado. Para melhorar o resultado o autor diminuiu os valores de pixels representativos da não ocorrência de focos de incêndio, utilizando amostras onde os números de focos de incêndio eram maiores. Com isso o autor melhorou muito o modelo, obtendo a acurácia de 80% para alguns anos. No entanto, essa estratégia, do ponto de vista prático, fez com que o modelo não fosse efetivamente coerente em toda a área de estudo.

2.4. Redes Neurais

Uma rede neural é constituída de unidades de processamento simples que trabalham em paralelo, armazenando conhecimento experimental e tornando-o disponível para uso. Se assemelhando ao cérebro em dois aspectos: o conhecimento adquirido pela rede pelo processo de aprendizagem e as forças de conexão entre os neurônios, conhecidas como pesos sinápticos, que armazenam o conhecimento adquirido (HAYKIN, 1998). Ela se difere de sistemas de especialistas, os quais oferecem ao usuário uma resposta definitiva, caso as características avaliadas corresponderem àquelas codificadas nas bases das regras. Na rede neural é realizada uma análise das informações e é fornecida uma estimativa de probabilidade de que o dado é compatível com as características que foi treinado a reconhecer, podendo a chegar a uma acurácia de 100%, dependente da experiência adquirida pela rede (CANNADY, 1998 apud ALMEIDA & VIERA, 2015).

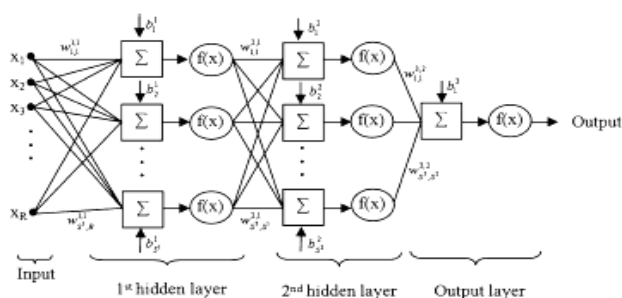
Para utilização de uma rede é necessário definir sua arquitetura, ou seja, como está estruturado o número de entradas, camadas e nós ocultos e o número de saída dos nós. Na figura 1 tem-se um exemplo de uma rede *Perceptron* multicamadas, com duas camadas escondidas, sendo que cada entrada (x) é multiplicada por um peso (w) e a soma das entradas com os pesos e o bias (b) resulta em

uma função de transferência ($f(x)$), que se transforma em entrada para a camada seguinte. O índice S representa o número de neurônios. Portanto, cada camada manda sinais para a próxima camada, como se fossem sinapses. A quantidade de neurônios e camadas pode ser modificada para melhorar o desempenho do modelo, sendo que para se saber a melhor arquitetura é necessário se fazer experimentos, não havendo uma fórmula pronta de como se fazer.

Para aprendizagem da rede é utilizado um processo iterativo de ajustes aplicados a seus pesos sinápticos e bias, tornando a rede mais instruída após cada interação. Para isso, é utilizado um algoritmo de aprendizagem. Neste processo são selecionados dados históricos como saída e as variáveis escolhidas como entrada. Após o treinamento a rede funciona como modelo de previsão, onde são inseridas as variáveis e a saída é calculada automaticamente.

Este método tem sido aplicado na resolução de problemas diversos. Um exemplo é o aplicado por Almeida & Viera (2015) onde há a simulação de sistemas de corte de folhas para criação de um ambiente onde é possível conhecer o rendimento de produção a partir dos dados de defeitos presentes na folha e a produção requerida. Enquanto Cavalcante et al (2015) construíram uma rede para previsão de ajustes para o consumo de produtos químicos utilizados no tratamento de água.

Figura 1- Rede Neural Perceptron Multicamadas.



Fonte: Goldarag, 2016.

Goldarag (2016) também utilizou a abordagem de rede neural utilizando as mesmas variáveis para regressão logística. Esse autor fez um experimento variando a sua quantidade de neurônios desde 5 até 65, conseguindo uma acurácia máxima de 95%, a qual foi atingida com 20 neurônios, sendo que o autor teve um trabalho computacional desnecessário fazendo estes testes com até 65 neurônios. Outro processo de teste que o autor fez foi o uso de 11 algoritmos de treinamento, sendo que o que obteve o melhor desempenho foi o Levenberg-Marquardt, com acurácia de 91%. Comparando os dados de pontos em que houve incêndio com os dados em que teve alto risco de incêndios, todos os pontos foram contemplados.

Alonso-Betanzos et al (2003) também utilizou redes neurais para previsão de riscos de ocorrência de incêndio, mas ele utilizou apenas dados meteorológicos, empregando também o algoritmo Levenberg-Marquardt para treinamento. Este último autor realizou vários testes com as melhores estruturas, obtendo o melhor resultado com 9 neurônios e atingindo uma acurácia de 80%, na validação com dados históricos de ocorrência de incêndio.

2.5. Fórmula de Monte Alegre

Desenvolvida por Soares (1972) a partir de dados meteorológicos e de ocorrência de incêndios da Fazenda Monte Alegre, no município de Telêmaco Borba, PR, durante 7 anos (1965 a 1971). Esta foi amplamente utilizada por empresas e instituições florestais para estimar o grau de risco de incêndios. Consiste numa fórmula simples que requer apenas duas variáveis meteorológicas: umidade relativa do ar e precipitação. É um índice acumulativo que requer valores diários de umidade relativa do ar e quantidade de precipitação. Sua fórmula básica é apresentada na equação 2.

$$FMA = \sum_{n=1}^n \frac{100}{H} \quad (2)$$

Onde H é a umidade relativa do ar medida às 13:00 h e n é o número de dias sem chuva. Apresenta algumas restrições referentes à precipitação diária, como mostrado na tabela 3. O perigo de incêndio do dia é indicado através de uma escala de cinco níveis: Nulo ($FMA \leq 1,0$); Pequeno ($1,1 \leq FMA \leq 3,0$); Médio ($3,1 \leq FMA \leq 8,0$); Alto ($8,1 \leq FMA \leq 20,0$) e Muito Alto ($FMA > 20,0$).

