UNIVERSIDADE FEDERAL DE MATO GROSSO INSTITUTO DE FÍSICA PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA AMBIENTAL

MODELAGEM E ANÁLISE DE CENÁRIOS ATRAVÉS DA CLASSIFICAÇÃO DO ÍNDICE DE VARIÂNCIA DO CAMPO TERMAL URBANO EM ILHAS DE CALOR URBANO EM CUIABÁ UTILIZANDO FLORESTA RANDÔMICA

ALBERTO SALES E SILVA

PROF. DR. RAPHAEL SOUZA ROSA GOMES Orientador

> Cuiabá, MT Outubro/2023

UNIVERSIDADE FEDERAL DE MATO GROSSO INSTITUTO DE FÍSICA PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA AMBIENTAL

MODELAGEM E ANÁLISE DE CENÁRIOS ATRAVÉS DA CLASSIFICAÇÃO DO ÍNDICE DE VARIÂNCIA DO CAMPO TERMAL URBANO EM ILHAS DE CALOR URBANO EM CUIABÁ UTILIZANDO FLORESTA RANDÔMICA

ALBERTO SALES E SILVA

Tese apresentada ao Programa de Pós-Graduação em Física Ambiental da Universidade Federal de Mato Grosso, como parte dos requisitos para obtenção do título de Doutorado em Física Ambiental.

PROF. DR. RAPHAEL SOUZA ROSA GOMES Orientador

Cuiabá, MT Outubro/2023

Dados Internacionais de Catalogação na Fonte.

S586m Silva, Alberto Sales e.

Modelagem e análise de cenários através da classificação do índice de variância do campo termal urbano em ilhas de calor urbano em Cuiabá utilizando floresta randômica [recurso eletrônico] / Alberto Sales e Silva. -- Dados eletrônicos (1 arquivo : 143 f., pdf). -- 2024.

Orientador: Raphael Souza Rosa Gomes.

Tese (doutorado) - Universidade Federal de Mato Grosso, Instituto de Física, Programa de Pós-Graduação em Física Ambiental, Cuiabá, 2024.

Modo de acesso: World Wide Web: https://ri.ufmt.br. Inclui bibliografia.

1. fenômeno UHI. 2. ilhas de calor urbano. 3. UTFVI. 4. floresta aleatória. 5. modelagem espaço-temporal. I. Gomes, Raphael Souza Rosa, *orientador*. II. Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

Permitida a reprodução parcial ou total, desde que citada a fonte.



MINISTÉRIO DA EDUCAÇÃO UNIVERSIDADE FEDERAL DE MATO GROSSO PRÓ-REITORIA DE ENSINO DE PÓS-GRADUAÇÃO PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA AMBIENTAL

FOLHA DE APROVAÇÃO

TÍTULO: MODELAGEM E ANÁLISE DE CENÁRIOS ATRAVÉS DA CLASSIFICAÇÃO DO ÍNDICE DE VARIÂNCIA DO CAMPO TERMAL URBANO EM ILHAS DE CALOR URBANO EM CUIABÁ UTILIZANDO FLORESTA RANDÔMICA

AUTOR: DOUTORANDO ALBERTO SALES E SILVA

Tese defendida e aprovada em 08 de março de 2024.

COMPOSIÇÃO DA BANCA EXAMINADORA

- 1. **Prof. Dr. RAPHAEL DE SOUZA ROSA GOMES** (Presidente Banca / **ORIENTADOR**) INSTITUIÇÃO: Universidade Federal de Mato Grosso
- 2. **Profa. Dra. FLÁVIA MARIA DE MOURA SANTOS** (Membro Interno) INSTITUIÇÃO: Universidade Federal de Mato Grosso
- 3. **Prof. Dr. SÉRGIO ROBERTO DE PAULO** (Membro Interno) INSTITUIÇÃO: Universidade Federal de Mato Grosso
- 4. **Prof. Dr. JOSIEL MAIMONE DE FIGUEIREDO** (Membro Externo) INSTITUIÇÃO: Universidade Federal de Mato Grosso
- 5. **Profa. Dra. ANGELA SANTANA DE OLIVEIRA** (Membro Externo) INSTITUIÇÃO: Instituto Federal de Mato Grosso

Cuiabá, 08/03/2024.



Documento assinado eletronicamente por RAPHAEL DE SOUZA ROSA GOMES, Docente da Universidade Federal de Mato Grosso, em 08/03/2024, às 17:37, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de



Documento assinado eletronicamente por JOSIEL MAIMONE DE FIGUEIREDO, Docente da Universidade Federal de Mato Grosso, em 08/03/2024, às 17:40, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.



Documento assinado eletronicamente por FLAVIA MARIA DE MOURA SANTOS, Docente da Universidade Federal de Mato Grosso, em 08/03/2024, às 17:44, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.



Documento assinado eletronicamente por **SERGIO ROBERTO DE PAULO**, **Docente da Universidade Federal de Mato Grosso**, em 10/03/2024, às 09:51, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do <u>Decreto nº 10.543</u>, de 13 de novembro de 2020.



Documento assinado eletronicamente por **ANGELA SANTANA DE OLIVEIRA**, **Usuário Externo**, em 11/03/2024, às 12:07, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.



Documento assinado eletronicamente por MARCELO SACARDI BIUDES, Coordenador(a) de Programas de Pós-Graduação em Física Ambiental - IF/UFMT, em 11/03/2024, às 12:15, conforme horário oficial de Brasília, com fundamento no § 3° do art. 4° do Decreto nº 10.543, de 13 de novembro de 2020.



A autenticidade deste documento pode ser conferida no site http://sei.ufmt.br/sei/controlador_externo.php?
acesso_externo=0, informando o código verificador 6661889 e o código CRC 2159C146.

Referência: Processo nº 23108.014387/2024-64 SEI nº 6661889

DEDICATÓRIA

Dedico a minha família que me deu todo suporte neste grande desafio. À minha esposa, Nádia Regina, pelo incentivo e aos meus filhos pela compreensão durante esta jornada.

AGRADECIMENTOS

Agradeço ao Prof. Orientador Dr. Raphael Souza pelos direcionamentos e disposição em contribuir para o desenvolvimento deste trabalho. Meus agradecimentos e homenagem ao prof. Dr. José de Souza Nogueira (in memoriam), por toda a sua dedicação ao programa de pós-graduação.

Aos professores deste programa pelo empenho em transmitir os conhecimentos e pela dedicação primorosa e aos colegas estudantes do programa que de alguma forma colaboraram com a construção do conhecimento nos momentos mais difíceis.

Um agradecimento especial para a minha família que esteve do meu lado e sempre dando apoio nesta grande busca do conhecimento. Agradeço-os imensamente por compreenderem o quão importante foi e continuará sendo esta jornada.

Sumário

LISTA DE FIGURAS	X
LISTA DE TABELAS	XIII
LISTA DE ABREVIATURAS	XV
LISTA DE SÍMBOLOS	.XVII
RESUMO	XVIII
ABSTRACT	XIX
1. INTRODUÇÃO	1
1.1. Problemática	1
1.2. Justificativa	
2. REVISÃO BIBLIOGRÁFICA	4
2.1. Clima Urbano	4
2.2. Ilhas De Calor	4
2.2.1. Os Fatores Que Afetam A Intensidade De Ilhas De Calor Urbanas	5
2.3. Monitoramento De Ilhas De Calor Com Landsat E Era5	
2.4. Análise Da Superfície Urbana	7
2.4.1. Classificação Da Cobertura Do Solo Utilizando Índices Espectrais	
2.4.1.1. Índice de Vegetação por Diferença Normalizada (NDVI)	
2.4.1.2. Temperatura da Superfície do Solo (LST)	
2.4.1.3. Índice por Diferença Normalizada de Áreas Construídas (NDBI)	
2.4.1.4. Umidade Relativa do Ar (UR)	
2.4.1.5. Índice de Calor Urbano (UHI)	
2.4.1.6. Índice de Variância do Campo Termal Urbano (UTFVI)	
2.4.1.7. Índice de Construção por Diferenças Normalizadas (NDBI)	
2.5. Árvores De Decisão	
2.6. Floresta Aleatória (Random Forest)	
3. MATERIAL E MÉTODOS	17
3.1. Área De Estudo - Cidade De Cuiabá-MT	
3.2. Dados Meteorológicos E Pré-processamento	
3.2.1. Coleção Das Imagens Obtidas Dos Satélites Landsat 7 E 8	
3.3. Definição Das Variáveis Da Amostra	
3.4. Definição Dos Quadrantes Na Rosa Dos Ventos	
3.5. Seleção Das Imagens De Satélites	
3.6. Engenheria Dos Dados.	
3.7. Procedimento Dos Cálculos Pertinentes Aos Índices Espectrais	
3.8. O Modelo Do Algoritmo Floresta Aleatória	
3.8.1. Parametrização Do Algoritmo Para Classificação Da Cobertura Do Sol	o A
Partir De Imagens De Satélite	
3.8.2. Parametrização Do Algoritmo Para Classificação Dos Dados Para O	
Fenômeno UHI	26
3.8.2.1. Métricas Entropia e Gini	
4. APRESENTAÇÃO E ANÁLISE DOS RESULTADOS	28
4.1. Pontos Geográficos.	
4.2. Caracterização Do Regime De Vento Em Cuiabá-MT	
4.2.1. Estação Seca E Quente	
4.2.1.1. Comportamento dos Ventos Locais Na Estação Seca e Quente	
4.2.2. Estação Quente E Úmida – Verão	
4.2.2.1. Comportamento dos Ventos Locais Na Estação Quente e Úmida	38
4.3. Caracterização Dos Níveis De Vegetação Em Cuiabá-MT	
1.3. Caracterização Dos raveis De vegetação Enii Caraba-ivi i	тЭ

4.5.1. Alianse Da Expansão Do Espaço Otoano Com O Oso Do NDVI, NDBI	
LST	
4.3.1.1. Classificação das coberturas do solo com o uso do algoritmo Flores	
Aleatória (Random Forest)	49 5 1
4.4. Caracterização Do Campo Termal Urbano De Cuiabá-MT	31
4.4.1. Distribuição Do Índice De Variância Do Campo Termal Por Pontos	<i>5</i> 1
Geográficos De Cuiabá-MT	34 551
Urbano (Fenômeno UHI)	
4.5.1. Análise Dos Impactos Das Variáveis Microclimáticas Sobre O Índice Do	37
Variância Do Campo Termal Urbano Na Superfície Do Solo	5 50
4.5.1.1. Indicadores de Robustez do Modelo Do Fenômeno UHI na superfície	
do Solodo	
4.5.1.2. Caracterizando a importância das variáveis Microclimáticas na flor	
aleatóriaaleatória	
4.5.2. Análise Dos Impactos Das Varíaveis Microclimáticas Sobre O Índice Do	0 <i>3</i>
Variância Do Campo Termal Urbano A 2 Metros De Altura	5 69
4.5.2.1. Indicadores de Robustez do Modelo Do Fenômeno UHI a 2 metros	
AlturaAltura	
4.5.2.2. Caracterizando A Importância Das Variáveis Microclimáticas Na	/ 0
Floresta Aleatória	74
4.5.3. Análise Sintética Dos Impactos Das Variáveis Microclimáticas Sobre O	/ ¬
Índice De Variância Do Campo Termal Urbano	76
4.5.4. Análise Dos Impactos Das Variáveis Microclimáticas Sobre O Índice Do	/ O
Variância Do Campo Termal Urbano Em Função Do NDVI	
4.5.4.1. Indicadores de Robustez do Modelo Para o Fenômeno UHI na	01
Superfície do Solo em função do NDVI e Demais Variáveis MicroClimátic	as 83
4.5.4.2. Caracterizando A Importância Das Variáveis Climáticas Na Florest	
<u> </u>	86
4.5.5. Análise Dos Impactos Das Variáveis Microclimáticas Sobre O Índice Do	
Variância Do Campo Termal Urbano A 2 Metros De Altura Em Função Do NI	
,	
4.5.5.1. Indicadores de Robustez do Modelo Para o Fenômeno UHI A 2 Me	etros
de Altura em função do NDVI e Demais Variáveis MicroClimáticas	89
4.5.5.2. Caracterizando A Importância Das Variáveis Microclimáticas Na	
Floresta Aleatória	91
4.6. Discussão	94
4.6.1. Expansão Da Ilha De Calor Em Cuiabá	94
4.6.2. Análise Estátistica Da Classificação Do Fenômeno UHI	95
4.6.3. Análise Da Correlação Entre UTFVI Na Superfície Do Solo E A 2 Metr	os99
4.6.4. Análise Dos Resultados Da Classificação Do Fenômeno UHI	103
4.6.4.1. UTFVI na superfície do solo	
4.6.4.2. UTFVI na superfície do solo E O NDVI	105
4.6.4.3. UTFVI a 2 metros de altura	
4.6.4.4. UTFVI a 2 metros de altura e o NDVI	
4.6.5. Distribuição Espacial Da Temperatura Da Superfície Do Solo E Do Ar	
5. CONCLUSÕES	
5.1. Contribuições	
5.2. Trabalhos Futuros	
6. REFERÊNCIAS BIBLIOGRÁFICAS	
7. APÊNDICES	125

7.1. Análise Dos Impactos Das Variáveis Microclimáticas Sobre O Índice De	
Variância Do Campo Termal Na Superficie Do Solo	125
7.2. Análise Dos Impactos Da Variáveis Microclimáticas Sobre O Índice De	
Variância Do Campo Termal Urbano A 2 Metros De Altura	127
7.3. Classificação Das Variáveis Microclimáticas Em Função Do Índice De Variâ	ìncia
Do Campo Termal Urbano (feNômeno UHI) Para Os Anos 2005, 2010, 2015 E 2	2020
	131
7.4. Classificação Das Variáveis Microclimáticas Em Função Do Índice De Variâ	
Do Campo Termal Urbano A 2 Metros DE ALTURA Para Os Anos 2005, 2010, 2	2015
E 2020	135

Lista de Figuras

Figura 1: Estrutura básica de uma árvore de decisão	14
Figura 2: Classificador de Floresta Aleatória.	15
Figura 3: Localização da cidade de Cuiabá-MT (Brasil). Fonte: Google Earth Engine	
(acessado em 22/11/23)/(PEREIRA; GONÇALVES, 2019)	
Figura 4: Climograma de Cuiabá-MT para os anos de 2000, 2005, 2010, 2015 e 2020	
com médias mensais das variáveis microclimáticas Temperatura do Ar e precipitação.	
Fonte: Autor.	
Figura 5: Fluxograma para estimar o impacto das variáveis (LST-Temperatura da	
Superfície do Solo). Baseado em (ULLAH et al., 2022)	21
Figura 6: Rosa dos ventos com os respectivos quadrantes (1 a 8)	
Figura 7: Entropia demonstrando o grau de pureza da informação (conjunto de dados	
1 Igura / . Emacpia demonstrata e gran de pareza da miormação (conjunto de dados	
Figura 8: Demonstração da estimação da importância das variáveis usando a impurez	, :a
Gini	2.7
Figura 9: Mapa de Cuiabá com a indicação numérica dos respectivos pontos de coleta	
em referência à Tabela 4. Fonte: Google Maps	
Figura 10: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m c	
as respectivas temperaturas do ar para a estação seca e quente do ano 2000, a) às 09hs	
	32
Figura 11: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m c	
as respectivas temperaturas do ar para a estação seca e quente do ano 2005, a) imager	
de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs	
Figura 12: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m c	
as respectivas temperaturas do ar para a estação seca e quente do ano 2010, a) imager	
de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs	
Figura 13: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m c	
as respectivas temperaturas do ar para a estação seca e quente do ano 2015, a) imager de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs	
Figura 14: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m c	
· · · · · · · · · · · · · · · · · · ·	
as respectivas temperaturas do ar para a estação seca e quente do ano 2020, a) imager do 10hg, b) imagem do 11hg, a) imagem do 12hg a d) imagem do 15hg	
de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs	
Figura 15: Rosa dos ventos apresentando a direção e velocidade dos ventos a 10 m co	
as respectivas temperaturas do ar para a estação úmida e quente do ano 2000, a) imag	
de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs	
Figura 16: Rosa dos ventos apresentando a Direção e Velocidade dos ventos a 10 m c	
as respectivas temperaturas do ar para a estação úmida e quente do ano 2005, a) imag	-
de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs	
Figura 17: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m c	
as respectivas temperaturas do ar para a estação úmida e quente do ano 2010, a) imag	-
de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs	
Figura 18: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m c	
as respectivas temperaturas do ar para a estação úmida e quente do ano 2015, a) imag	-
de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs	
Figura 19: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m c	
as respectivas temperaturas do ar para a estação úmida e quente do ano 2020, a) imag	-
de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs	
Figura 20: Os valores LST, NDVI e NDBI com variação espaço-temporal em Cuiabá	
período de 2000 (a-c), 2005 (d-f), 2010 (g-i), 2015 (j-i) e 2020 (m-o)	48