Tabela 3 – Modificação no cálculo conforme a precipitação ocorrida no dia

Chuva do dia (mm)	Modificação no cálculo
$\leq 2,4$	Nenhuma
2,5 a 4,9	Abater 30% na FMA acumulada e somar (100/H) do dia
5,0 a 9,9	Abater 60% na FMA acumulada e somar (100/H) do dia
10,0 a 12,9	Abater 80% na FMA acumulada e somar (100/H) do dia
$\geq 12,9$	Interromper o cálculo e recomeçar no dia seguinte

Fonte: Nascimento, Amaral & Pontes, 2013

Soares, 1998, avaliou o desempenho do índice e comparou com outros índices existentes no mundo, aplicando-o em três regiões: no Distrito Florestal de Monte Alegre (o local aonde foi desenvolvido), norte do estado de Santa Catarina e no Distrito de Suwannee (no estado da Florida, U.S.A). Neste estudo ele verificou no primeiro caso que pelo menos 50% dos incêndios ocorreram quando o grau de risco estava alto, já no segundo 87% dos incêndios ocorreram quando estava alto e no terceiro teve uma correspondência de apenas 24,3% para risco muito alto e 52,6% para alto risco.

Nascimento, Amaral & Pontes (2013) avaliaram o desempenho da FMA em cinco municípios do estado do Pará: Paragominas, Conceição do Araguaia, Itaituba, Parauapebas e Marabá, entre os anos de 2010 e 2011. Verificando o desempenho do índice em cada cidade os autores repararam que existe uma tendência dos focos de incêndio de se concentrarem nas classes de perigo alto e muito alto, tendo 91% de correspondência entre os dados de incêndio na frequência média, porém houve casos onde ocorreram incêndios quando o índice estava nulo.

III. DISCUSSÕES

Comparando as abordagens apresentadas nota-se que a abordagem multicritério, é a que possui um desempenho pior em relação às outras, já que necessita muito dos conhecimentos do pesquisador na atribuição de pesos, por sua vez é a mais simples e rápida, caso a necessidade seja de apenas se ter uma ideia das áreas com maior risco, sendo o desenvolvimento bem mais rápido. Por outro lado, a abordagem Fuzzy é um melhoramento da multicritério já que ela não define limite entre as classes, mas também depende do operador para definir as funções a serem utilizadas e o intervalo dos pesos, além de possuir melhor acurácia com relação às outras.

A Fórmula de Monte Alegre é um índice simples e fácil de ser aplicado, caso se tenha disponível dados de estações meteorológicas. No entanto, por ser muito genérica se comporta de forma diferente em diferentes regiões, além de que ocorrem casos de incêndio quando o grau de perigo é nulo. Esse fato pode ser explicado devida a não consideração de outros fatores de ocorrência de incêndio, tais como do uso do solo e de fatores socioeconômicos.

Os métodos de regressão logística e de redes neurais definem os pesos das classes independentemente do pesquisador, comparando os dados já existentes de dados de incêndio para produzir os modelos. O método de regressão logística faz uso de análises estatísticas e o método de redes neurais que utiliza algoritmos computacionais que imitam o cérebro humano. Segundo os trabalhos pesquisados a regressão logística não apresentou uma acurácia muito boa, principalmente no trabalho de Goldarag (2006), onde se faz a comparação entre os dois métodos. O método de Redes Neurais foi o que apresentou melhor desempenho dentre todas as técnicas, mas é uma técnica mais complicada que necessita de conhecimentos computacionais, sendo que para apresentar o melhor resultado deve-se testar as melhores arquiteturas, o que torna o processo trabalhoso.

Um outro fator observado na modelagem do risco de incêndio foi o uso de variáveis estáticas e dinâmicas, sendo que os modelos com melhor acurácia foram os que utilizaram variáveis meteorológicas. Isso se explica pelo fato de que a ocorrência de incêndios é muito influenciada por esses tipos de variáveis. Adicionalmente, nas pesquisas onde essas variáveis não foram utilizadas recomendou-se o emprego das mesmas em trabalhos futuros. Por meio do método de Rede Neural após o treinado poder ser usado na produção de cenários e predição, ele pode ser alimentado diariamente com dados meteorológicos para conhecer os riscos diários.

IV. CONCLUSÃO

A escolha do método depende do objetivo do trabalho, e do tempo disponível para se fazer essa modelagem. Além disso, deve-se considerar os conhecimentos dos profissionais que vão fazer essa modelagem. Após a elaboração dos mapas de risco pode-se aplicar estes nas políticas públicas de prevenção e combate de incêndios.

Vale ressaltar finalmente que faltam em muitos casos ou metodologias aplicadas a devida validação para a consagração do método proposto.

V. REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, L.F. VIERA, W.L.M. Redes neurais artificiais aplicadas ao treinamento e auxílio à tomada de decisão no processo de fabricação de vidros planos. *Revista Sodebras* [on line]. v. 10, n.114, Jun./2015, p. 49-53. ISSN 1809-3957. Disponível em: <<http://www.sodebras.com.br/edicoes/N114.pdf>>. Acesso em 11 out. 2017.
- ALONSO-BETANZOS, Amparo et al. An intelligent system for forest fire risk prediction and fire fighting management in Galicia. *Expert systems with applications*, v. 25, n. 4, p. 545-554, 2003.
- AYANZ, J.S.M et al. The European Forest Fire Information System (EFFIS). European Commission. **JIR Institute for Environment and sustainability**, 2003.
- BERGONSE, Rafaelo; BIDARRA, João. Probabilidade bayesiana e regressão logística na avaliação da susceptibilidade à ocorrência de incêndios de grande magnitude. *Finisterra*, v. 45, n. 89, 2010.
- BONAZOUNTAS, Marc et al. A decision support system for managing forest fire casualties. **Journal of Environmental Management**, v. 84, n. 4, p. 412-418, 2007.
- CANNADY, J. Artificial Neural Networks for Misuse Detection. **National Information System Security Conference**. School of Computer and Information Sciences Nova Southeastern University, Fort Lauderdale, Out-1998.
- CAVALCANTE et al. Redes neurais artificiais como ferramenta de predição dos ajustes no consumo dos produtos químicos da ETA de Palmeira dos Índios. *Revista Sodebras* [on line]. v. 10, n.116, p. 183-187, Ago./2015. ISSN 1809-3957. Disponível em: <<http://www.sodebras.com.br/edicoes/N116.pdf>>. Acesso em 10 out. 2017.
- CHUVIECO, Emilio et al. Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. **Ecological Modelling**, v. 221, n. 1, p. 46-58, 2010.
- CHUVIECO, Emilio; CONGALTON, Russell G. Application of remote sensing and geographic information systems to forest fire hazard mapping. **Remote sensing of Environment**, v. 29, n. 2, p. 147-159, 1989.
- DEL HOYO, Lara Vilar; ISABEL, M. Pilar Martín; VEGA, F. Javier Martínez. Logistic regression models for human-caused wildfire risk estimation: analysing the effect of the spatial accuracy in fire occurrence data. **European Journal of Forest Research**, v. 130, n. 6, p. 983-996, 2011.
- FERREIRA, Matheus Pinheiro; KOPROSKI, Leticia; ZANOTTA, Daniel Capella. Uma abordagem fuzzy no zoneamento de risco de incêndio. **Anais XV Simpósio Brasileiro de Sensoriamento Remoto-SBSR**, Curitiba, PR, Brasil, v. 30, p. 4555, maio/2011.
- GOLDARAG, Y. Jafari; MOHAMMADZADEH, Ali; ARDAKANI, A. S. Fire Risk Assessment Using Neural Network and Logistic Regression. **Journal of the Indian Society of Remote Sensing**, v. 44, n. 6, p. 885- 894, 2016.

- GUO, Futao et al. Wildfire ignition in the forests of southeast China: Identifying drivers and spatial distribution to predict wildfire likelihood. **Applied Geography**, v. 66, p. 12-21, 2016.
- HAYKIN, S. **Neural Networks - A comprehensive Foundation**. New Jersey, USA: Ed. Prentice Hall, 1998. ISBN:013273301
- HERNANDEZ-LEAL et al, A. Fire risk assessment using satellite data. **Advances in Space research**, v. 37, n. 4, p. 741-746, 2006.
- JAISWAL, Rajeev Kumar et al. Forest fire risk zone mapping from satellite imagery and GIS. **International Journal of Applied Earth Observation and Geoinformation**, v. 4, n. 1, p. 1-10, 2002.
- NASCIMENTO, Janaína Corrêa do. AMARAL, Renato Silva do. PONTES, Altem Nascimento. Determinação do grau de perigo de incêndios por meio da fórmula de monte alegre em municípios do estado do Pará, Brasil. **Revista Sodebras** [on line]. v. 8, n.94, p. 24-28, out./2013. ISSN 1809-3957. Disponível em: <<http://www.sodebras.com.br/edicoes/N94.pdf>>. Acesso em 11 out. 2017.
- NUNES, José Renato Soares; SOARES, Ronaldo Viana; BATISTA, Antônio Carlos. FMA+-um novo índice de perigo de incêndios florestais para o Estado do Paraná, Brasil. **Floresta**, v. 36, n. 1, 2006.
- OLIVEIRA, André; NERO, Marcelo. Application of Fuzzy Logic in Prediction of Fire in João Pessoa City- Brazil. In: **Geo-Informatics in Resource Management and Sustainable Ecosystem**. Springer, Berlin, Heidelberg, p. 323-334, 2013.
- RIBEIRO, L.et al. Zoneamento de Riscos de Incêndios Florestais para a Fazenda Experimental do Canguiri, Pinhais (PR). **Floresta**, Curitiba, PR, v. 38, n. 3, jul./set de 2008.
- SAATY, T. L. A scaling method for priorities in hierarchical structures. **Journal of Mathematical Psychology**, v.15, n. 3, p. 234-281, 1977.
- SARAGIOTO, S.R. P. PEREIRA, W.N. A. Lógica “fuzzy” aplicada ao controlador de velocidade de uma linha de montagem de eixos de veículos. **Revista Sodebras** [on line]. v. 7, n.83, p. 29-35, nov./2012. ISSN 1809-3957. Disponível em: <<http://www.sodebras.com.br/edicoes/N83.pdf>>. Acesso em 11 out. 2017.
- SILVA, Isaac da. DB; PONTES JR, Antonio CF. Elaboração de um fator de risco de incêndios florestais utilizando lógica fuzzy. **Biomatemática**, v. 21, p. 113-128, 2011.
- SOARES, Ronaldo Viana. Desempenho da "fórmula de monte alegre" índice brasileiro de perigo de incêndios florestais. **Cerne**, V.4, N.1, p.087-099, Lavras-MG, 1998.
- STEINER, I. PORTELLA, J.A.Z. PACHECO, M.A Lógica fuzzy e séries temporais integradas à tomada de decisão de investimentos. **Revista Sodebras** [on line]. v. 8, n.91, p. 50-57, Jul./2013. ISSN 1809-3957. Disponível em: <<http://www.sodebras.com.br/edicoes/N91.pdf>>. Acesso em 11 out. 2017.
- ZADEH, Lotfi A. Fuzzy sets. **Information and control**, v. 8, n. 3, p. 338-353, 1965.

VI. COPYRIGHT

Direitos autorais: Os autores são os únicos responsáveis pelo material incluído no artigo.