Figura 21: Imagens de Cuiabá classificadas com algoritmo <i>smileRandomForest</i>
utilizando imagens Landsat 7 [a) 2000, b) 2005 e c) 2010] e Landsat 8 [d) 2015 e e)
2020]50
Figura 22: Mapas do UTFVI e respectivos NDVI de Cuiabá em 2000-202054
Figura 23: Floresta Randômica, com suas árvores de decisão, resultante para amostra do
treinamento dos dados do ano 2000 com profundidade 6. As folhas (círculos) resultantes
demonstram a classificação do Fenômeno UHI (Forte, Muito Forte e Extremo) para
UTFVI na superfície do solo
1
Figura 24: Classificação da importância das variáveis independentes para UTFVI na
superfície do solo referentes aos dados amostrais dos anos 2000-2020
Figura 25: Médias do descréscimo da acurácia e índice Gini referentes ao período 2000-
2020 para UTFVI em relação à Temperatura da Superfície do Solo
Figura 26: Médias do descréscimo do índice Gini referentes ao período 2000-2020 para
UTFVI em relação à Temperatura da Superfície do Solo
Figura 27: Floresta Randômica, com suas árvores de decisão, resultante para amostra do
treinamento dos dados do ano 2000 com profundidade 5. As folhas (círculos) resultantes
demonstram a classificação do Fenômeno UHI (Forte, Muito Forte e Extremo) para a
UTFVI a 2 metros de altura73
Figura 28: Classificação da importância das variáveis independentes para UTFVI a 2
metros de altura referentes aos dados amostrais do ano 2000-2020
Figura 29: Médias do descréscimo da acurácia referentes ao período de 2000-2020 para
UTFVI em relação à Temperatura do Ar (2m)
Figura 30: Médias do descréscimo do Índice Gini referentes ao período de 2000-2020
para UTFVI em relação à Temperatura do Ar (2m)
Figura 31: Visão parcial da Floresta Aleatória para o Fenômeno UHI na superfície do
solo da síntese de todos os períodos entre 2000-2020). As categorias normal e anormal
, ,
se referem aos níveis UTFVI onde anormal é a união dos níveis Extremo, Forte e Muito
Forte e a categoria normal é a união dos níveis nenhum, fraco e médio
Figura 32: Visão parcial da Floresta Aleatória para o Fenômeno UHI a 2 metros de
altura da síntese de todos os períodos (2000-2020). As categorias normal e anormal se
referem aos níveis UTFVI onde anormal é a união dos níveis Extremo, Forte e Muito
Forte e a categoria normal é a união dos níveis nenhum, fraco e médio79
Figura 33: Classificação da importância das variáveis independentes para UTFVI na
superfície do solo referente à síntese dos dados amostrais dos períodos 2000-202080
Figura 34: Classificação da importância das variáveis independentes para UTFVI a 2
metros de altura referente à síntese dos dados amostrais dos períodos 2000-202080
Figura 35: Floresta Randômica, com suas árvores de decisão, resultante para amostra do
treinamento dos dados do ano 2000 com profundidade 5. As folhas (círculos) resultantes
demonstram a classificação do Fenômeno UHI (Forte, Muito Forte e Extremo) em
função das variáveis microclimáticas Umidade Relativa do Ar, NDVI, Velocidade e
Direção dos Ventos locais para UTFVI em relação à superfície do solo
Figura 36: Classificação da importância das variáveis independentes para UTFVI em
relação à superfície do solo referentes aos dados amostrais do período de 2000-2020 86
Figura 37: Médias do descréscimo da acurácia referentes ao período 2000-2020 para
UTFVI em relação à Temperatura da Superfície do Solo
Figura 38: Médias do descréscimo Gini referentes ao período 2000-2020 para UTFVI
em relação à Temperatura da Superfície do Solo
Figura 39: Classificação da importância das variáveis independentes para UTFVI a 2
metros de altura resultantes da classificação dos dados amostrais do período de 2000-
202091

Figura 40: Médias do descréscimo da acurácia referentes ao período 2000-2020 para
UTFVI a 2 metros de altura
Figura 41: Médias do descréscimo da acurácia e índice Gini referentes aos anos 2000-
2020 para UTFVI em relação da Temperatura do Ar (2 metros)
Figura 42: Árvores de decisão, resultante para amostra do treinamento dos dados do ano
2000 com profundidade de 5 níveis. As folhas resultantes demonstram a classificação do
Fenômeno UHI a 2 metros de altura (Forte, Muito Forte e Extremo) em função das
variáveis microclimáticas Umidade Relativa do Ar, NDVI, Velocidade e Direção dos
Ventos locais
Figura 43: Séries das Temperaturas diárias da superfície do solo e das temperaturas do
ar (2 metros) e UTFVI da superfície do solo e a 2 metros para ano 2000
Figura 44: Séries das Temperaturas diárias da superfície do solo e das temperaturas do
ar (2 metros) e UTFVI da superfície do solo e a 2 metros para ano 2010101
Figura 45: Séries das Temperaturas diárias da superfície do solo e das temperaturas do
ar (2 metros) e UTFVI da superfície do solo e a 2 metros para ano 2015102
Figura 46: Séries das Temperaturas diárias da superfície do solo e das temperaturas do
ar (2 metros) e UTFVI da superfície do solo e a 2 metros para ano 2020103
Figura 47: Temperatura da Superfície do Solo em relação à Direção dos Ventos
denotada pelos quadrantes de 1 a 8. O traço dentro da caixa indica o valor mediano dos
dados das temperaturas da superfície do solo para os anos de 2000-2020109
Figura 48: Temperatura do Ar (2m) em relação à Direção dos Ventos denotada pelos
quadrantes de 1 a 8. O traço dentro da caixa indica o valor mediano dos dados das
temperaturas do ar para a cidade de Cuiabá para os anos de 2000-2020109
Figura 49: Floresta Aleatória para Fenômeno UHI - ano 2005
Figura 50: Floresta Aleatória para Fenômeno UHI - ano 2010
Figura 51: Floresta Aleatória para Fenômeno UHI - ano 2015
Figura 52: Floresta Aleatória para Fenômeno UHI - ano 2020
Figura 53: Floresta Aleatória para Fenômeno UHI a 2 metros de altura para o ano de
2005
Figura 54: Floresta Aleatória para Fenômeno UHI a 2 metros de altura para o ano de
2010
Figura 55: Floresta Aleatória para Fenômeno UHI a 2 metros de altura para o ano de
2015 com profundidade 5
Figura 56: Floresta Aleatória para Fenômeno UHI a 2 metros de altura para o ano de
2020 com profundidade 5

Lista de tabelas

Tabela 1: Caracteristicas das Imagens Landsat. Fonte: USGS	6
Tabela 2: Os limites dos Índices de Avaliação Ecológica e Fenômeno de Ilha de Calor	
	12
Tabela 3: Quantificação das imagens coletadas dos satélites Landsat 7 e 8	21
	28
Tabela 5: Velocidades médias dos ventos e suas direções em Cuiabá-MT durante a	
estação de quente e seca para os anos 2000-2020.	30
Tabela 6: Velocidade e Direção dos Ventos médio em Cuiabá-MT, com suas respectiva	
	18 .38
1 , 1	
Tabela 7: Estatísticas descritivas de LST, NDVI e NDBI para Cuiabá de 2000-2020	
Tabela 8: Proporção da área vegetada em Cuiabá-MT	
Tabela 9: Proporção das áreas construídas em Cuiabá para os anos 2000-2020	
Tabela 10: Classificação obtida com o uso do algoritmo <i>smileRandomForest</i> (GEE)	
Tabela 11: Acurácia geral calculada para as imagens dos anos 2000-2020 (Figura 21)	.51
Tabela 12: Proporção das áreas afetadas, em km² pela condição do UTFVI em 2000-	
2020	53
Tabela 13: Distribuição do Índice de Variância do Campo Termal Urbano por pontos	
, 1 1	55
Tabela 14: Distribuição do Índice de Variância do Campo Termal Urbano por pontos	
,	.55
Tabela 15: Distribuição do Índice de Variância do Campo Termal Urbano por pontos	55
· · · · · · · · · · · · · · · · · · ·	.56
	.50
Tabela 16: Medidas descritivas das variáveis independentes quantitativas. UTFVI	~ 0
	.58
Tabela 17: Matriz de Confusão e estatísticas da variável Fenômeno UHI para os dados	
	.59
Tabela 18: Estatísticas da variável Fenômeno UHI para os dados amostrais de	
1	60
Tabela 19: Indicadores preditivos referentes aos dados amostrais de teste para os anos	
2000-2020	61
Tabela 20: Matriz de Confusão e estatísticas da variável Fenômeno UHI a 2 metros de	;
	69
Tabela 21: Estatísticas da variável Fenômeno UHI para os dados amostrais de	
treinamento e teste para o Fenômeno UHI a 2 metros de altura dos anos de 2000-2020	i
Tabela 22: Indicadores preditivos referentes aos dados amostrais de treino para os anos	, O
2000-2020	
Tabela 23: Matriz de Confusão e estatísticias da variável Fenômeno UHI para os dado	. / 1
1	
amostrais do ano 2000	81
Tabela 24: Estatísticas da variável Fenômeno UHI para os dados amostrais de	_
treinamento e teste para o Fenômeno UHI na superfície do solo dos anos de 2000-2020	
	82
Tabela 25: Indicadores preditivos dos dados amostrais do período 2000-2020	
Tabela 26: Matriz de Confusão e estatísticias da variável Fenômeno UHI a 2 metros de	e
altura para os dados amostrais do 2000	88
Tabela 27: Estatísticas da variável Fenômeno UHI a 2 metros de altura para os dados	
amostrais de treinamento e teste dos anos de 2000-2020.	89
Tabela 28: Indicadores preditivos referentes à validação (Teste) dos dados amostrais	
para o período 2000-2020	90
r r v	- 0

Tabela 29: Avaliação estatística sumarizada da classificação dos dados para os anos
2000-2020 em relação à UTFVI a 2 metros de altura96
Tabela 30: Avaliação estatística sumarizada da classificação dos dados para os anos
2000-2020 em relação à UTFVI a 2 metros de altura
Tabela 31: Avaliação estatística sumarizada da classificação dos dados para os anos
2000-2020 em relação à UTFVI na Superfície do Solo em função do NDVI e demais
variáveis microclimáticas
Tabela 32: Avaliação estatística sumarizada da classificação dos dados para os anos
2000-2020 em relação à UTFVI a 2 metros de altura em função do NDVI e demais
variáveis microclimáticas98
Tabela 33: Matriz de Confusão e estatísticas da variável Fenômeno UHI para os dados
amostrais do ano 2005
Tabela 34: Matriz de Confusão e estatísticas da variável Fenômeno UHI para os dados
amostrais do ano 2010
Tabela 35: Matriz de Confusão e estatísticas da variável Fenômeno UHI para os dados
amostrais do ano 2015127
Tabela 36: Matriz de Confusão e estatísticas da variável Fenômeno UHI para os dados
amostrais do ano 2020127
Tabela 37: Matriz de Confusão e estatísticas da variável Fenômeno UHI a 2 metros de
altura para os dados amostrais do ano 2005
Tabela 38: Matriz de Confusão e estatísticas da variável Fenômeno UHI a 2 metros de
altura para os dados amostrais do ano 2010
Tabela 39: Matriz de Confusão e estatísticas da variável Fenômeno UHI a 2 metros de
altura para os dados amostrais do ano 2015
Tabela 40: Matriz de Confusão e estatísticas da variável Fenômeno UHI a 2 metros de
altura para os dados amostrais do ano 2020

LISTA DE ABREVIATURAS

Km - Quilômetros

Km2 – Quilômetros quadrados

L - Leste

LST – Temperatura da Superfície do Solo

m - Metros

MAE – Erro Médio Absoluto

mm – Milímetros

MT - Mato Grosso

MSE – Erro Médio Quadrático

N - Norte

NDBI – Índice de Área Construída por Diferença Normalizada

NDVI – Índice de Vegetação por Diferença Normalizada

NE – Nordeste

NNE - Nor-Nordeste

NNO - Nor-Nordeste

NO – Noroeste

O - Oeste

R² – R-Quadrado

RF – Random Forest (Floresta Aleatória)

RH – Umidade Relativa do Ar

RMSE – Raiz do Erro Médio Quadrático

S-Sul

skinTemp – Temperatura da Superfície do Solo

SO - Sudoeste

SSE - Su-Sudeste

SVM – Máquinas Vetorias de Suporte

Temp2m – Temperatura do Ar (2m)

TOA – Reflectância do Topo da Atmosfera

UHI – Ilhas de Calor Urbano

UHIA – Ilhas de Calor Urbano Atmosférico

UHIS – Ilhas de Calor Urbano de Superfície

UR – Umidade Relativa do Ar

UTFVI – Índice de Variância do Campo Termal Urbano

WINDQUAD – Direção dos Ventos (Quadrante)

WINDSPEED – Velocidade dos Ventos

LISTA DE SÍMBOLOS

- % Percentual
- °C Graus Celsius
- dε Efeito da Distribuição Geométrica das Superfícies Naturais e Reflexões Internas
- εv Emissividade da Vegetação
- F_v Vegetação Fracionada
- K Kelvin
- K₁ Constante de Calibração (Landsat)
- K₂ Constante de Calibração (Landsat)
- L_{λ} Radiância Espectral
- $T_{\text{a}}-Temperatura\ do\ Ar$
- T_d Temperatura do Orvalho
- $T_{\rm B}\,$ Temperatura do Brilho
- $Wm^{\text{-}2}sr^{\text{-}1}mm^{\text{-}1} Radiância Espectral$
- λ Tamanho Efetivo da Onda

RESUMO

Silva, A. S. Modelagem e análise de cenários através da classificação do Índice de Variância do Campo Termal Urbano em ilhas de calor urbano em Cuiabá utilizando Floresta Aleatória. 2024, 143f. Tese (Doutorado em Física Ambiental) — Instituto de Física, Universidade Federal de Mato Grosso.

Ilhas de Calor Urbanas (UHI) têm em seu contexto uma complexidade bastante peculiar em sua formação que é fortemente influenciada pela atuação humana. Retenção de calor, barreiras não naturais que impedem os ventos de seguirem o caminho natural, mudanças na geomorfologia causam enormes impactos na vegetação local, na sociedade em relação a saúde e economia. Viabilizar uma forma de compreensão da ocorrência do Fenômeno de Ilhas de Calor Urbanas (IHU) e de como ela impacta nas variáveis microclimáticas Temperatura do Solo, Temperatura do Ar, Direção e Velocidade dos Ventos Locais, Umidade Relativa do Ar e o Índice por Diferenças Normalizada da Vegetação (NDVI) é o objetivo deste estudo que utiliza uma abordagem de análise de dados através de Floresta Aleatória (Random Forest - RF) com o objetivo de avaliar e analisar a influência do Índice de Variância do Campo Termal Urbano (UTFVI) nas variáveis microclimáticas e no índice espectral NDVI e buscar classificar quais as tendências que podem ocorrer diante de circunstâncias e comportamento destas variáveis microclimáticas bem como do NDVI. Para alcancar os objetivos, o modelo é treinado com dados advindos de satélite Landsat 7 e 8 para classificação das imagens e obtenção da predominância da cobertura do solo e dados de reanálise ERA5 que fornece informações sobre as variáveis microclimáticas. Num primeiro momento foi realizado análise espacial da região de interesse com o uso do software Google Earth Engine de forma a obter a espacialização da imagem bem como a temporalidade dos dados amostrais e assim poder realizar a obtenção dos índices espectrais relacionados ao fenômeno da formação de ilhas de calor urbano, bem como dados microclimáticos. Durante a experimentação, o modelo foi gerado com uma amostra de dados reduzida a níveis de indicação do UTFVI para a determinação do acontecimento do Fenômeno UHI (forte, muito forte e extremo) para balancear a taxa de falsos negativos e falsos positivos e assim poder gerar um modelo preditivo mais preciso. A obtenção dos resultados demonstram a importância do aprendizado de máquina para melhorar a detecção dos fatores que levam a condições perigosas da existência do Fenômeno UHI e consequente auxílio na mitigação dos riscos inerentes a Ilhas de Calor Urbanas. Este trabalho contribui no sentido de identificar quais as localizações da área em estudo tendem a ter uma maior intensidade no Fenômeno UHI em função de variáveis microclimáticas, ou mesmo quais locais são propensos a ter baixa intensidade no Fenômeno UHI e, consequentemente, proporcionar aos gestores públicos uma melhor percepção de como resolver os problemas causados pela formação de ilhas de calor urbano.

Palavras-chave: Fenômeno UHI, ilhas de calor urbano, UTFVI, floresta aleatória, modelagem espaço-temporal.

ABSTRACT

Silva, A. S. Modeling and analysis of scenarios through the classification of the Urban Thermal Field Variance Index in urban heat islands in Cuiabá using Random Forest. 2024, 143f. Thesis (Doctorate in Environmental Physics) – Institute of Physics, Federal University of Mato Grosso.

Urban Heat Islands (UHI) have in their context a very peculiar complexity in their formation that is strongly influenced by human action. Heat retention, unnatural barriers that prevent winds from following the natural path, changes in geomorphology cause huge impacts on local vegetation, society in relation to health and economy. Provide a way of understanding the occurrence of the Urban Heat Island Phenomenon (IHU) and how it impacts microclimatic variables: Soil Temperature, Air Temperature, Direction and Speed of Local Winds, Relative Air Humidity and the Normalized Difference Index of Vegetation (NDVI) is the objective of this study that uses a data analysis approach through Random Forest (Random Forest – RF) with the objective of evaluating and analyzing the influence of the Urban Thermal Field Variance Index (UTFVI) on the variables microclimatic variables and the NDVI spectral index and seek to classify what trends may occur given the circumstances and behavior of these microclimatic variables as well as the NDVI. To achieve the objectives, the model is trained with data from the Landsat 7 and 8 satellite to classify images and obtain the predominance of land cover and ERA5 reanalysis data that provides information on microclimatic variables. Initially, a spatial analysis of the region of interest was carried out using the Google Earth Engine software in order to obtain the spatialization of the image as well as the temporality of the sample data and thus be able to obtain spectral indices related to the phenomenon of island formation. of urban heat, as well as microclimatic data. During the experiment, the model was generated with a data sample reduced to UTFVI indication levels to determine the occurrence of the UHI Phenomenon (strong, very strong and extreme) to balance the rate of false negatives and false positives and thus be able to generate a more accurate predictive model. Obtaining the results demonstrate the importance of machine learning to improve the detection of factors that lead to dangerous conditions in the existence of the UHI Phenomenon and consequent assistance in mitigating the risks inherent to Urban Heat Islands. This work contributes towards identifying which locations in the area under study tend to have a greater intensity in the UHI Phenomenon due to microclimatic variables, or even which locations are likely to have low intensity in the UHI Phenomenon and, consequently, provide public managers with a better understanding of how to solve the problems caused by the formation of urban heat islands.

Keywords: UHI phenomenon, urban heat islands, UTFVI, random forest, spatiotemporal modeling.