APÊNDICE F

Artigo:

FIRE RISK MODELING USING ARTIFICIAL NEURAL NETWORKS



Fire Risk Modeling Using Artificial Neural Networks

Luiza Cintra Fernandes^{1(✉)}, Rosangela S. C. Cintra²,
Marcelo Antonio Nero¹, and Plínio da Costa Temba¹

¹ Federal University of Minas Gerais (UFMG),
Belo Horizonte, MG 31270-901, Brazil
luizacintra@gmail.com, marcelo-nero@ufmg.br,
temba.mobile@gmail.com

² Nacional Institute for Space Research (INPE),
S.J. Campos, SP 1227-010, Brazil
rocintra@gmail.com

Abstract. Forest fires cause many changes in environment and in climate, becoming a huge concern related with environment, as your prevention and control. The fire risk calculation supports the planning of activities to prevent forest fire, as it determines the probability of fire occurrence in certain place. This article has the aim of mapping fire risk areas of Belo Horizonte, one of the most populous cities from Brazil, located in the Minas Gerais State, in the Southeast Region of Brazil. The proposed modeling is to create an artificial neural network with supervised training. A neural network to do the prediction of most propitious fire areas is expected, where it can be introduced the input variables at any period that desire to be determined. This estimate will provide the outline of priority areas for prevention activities and allocation of brigade teams, seeking to minimize possible damages caused by fires.

Keywords: Fire risk · Artificial Neural Networks · Inverse problems

1 Introduction

Forest Fires are one of the major concerns related with the environmental, since they devastated great portions of forests, fields and agricultural regions. Among your main impacts are: the fauna and flora destruction, erosion, soil degradation, trees weakness, habitats destruction, human health damages and landscape beauty reduction [1]. In addition of causing the increase of gases in the atmosphere, such as: CO₂ (dioxide of carbon), CH₄ (methane gas) and NO_x (nitrous gases), that are responsible for increasing the greenhouse effect and causing respiratory problems [2].

The fire risk calculation consists in verify the probability of fire occurring in determined place, that is, the chance of ignition [3]. So it is possible to do a critical evaluation of fire occurrence, and do the correct monitoring and prevention of burned areas, using objective tools [4].

The importance of determining areas with higher risk of fires is related with activities planning for your prevention. So, it is possible to know the seasons that most

occur the forest fires, allowing controlled burns, combat and prevention of it, notify the population about the risks and do environmental education programs and, in extreme cases, limited the access in this areas [5].

The variables used in the fire risk calculation can be of two types: some derived from factors that do not change in a short period of time, called statics, and those who vary in shorter time period, called dynamics. The first ones include the slope, types of vegetation, types of soil, road and settlement proximity. As the dynamic variables, can be cited the meteorological ones, such as, temperature, wind intensity and direction, humidity, among others [6].

There is a lot of approaches used in the fire risk modelling, statistics methods such as logistic regression, quantitative and qualitative methods, based on expert knowledge, such as multicriteria analysis. Besides artificial intelligence methodologies, for example Artificial Neural Networks (ANN) and Fuzzy Logic, that has been used for this and returning accurate results [7–12].

The Southeast Region of Brazil presents a lot of fire spots in the winter, when the weather is more dry, which coincide with the registration of the air particle material concentrations increase, as observed in the air monitoring stations of the State Foundation of Environmental (in portuguese: Fundação Estadual do Meio Ambiente - FEAM) of Minas Gerais State.

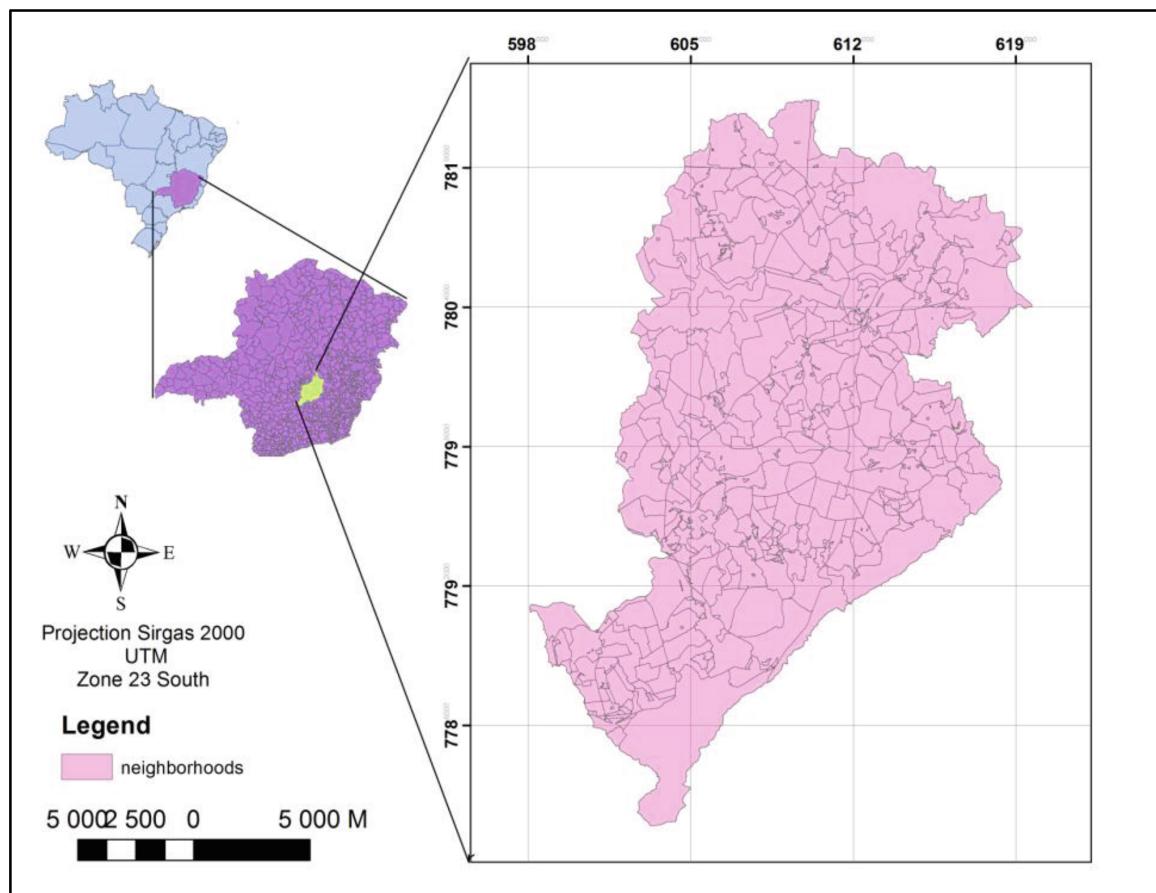


Fig. 1. Study area location in Brazil

Then, the aim of this article is to mapping fire risk areas of Belo Horizonte, city located in the Minas Gerais State, in the Southeast Region of Brazil. The proposed modeling is to create an artificial neural network with supervised training. A neural network to do the prediction of most propitious fire areas is expected, where it can be introduced the input variables at any period that desire to be determined. This estimate will provide the outline of priority areas for prevention activities and allocation of brigade teams, seeking to minimize possible damages caused by fires.

2 Study Area

With 330.9 km² of area, the Belo Horizonte city (Fig. 1) has a population of 2.4 million in the 2010 census and a density of 7167 hab/km². Has a GDP per capita of approximately 35000 brazilian real, and a HDI of 0.81 [13]. It is inserted in the third biggest metropolitan region of Brazil, named [14] as a “Ecological Tension Area”. It is a transition area between two important Brazilian biomes: the “Cerrado” and the “Mata Atlântica”, both of them considerate biodiversity “hotspots” that demanding concrete measures of protecting. It is a very afforest city, 82.7% of urban households in public roads have afforestation. The original vegetation is basically Brazilian savanna and sub-tropical forest, presents in the municipal parks around the city.

The weather in the region is defined by subtropical semi-humid [15] with a defined dry season with duration of 4 to 5 months, between April and September. Approximately 80% of annual precipitation occurs in the rainy season. [16]. The region presents seasons of low humidity, especially in the winter, due the presence of the anticyclones that brings dry air mass and decrease the temperature and humidity. In a study [17] have observed that the relative air humidity is decreasing in the city. This low humidity became the region propitious to forest fires.

3 Materials and Methods

3.1 Input Variables

Slope. It affects the wind behavior and direction, the fires propagates faster in areas with higher slope, affecting the rate and direction of the fire [18]. The slope was calculated from the function terrain, from raster R package, were the input elevation data used was from SRTM (Shuttle Radar Topography Mission), available from the USGS website [19], the input data have 30 m of resolution.

Aspect. The aspect describes the terrain orientation, that is important because it get to know the faces with direct influence of sun direct heat, which affects the dry rate of the combustible material [20]. The calculation of aspect was also done with the terrain function from the SRTM.

Land Cover and Use. It comprehends the characterization of land cover and your locations. So it is possible to know the distribution and quantifying the vegetation areas, the fields, agriculture, urban area, among others [21]. The vegetation affects the

kind and properties of the combustible material and the weather conditions associated. Also this variable is important to compute the human influence from the roads and settlement proximities influence.

The map used it was produced by Prodabel (Informatics and Information Company of the city of Belo Horizonte) in 2012. The Land Use classification was extract with reference in orthorectified images, from visual interpretations and manual practices for thematic classes. The classes from the original image, that had 3 levels of details, were grouped in 29 classes, showed in Table 1. The map was transformed in raster with a cell size of 100 m, in ArcGIS software.

Table 1. Classes of Land Use map

Class	Value
Water	1
bare soil	2
Rock outcrop	3
Buildings high density	4
Buildings low density	5
Buildings medium density	6
Buildings very high density	7
Buildings very low density	8
Roads	9
Railroad	10
Subway	11
Urban road	12
Central public roadway	13
Parks	14
Fields	15
Plaza	16
Tree-shrub vegetation	17
Undergrowth	18
Mining	19
Leisure area	20
Big equipment	21
Landfill	22
Industry	23
Allotment high ordering	24
Allotment low ordering	25
Cultivation areas	26
Fields	27
Woods	28
Reforest	29

Meteorological Variables. The meteorological variables were collected from the INMET (National Institute of Meteorology of Brazil) from automatic and manual stations. Monthly maps, from 2014 to 2016, of the parameters were interpolate using Thin Plate Spline function, from `rgdal` package in R.

Average Temperature. The temperature represents the internal energy of movement in atoms and molecules, this heat energy is important in the evaporative phase of combustion [22]. Higher temperatures are more susceptible to fires, predisposes them to ignition, in addition reduce the humidity of the material [23]. The data used is in Celsius degrees units.

Air Pressure. The climate conditions changes dependent of the air pressure. Higher pressure, have higher temperatures and lower humidity, leaving the weather susceptible to fires [23]. The data used is in hPa.

Air Relative Humidity. Represents the relation between the amount of water in the air and the maximum amount that could have in the same temperature and pressure. [22]. As higher the relative humidity, higher the combustible humidity absorption [23]. The data used is in percentage.

Wind Speed. It is the movement of air at a horizontal distance per unit of time. It can accelerate the water evaporation in the combustible material, becoming it dry and flammable. The wind also increases the oxygen supply, increasing the combustion conditions [23]. The unit in the data used is in m/s.

Radiation. Represents the great amount of energy from the sun, whose energy is received by Earth in the form of electromagnetic waves. This type of radiation is the primary source of energy that the globe has, and its variable distribution in time and space is the generator of all atmospheric processes [23]. The unit in the data used is in Kj/M^2 .

3.2 Output Variable: Fire Data

Active fire products have information about the location and timing of fires that are burning at the time of the satellite overpass, usually in form of lists of fire pixel locations and dates. These observations are then summarized at coarse spatial resolutions over daily time periods.

In this article were used two active fire products. The first is the Near real-time (NRT) Moderate Resolution Imaging Spectroradiometer (MODIS) Thermal Anomalies/Fire locations- Collection 6, that use a swath products (MOD14/MYD14), with resolution of 1 km. The second is de Near real-time (NRT) Suomi National Polar-orbiting Partnership (Suomi-NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) Active Fire detection product, that has a resolution of 375 m.