1. INTRODUÇÃO

Os impactos causados pela interferência humana no meio ambiente têm sido muito importantes e a comunidade científica mundial tem envidado esforços para tentar compreender qual é a profundidade dos impactos diretos na ecologia e analisálos é de suma importância, uma vez que eles interferem nos mais variados setores da sociedade, tais como agricultura, economia e saúde. Um dos aspectos destes impactos mais comumente estudado é a formação de ilhas de calor cujas características antropogênicas são antagônicas, se por um lado o ser humano prospera vivendo em comunidade por outro há uma inevitável decadência ecológica com a vegetação sendo substituida por edifícios, asfaltos e concretos causando aumento significativo na temperatura local.

O que define Ilhas de Calor Urbanas (UHI) é a diferença da temperatura, da superfície do solo ou do ar, existente entre a área urbana e a área não urbana (o ambiente rural em torno da cidade). De acordo com a ocorrência espacial elas podem ser classificadas como Ilhas de Calor Urbanas de superfície (UHIS) que se baseia na Temperatura da Superfície do Solo ou Ilhas de Calor Urbanas atmosférico (UHIA) cuja relação é com a Temperatura do Ar.

Analisar o contexto de ilhas de calor pode ser bastante complexo, mas compreender alguns aspectos de como variáveis microclimáticas como a Umidade Relativa do Ar, Índice por Diferença Normalizada da Vegetação (NDVI), Direção e Velocidade dos Ventos Locais, Temperatura da Superfície do Solo e Temperatura do Ar podem colaborar na formação de um abordagem eficaz para revelar os efeitos nocivos derivados deste fenômeno. Assim, fazer uso de classificação de dados com o uso de Floresta Aleatória pode contribuir para a compreensão do comportamento climático, uma vez que com ela é possível verificar quais as influências que as variáveis microclimáticas têm entre si ou uma sobre a outra.

1.1. PROBLEMÁTICA

A realização de pesquisas que abordam questões climáticas e que envolvem áreas geográficas extensas podem ser bastante trabalhosas e custosas devido à necessidade de se realizar obtenção de dados em diversos locais. A origem de dados baseada em espacialização geográfica devido às suas características inerentes impõem certa obstaculização para a detecção e captação de dados micrometeorológicos (Temperaturas do Ar e do Solo, Umidade Relativa, Velocidade e Direção dos Ventos, NDVI, Pressão Atmosférica e etc) já que captura de dados micrometeorológicos pode necessitar de repetidos equipamentos posicionados em diversos locais em uma região de estudo.

Fazer o uso de dados de satélite é um boa alternativa aos empecilhos em função do dimensionamento geográfico de uma área de estudo, assim por meio de conjunto de dados de reanálise ERA5, que combina modelos climáticos com dados de satélites, (observacionais e sensores de solo) e de imagens de satélites (Landsat 7 e 8) será possível a realização de estudos das variáveis microclimáticas com análise espaço-temporal a cerca de Ilhas de Calor Urbanas em regiões tropicais, especificamente em Cuiabá-MT.

1.2. JUSTIFICATIVA

A classificação de dados é uma técnica utilizada em algoritmos de aprendizado de máquina muito empregada tanto na predição de eventos quanto em tomadas de decisões. Com grande aceitação em várias áreas do conhecimento a Floresta Aleatória (Random Forest), que emprega o uso de classificação de dados, além de regressão, é utilizada para analisar cenários dada uma variedade de dados. No contexto climático, é muito utilizada no campo da previsão do clima.

Este trabalho tem como objetivo utilizar a classificação de dados para analisar os efeitos do Índice de Variância do Campo Termal Urbano (UTFVI) na cidade de Cuiabá e verificar quais variáveis microclimáticas, especificamente

Temperatura da Superfície do Solo (LST), Temperatura do Ar, Umidade Relativa do Ar e Direção e Velocidade dos Ventos Locais têm maiores influências no fenômeno de Ilhas de Calor Urbanas (UHI), além do NDVI.

E como objetivos específicos deste estudo, que colaboraram para se alcançar o objetivo geral acima, temos:

- A realização de um levantamento espaço-temporal dos dados a cerca da área em estudo para obtenção dos índices espectrais NDVI, Índice por Diferença Normalização de Áreas Construídas (NDBI), LST e UTFVI com o propósito de avaliar a área urbana de Cuiabá a fim de compreender a relação da vegetação com áreas construídas;
- Compreensão de como se comporta os ventos locais no que diz respeito à direção e velocidade quando há uma intensificação da temperatura do solo ou do ar;
- Realizar classificação da cobertura do solo com o uso de Floresta Aleatória para descobrir a abrangência de áreas construídas e áreas vegetadas para buscar uma correlação com os índices espectrais NDVI, LST e UTFVI, e, por fim;
- 4. A análise de cenários com base nos resultados da predição dos dados classificados para compreender a influência das variáveis microclimáticas no Índice de Variância do Campo Termal Urbano no sentido da formação do fenômeno Ilhas de Calor Urbano.

De maneira geral, esta pesquisa irá realizar análises através de classificação de dados com o uso de Floresta Aleatória para uma melhor compreensão da variância do campo termal ao nível da superfície do solo e a 2 metros de altura.

2. REVISÃO BIBLIOGRÁFICA

Nesta revisão serão abordados todos os conceitos que auxiliará o desenvolvimento deste trabalho aplicados ao local de estudo. A contextualização do clima urbano, algoritmos de árvore de decisão e classificação com o uso de Floresta Aleatória (Random Forest) e técnicas de aprendizagem de máquina, dados e imagens de satélites Landsat 7 e 8, além de dados de reanálise provenientes do ERA5.

2.1. CLIMA URBANO

O clima urbano, notoriamente estudado desde 1820, teve como campo de pesquisa inicial a análise do calor da cidade de Londres em relação ao seu entorno onde se atribuiu a causa da diferença de calor entre a cidade e a área rural como sendo a maior absorção da radiação solar em função da cobertura do solo e a diminuição da evaporação (MILLS, 2008). (STEWART; OKE, 2012) confirmaram muitas das características temporais do efeito urbano na temperatura do ar observadas nas pesquisas iniciadas por Luke Howard (HOWARD, 2012) e afirmam que o clima urbano caracteriza-se pela anormalidade das variáveis Temperatura do Solo, da Temperatura do Ar, Umidade Relativa do Ar, Precipitação, comportamento dos ventos locais (Direção e Velocidade) e outros fatores.

2.2. ILHAS DE CALOR

Uma Ilha de Calor Urbana se caracteriza, basicamente, pela alteração na cobertura do solo em função do processo de urbanização onde o uso de materiais impermeáveis como concreto e asfalto é imperativo e esse processo influencia diretamente no aumento da temperatura nas áreas urbanas. Além disso, implica em importantes desafios para o planejamento urbano e melhorias ambientais, haja vista a grande quantidade de energia demandada, políticas climáticas necessárias e

atenuação dos seus efeitos na saúde pública. É definida pela diferença da temperatura superfície do solo da área urbana com a área não-urbana (LEHOCZKY et al., 2017) e (O'MALLEY et al., 2014; SOUTO; COHEN, 2021; VILLANUEVA-SOLIS, 2017).

De acordo com (SHAHMOHAMADI et al., 2011; SOUTO; COHEN, 2021; TOMLINSON et al., 2011; VILLANUEVA-SOLIS, 2017; WICKI; PARLOW; FEIGENWINTER, 2018) monitorar as consequências e prover alguma solução para mitigar os efeitos das ilhas de calor é de interesse das pesquisas realizadas e das atualmente em desenvolvimento que lançam mão de diversas tecnologias com o objetivo de compreender como as variáveis microclimáticas (Temperatura do Ar, Umidade Relativa do ar, Temperatura do Solo, Ventos e Precipitação) atuam para aumentar ou diminuir a percepção da intensidade do calor.

2.2.1. Os Fatores Que Afetam A Intensidade De Ilhas De Calor Urbanas

A formação de Ilha de Calor Urbana é tida como as diferenças no balanço de energia da superfície entre a área urbana e o seu entorno. Segundo (ERELL; PEARLMUTTER; WILLIAMSON, 2012) é o balanço de energia em um dado momento que determina se o resfriamento ou o aquecimento ocorre em uma superfície e a taxa na qual esses processos ocorrem.

É interessante observar que, embora o seu surgimento sejam relatado durante o dia, o fenômeno se destaca mais comumente no período noturno (particularmente em cidades densas de baixa latitude com um grande fluxo de calor antropogênico). Devido a incidência maior ser no período noturno, este fenômeno é também conhecimento UHI noturno que é formado pelo resultado do resfriamento relativamente rápido nas áreas rurais no final da tarde e início da noite, em comparação com o resfriamento mais lento na cidade (ERELL; PEARLMUTTER; WILLIAMSON, 2012).

Como as temperaturas começam a cair no final da tarde e início da noite, a taxa de resfriamento na cidade costuma ser menor do que na área rural. Como resultado, uma diferença de temperatura é formada, normalmente atingindo um

máximo várias horas após o pôr do sol. A tendência se inverte logo após o nascer do sol, quando o aquecimento urbano ocorre em ritmo mais lento do que no entorno rural (ERELL; PEARLMUTTER; WILLIAMSON, 2012). Alguns fatores que potencializam o surgimento de UHI: (i) A forma urbana com suas dimensões e espaçamentos; (ii) a densidade das áreas construídas; (iii) superfícies impermeáveis; (iv) vegetação; (v) as propriedades dos materiais urbanos; (vi) o clima; (vii) localização geográfica; e (viii) atividade humana.

2.3. MONITORAMENTO DE ILHAS DE CALOR COM LANDSAT E ERA5

Para análise de dados desta pesquisa foram utilizadas imagens adquiridas de dois satélites, Landsat 7 ETM+ e Landsat 8 OLI/TRS com suas características apresentadas na Tabela 1, ambos com resoluções espaciais de 60m e 100m, respectivamente e são fornecidas livremente pela USGS (U.S. Geological Survey). O Landsat 7 é dotado de oito bandas espectrais, incluindo bandas termais (banda 6), já o Landsat 8 possui 8 bandas sendo que duas são bandas de infravermelho termais (bandas 10 e 11), os quais são utilizados para calcular a Temperatura da Superfície do Solo. As demais são utilizadas para obtenção de NDVI, NDWI dentre outras informações (NUGRAHA; GUNAWAN; KAMAL, 2019).

Tabela 1: Características das Imagens Landsat. Fonte: USGS

Resolução	Landsat 7 ETM+	Landsat 8 OLI/TIRS
30 m	Banda 1 – Azul	Banda 1 – Aerosol Costeiro
	Banda 2 – Vermelha	Banda 2 – Azul
	Banda 3 – Verde	Banda 3 – Vermelha
	Banda 4 – NIR	Banda 4 – Verde
	Banda 5 – SWIR-1	Banda 5 – NIR
	Banda 7 - SWIR-2	Banda 6 – SWIR-1
		Banda 7 – SWIR-2
		Banda 9 – Cirrus
15 m	Banda 8 – Pan	Banda 8 – Pan
60 m	Banda 6 - TIR	
100 m		Banda 10 – TIR-1
_		Banda 11 – TIR-2

O ERA5¹ por ser um conjunto de dados de reanálise combina modelo de dados com observações de todo o mundo e fornece um conjunto de dados consistente aos quais se aplicam as leis da física. Este princípio, também chamado de assimilação de dados, é baseado em uma metodologia numérica de centros de previsão do clima. O ERA5 fornece dados mensais e diários (por dia e hora) disponíveis desde 1940, já com os devidos tratamentos e predições (HERSBACH et al., 2020) através de um banco de dados que combina observações passadas com modelos para gerar séries temporais consistentes de múltiplas variáveis climáticas.

2.4. ANÁLISE DA SUPERFÍCIE URBANA

Analisar os índices de cobertura do solo observando a vegetação urbana e da Temperatura da Superfície do Solo é importante pois influenciam fortemente a Temperatura do Ar e a Umidade Relativa do Ar, fatores determinantes do clima urbano.

2.4.1. Classificação Da Cobertura Do Solo Utilizando Índices Espectrais

Com a expansão das áreas urbanas naturalmente há uma redução significativa na vegetação e isto minimiza os efeitos naturais de resfriamento do sombreamento e da evapotranspiração do solo e da vegetação. Observar a evolução dos índices de vegetação em uma escala local e associá-los ao comportamento do clima urbano é de suma importância, pois assim pode-se analisar a relação intrísica que há entre a cobertura do solo e o aumento da temperatura do solo.

2.4.1.1. Índice De Vegetação Por Diferença Normalizada (NDVI)

O NDVI foi proposto por (ROUSE et al., 1974) com o objetivo analisar as mudanças ocorridas na cobertura do solo em função da vegetação. O procedimento aplicado é uma separação da vegetação do brilho do solo através de dados obtidos de satélites, como Landsat e Modis. Como a cobertura do solo muda constantemente

¹ https://climate.copernicuUHIs.eu/climate-reanalysis

devido a sazonalidade do próprio desenvolvimento e ou atividade da vegetação, este índice permite comparações de dados temporais e espaciais das condições da vegetação (NERY; MOREIRA; FERNANDES, 2014)

Conforme proposto por (ROUSE et al., 1974) a equação geral do NDVI é calculado pela diferença da reflectância e pela divisão da soma entre a faixa do infravermelho médio e a faixa do vermelho (NERY; MOREIRA; FERNANDES, 2014) e é estruturada da seguinte forma:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

onde: NIR = comprimento de onda do infravermelho próximo e RED = comprimento de onda do vermelho.

O índice **NDVI** tem o seu espectro em um intervalo entre -1 e 1 e neste intervalo os valores que estiverem próximo a -1 representam corpos d'água e ou nuvens, os valores próximos a zero indicam que o solo está exposto, valores positivos entre 0,6 e 0,80 representam a presença de vegetação.

2.4.1.2. Temperatura Da Superfície Do Solo (LST)

Segundo (BENALI et al., 2012), LST é uma variável indicativa do balanço de energia de superfície líquido impulsionado pela emissão de superfície de radiação de onda longa. Os fluxos de calor latente e sensível são fortemente influenciados pela temperatura da superfície e a distribuição de energia entre eles é governada pelo teor de umidade, tipo de superfície, velocidade dos ventos e emissividade.

A extração dos dados relacionados à Temperatura da Superfície do Solo se dá através das medições do espectro do infravermelho termal (TIR) feitas por sensores terrestres, aéreos ou por sensores de satélites. A recuperação do LST a partir de banda termal Landsat 8 OLI (banda 10) pode ser feita pelo seguinte algoritmo (GUHA et al., 2018):

$$L_{\lambda} = 0.0003342 * DN + 0.1 \tag{1}$$

onde L_{λ} é a radiância espectral em Wm⁻²sr⁻¹mm⁻¹.

$$T_{B} = \frac{K_{2}}{\ln\left(\binom{K_{1}}{L_{\lambda}} + 1\right)} \tag{2}$$

onde T_B é a temperatura do brilho em Kelvin (K), onde L_λ é a radiância espectral em Wm⁻²sr⁻¹mm⁻¹; K_2 e K_1 são constantes de calibração. Para Landsat 8 OLI, K_1 é 774,89 e K_2 é 1321,08.

De acordo com (GUHA et al., 2018; SOBRINO; JIMENEZ; PAOLINI, 2004) a emissividade é calculada usando método dos limites do NDVI. (CARLSON; RIPLEY, 1997) disponibilizou a equação para o cálculo da cobertura da vegetação fracionada. Sendo ela:

$$F_{v} = \left(\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}}\right) \tag{3}$$

onde $NDVI_{min}$ é o valor mínimo do NDVI (0,2) onde os pixels são considerados como solo nu e $NDVI_{max}$ é o valor máximo do NDVI (0.5) onde os pixels são considerados vegetação saudável.

de é o efeito da distribuição geométrica das superfícies naturais e reflexões internas. Para superfícies heterogêneas e onduladas, o valor de $d\varepsilon$ pode ser de 2%.

$$d\varepsilon = (1 - \varepsilon_s)(1 - F_v)F\varepsilon_v \tag{4}$$

onde ε_v é a emissividade da vegetação, ε_s é a emissividade do solo, F_v a vegetação fracionada e F é um fator de forma cuja média é 0,55 (SOBRINO; JIMENEZ; PAOLINI, 2004).

$$\varepsilon = \varepsilon_{\nu} F_{\nu} + \varepsilon_{s} (1 - F_{\nu}) + d\varepsilon \tag{5}$$

onde ε é a emissividade. A partir das equações (4) e (5), ε pode ser determinado pela seguinte equação:

$$\varepsilon = 0.004 * Fv + 0.986 \tag{6}$$

E finalmente, de acordo com (WENG; LU; SCHUBRING, [s.d.]) o LST deriva da seguinte equação:

$$LST = \frac{T_B}{1 + \left(\lambda k \frac{T_B}{hc}\right) \ln \varepsilon}$$
 (7)

onde λ é o tamanho da onda efetiva (10,9 mm para a banda 10 nos dados Landsat 8), k é a constante de Boltzmann (1,38 * 10⁻²³ J/K), h é a constante de Plank (6,626 x 10⁻³⁴ Js), c é a velocidade da luz no vácuo (2,998 x 10⁻⁸ m/s) e ε é a emissividade.

2.4.1.3. Índice Por Diferença Normalizada De Áreas Construídas (NDBI)

O Índice por Diferença Normalizada de Áreas Construídas é um indicador que combina as bandas 3 (vermelho) com a banda 5 (infra-vermelho de ondas curtas) e a banda 1 (azul) com a banda 4 (infra-vermelho próximo), para revelar variações do solo para a região de interesse utilizando o satélite Landsat 7 e as bandas B6 e B4 combinadas e B5 e B2 também combinadas para se utilizar dados de imagens do satélite Landsat 8 (ALBONWAS; AL-KHAKANI, 2022). Este indicador é usado para capturar as variações do solo livre de vegetação, é pode ser calculado utilizando as seguintes equações as quais se referem ao satélite Landsat 7 (8) e Landsat 8 (9):

$$NDBI = \frac{(B5+B3)-(B4+B1)}{(B5+B3)+(B4+B1)}$$
(8)

$$NDBI = \frac{(B6+B4)-(B5+B2)}{(B6+B4)+(B5+B2)}$$
(9)

2.4.1.4. Umidade Relativa Do Ar (UR)

A Umidade Relativa do AR (UR) é uma medida de quanto vapor de água existe em uma mistura de água e ar em comparação com a quantidade máxima possível. A UR é uma relação entre a relação de umidade de uma mistura particular de água e ar em comparação com a relação de umidade de saturação a uma determinada temperatura (CALLAHAN; ELANSARI; FENTON, 2019). (KONG; SINGH, 2016) diz que a UR do ar é definida como a razão entre a pressão de vapor

do ar e sua pressão de vapor de saturação. (LAWRENCE, 2005) diz que a Umidade Relativa do Ar e a Temperatura do Ponto de Orvalho são dois indicadores amplamente usados da quantidade de umidade no ar. De acordo com (HOWELL; DUSEK, 1995), a seguinte equação deriva a Umidade Relativa do Ar.

$$UR = ((17,269 * Td) / (273.3 + Td)) - ((17.269 * Ta) / (237.3 + Ta))$$
 (10)

onde,

Ta se refere a Temperatura do Ar; *Td* se refere a temperatura do orvalho

2.4.1.5. Índice De Calor Urbano (UHI)

O termo Ilha de Calor Urbana (UHI) refere-se ao aumento das temperaturas urbanas nos últimos anos denotado pela alta carga térmica, problemas relacionados à ventilação urbana e ao aumento do uso de asfalto e concreto (ENTERIA; SANTAMOURIS; EICKER, 2021). A contribuição do calor antropogênico varia consoante a localização e a estação: durante o verão tem uma contribuição insignificante enquanto que durante o inverno pode ter uma contribuição significativa (FILHO et al., 2017). A intensidade do UHI é dada pela diferença entre a LST das áreas urbanas e da áreas não-urbanas usando as seguintes equações (NGUYEN, 2020; WALEED et al., 2023):

$$LST > \mu + 0.5 * \delta \tag{11}$$

$$0 < LST \le \mu + 0.5 * \delta \tag{12}$$

A equação (11) se refere à área urbana e a equação (12) à área não-urbana, onde μ e δ são a média e o desvio padrão da LST da área em estudo, respectivamente.

$$UHI_{N} = (T_{S} - T_{M})/T_{SM} \tag{13}$$

Onde: UHI_N é o UHI normalizado, T_S é a LST, T_M é a média de LST da área em estudo e T_{Std} é o desvio padrão da LST da área em estudo.

2.4.1.6. Índice De Variância Do Campo Termal Urbano (UTFVI)

O efeito UHI é um fator importante em estudos ambientais e de saúde pública e dada a importância dos efeitos causados há alguns índices de conforto térmico disponíveis para a avaliação dos impactos do UHI na qualidade de vida urbana e neste caso o UTFVI foi usado nesta análise para medir o Fenômeno UHI conforme a equação abaixo (FILHO et al., 2017; NGUYEN, 2020):

$$UTFVI = \frac{T_s - T_M}{T_S} \tag{13}$$

onde T_s é a LST do local (°C ou °K) e T_m é a média da LST de toda área em estudo.

A Tabela 2 apresenta os valores das respectivas categorias quanto ao Fenômeno de Ilha De Calor Urbana (UHI) e cada uma destas categorias com o seu correspondente Índice de Avaliação Ecológica (LIU; ZHANG, 2011).

Tabela 2: Os limites dos Índices de Avaliação Ecológica e Fenômeno de Ilha de Calor Urbano.

UTFVI	Fenômeno UHI	Índice de Avaliação Ecológica
< 0,000	Nenhum	Excelente
0,000 - 0,005	Fraco	Bom
0,005 - 0,010	Médio	Normal
0,010 - 0,015	Forte	Ruim
0,015 - 0,020	Muito Forte	Pior
> 0.020	Extremo	Pior caso

2.4.1.7. Índice De Construção Por Diferenças Normalizadas (NDBI)

O NDBI permite mapear pontos de áreas construídas manufaturadas e juntamente com o NDVI é utilizado para examinar as variações espaciais e temporais da LST (KUMAR; SHEKHAR, 2015; RANAGALAGE; ESTOQUE; MURAYAMA, 2017). Este índice é calculado com base na diferença entre as bandas de infravermelho ondas de curtas (SWIR) e o infravermelho próximo (NIR) e tem como limites valores entre -1 e +1. Valores negativos representam corpo d'água enquanto valores altos acima de 0,40 representam áreas construídas, valores baixo representam vegetação. Exemplificando a equação temos:

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR} \tag{14}$$

2.5. ÁRVORES DE DECISÃO

Conhecidas como árvores de regressão e classificação (CART), são algoritmos de aprendizado de máquinas supervisionados para classificação e regressão de problemas. Basicamente, é um classificador com duas etapas: aprendizado e classificação (validação). Na fase de aprendizado, ela aprende a gerar uma árvore de decisão de um conjunto de amostras de treinamento que foram classificados. Na fase de classificação, a árvore obtida na fase de aprendizado é utilizada para classificar dados ainda não classificados. A Figura 1 apresenta um fluxograma com suas quatro partes: nós raiz, nós internos, ramificações e nós folhas (LAN et al., 2020).

O nó raiz inicia a árvore de decisão e a partir dele a população começa a se dividir de acordo com várias características, os nós internos realizam a tomada de decisão com base em valores condicionantes que disparam o processo de tomada de decisão pelos nós internos, levando a árvore a gerar os nós folhas que represetam todas as possíveis saídas do conjunto de dados.

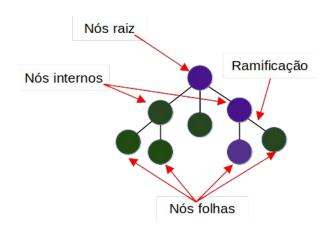


Figura 1: Estrutura básica de uma árvore de decisão

Como uma Máquina Vetorial de Suporte (SVM), as árvores de decisão são algoritmos versáteis de aprendizado de máquina que consistem em conjuntos de regras organizados hierarquicamente. É uma estrutura recursiva simples para representar um procedimento de decisão no qual uma instância futura é classificada em classes pré-definidas presentes e tenta dividir as observações em subgrupos mutuamente exclusivos. Cada parte em uma árvore corresponde a um ou mais registros do conjunto de dados original.

Os nós superiores são nomeados como o nó raiz (sem link de entrada) e representam todas as linhas no conjunto de dados fornecido. Os outros nós são nomeados como nós internos ou de decisão (apenas um link de entrada) usados para testar um atributo (HUSSEIN ABDULZHRAA AL-SAGHEER; ALHARAN; AL-HABOOBI, 2017).

As árvores de decisão também são os componentes fundamentais da Floresta Aleatória, que estão entre os mais poderosos algoritmos de aprendizado de máquina disponíveis atualmente.

2.6. FLORESTA ALEATÓRIA (RANDOM FOREST)

A Floresta Aleatória (RF) é um algoritmo de aprendizado de máquina não paramétrico baseado no conceito de árvores de decisão, elas são métodos estatísticos

que podem ser aplicados à problemas de regressão, bem como problemas de classificação de duas classes ou de multi-classes (GENUER et al., 2017). Juntamente com a árvores de decisão, Floresta Aleatória é provalmente o modelo de aprendizagem de máquina não-linear mais amplamente utilizado, uma vez que proporciona boa performance em modelos de classificação e regressão (CORREIA; PEHARZ; DE CAMPOS, 2020), superando outros algoritmos de classificação como regressão logística, máquina vetoriais de suporte (SVM) e redes neurais.

Uma Floresta Aleatória é um classificador que consiste em uma coleção de classificadores estruturados em árvore $\{h(x, k), k = 1,...\}$ onde $\{k\}$ são vetores aleatórios independentes e distribuídos de forma idêntica e cada árvore dá um voto unitário para a classe mais popular na entrada x (BREIMAN, 2001).

Algumas das suas aplicações no contexto climático foram apresentadas nos trabalho de (VENTER et al., 2020) que realizou um mapeamento, com modelo de regressão, da Temperatura do Ar e no trabalho de (DOS SANTOS, 2020) no qual foi feito uma estimativa espaço-temporal da Temperatura do Ar. Há também trabalhos em que a RF foi utilizada na detecção de objetos (GALL; RAZAVI; VAN GOOL, 2012; LIU; XIONG, 2012).

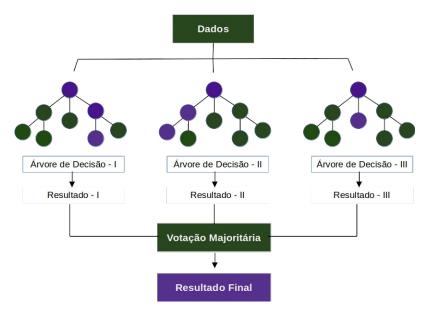


Figura 2: Classificador de Floresta Aleatória

Este tipo de algoritmo é apropriado para tratar alta dimensionalidade na modelagem de dados porque pode lidar com valores ausentes e ainda pode lidar com dados contínuos, categóricos e binários. O esquema *bootstrapping* e *ensemble* proporciona robustez o suficiente para superar os problemas de ajuste excessivo e, portanto, não há necessidade de diminuir o tamanho da profundidade das árvores. Além da alta precisão de previsão, a Floresta Aleatória é eficiente, interpretável e não paramétrico para vários tipos de conjuntos de dados (ALI et al., 2012).

O uso de RF pode ser aplicado tanto para classificação como para regressão e também para análise da importância das variáveis e a proximidade dos dados (ENGLUND et al., 2012). Segundo Verikas et al., quando aplicado em um conjunto de dados, cada árvore que compõem a RF é treinada usando uma amostra bootstrap dos dados que é uma técnica de reamostragem estatística que envolve amostragem aleatória de um conjunto de dados com substituição e, claro, permite descrever as amostras a serem extraídas dos dados de treinamento com substitutos (EFRON, 1979). Para cada árvore, os dados não usados para treinamento, os dados *out of bag* (OOB), podem ser usados para testar o desempenho da generalização (erro OOB) e também podem ser utilizados para estimar a importância da variável (ENGLUND et al., 2012; VERIKAS; GELZINIS; BACAUSKIENE, 2011).

3. MATERIAL E MÉTODOS

Para a realização deste trabalho foi definido um estudo com levantamento de dados sobre os índice de Vegetação, Umidade Relativa do Ar, Temperatura do Ar, Temperatura da Superfície do Solo e Ventos (direção e velocidade). E para tanto, foi empregado o sensoriamento remoto com o uso de imagens dos satélites Landsat 7 e 8, a fim de calcular a área de cobertura de solo, calor da superfície do solo e dados de reanálise ERA5 para processamento das variáveis microclimáticas: Umidade Relativa do Ar, Temperatura do Ar, Ventos e incluindo também Temperatura da Superfície do Solo.

Nas seguintes seções serão descritos a localização da área de estudo, dados meteorológicos e algoritmos utilizados para avaliação do impacto dos índice de variância de campos termais urbanos, objeto de estudo deste trabalho.

3.1. ÁREA DE ESTUDO - CIDADE DE CUIABÁ-MT

A cidade de Cuiabá, capital do estado de Mato Grosso, está situada entre as coordenadas geográficas 15°50' de latitude sul e 50°10' de longitude oeste, no centrooeste brasileiro. A Figura 3 apresenta a localização da cidade de Cuiabá, a qual está inserida no estado de Mato Grosso e no Brasil.

De acordo com a classificação Koppen (BECK et al., 2018), seu clima é do tipo AW, tropical semi-úmido com temperaturas variando entre 30°C e 36°C, com estações bem definidas: (i) seca e quente (outono-inverno) abrangendo os meses de maio a setembro; e (ii) e chuvosa e quente (primavera-verão) que abrange os meses de outubro a abril (ALCARDE ALVARES et al., 2013; VILANOVA; MAITELLI, 2009). As medições de precipitação obtidas pelo sítio do INMET, considerando os anos de 2000, 2005, 2010, 2015 e 2020, para Cuiabá direcionam para um regime pluviométrico com estações bem definidas: (i) seca e quente (de maio a setembro); e outra chuvosa e quente (outubro a abril), conforme a Figura 4.

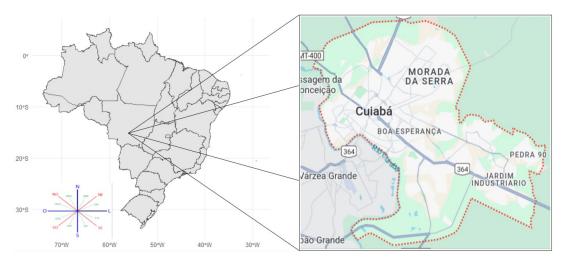


Figura 3: Localização da cidade de Cuiabá-MT (Brasil). Fonte: Google Earth Engine (acessado em 22/11/23)/(PEREIRA; GONÇALVES, 2019).

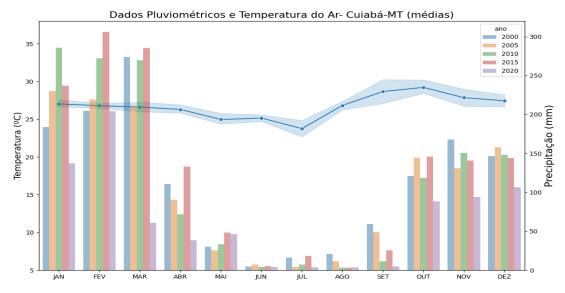


Figura 4: Climograma de Cuiabá-MT para os anos de 2000, 2005, 2010, 2015 e 2020 com médias mensais das variáveis microclimáticas Temperatura do Ar e precipitação. Fonte: Autor

Com um clima típico das cidades tropicais, Cuiabá é um centro urbano conhecido como uma dos mais quentes do Brasil devido às suas características geográficas, pouca ventilação e baixa altitude. A pouca ventilação, em função da

baixa frequência e velocidade dos ventos, faz com que a influência antropogênica na Temperatura do Ar seja mais perceptível (DUARTE, 2015).

3.2. DADOS METEOROLÓGICOS E PRÉ-PROCESSAMENTO

O uso de estações fixas, transectos móveis e sensoriamento remoto são tecnologias utilizadas para medição e monitoramento de UHI e o uso destas tecnologias depende da amplitude da pequisa (escala), recursos e do contexto analítico (WERNECK, 2022).

Tipicamente, as UHIs são investigadas utilizando dados da temperatura próximo a superfície obtidos através de estações climáticas automáticas, ocorre que nem sempre as mesmas estão disponíveis, seja por falta de investimento público ou privado ou por questões inerentes à localização geográfica da área de interesse. Para a realização de estudo acerca de UHIs se faz necessário a distribuição uniforme de estações meteorológicas, automáticas ou não, o que leva a inviabilidade de pesquisas que tenham a necessidade de analisar informações sobre variáveis micrometeorológicas de grandes áreas territoriais (RAVANELLI et al., 2018).

Assim, realizar estudos com o uso de dados provenientes de satélites se torna mais prático ao mesmo tempo em que proporciona redução do custo operacional e financeiro que envolve uma pesquisa em área geográficas de grande escala, uma vez que dados e imagens provenientes de satélites estão disponíveis em sítios especializados na Internet.

A praticidade da obtenção de informações disponibilizadas por satélites impulsionou muitas pesquisas sobre questões climáticas (DO NASCIMENTO et al., 2022; FERREIRA; JÚNIOR, 2020). Mas, o acesso às informações advindas dos satélites podem não ser de simples manuseio pois há uma necessidade de conhecimentos técnicos mais específicos para lidar com algoritmos, tipo de dados, tratamento dos dados por eventuais perdas ou danos na obtenção dos mesmos e ainda o fato de haver uma diversidade muito grande de ferramentas. Apesar do avanço e

das facilidades proporcionadas por estas ferramentas, ainda há complexidades a serem enfrentadas no domínio do processamento da informação.

A captação dos dados dos satélites Landsat 7 e 8 se dá em um periodicidade revisitação de 16 dias com horário de passagem no território brasileiro entre 10 e 12 horas da manhã e como destacado na Tabela 1 as resoluções das imagens variam de acordo com a banda a ser utilizada, sendo a máxima de 100 metros e a mínima de 15 metros. Os dados de reanálise ERA5 são disponibilizados com valores diários de 24 horas e os mesmos podem ser sintetizados em médias diárias ou mensais.

O acesso às informações dos índices espectrais (LST e NDVI) obtidos das imagens dos satétites Landsat 7 e 8 e do dados de reanálise do produto ERA5 (Temperatura da Superfície, Temperatura do Ar, Umidade Relativa do Ar, Direção dos Ventos, Velocidade dos Ventos, Precipitação e Evaporação), além das coordenadas geográficas de latitude e longitude, foi feito através da plataforma Google Earth Engine (https://code.earthengine.google.com) que permite ao usuários executar análises geoespaciais na infraestrutura Google utilizando Javascript ou Python. Estas informações podem ser imagens ou arquivos de dados tabelados, no formato comma-separated values — CSV. Para o processamento destas informações foram utilizadas as linguagens de programação R e Python.

A Figura 5 apresenta um fluxograma das etapas a serem realizadas, as quais indicam uma preparação dos dados obtidos em relação a área de estudo para serem utilizados como informações de entrada no processamento para análise posterior. As cores destacadas identificam as fases de cada etapa, onde o verde indica a origem das imagens e dos dados que são utilizados para a derivação dos índices espectrais (marrom) bem como na classificação dos dados e das próprias imagens (azul), a cor laranja indica a fase de análise estatística dos dados amostrais.