The fire points were spatialized according to the coordinates, and then convert to raster, with 1 km of resolution, where the value 1 represents fire occurrence, and the value 0, represents non-occurrence.

3.3 Data Arrangement

All the data were set in the same projection, Sirgas 2000, in the same extent and resolution, using R `rgdal` package. Resulted in a regular grid of 100×100 m resolution, with 313 rows and 217 columns.

The months of August, September and October were select because of the fire frequency. For each pixel of the fire map were related each of 8 parameters, generating the samples to input in the ANN. As the non-occurrence samples (background pixels) were much larger than the occurrence samples, the area were partitioned in 1 km 1 km squares and background samples were selected randomly.

The data from 2014 and 2015 were used to fit the ANN, and the samples were divided in: train (60%), validate (20%) and test (20%).

The data from 2016 was used to evaluate the ANN, being used the pixels of the whole area.

3.4 The Artificial Neural Network

An Artificial Neural Network (ANN) consists in simple process units that works in parallel, storing experimental knowledge and becoming available to use in modeling of acquired knowledge. This technique seeks to reproduce the human brain behavior in two aspects: the acquired knowledge for the network, done from a learning process and the connection strength between neurons, known as synaptic weight, which stores the knowledge (Haykin, 2001). Between the ANN benefits are the possibility of non-linearity, the fact of mapping inputs and outputs, from supervised training, that also allows the ANN to be adaptable, modifying your synaptic weights according to changes in the environment [24].

There are many architectures for ANN, from that the Multilayer feed-forward Perceptron, is one of the most used, as can be seen in the figure (Fig. 2), that has two hidden layers, where each input (x) is multiplied by an appropriate weight (w) and the sum of these weighted (Eq. 1) inputs and the bias (b) results in the input transfer function $f(x)$ (Eq. 2), becomes the input of following layer. The input layer is the source node, that provide the initial signal and the output layer is the response for the input [24].

$$u_i = \sum_{j=1}^m w_{ij}x_j \quad (1)$$

$$y_i = f(u_i + b_i) \quad (2)$$

The transfer function is important because restricts the amplitude of the signal at the output of the neuron as a function of the arguments and can define the linearity or not of the network output. The used in this article are the sigmoid-logistic (Eq. 4) and the linear (Eq. 3).

$$f(v) = nv \quad (3)$$

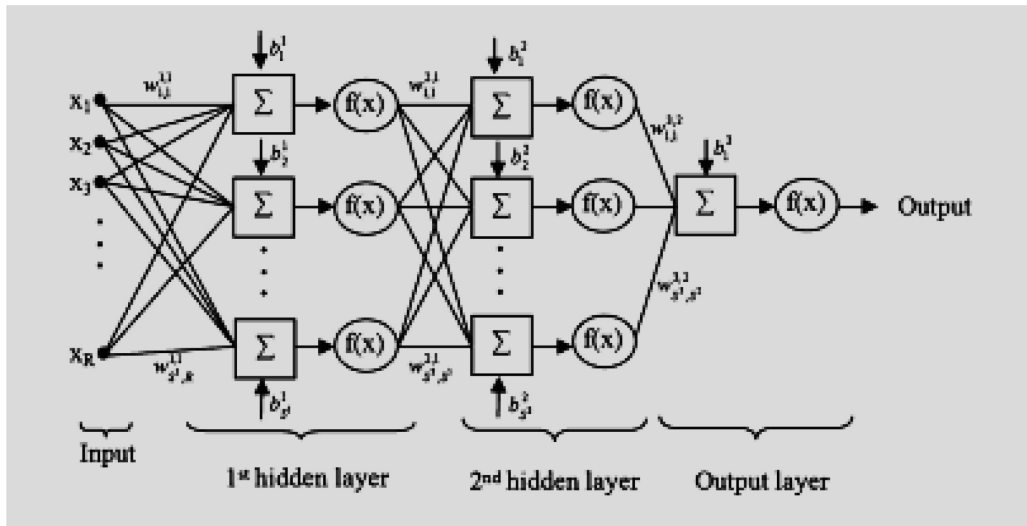


Fig. 2. Multilayer feed-forward perceptron [10]

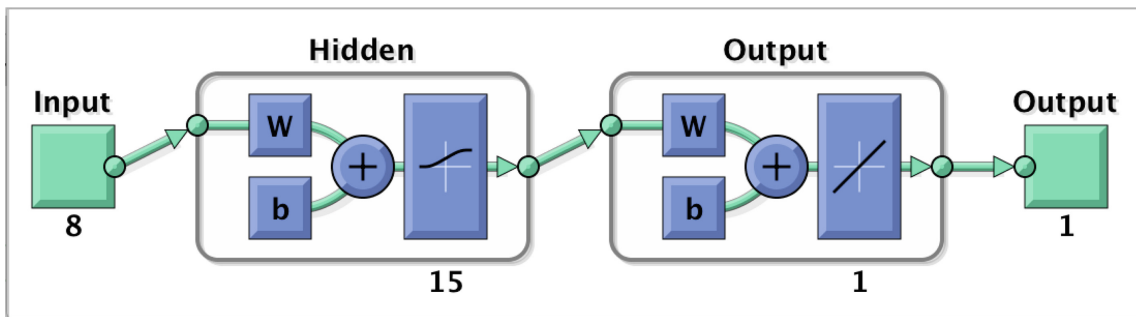


Fig. 3. ANN results diagram.

$$f(v) = \frac{1}{1 + \exp(-\beta v)} \tag{4}$$

The training is very important to the network learning and improve your performance. It is used an interactive process to adjust weights and bias. Generally the data are divided in 3 subsets. The first one, training subset, consists in the knowing patterns to adjust the network parameters, the validation has the aim of verify the capacity of generalization of the network and the third, the test, tests the behavior of the ANN with new data [25]. The supervised learning is the training where is given a desired output and the parameters are adjusted based on that [26].

To training the multilayer perceptron the most used algorithm is the backpropagation, which is based in the Delta rule. This rule has the aim of minimizing the cost function of error (Eq. 5), where a delta (Eq. 6) based in the error computed is used to update the weights.

$$\varepsilon(\mathbf{n}) = \frac{1}{2} e_k^2(\mathbf{n}) \tag{5}$$

$$\Delta w_{kj} = \eta e_k(\mathbf{n}) x_j(\mathbf{n}) \tag{6}$$

The backpropagation algorithm consists of two steps: the step forward and the step backward, where the error is backpropagated from the output layer to the input layer. And the delta is given by the gradient from the transfer function.

$$\Delta w_{jk}(\mathbf{n}) = -\eta \frac{\partial \varepsilon(\mathbf{n})}{\partial w_{jk}(\mathbf{n})} \tag{7}$$

The Levenberg-Marquardt algorithm, the one used in this article, despite backpropagation, that has a steepest descent algorithm, is an approximation to Newton’s method, so the delta is given as Eq. 8. This algorithm is used to speed up convergence of the backpropagation learning algorithm.

$$\Delta w = (J^T J + \mu I)^{-1} J^T e \tag{8}$$

Where w is the weight vector, I is the identity matrix, μ is the combination coefficient, J the Jacobian matrix (Eq. 9), and e , the error vector.

$$J = \begin{bmatrix} \frac{\partial e_{11}}{\partial w_1} & \frac{\partial e_{11}}{\partial w_N} \\ \vdots & \vdots \\ \frac{\partial e_{1M}}{\partial w_1} & \frac{\partial e_{1M}}{\partial w_N} \\ \vdots & \vdots \\ \frac{\partial e_{P1}}{\partial w_1} & \frac{\partial e_{P1}}{\partial w_N} \\ \vdots & \vdots \\ \frac{\partial e_{PM}}{\partial w_1} & \frac{\partial e_{PM}}{\partial w_N} \end{bmatrix} \cdot \dots \cdot \begin{bmatrix} e_{11} \\ \vdots \\ e_{1M} \\ \vdots \\ e_{P1} \\ \dots \\ e_{PM} \end{bmatrix} \tag{9}$$

Where P is the number of training patterns, M is the number of outputs, and N is the number of weights. Elements in error vector are calculated from Eq. 10.

$$e_{PM} = d_{PM} - o_{PM} \tag{10}$$

Where d is the desired output and o the observed output. These algorithms were implemented in neural network toolbox of matlab.

4 Results and Discussion

After testing a lot of architectures with different number of neurons in the hidden layer and the one that have the least error was with 15 neurons. So the final ANN had 8 neurons in the input layer, 15 neurons in the hidden layer and 1 output. The transfer

function used was the logistic-sigmoid in the first layer and the linear function in the second layer, and it was trained with Levenberg-Marquardt algorithm.

From the resulted output the values next to 1 were considered high risk, and the next to 0 low risk. In table it can be seen the classification of pixels from the data used to fit the ANN. What can be seen is that the majority were classified as high risk. But the fire occurrence points it was in the area of most risk, which is a good result (Table 2).

Table 2. Results from ANN fitting data

Percentage	Risk	Pixels	Fire occurrence pixels
0–0.25	Low	3.84%	0.05%
0.25–0.5	Medium	2.83%	0.36%
0.5–0.75	High	9.01%	6.95%
0.75–1	Very high	84.33%	92.64%

In 2016 data, that wasn't used in ANN fitting, it can be seen that almost all the pixels were classified as high and very high. And all the fire occurrences were in the Very high classification (Table 3).

Table 3. Results for model validation

Date	Percentage	Risk	Pixels	Fire occurrence pixels
August/2016	0–0.25	Low	0.0%	0.0%
	0.25–0.5	Medium	0.9%	0.0%
	0.5–0.75	High	4.3%	0.0%
	0.75–1	Very high	94.8%	100.0%
September/2016	0–0.25	Low	0.0%	0.0%
	0.25–0.5	Medium	0.0%	0.0%
	0.5–0.75	High	3.8%	0.0%
	0.75–1	Very high	96.2%	100.0%
October/2016	0–0.25	Low	0.0%	0.0%
	0.25–0.5	Medium	6.1%	0.0%
	0.5–0.75	High	8.1%	0.0%
	0.75–1	Very high	85.8%	100.0%

So the network could recognize the fire points very well, but maybe is whereas almost all the area with high risk. This can be due the small variation of meteorological parameters in the whole area, as they are the majority of the inputs (5 inputs). As were taken just months with lower precipitation, when is expected higher fire risks. Also the area is urban and in the literature is seen that the proximity with urban areas has higher risk because of the human influence.

[10] did a similar approach with a neural network with 12 variables (slope, aspect, land cover, NDVI, air temperature, day with precipitation, duration of Sunshine, prevailing wind percent, relative humidity, height, road and city distance), being 6 dynamic and 6 static in an area of 268 km², predicting 95% of the fire points. [11] uses in their network only meteorological parameters in an area of 360 km², and had a accuracy of 80%. [12]) used as inputs NDVI images, that are very sensitive to vegetation changes, and had a Mean Square Error of 0.07.

5 Conclusion

Using Artificial Neural Networks to predict fire risk is a methodology that results in good responses, but is necessary to be careful with the chosen of input variables. In this study the network could find very well the fire points, but didn't recognize with great accuracy points with low risk.

In future studies could be added more static variables, as NDVI images and distance to roads and settlement. Also would be interesting study a longer period, to see seasonal changes in meteorological variables.