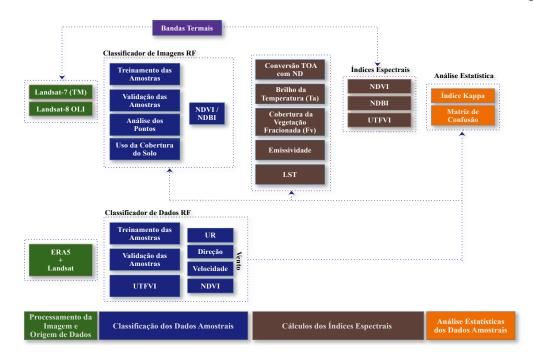


Figura 5: Fluxograma para estimar o impacto das variáveis (LST–Temperatura da Superfície do Solo). Baseado em (ULLAH et al., 2022)

3.2.1. Coleção Das Imagens Obtidas Dos Satélites Landsat 7 E 8

A Tabela 3 apresenta a coleção de imagens de satélites coletadas e utilizadas neste trabalho. As imagens utilizadas de ambos os satélites (Landsat 7 e 8) foram obtidas da coleção 2, nível/camada 1 que consistem em Números Digitais (DN) em escalas quantizadas e calibradas que representam os dados da imagem multiespectral. Esses dados de nível 1 podem ser redimensionados para refletância e/ou radiância do Topo da Atmosfera (TOA) e os dados da banda térmica podem ser redimensionados para a temperatura de brilho TOA. Os coeficientes de calibração são extraídos dos metadados da imagem (CHANDER; MARKHAM; HELDER, 2009).

Tabela 3: Quantificação das imagens coletadas dos satélites Landsat 7 e 8.

Ano	Landsat 7	Landsat8
2000	18	-
2005	22	-
2010	18	-
2015	-	19
2020	-	22

3.3. DEFINIÇÃO DAS VARIÁVEIS DA AMOSTRA

O conjunto de dados da amostra é formado por 4 variáveis, sendo 01 variável dependente (categórica), identificada como Índice de Variância do Campo Termal Urbano (UTFVI) e 4 variáveis independentes que são variáveis microclimáticas correlacionadas às características proeminentes do clima urbano. De acordo com (KAUR; MITTAL, 2021), uma variável independente é a causa presumida da variável dependente (o efeito presumido) e ela explica ou leva em conta a variação na variável dependente. As variáveis independentes a serem utilizadas nesta pesquisa são as seguintes: Velocidade dos Ventos, Direção dos Ventos, Umidade Relativa do Ar e NDVI.

3.4. DEFINIÇÃO DOS QUADRANTES NA ROSA DOS VENTOS

A representação da Direção dos Ventos é dada pelo artefato gráfico Rosa dos ventos o qual determina a Direção dos Ventos em graus. Assim, a Direção dos Ventos ao norte é definida como sendo 0° (zero grau), ao leste 90°, ao sul 180° e a oeste 270°. Utilizar a Direção dos Ventos exata torna bastante complexo a determinação dos seus efeitos sobre outras variáveis microclimáticas e, portanto, neste trabalho será definido a utilização de quadrantes o que poderá facilitar a descoberta dos efeitos em tais variáveis.

A Figura 6 apresenta a Rosa dos Ventos com os seus respectivos quadrantes, sendo o quadrante 1 compreendido pelas direções entre 0° e 45°, o quadrante 2 correspondendo as direções entre 45° e 90°, o quadrante 3 abrange as direções dos ventos entre 90° e 135°, o quadrante 4 guarda similitude com as direções dos ventos entre 135° e 180°, o quadrante 5 é análogo às direções dos ventos entre 180° e 225°, direções dos ventos entre 225° e 270°, 270° e 315° e 360° compreendem, respectivamente, os quadrantes 6, 7 e 8.

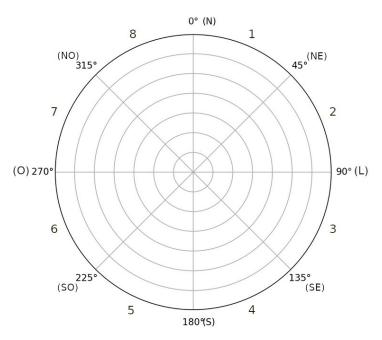


Figura 6: Rosa dos ventos com os respectivos quadrantes (1 a 8).

3.5. SELEÇÃO DAS IMAGENS DE SATÉLITES

A seleção das imagens utilizadas foram obtidas a partir dos satélites Landsat 7 e 8 restringindo as coordenação de localização latitude e longitude ao local de estudo deste trabalho, a cidade de Cuiabá-MT. Frisa-se que estas coordenadas margeam os limites retornando um polígono que retrata toda a área urbana desta cidade.

3.6. ENGENHERIA DOS DADOS

Para o tratamento dos dados obtidos dos satélites Landsat 7 e 8 bem como dados de reanálise originados do Era5 foram utilizadas a linguagem de programação Python e ferramentas específicas para manipulação e manutenção de dados Pandas e Numpy. Estas ferramentas fornecem funcionalidades de manutenção dos dados propriamente dita utilizando o Pandas, bem como manipulação de uma massa de dados através de vetorização, neste caso utiliza-se o Numpy.

Nem sempre as informações advindas de imagens de satélites são adequadas e completas, assim se faz necessário realizar verificação de dados faltantes ou mesmo dados com falhas. Assim, eventuais falhas e ou duplicação de dados devem ser tratados para se evitar ambiguidades das informações. Neste trabalho, houve a necessidade de se analisar inconsistências dos dados realizando correções pontuais como erro na separação dos valores numéricos nos dados, realizar a derivação de algumas informações tais como UTFVI, cálculo das direções dos ventos em função da plotagem das informações dos respectivos quadrantes na Rosa dos Ventos.

3.7. PROCEDIMENTO DOS CÁLCULOS PERTINENTES AOS ÍNDICES ESPECTRAIS

Conforme a quantificação das imagens coletadas (vide Tabela 3) se fez necessário obter uma média da imagens e assim obter as respectivas médias dos índices espectrais, NDVI por exemplo, através de um procedimento de redução aplicado ao tamanho da região de interesse.

Para se calcular os índices espectrais é importante utilizar as bandas que compõem uma imagem. O índice espectral NDVI, utilizando satélite Landsat 8, é calculado a partir da banda B5 (infravermelho próximo) e a banda B4 (vermelho) onde se subtrai a banda B4 da banda B5 e divide pela banda B5 adicionada à banda B4. Em relação ao uso das imagens do satélite Landsat 7, utilizou-se as bandas B3 (vermelha) e B4 (infravermelho próximo).

Ressalta-se que o cálculo para o índice espectral Temperatura da Superfície do Solo utiliza a banda B10 para imagens derivadas do Landsat 8 e a banda B6 para imagens obtidas do Landsat 7 e se dá conforme expresso na seção 2.4.1.

3.8. O MODELO DO ALGORITMO FLORESTA ALEATÓRIA

Segundo (ALI et al., 2012; BREIMAN, 2001), o algoritmo da Floresta Aleatória é simples e se baseia na utilização de várias árvores de decisão cujo

objetivo é apresentar resultados com a média calculada a partir da saída das árvores individuais. Tecnicamente, o algoritmo emprega a redução da variância em uma função de previsão de estimativa ("agregação bootstrap"), onde dado um conjunto de treinamento de tamanho n a agregação gera m novos conjuntos de treinamentos, cada um com um tamanho distinto, por amostragem do conjunto de treinamento original.

Em (ASSCHE et al., 2004) é apresentada a seguinte estrutura básica do algoritmo

- 1. Para um dado conjunto de dados D faça:
- a. construa um subconjunto de dados D_i , por amostragem com substituição do conjunto de dados D
- b. gere uma árvore de decisão T_i a partir de D_i usando um conjunto de recursos restrito, aleatoriamente.
- c. faça predições/classificações de acordo com a votação majoritária do conjunto de k árvores.

Definição:

$$\{DT(x,\theta_k)\}$$
 $X = 1$

onde x é o vetor de entrada e θ_k denota um vetor aleatório, que é amostrado de forma independente, mas com a mesma distribuição que o θ_k , . . . , θ_{k-1} . Amostras de bootstrap T são inicialmente derivadas dos dados de treinamento.

3.8.1. Parametrização Do Algoritmo Para Classificação Da Cobertura Do Solo A Partir De Imagens De Satélite

Uma vez obtida as imagens da área de interesse realiza-se a indicação das coordenadas para cada tipo de cobertura de solo a ser classificada. Neste caso, há uma seleção das coordenadas para pontos que deverão ser reconhecidos como água, área construída, área vegetada e solo nu com o fito de treinar o algoritmo. Indicando qual o percentual dos dados do satélite, em relação às "bandas" que referenciam cada

tipo de classificação do solo, a serem utilizados como treinamento do algoritmo (70%) e para validação (30%) e qual a quantidade de árvores que a Floresta Aleatória deverá utilizar para o treinamento, neste caso foi atribuído 500 árvores (número de estimadores).

3.8.2. Parametrização Do Algoritmo Para Classificação Dos Dados Para O Fenômeno UHI

A classificação dos dados em relação ao Fenômeno UHI é comum ao processo de parametrização da classificação de imagens com um adendo de que a construção das árvores se baseia puramente em dados e com representação gráfica (Figuras 2 e 23) onde se define a profundidade da árvore, ou seja o número de divisões que cada árvore pode fazer, que no contexto deste trabalho foi considerada 6. Há ainda uma necessidade de se informar qual a variável que será utilizada como objetivo, variável depentende, a ser alcançado dada uma votação majoritária.

Uma parametrização que auxilia na construção das árvores na Floresta Aleatória é o uso de critérios Gini ou Entropia que são medidas que determina em quais características uma árvore deve ser dividida. Exemplificando, os dados amostrais utilizados contém as variáveis microclimáticas Direção dos Ventos, Velocidade dos Ventos e Umidade Relativa do Ar com base nas quais a divisão da árvore deve ser feita, incluindo o NDVI. O algoritmo faz a divisão usando estas variáveis e escolhe aquele que resulta em uma entropia mais baixa ou uma menor impureza (Gini) após a divisão como o recurso a ser dividido, descartando o outro.

3.8.2.1. Métricas Entropia E Gini

Entropia determina como uma árvore de decisão deve escolher a divisão dos dados amostrais e é uma métrica que mede a impureza ou incerteza em um grupo de observações. A Figura 7 apresenta a pureza de um determinado conjunto.

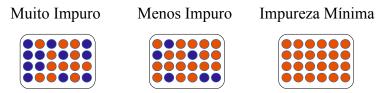


Figura 7: Entropia demonstrando o grau de pureza da informação (conjunto de dados).

De maneira geral, o índice de Gini é uma medida de variância, exemplificado na Figura 8 com a indicação de percentual. Quanto maior a variância, mais incorreta será a classificação. Portanto, valores mais baixos do Índice de Gini proporcionam uma melhor classificação.

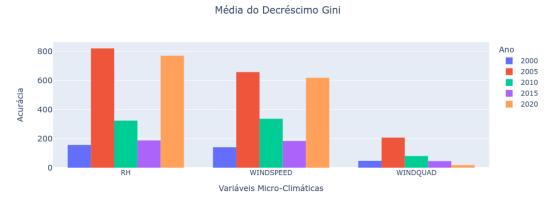


Figura 8: Demonstração da estimação da importância das variáveis usando a impureza Gini.

4. APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Para esta pesquisa atingir o objetivo esperado foram empregados métodos analíticos, com base em algoritmos de aprendizado de máquina supervisionados e o uso de diferentes fontes de dados e imagens coletadas de satélites Landsat 7 e 8 e dados de reanálise ERA5. Os procedimentos foram assim determinados: (i) determinação dos pontos geográficos; (ii) classificação da cobertura do solo; (ii) coleta dos dados das variáveis microclimáticas (Vento, Temperatura do Ar, Umidade Relativa do Ar, Temperatura da Superfície do Solo, Índice de Intensidade do Calor, Índice de Variância do Campo Termal Urbano) e, por fim; (iii) análise dos dados.

4.1. PONTOS GEOGRÁFICOS

Os pontos geográficos, localizados na cidade de Cuiabá-MT, foram determinados com base nas características da cobertura do solo. Assim, há pontos geográficos dispostos em locais de cobertura de solo com vegetação, corpos d'água e áreas construídas, conforme Tabela 4.

Tabela 4: Localização dos pontos das coletas de dados em Cuiabá-MT

	Localização dos Pontos para Coleta de Dados						
Id.	Local	Lat.	Long.	Tipo Cobertura do Solo			
1	Coxipó	-15.6320	-56.0566	Construção, Concreto e Asfalto			
2	Tijucal	-15.6257	-56.0283	Construção, Concreto e Asfalto			
3	Três Lagoas	-15.5705	-56.0382	Corpo D'água			
4	Morada da Serra II	-15.5629	-56.0481	Construção, Concreto e Asfalto			
5	UFMT	-15.6096	-56.0660	Construção, Concreto e Asfalto			
6	Aecim Tocatins	-15.6033	-56.1222	Construção, Concreto e Asfalto			
7	Centro	-15.5974	-56.0979	Construção, Concreto e Asfalto			
8	Alphaville	-15.5952	-56.0647	Construção, Concreto e Asfalto			
9	Parque Tia Nair	-15.5943	-56.0580	Vegetação/Corpo D'água			
10	Parque Mãe Bonifácia	-15.5799	-56.1056	Vegetação			
11	Parque das Águas	-15.5678	-56.0795	Vegetação/Corpo D'água			
12	Centro Político Administrativo	-15.5687	-56.0732	Construção, Concreto e Asfalto			

A Figura 9 apresenta o mapa da região urbana da cidade de Cuiabá com os referidos pontos de coleta (localizações) de dados com os seus respectivos numeradores, nela pode-se identificar os mais variados tipos de cobertura de solo.

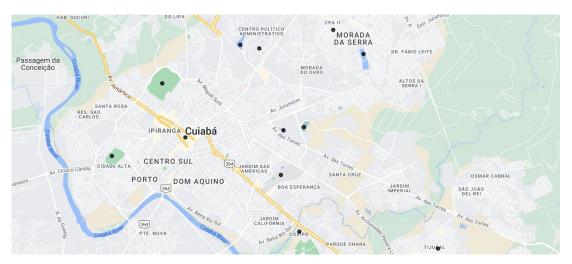


Figura 9: Mapa de Cuiabá com a indicação numérica dos respectivos pontos de coleta em referência à Tabela 4. Fonte: Google Maps

4.2. CARACTERIZAÇÃO DO REGIME DE VENTO EM CUIABÁ-MT

Alguns estudos sobre o regimento dos ventos em Cuiabá foram realizados por (FRANCO et al., 2013) e (OLIVEIRA et al., 2012). (FRANCO et al., 2013) definem que Cuiabá tem um regime de vento onde a direção predominante é norte (N) e nordeste (NE) durante boa parte do ano e sul (S) durante o período de inverno e ainda cabe ressaltar que a Velocidade dos Ventos é reduzida, porém pode ocorrer rajadas de ventos de até 35,9 m/s.

Nesta seção será feito uma análise da predominância da Direção e Velocidade dos Ventos locais em Cuiabá-MT de forma a especificar quais os impactos que estas variáveis sofrem de acordo com a variação da Temperatura do Ar em função dos horários. Neste trabalho será adotado que o direcionamento define de onde o vento se origina e para onde vai.

4.2.1. Estação Seca E Quente

A Tabela 5 apresenta o percentual das ocorrências de ventos durante a estação e quente (21 de junho a 21 de setembro), de acordo com os dados obtidos a partir do ERA5, para os anos de 2000, 2005, 2010, 2015 e 2020. Os dados referentes às localizações citadas na Tabela 4 apontam que os ventos ocorridos para os respectivos anos se originam do sul em direção ao norte e têm uma predominância que varia de 23% a 32%, aproximadamente.

Especificando as direções dos ventos para estes anos temos que para o ano 2000 houve uma maior ocorrência para o norte (N) com 25,92%, para o su-sudeste (SSE) com aproximadamente 18,43%, para o sul (S) com 15,62% e nor-noroeste (NNO) com 07,50%. Para o ano 2005, a maior incidência dos ventos foram para o norte com 23,77%, para o sul com 19,39%, para o su-sudeste com 16,19% e nor-nordeste (NNE) com 11,70%. No ano de 2010, a maior ocorrência foi para a direção norte com os dados apontando 26,01%, ocorrências para su-sudeste, sul e nor-nordeste com, respectivamente, 19,56%, 17,09% e 12,60%.

Tabela 5: Velocidades médias dos ventos e suas direções em Cuiabá-MT durante a estação de quente e seca para os anos 2000-2020.

Ano	Direção	Ocorrência (%)	Vel. Mín.	Vel. Máx.	Vel. Média	
	S	18,43		22.50	8,12	
2000	NNE	17,30	0,06			
2000	N	16,12	0,00	22,50		
	SSO	15,62				
	SSO	19,39				
2005	NNE	18,08	0,02	22,80	8,47	
	S	16,19				
	NNE	20,38	0,14	23,43	8,17	
2010	S	19,56				
2010	SSO	17,09				
	NE	12,60				
	NNE	20,55		20,73	7,44	
2015	N	14,82	0,07			
2013	S	14,72	0,07			
	NE	12,02				
	NNE	25,28		24,13	7,93	
2020	NE	15,53	0.22			
2020	S	13,13	0,22			
	N	10,48				

Os dados para o ano 2015 apontam que os ventos no sentido norte tiveram uma ocorrência de 29,17%, 14,72% para o su-sudeste, 12,02% nor-nordeste e sul com 08,41%, enquanto que para o ano 2020 houve uma maior incidência dos ventos para o norte com 31,88%, para o nor-nordeste com 15,53%, su-sudeste com 13,13% e finalmente para o sul com 10,10%. Registra-se que na Tabela 5 apresenta as quatro maiores incidências do direcionamento dos ventos, há incidências para as demais direções, muito embora com menores frequências e cabe ressaltar que para cada ponto referente as coordenadas dos locais escolhidos há informações sobre Temperatura do Ar, Umidade Relativa do Ar, Temperatura da Superfície do Solo, Direção e Velocidade dos Ventos, NDVI e Precipitação.

4.2.1.1. Comportamento Dos Ventos Locais Na Estação Seca E Quente

Analisando o comportamento da direção e velocidade dos ventos que ocorrem na estação do "inverno" com suas respectivas temperaturas do ar em função do horário da coleta dos dados (9hs, 11hs, 13hs e 15hs) realizada diariamente no ano de 2000, é possível perceber claramente a mudança de Direção dos Ventos. Pode-se observar na Figura 10a, com dados referente ao horário das 09 hs da manhã, que há uma concentração da incidência dos ventos no quadrante I e III.

No quadrante I há Temperatura do Ar mais alta, em torno de 22° a 24°C, e ventos de no máximo 10 km/h e no quadrante III há Temperatura do Ar mais fria em torno de 18°C e com ventos mais fortes indo até os 17 km/h. Frisa-se que a Figura 10b, a qual se refere ao horário de 11hs, apresenta um comportamento dos ventos locais similar ao da 09hs, com ventos fortes e Temperatura do Ar um pouco mais quente, indo para a direção sul e os mais atenuados na direção nor-nordeste, com as respectivas temperaturas do ar entre 12 e 26° C.

Na Figura 10c há uma demonstração de que houve uma alteração no comportamento dos ventos locais, onde há um direcionamento deste mais especificamente ao sul para as temperaturas do ar mais baixas (variando de 12,5°C a 25°C) e no sentido inverso (norte) a ocorrência de temperaturas do ar mais altas, entre 25°C e 32°C. Para a imagem da Figura 10d há uma repetição de similaridade do

comportamento dos ventos e da Temperatura do Ar com a imagem da Figura 10c, incidindo ventos ao sul com Temperatura do Ar mais baixa e na medida que a Temperatura do Ar aumenta há uma clara mudança na Direção dos Ventos para o norte com ventos mais espalhados indo na direção nor-nordeste e nor-nordeste, observando que a temperatura do atingiu no horário das 15hs a máxima de 35°C.

No ano de 2005, a direção e velocidade dos ventos para os horários coletados das 09hs e 11hs (Figuras 11a e 11c) tiveram um comportamento semelhante ao do ano 2000, diferenciando apenas a Temperatura do Ar que para o ano de 2005 foram mais amenas nestes horários, as temperaturas do ar mais frias se concentraram na direção sul e mais altas na direção nor-nordeste. Quanto aos demais dados atribuídos aos horários de 13hs e 15hs, tiveram um direcionamento ao sul para as temperaturas do ar mais frias, enquanto as as variações das temperaturas do ar mais alta tiveram o direcionamento ao nor-nordeste (Figura 11c) e com um maior espalhamento no sentido norte (Figura 11d), com maior incidência no sentido nor-nordeste.

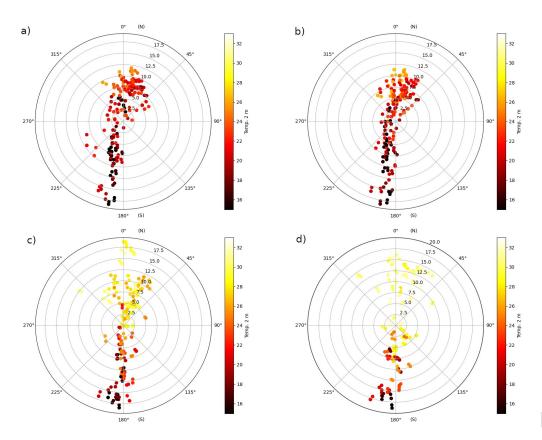


Figura 10: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m com as respectivas temperaturas do ar para a estação seca e quente do ano 2000, a) às 09hs, b) às 11hs, c) às 13hs e d) às 15hs.

Em relação ao ano de 2010, a Figura 12 demonstra que houve apenas uma discrepância no comportamento dos ventos locais somente para o horário das 15 hs onde a temperatura do air foi mais quente que os anos anteriores, muito embora o vento tem o direcionamento concentrado no quadrante referente à direção nornordeste.

Para o ano de 2015, Figura 13, pode se observar que houve o acontecimento de ventos mais fortes com média em torno de 6 km/h é há também uma semelhança no comportamento da variável microclimática Temperatura do Ar com o ano de 2005, já com as variáveis Direção e Velocidade dos Ventos há uma variação positiva em relação aos demais anos (2000, 2005 e 2010).

Observando os dados coletados para o ano 2020, pode-se observar que a Direção dos Ventos se mantém na direção nor-nordeste, com velocidades máxima entre 10 e 12 km/h para os respectivos horários de 09 e 11 hs e para os horários das 13 e 15 hs há a ocorrência de ventos mais fortes chegando a 20 km/h e com a Temperatura do Ar mais alta em relação aos anos anteriores.

Ressalta-se que neste ano houve um grande impacto de queimadas no entorno da região de Cuiabá, com grande queimadas ocorrendo, em agosto e setembro deste ano, no pantanal mato-grossense e as típicas queimadas no Parque Nacional de Chapada de Guimarães, o que pode ter contribuido de alguma maneira para o comportamento dos ventos e das temperaturas do ar, como se pode observar na Figura 14.

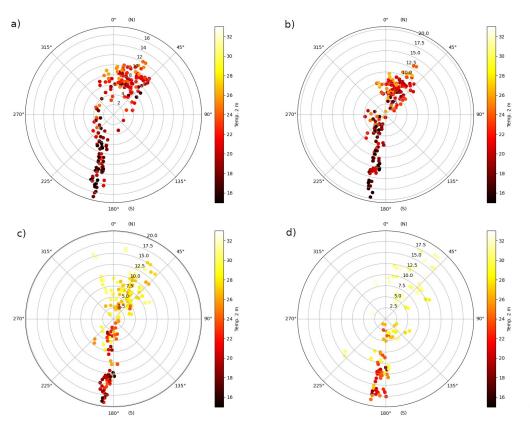


Figura 11: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m com as respectivas temperaturas do ar para a estação seca e quente do ano 2005, a) imagem de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs.

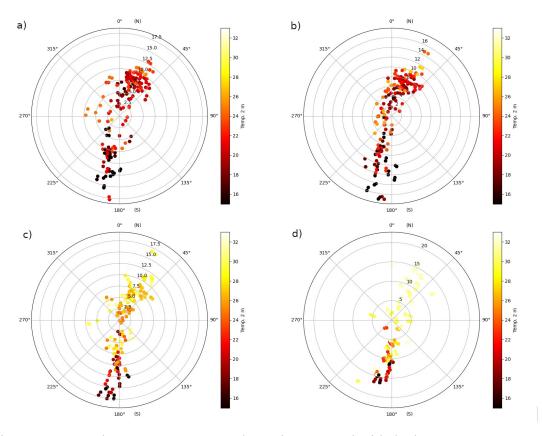


Figura 12: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m com as respectivas temperaturas do ar para a estação seca e quente do ano 2010, a) imagem de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs.

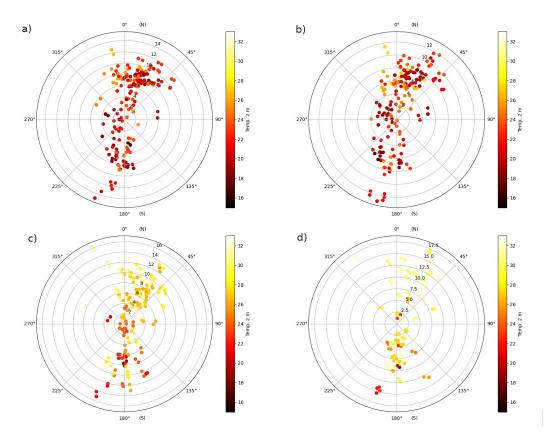


Figura 13: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m com as respectivas temperaturas do ar para a estação seca e quente do ano 2015, a) imagem de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs.

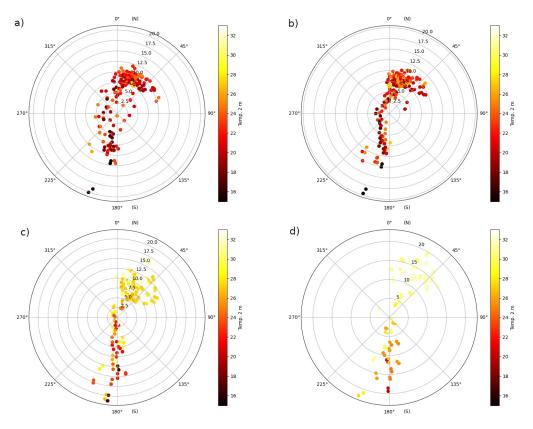


Figura 14: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m com as respectivas temperaturas do ar para a estação seca e quente do ano 2020, a) imagem de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs.

4.2.2. Estação Quente E Úmida – Verão

A Tabela 6 apresenta o percentual das ocorrências de ventos durante a estação de verão (21 de dezembro a 21 de março), referentes aos anos 2000, 2005, 2010, 2015 e 2020, conforme dados obtidos a partir do ERA5. As localizações para os quais estes dados se referem são as citadas na Tabela 4 e na Tabela 6 podemos observar que os ventos ocorridos para os respectivos anos se originam do sul em direção ao norte e têm uma predominância numa proporção que varia de 29% a 32%.

Tabela 6: Velocidade e Direção dos Ventos médio em Cuiabá-MT, com suas respectivas maiores frequências, durante a estação de verão para os anos 2000-2020.

Ano	Direção	Ocorrência (%)	Vel. Mín.	Vel. Máx.	Vel. Média
	N	29,47		17.60	6,38
2000	NNE	16,86	0,12		
2000	NNE	12,66	0,12	17,69	
	NO	7,76			
	N	30,13			
2005	NNO	19,67	0,01	21,19	6,85
	NNE	11,05			
	N	32,57	0,17	18,78	6,45
2010	NNE	19,37			
2010	NNO	15,02			
	NO	6,77			
	N	29,37		20,73	7,44
2015	NNO	15,84	0.07		
2013	NNE	11,17	0,07		
	NO	7,59			
	N	30,14		20,31	
2020	NNO	16,01	0.07		6,27
2020	NNE	13,39	0,07		
	NE	6,53			

4.2.2.1. Comportamento Dos Ventos Locais Na Estação Quente E Úmida

As imagens retratadas nas Figuras 15, 16, 17, 18 e 19 que se referem ao período do verão dos anos 2000, 2005, 2010, 2015 e 2020, respectivamente, demonstram que os ventos têm os seus comportamentos semelhantes entre si para os referidos anos com a grande maioria das ocorrências de ventos em direção ao norte, apenas para o ano de 2005 houveram, eventualmente, ventos no sentido sul e ainda as temperaturas do ar dos anos 2000, 2010, 2015 e 2020 tiveram menores valores, proporcionalmente, em relação ao ano de 2005.

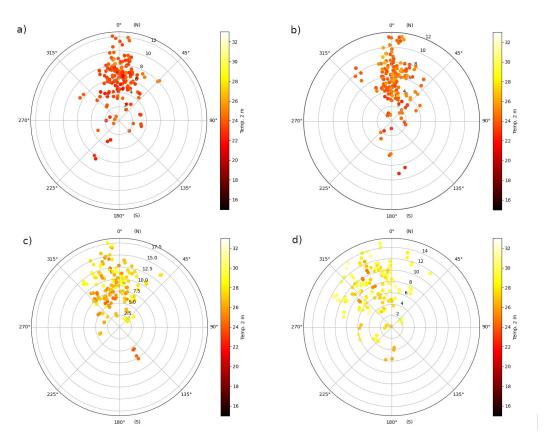


Figura 15: Rosa dos ventos apresentando a direção e velocidade dos ventos a 10 m com as respectivas temperaturas do ar para a estação úmida e quente do ano 2000, a) imagem de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs.

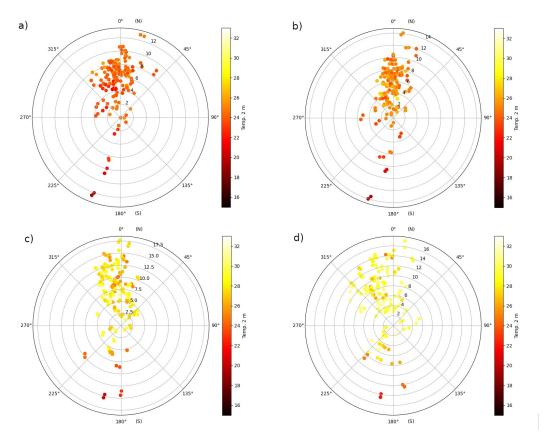


Figura 16: Rosa dos ventos apresentando a Direção e Velocidade dos ventos a 10 m com as respectivas temperaturas do ar para a estação úmida e quente do ano 2005, a) imagem de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs.

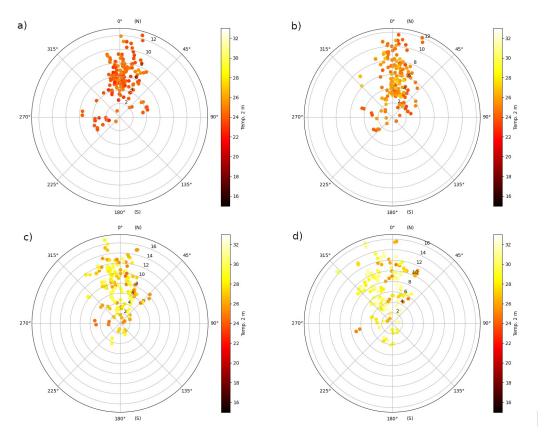


Figura 17: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m com as respectivas temperaturas do ar para a estação úmida e quente do ano 2010, a) imagem de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs.

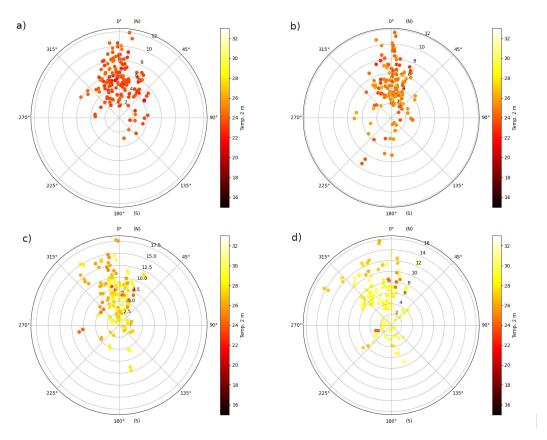


Figura 18: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m com as respectivas temperaturas do ar para a estação úmida e quente do ano 2015, a) imagem de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs.

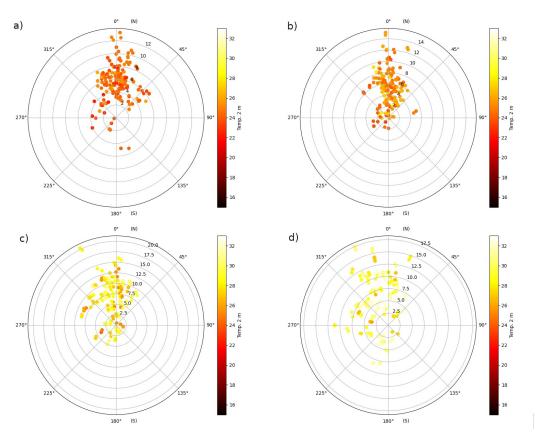


Figura 19: Rosa dos ventos apresentando a Direção e Velocidade dos Ventos a 10 m com as respectivas temperaturas do ar para a estação úmida e quente do ano 2020, a) imagem de 09hs, b) imagem de 11hs, c) imagem de 13hs e d) imagem de 15hs.

4.3. CARACTERIZAÇÃO DOS NÍVEIS DE VEGETAÇÃO EM CUIABÁ-MT

A mudança na cobertura do solo em ilhas de calor urbano é fator de grande importância na análise microclimática destas ilhas. O Índice de Vegetação por Diferenças Normalizadas (NDVI – *Normalized Difference Vegetation Index*) é bastante utilizado como indicador da qualidade da vegetação existente e tem forte correlações com fatores climáticos e ambientais (KANG et al., 2014). Caracterizar os níveis de vegetação e monitorá-los permitirá detectar mudanças no ecossistema urbano o que é de grande valia para o estudo do clima urbano.

Como consequência da expansão das áreas urbanas, a diminuição da cobertura do solo pode ser observado de uma perspectiva de indicador de áreas construídas. Este indicador trata-se do Índice de Áreas Construídas por Diferenças Normalizadas (NDBI – *Normalized Difference Vegetation Index*) que é apropriado para apontar informações sobre a extensão e a mudança dos espaços urbanos (LUZ et al., 2019).

4.3.1. Análise Da Expansão Do Espaço Urbano Com O Uso Do NDVI, NDBI E LST

O índice NDVI é amplamente utilizado para análise da expansão dos espaços urbanos, tais valores indicam características do solo urbano e rural, apontado um certo grau de saúde da vegetação. O NDVI se limita a valores entre um negativo (-1) a um positivo (+1). Áreas de rocha estéril, areia geralmente apresentam valores de NDVI muito baixos (por exemplo, 0,10 ou menos). Vegetação esparsa, como arbustos e gramíneas ou culturas em processo natural de envelhecimento, pode resultar em valores moderados de NDVI (aproximadamente 0,20 a 0,50). Já valores altos de NDVI (aproximadamente 0,60 a 0,90) correspondem a vegetação densa, comumentemente encontradas em florestas temperadas e tropicais ou mesmo culturas em seu estágio de crescimento máximo.

A Figura 20 apresenta os mapas dos índices espectrais LST, NDVI, NDBI de Cuiabá em 5 períodos de 2000 a 2020 e a Tabela 7 demonstra os respectivos valores sumarizados. No ano 2000 os valores para a Temperatura da Superfície do Solo (LST) variaram de 28,31 °C a 29,77 °C, com a média estabelecida em 29,05°C, para o índice espectral NDVI a faixa de valores foi de -0,5107 para o mínimo e o máximo de 0,7522, com a respectiva média em 0,3024. O índice NDBI teve a sua faixa variando de -1.0061 a 0,6370 com uma média de 0,3173. Em 2005 percebe-se que para o LST e o NDVI houve um ligeiro aumento nos valores mínimos, médio e máximo.

O mesmo acontecendo com o NDBI, com exceção do valor máximo que apresentou um valor mais baixo em relação ao ano 2000. Os anos de 2010, 2015 e 2020 tiveram os índices mantendo um patamar com valores semelhantes para a

Temperatura da Superfície do Solo, para o índice de vegetação por diferenças normalizadas e índice de áreas construídas por diferenças normalizadas.