References

1. Roy, P.S.: Forest fire and degradation assessment using satellite remote sensing and geographic information system. In: Sivakumar, M.V.K., Roy, P.S., Harmsen, K., Saha, S.K. (eds.) *Satellite Remote Sensing and GIS Applications in Agricultural Meteorology*, pp. 361–400. World Meteorological Organisation, Geneva (2004)
2. Ichoku, C., Kaufman, Y.J.: A method to derive smoke emission rates from MODIS fire radiative energy measurements. *IEEE Trans. Geosci. Remote Sens.* **43**, 2636–2649 (2005). <https://doi.org/10.1109/TGRS.2005.857328>
3. Hardy, C.C.: Wildland fire hazard and risk: problems, definitions, and context. *For. Ecol. Manage.* **211**, 73–82 (2005). <https://doi.org/10.1016/J.FORECO.2005.01.029>
4. Chuvieco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M.P., Vilar, L., Martínez, J., Martín, S., Ibarra, P., de la Riva, J., Baeza, J., Rodríguez, F., Molina, J.R., Herrera, M.A., Zamora, R.: Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. *Ecol. Modell.* **221**, 46–58 (2010). <https://doi.org/10.1016/j.ecolmodel.2008.11.017>
5. Nunes, R.S.: FMA + - um novo índice de perigo de incêndios florestais para o estado do Paraná - Brasil. *Floresta* **36**, 75–91 (2006)
6. López, A.S., San-Miguel-Ayanz, J., Burgan, R.E.: Integration of satellite sensor data, fuel type maps and meteorological observations for evaluation of forest fire risk at the pan-European scale (2002)
7. Ferreira, M.P., Koproski, L., Zanotta, D.C.: Uma abordagem fuzzy no zoneamento de risco de incêndio. In: XV Simpósio Brasileiro de Sensoriamento Remoto - SBSR, pp. 4555–4562. Curitiba, PR (2011)
8. Luiz de Sá de Oliveira, A.: Modelagem espacial de predição de riscos de incêndios com lógica fuzzy, comparação e validação (2013)
9. da Silva, I.D.B., Pontes, A.C.F.J.: Elaboração de um Fator de Risco de Incêndios Florestais utilizando Lógica Fuzzy. **21**, 113–128 (2011)

10. Goldarag, Y.J., Mohammadzadeh, A., Ardakani, A.S.: Fire risk assessment using neural network and logistic regression. *J. Indian Soc. Remote Sens.* **44**, 885–894 (2016). <https://doi.org/10.1007/s12524-016-0557-6>
11. Alonso-Betanzos, A., Fontenla-Romero, O., Guijarro-Berdiñas, B., Hernández-Pereira, E., Paz Andrade, M.I., Jiménez, E., Soto, J.L.L., Carballas, T.: An intelligent system for forest fire risk prediction and fire fighting management in Galicia. *Expert Syst. Appl.* **25**, 545–554 (2003). [https://doi.org/10.1016/S0957-4174\(03\)00095-2](https://doi.org/10.1016/S0957-4174(03)00095-2)
12. Maeda, E.E., Formaggio, A.R., Shimabukuro, Y.E., Arcoverde, G.F.B., Hansen, M.C.: Predicting forest fire in the Brazilian Amazon using MODIS imagery and artificial neural networks. *Int. J. Appl. Earth Obs. Geoinf.* **11**, 265–272 (2009). <https://doi.org/10.1016/j.jag.2009.03.003>
13. IBGE, (Instituto Brasileiro de Geografia e Estatística): Panorama Belo Horizonte. <https://cidades.ibge.gov.br/brasil/mg/belo-horizonte/panorama>
14. Veloso, H.P., Rangel Filho, A.L.R., Lima, J.C.: Classificação da vegetação brasileira, adaptada a um sistema universal (1991)
15. Nimer, E.: Climatologia do Brasil (1979)
16. Lucas, T.P.B., Abreu, M.L.: Caracterização climática dos padrões de ventos associados a eventos extremos de precipitação em Belo Horizonte - MG. *Cad. Geogr.* **14**, 135–152 (2004)
17. da Franca, R.R.: Anticiclones e umidade relativa do ar: um estudo sobre o clima de Belo Horizonte (2009)
18. Vadrevu, K.P., Eaturu, A., Badarinath, K.V.S.: Fire risk evaluation using multicriteria analysis—a case study. *Environ. Monit. Assess.* **166**, 223–239 (2010). <https://doi.org/10.1007/s10661-009-0997-3>
19. USGS (U.S. Department of the Interior U.S. Geological Survey): EarthExplorer
20. Juvanhol, R.S.: Modelagem da vulnerabilidade à ocorrência e propagação de incêndios florestais (2014)
21. Prudente, T.D.: Geotecnologias Aplicadas ao mapeamento de risco de incêndio florestal no parque nacional da chapada dos veadeiros e área de entorno (2010)
22. Hamadeh, N., Karouni, A., Daya, B., Chauvet, P.: Using correlative data analysis to develop weather index that estimates the risk of forest fires in Lebanon & Mediterranean: assessment versus prevalent meteorological indices. *Case Stud. Fire Saf.* **7**, 8–22 (2017). <https://doi.org/10.1016/j.csfs.2016.12.001>
23. Liu, D., Zhang, Y.: Research of regional forest fire prediction method based on multivariate linear regression. *Int. J. Smart Home.* **9**, 13–22 (2015). <https://doi.org/10.14257/ijsh.2015.9.1.02>
24. Haykin, S.: *Redes Neurais: Princípios e prática*. Bookman, Porto Alegre (2001)
25. Anochi, J.A., de Campos Velho, H.F.: Optimization of feedforward neural network by multiple particle collision algorithm. In: 2014 IEEE Symposium on Foundations of Computational Intelligence (FOCI), pp. 128–134. IEEE, Orlando (2014)
26. Goulart, D.A., Tacla, M.A., Marback, P.M.F., Solé, D., Paranhos, A., Perez, H.B., de Freitas, D., Sato, E.H.: Redes neurais artificiais aplicadas no estudo de questionário de varredura para conjuntivite alérgica em escolares. *Arq. Bras. Oftalmol.* **69**, 707–713 (2006). <https://doi.org/10.1590/S0004-27492006000500017>