Cabe ressaltar que o índice NDVI aponta para um aumento na vegetação, ou pelo menos no âmbito da saúde dela, nos anos de 2015 e 2020. Por serem antagônicos, os índices NDBI e NDVI demonstram que nos períodos citados houve uma acentuada influência entre si dada a característica inversamente proporcional. É possível observar que quando houve um aumento no índice de vegetação, houve uma queda na percepção de áreas construídas e para se ter uma melhor convicção se faz necessário realizar classificação através de algoritmos especializados como por exemplo o de Floresta Aleatória, já que as imagens dos satélites Landsat 7 e 8 não apresentam uma diferenciação perceptível nos contextos dos índices.

Tabela 7: Estatísticas descritivas de LST, NDVI e NDBI para Cuiabá de 2000-2020

Índices Espectrais									
	LST			LST NDVI		NDBI			
Ano	Mín.	Médio	Máx.	Mín.	Médio	Máx.	Mín.	Médio	Máx.
2000	28,31	29,05	29,77	-0.5107	0.3024	0.7522	-1.0061	0.3173	0.6370
2005	29,20	29,87	30,65	-0,6875	0.3410	0.7826	-1.1035	0,3380	0,6279
2010	29,56	29,90	30,60	-0,5306	0.2988	0,7601	-1,2398	0,3064	0,6616
2015	29,71	30,18	30,93	-0,7547	0,3196	0,8280	-1.0942	0,3403	0,7485
2020	29,72	30,15	30,84	-0,8553	0,2925	0,8153	-0,9664	0,3154	0,6055

Em todas as imagens que compõem a Figura 20, as áreas com valores de NDVI mais altos foram encontradas principalmente em parques com vegetação mais densas e reservas florestais como o Parque Mãe Bonifácia e o Parque Massai Okamura e nas áreas mais urbanizadas e nas áreas de solo nu os valores para o NDVI são mais baixos. A partir de 2005 o solo descoberto começou a representar uma área significativa no nordeste da cidade de Cuiabá, o que revela a causa dos baixos valores de NDVI nessa área e um aumento inversamente proporcional nos valores do NDBI.

Em 2010, 2015 e 2020, áreas com baixos valores de NDVI foram encontradas na região sudeste de Cuiabá muito disso devido à influência da expansão da própria cidade e ainda pelo aumento da Temperatura da Superfície do Solo, que

em 2015 e 2020 teve um acentuado aumento na região sudeste, o que coincede com a diminuição do NDVI e com o aumento do NDBI. Ainda na esteira da discussão sobre o NDBI, a Figura 20 demonstra que a sua maior incidência se dá na região centralizada da cidade de Cuiabá.

Pode-se observar na Figura 20 que para o ano 2015 o NDBI está bastante constrastado com os demais anos, percebe-se que há uma maior intensidade na região centralizada, o que coincide com o advento da Copa do Mundo de 2014. Período em que houve a realização de muitas obras, principalmente na área da mobilidade com a restruturação das principais vias da cidade e readequação de vias para implantação à época do sistema de transporte de veículo leve sobre trilho (VLT).

A análise da dinâmica da ocupação para o espaço urbano da cidade de Cuiabá para as últimas duas décadas demonstram um declínio natural da vegetação urbana, muito em função da própria expansão da área urbana com o aparecimento de novos empreendimentos imobiliários e de áreas construídas de forma desordenadas. Neste período houve um aumento da área construída como se pode observar na Figura 20. Na Tabela 8 há uma demonstração da extensão da área de vegetação com as proporções calculadas em kilômetros quadrados. Com relação ao contexto da vegetação densa tem-se que a área ficou estabelecida em torno de uma média de 0,28 km² e de vegetação moderada, composta de arbustos e pequenas árvores, com uma extensão de 42,23 km², muito por conta das reservas e parques existentes na cidade.

Para as demais condições das vegetações houve um decréscimo em torno de 15% a partir do ano 2000, havendo um impacto maior para o ano de 2010, cujas condições das áreas vegetadas moderada e densa sofreu um descréscimo ainda maior em relação aos demais anos estudados. Um comparativo com o ano 2000 demonstra que a perda ocorrida foi em torno de 50% para a vegetação tipificada como moderada e a vegetação diminui em cerca de 98%, voltando a prosperar a partir do ano 2015 conforme demonstrada na Tabela 8.

Tabela 8: Proporção da área vegetada em Cuiabá-MT

	Área NDVI (km²)					
Ano	0,00 – 0,20 (baixa)	0,21 – 0,40 (esparsa)	0,41 – 0,60 (moderada)	> 0,61 (densa)	Área Total	
2000	59,14	127,31	42,23	0,28	229,06	
2005	55,87	115,60	56,05	1,19	228,71	
2010	92,78	116,17	20,74	0,02	229,71	
2015	34,86	108,81	76,00	10,09	229,76	
2020	49,66	108,44	64,53	7,16	229,79	
Média	58,46	115,26	51,91	3,74	229,40	

Dados do NDBI, conforme a Tabela 9, demonstra que o tamanho das áreas construídas apresentou uma grande variação entre os anos 2000 e 2020, se estabelecendo com um acréscimo de aproximadamente 50% em áreas consideradas impermeáveis.

Tabela 9: Proporção das áreas construídas em Cuiabá para os anos 2000-2020

NDBI (km²)				
Ano	Área			
2000	44,86			
2005	88,83			
2010	51,06			
2015	89,21			
2020	67,61			

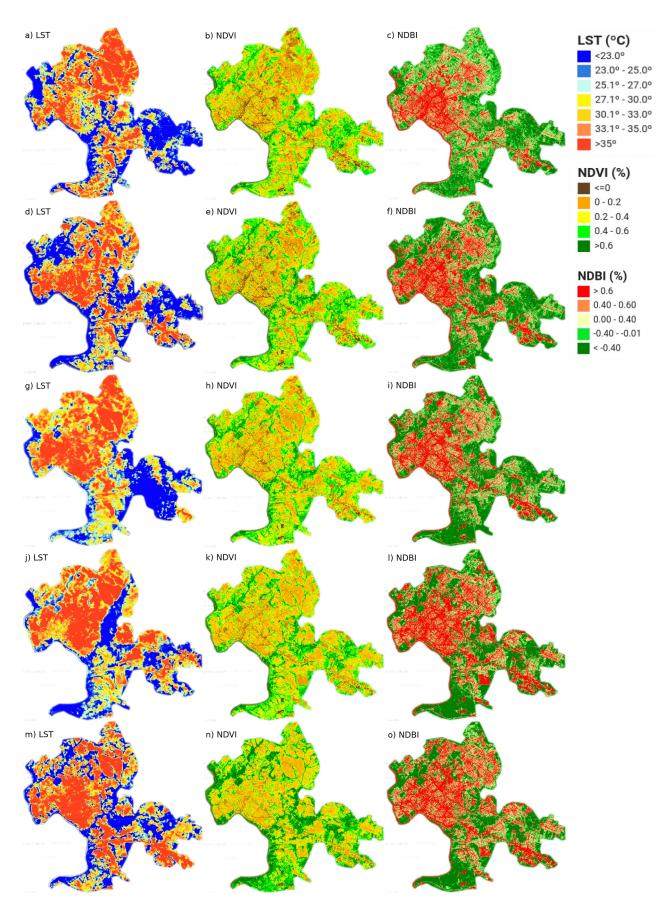


Figura 20: Os valores LST, NDVI e NDBI com variação espaço-temporal em Cuiabá no período de 2000 (a-c), 2005 (d-f), 2010 (g-i), 2015 (j-i) e 2020 (m-o).

4.3.1.1. Classificação Das Coberturas Do Solo Com O Uso Do Algoritmo Floresta Aleatória (Random Forest)

Como forma de melhorar os resultados da classificação da cobertura do solo foi necessário utilizar um algoritmo de aprendizado de máquina supervisionado, pois os cálculos empregados nos resultados da seção anterior foram realizados sem uma classificação através de algoritmos especializados. Dentre vários existentes foi escolhido o algoritmo de Floresta Aleatória, mais especificamente um fornecido pela ferramenta *Google Earth Engine* conhecido como *smileRandomForest* que como todo algoritmo deste tipo usa uma estrutura de tomada de decisão onde o objetivo é calcular e encontrar a melhor árvore que traz os melhores resultados através de uma votação e assim possibilitar a classificação da imagem obtida.

Para realizar esta classificação houve um treinamento das árvores, ao todo 500 árvores foram aplicadas neste treinamento, utilizando 70% dos dados, para condicionar o algoritmo à melhor classificação possível, com os dados restantes houve uma aplicação de teste para validação dos dados. Para este treinamento e validação, foi utilizado uma Floresta Aleatória com uma quantidade de 500 árvores de decição, o que se refere ao tamanho da floresta em si.

Tabela 10: Classificação obtida com o uso do algoritmo *smileRandomForest* (GEE)

		Tipo de Cobertura de Solo (km²)						
Ano	Água	Áreas Construídas	Vegetação	Solo Nu	Acurácia (%)			
2000	15,29	83,47	121,74	12,09	87,58			
2005	14,64	84,53	115,84	17,57	91,45			
2010	16,94	92,49	80,55	42,60	94,56			
2015	7,67	102,54	87,38	34,99	95,65			
2020	14,14	99,35	77,50	41,60	95,44			

A Tabela 10 apresenta as proporções para quatro tipos de coberturas de solo, sendo corpo d'água, áreas construídas, vegetação e solo nu após a realização do algoritmo de classificação supervisionada, onde temos um resultado da performance do algoritmo acima de 80% de acurácia, implicando numa precisão bastante alta em

relação às previsões totais. Além da constante diminuição da área vegetada, onde em 2000 havia 121,74 km² e em 2020 o tamanho da respectiva área diminui para o patamar de 77,50 km², um descréscimo de 36,33% nestes 20 anos, houve um preocupante aumento da proporção do tipo de cobertura solo nu que em 2000 era de 12,09 km² evoluindo para um total de 41,60 km², o que implica em um aumento expressivo de 244,08%, no mesmo sentido o tipo de cobertura de solo áreas construídas também teve um aumento de 19,02%, quando comparado ao ano 2000.

A Figura 21 apresenta as imagens classificadas e se referem aos respectivos anos 2000, 2005, 2010, 2015 e 2020. Nela pode-se observar uma evidente melhora em relação às imagens apresentadas na Figura 20 no que se refere ao NDBI, uma vez que foram aplicadas em cada uma o algoritmo de classificação de Floresta Aleatória (*smileRandomForest*) fornecido pela ferramenta Earth Engine.

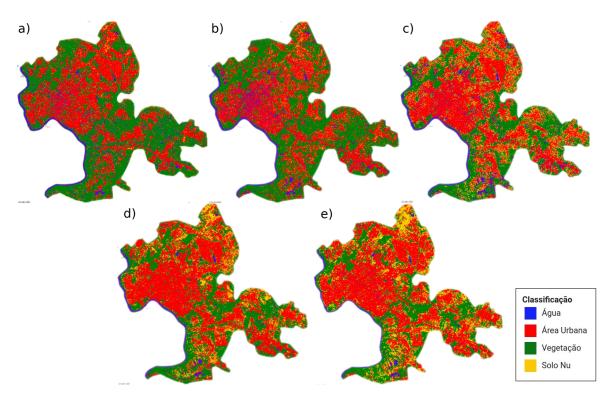


Figura 21: Imagens de Cuiabá classificadas com algoritmo *smileRandomForest* utilizando imagens Landsat 7 [a) 2000, b) 2005 e c) 2010] e Landsat 8 [d) 2015 e e) 2020].

Nas imagens da Figura 21 há uma nítida demonstração da evolução da interferência humana com o recrudescimento das áreas construídas, representada pela cor vermelha, há também um aumento da cobertura de solo nu e consequente diminuição da área de vegetação, de acordo com os dados da Tabela 10. Cabe salientar que por conta da tecnologia dos satélites Landsat 7 e 8 há uma diferença significativa na apresentação e até mesmo na classificação das imagens dos respectivos satélites. Pode-se observar que as imagens de 2015 e 2020, por serem originadas do satélite Landsat 8, estão bem mais caracterizadas. É importante salientar que, devido às características das imagens obtidas, o algoritmo se confunde e classifica alguns pontos da imagens errôneamente, mas nada significante a ponto de interferir na própria classificação já que a acurácia calculada pelo algoritmo chega a 95% em alguns casos, como demonstrado na Tabela 11 o que é bastante significativo.

Tabela 11: Acurácia geral calculada para as imagens dos anos 2000-2020 (Figura 21).

Ano	Acurácia (%)
2000	86,47
2005	89,33
2010	92,85
2015	93,52
2020	95,23

4.4. CARACTERIZAÇÃO DO CAMPO TERMAL URBANO DE CUIABÁ-MT

O fenômeno da Ilha de Calor Urbana pode ser explicado por meio da análise da formação do seu campo termal e o Índice de Variância do Campo Termal Urbano (UTFVI) é o mais amplamente utilizado e mais apropriado para descrever o efeito da UHI com bastante precisão (KAFY, 2019; KAFY et al., 2021; NAIM; KAFY, 2021; TOMLINSON et al., 2011) . Devido aos padrões da formação ilha de calor ter uma relação direta com a cobertura do solo sabe-se que as áreas urbanas tem uma maior

concentração de UTFVI em relação às áreas rurais circundantes (TESFAMARIAM; GOVINDU; UNCHA, 2023).

Os efeitos nocivos causados pelo UTFVI incluem impactos adversos nos padrões dos ventos locais, diminuição da umidade, prejudicando a qualidade do ar, redução no conforto, aumento da taxa de mortalidade e influencia em prejuízos mesmos que indiretos na economia, problemas ambientais, dentre outros fatores (SEJATI; BUCHORI; RUDIARTO, 2019; TESFAMARIAM; GOVINDU; UNCHA, 2023). A prevenção dos efeitos indesejados do UTFVI e a sua identificação podem levar ao conhecimento de possíveis zonas de ondas de calor, o que pode auxiliar em políticas públicas para mitigação da formação de ilhas de calor (NAJAFZADEH et al., 2021).

A Figura 22 mostra o resultado da identificação do Fenômeno UHI em Cuiabá para os períodos dos anos 2000, 2005, 2010, 2015 e 2020. Nos aspectos do índice de avaliação ecológica Fenômeno UHI (vide Tabela 2), a cidade é impactada fortemente pelo campo termal e experimenta três níveis de conforto térmico: áreas de extrema condição de calor (UTFVI > 0,02), áreas com condição considerada muito forte (0,015 < UTFVI < 0,020), áreas com condição nos limites entre 0,010 e 0,015, considerada forte na escala UTFVI. É possível observar, nesta figura, que em 2005 houve uma maior concentração do UTFVI na região central e nordeste da cidade de Cuiabá, na medida que cidade foi se expandindo, no decorrer dos anos, o Fenômeno UHI começou a impactar as regiões mais ao sul e sudeste.

Em 2000, os registros demonstram que 43,56% dos dias (159 dias) tiveram um campo termal no grau forte no que diz respeito ao Fenômeno de UHI, ou com índice de avaliação ecológica ruim, 5,75% dos dias (21) a cidade experimentou um grau muito forte e para a condição extrema foram dois dias, 10/05/2000 e 18/07/2000. Em 2005, foi constatado que em 62,46% dos dias (228) houve uma incidência do grau forte do Fenômeno de UHI, 27,67% dos dias foram afetados pela condição muito forte e 9,04% dos dias deste ano foi influenciado pelo grau extremo.

Para o ano 2010, houve uma redução nos dias com condição forte em relação ao ano de 2005 e foi constatado 41,36% dos dias nesta faixa. 7,94% dos dias

tiveram a condição estabelecida como muito forte e em apenas 3,5% dos dias a cidade experimentou a condição de grau extremo.

31,50% dos dias do ano de 2015 registram uma condição forte na escala do UTFVI, 6,02% para uma condição muito forte e em 4 dias do ano tiveram um grau extremo do UTFVI, estes dias foram 20/04/2015, 21/04/2015, 02/10/2014 e 15/12/2015. Já os graus UTFVI para o ano 2020 tiveram uma aumento em relação ao ano de 2015 e atingiram 56,43% dos dias com grau forte, 16,43% com grau muito forte e 7,12% dos dias com grau extremo na referida escala UTFVI. No geral, a maior parte de Cuiabá tem boas condições ecológicas, neste caso, as áreas com vegetação, as áreas circundades (rurais). No entanto, as áreas afetadas por condições muito fortes e extremas (UTFVI > 0,01) são bem significativas e estão relacionadas com a área central, as áreas residenciais e as áreas de solo descoberto.

Proporcionalmente à variação espacial de NDVI e LST, tipos de uso do solo e UHI, as áreas mais impactadas pela pior condição de calor em 2015 evoluiram significativamente nas regiões norte e sudeste.

Tabela 12: Proporção das áreas afetadas, em km² pela condição do UTFVI em 2000-2020

Ano	0,005 – 0,010 (média)	0,010 – 0,015 (forte)	0,015 – 0,020 (muito forte)	> 0,020 (extrema)	Área Total
2000	44,51	2,77	0,003	-	47,28
2005	41,28	1,19	0,92	0,02	43,41
2010	50,41	30,13	2,53	0,02	83,09
2015	44,59	20,53	1,62	0,01	66,75
2020	39,48	2,47	0,13	-	42,08

A cidade de Cuiabá por estar em uma região com um clima considerado tropical é bastante impactada pela Temperatura da Superfície do Solo e como se pode observar na Tabela 12 o campo termal formado pelas condições média, forte, muito forte e extrema atinge cerca de 18% do tamanho da área urbana da cidade que tem aproximadamente 232 km². Quando se observa a condição forte, é possível perceber que em 2010 e 2015 houve uma elevação drástica no tamanho da área do campo

termal, atingindo 30,13 km² e 20,53 km². O que pode vir a causar danos ecológicos, na saúde e até mesmo na economia local.

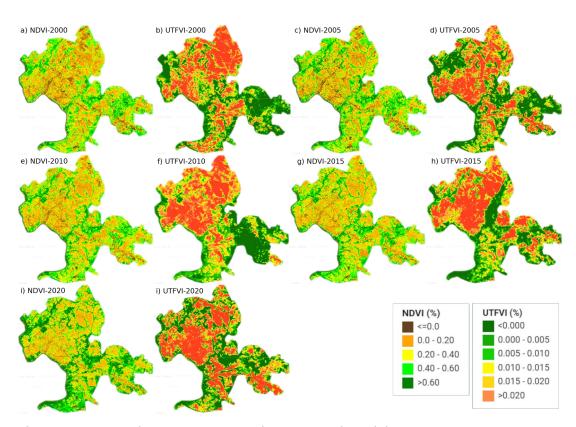


Figura 22: Mapas do UTFVI e respectivos NDVI de Cuiabá em 2000-2020.

4.4.1. Distribuição Do Índice De Variância Do Campo Termal Por Pontos Geográficos De Cuiabá-MT

Na seção anterior foram apresentados os níveis de UTFVI de forma abrangente para a cidade de Cuiabá-MT onde se verifica o tamanho da área afetada pelo Fenômeno UHI. No entanto, é importante uma análise de forma mais específica da situação das ilhas de calor neste cidade quanto aos pontos geográficos estabelecidos neste trabalho. A seguir nas Tabelas 13, 14 e 15 encontram-se a classificação dos níveis de UTFVI e suas respectivas distribuições para os pontos geográficos referenciados na Tabela 4.

Tabela 13: Distribuição do Índice de Variância do Campo Termal Urbano por pontos geográficos para o período 2000-2020.

	Fenômeno UHI Forte							
Local/Ano	2000	2005	2010	2015	2020			
Aecim Tocatins	87,32%	78,04%	77,42%	88,54%	74,30%			
Alphaville	87,32%	78,04%	77,42%	88,54%	74,30%			
Pq. Mãe Bonifácia	87,32%	78,04%	77,42%	88,54%	74,30%			
Centro	87,32%	78,04%	77,42%	88,54%	74,30%			
Coxipó	87,32%	78,04%	77,42%	88,54%	74,30%			
CPA	87,32%	78,04%	77,42%	88,54%	74,30%			
Morada da Serra	94,35%	69,30%	90,78%	85,19%	69,03%			
Pq. das Águas	87,32%	78,04%	77,42%	88,54%	74,30%			
Pq. Tia Nair	87,32%	78,04%	77,42%	88,54%	74,30%			
Tijucal	94,35%	69,30%	90,78%	85,19%	69,03%			
Três Lagoas	94,35%	69,30%	90,78%	85,19%	69,03%			
UFMT	87,32%	78,04%	77,42%	88,54%	74,30%			

Tabela 14: Distribuição do Índice de Variância do Campo Termal Urbano por pontos geográficos para o período 2000-2020.

	Fenômeno UHI Muito Forte							
Local/Ano	2000	2005	2010	2015	2020			
Aecim Tocatins	10,80%	16,59%	15,48%	11,46%	20,09%			
Alphaville	10,80%	16,59%	15,48%	11,46%	20,09%			
Pq. Mãe Bonifácia	10,80%	16,59%	15,48%	11,46%	20,09%			
Centro	10,80%	16,59%	15,48%	11,46%	20,09%			
Coxipó	10,80%	16,59%	15,48%	11,46%	20,09%			
CPA	10,80%	16,59%	15,48%	11,46%	20,09%			
Morada da Serra	5,30%	26,33%	6,45%	8,64%	19,18%			
Pq. das Águas	10,80%	16,59%	15,48%	11,46%	20,09%			
Pq. Tia Nair	10,80%	16,59%	15,48%	11,46%	20,09%			
Tijucal	5,30%	26,33%	6,45%	8,64%	19,18%			
Três Lagoas	5,30%	26,33%	6,45%	8,64%	19,18%			
UFMT	10,80%	16,59%	15,48%	11,46%	20,09%			

Tabela 15: Distribuição do Índice de Variância do Campo Termal Urbano por pontos geográficos para o período 2000-2020.

	Fenômeno UHI Extremo							
Local/Ano	2000	2005	2010	2015	2020			
Aecim Tocatins	1,88%	5,37%	7,10%	•	5,62%			
Alphaville	1,88%	5,37%	7,10%	ı	5,62%			
Pq. Mãe Bonifácia	1,88%	5,37%	7,10%	-	5,62%			
Centro	1,88%	5,37%	7,10%	-	5,62%			
Coxipó	1,88%	5,37%	7,10%	-	5,62%			
CPA	1,88%	5,37%	7,10%	-	5,62%			
Morada da Serra	0,35%	4,37%	2,76%	6,17%	11,79%			
Pq. das Águas	1,88%	5,37%	7,10%	-	5,62%			
Pq. Tia Nair	1,88%	5,37%	7,10%	-	5,62%			
Tijucal	0,35%	4,37%	2,76%	6,17%	11,79%			
Três Lagoas	0,35%	4,37%	2,76%	6,17%	11,79%			
UFMT	1,88%	5,37%	7,10%	_	5,62%			

Mesmo vegetados, alguns pontos geográficos, como os pontos localizados em parques, apresentaram resultados das classificações bastantes ruins. Neste trabalho foram desprezados os níveis de classificações de UTFVI nenhum, fraco e médio por não terem impactos negativos nos aspectos de Índice de Avaliação Ecológica ou Fenômeno UHI (vide Tabela 2). Outros pontos geográficos mesmo que afastados da região central da cidade, onde há uma provável maior incidência de vegetação, sofre um impacto considerável ou mesmo igual às áreas centrais da cidade. Para (FERREIRA; CARRILHO; MENDES, 2015) não basta apenas haver vegetação para se proporcionar uma diminuição do impacto do UTFVI, mas sim uma espécie arbórea indicada para cada região o que contribuiria para uma amenização dos efeitos nocivos de ilhas de calor.

4.5. ANÁLISE E CLASSIFICAÇÃO DO IMPACTO DO ÍNDICE DE VARIÂNCIA DO CAMPO TERMAL URBANO (FENÔMENO UHI)

Sabe-se que o Índice de Variância do Campo Termal Urbano (UTFVI) é amplamente utilizado para determinar os efeitos colaterais de uma Ilha de Calor Urbana (TOMLINSON et al., 2011). Um desses efeitos é o possível impacto causado nos padrões dos ventos locais, então realizar a classificação e análise, utilizando Florestas Randômicas, da influência que o UTFVI tem nos padrões dos ventos pode colaborar para a compreensão do comportamento dos ventos em um área urbanizada. Outro aspecto a ser estudado é o quão importante pode ser os impactos das variáveis microclimáticas na diminuição dos efeitos negativos da formação do UTFVI.

Ao utilizar o algoritmo RF é importante a preparação e adequação do conjunto de dados a ser analisado e classificado, pois a RF são algoritmos de aprendizado de máquina e nesse sentido se faz necessário realizar um treinamento do algoritmo. Para atender a esta necessidade, foi feita uma divisão no conjunto de dados em amostra de treinamento com 70% dos dados e amostra para teste (validação) com 30% dos dados, para a realização da classificação e posterior análise dos valores resultantes.

A partir de dados obtidos do sistema ERA5, foi estabelecido o uso das variáveis microclimáticas Temperatura da Superfície do Solo, Temperatura do Ar (2m), Umidade Relativa do Ar, Direção dos Ventos, Velocidade dos Ventos, a derivação da variável Índice de Variância do Campo Termal Urbano (UTFVI) e o NDVI a partir do LandSat.

Considerando a preparação dos dados levantados para a utilização da classificação das variáveis foram realizados ajustes na base de dados a fim de manter apenas os níveis de UTFVI significativos que são categorizados no aspecto de formação do Fenômeno UHI como (Forte, Muito Forte e Extremo) e no aspecto de índice de avaliação ecológica como ruim, pior e pior caso. Para as constituições das análises, as variáveis dependentes serão definidas de acordo com os objetivos das observações que neste caso são: (i) avaliar as influências ou impactos que as

variáveis microclimáticas Umidade Relativa do Ar, Direção e Velocidade dos Ventos têm sobre o Índice de Variância do Campo Termal Urbano; (ii) observar como o vento (direção e velocidade) se comporta de acordo com os níveis do UTFVI, NDVI, Umidade Relativa do Ar e Temperatura da Superfície do Solo.

Serão feitas também análise com relação à formação do UTFVI em função da Temperatura do Ar a dois metros aplicados aos mesmos contextos citados, ou seja, analisar quais as influências que as variáveis microclimáticas Umidade Relativa do Ar, Direção e Velocidade dos Ventos e o NDVI têm sobre o UTFVI a dois metros. Será feita também uma análise de qual o impacto que a Direção e Velocidade dos Ventos tem sobre as variáveis microclimáticas UTFVI a dois metros, Umidade Relativa do Ar, Temperatura do Ar a dois metros e o índice espectral NDVI.

A Tabela 16 apresenta as medidas descritivas sobre as variáveis microclimáticas (dependente e independentes). Ao todo serão utilizadas 9.162 observações que se referem aos anos 2000, 2005, 2010, 2015 e 2020. Neste trabalho a variável Índice de Variância do Campo Termal Urbano foi categorizada para fins de classificação preditiva quando do uso do algoritmo de aprendizado de máquina supervisionado Floresta Aleatória (*Random Forest*).

Tabela 16: Medidas descritivas das variáveis independentes quantitativas. UTFVI descreve o Fenômeno UHI.

Variável Microclimática	Média	Desv. Padrão	Valor Mín.	Valor Máx.
Temperatura do Solo (°C)	22,36	5,74	7,59	45,06
Temperatura do ar (°C)	24,01	4,27	11,74	37,90
Dir. dos Ventos (Quadrante)	3,03	2,45	0,00	7,00
Velocidade dos Ventos (Km/h)	4,71	3,07	0,15	23,16
Umidade Relativa do Ar (%)	0,49	0,03	0,36	0,54
UTFVI	0,01	0.01	-0,04	0,03
UTFVI a dois metros	0,02	0,01	0,01	0,05

4.5.1. Análise Dos Impactos Das Variáveis Microclimáticas Sobre O Índice De Variância Do Campo Termal Urbano Na Superfície Do Solo

Na etapa de treinamento dos dados, com o uso de 70% dos dados amostrais, foram obtidos os resultados demonstrados na Tabela 17 que representam os resultados da matriz de confusão em relação à variável Fenômeno UHI, gerada a partir da aplicação do algoritmo de Floresta Aleatória com um número de 500 árvores, para os dados amostrais referentes ao ano 2000. Nesta matriz os valores preditivos do modelo estão evidenciados em cinza e demonstram dados amostrais de treino e teste, onde serão avaliados a performance da classificação do modelo.

Com o uso de matriz de confusão pode-se avaliar o desempenho de modelos em especial na classificação dos resultados fornecidos pela Floresta Aleatória. Por ser basilar, a matriz de confusão fornece uma visão geral do desempenho do modelo proposto, onde se destaca a precisão, nuances de erros e acertos.

Do universo dos dados amostrais de treino que totaliza 1917 registros, 99,42% das classificações foram identificadas como corretas. Observando os resultados dos treinos aplicados na validação da amostra, seção Teste na Tabela 17, têm-se que para os dados amostrais de teste a taxa de classificações corretas foi de 99,18%, valor similar ao da taxa de treinamento. Ambos os valores Kappa dos dados amostrais para treino e teste sugerem um alto grau de concordância nas avaliações, respectivamente 0,9789 e 0,9482, sendo considerado uma concordância perfeita, ou seja, entre 82% e 100% dos dados são confiáveis, de acordo com (MCHUGH, 2012).

Tabela 17: Matriz de Confusão e estatísticas da variável Fenômeno UHI para os dados amostrais do ano 2000.

	Treino			Teste		
	Extremo	Forte	Muito Forte	Extremo	Forte	Muito Forte
Extremo	31	0	0	5	0	0
Forte	2	1703	9	1	772	6
Muito Forte	0	0	172	0	0	65
Acurácia:	0,9943	•		0,9918	•	
Kappa:	0,9789			0,9482		

Nas classificações realizadas nos dados amostrais de treino apresentadas pela matriz de confusão na Tabela 17, são evidenciados uma taxa de acerto para a

0,7364

predição do Fenômeno UHI Extremo de 93,93%, 100% como taxa de acerto para a predição do Fenômeno UHI Forte e aproximadamente 95,03% com taxa de acerto para a o Fenômeno UHI Muito Forte. Com relação aos dos amostrais de validação (teste) obteve-se um total de 842 registros classificados corretamnente, uma acurácia de 99,18%, sendo que a taxa de acerto para as predições do Fenômeno UHI Forte foi em torno de 100%, aproximadamente 83,33% dos dados de validação referentes ao Fenômeno UHI Extremo foram preditos corretamente e uma taxa de predição de 91,54% para valores referentes ao Fenômeno UHI Muito Forte.

Tabela 18: Estatísticas da variável Fenômeno UHI para os dados amostrais de treinamento e teste para os anos de 2000-2020.

uemamento	e teste para i	os anos de 20	00-2020.								
	Treinamento										
	2000 2005 2010 2015 2020										
Acurácia	0,9953	0,9000	0,9567	0,9895	0,9114						
Kappa	0,9763	0,7188	0,8561	0,9479	0,7222						
		•	•	•	•						
			Teste								
	2000 2005 2010 2015 2020										
Acurácia	0,9929	0,8785	0,9443	0,9927	0,8888						

0,8221

0,8855

Na Tabela 18, a qual se refere aos dados amostrais para todos os períodos, o grau de concordância, dado pelo índice Kappa, é considerado com um grau substancial (acima de 70%) na confiança para o treinamento dos dados, de acordo com os níveis de concordância apresentados por McHugh (MCHUGH, 2012). No Apêndice há uma complementação das matrizes de confusão e suas respectivas análises estatísticas em relação à confiabilidade dos dados para os demais anos, no caso 2005, 2010, 2015 e 2020.

0,6487

0,9559

Kappa

4.5.1.1. Indicadores De Robustez Do Modelo Do Fenômeno UHI Na Superficie Do Solo

A seguir são apresentadas, na Tabela 19, as estatísticas dos indicadores de robustez (sensibilidade, especificidade e acurária balanceada) dos modelos para as amostras de treino e teste do modelo. Estes indicadores manifestam resultados

insuficientes quanto a robustez na identificação correta das variáveis que correspondem aos Fenômenos UHI Extremo nos anos 2005, 2010 e 2020 com valores abaixo de 55%, resultados suficientes para Forte com todos acima de 98% e o Fenômeno UHI Muito Forte também apresentou uma taxa acima de 60%. Observando o balanceamento, no geral, há resultados suficientes com todos acima de 68%.

No quesito sensibilidade, os resultados alcançados são satisfatórios para a identificação de variáveis quantificadas e previstas corretamente como níveis Forte com mais de 99% na média, Muito Forte ficando numa média de 79% e Extremo em média 62%.

Tabela 19: Indicadores preditivos referentes aos dados amostrais de teste para os anos 2000-2020.

	Indicadores do Teste								
Extremo	2000	2005	2010	2015	2020				
Sensibilidade	0,8333	0,3611	0,5352	1,0000	0,3864				
Especificidade	1,0000	1,0000	1,0000	1,0000	1,0000				
Precisão	1,0000	1,0000	1,0000	1,0000	1,0000				
F1	0,9091	0,5306	0,6972	1,0000	0,5574				
Acurácia Balanceada	0,9167	0,6805	0,7676	1,0000	0,6932				
Forte	2000	2005	2010	2015	2020				
Sensibilidade	1,0000	0,9833	1,0000	1,0000	0,9977				
Especificidade	0,9221	0,5829	0,7352	0,9364	0,5907				
Precisão	0,9923	0,8727	0,9341	0,9918	0,8719				
F1	0,9961	0,9247	0,9659	0,9959	0,9306				
Acurácia Balanceada	0,9610	0,7831	0,8676	0,9682	0,7942				
Muito Forte	2000	2005	2010	2015	2020				
Sensibilidade	0,9296	0,6270	0,8311	0,9340	0,6608				
Especificidade	1,0000	0,9814	1,0000	1,0000	0,9959				
Precisão	1,0000	0,8967	1,0000	1,0000	0,9741				
F1	0,9635	0,7380	0,9077	0,9659	0,7875				
Acurácia Balanceada	0,9648	0,8042	0,9155	0,9670	0,8283				

Quanto a especificidade, que mede quão exata é a atribuição à classe positiva, neste caso um atribuição rotulada de uma Fenômeno UHI que se encontram nos níveis Extremo com 100%, Forte com uma média de 75,34% e Muito Forte com aproximadamente 100%, conforme a Tabela 19 em relação ao períodos estudados. A

taxa de 100% foi atribuída às classes Extremo e Muito Forte, pois o modelo restringe a composição da matriz de confusão com variáveis significativas. A acurácia balanceada, uma medida obtida pela média da sensibilidade e especificidade, é uma métrica que calcula o ajuste destes indicadores para determinar a capacidade preditiva dos resultados, é desejado que os resultados fiquem próximo de 1 (100%). A acurácia mínima calculada no modelo foi obtida na classe Extremo com um valor próximo de 81%.

Analisando as estatísticas resultantes do algoritmo que gera a matriz de confusão aplicados aos dados amostrais de testagem e validação referentes aos anos 2005, 2010, 2015 e 2020 ficam evidenciados que a proporção dos casos positivos identificados corretamente, foram satisfatórios para o Fenômeno UHI Forte, acima de 99%, em todos os anos estudados. O indicador sensibilidade evidencia que para o Fenômeno UHI Extremo a identificação positiva desta classe ficou abaixo de 55% nos períodos de 2000, 2005 e 2010, uma taxa abaixo do razoável na testagem dos dados. Este indicador postula o quão apto o modelo é para detectar eventos na classe positiva, ou seja, os casos positivos que foram identificados corretamente.

Quanto a Acurácia Balanceada, a mesma revelou resultados substanciais para para a identificação do Fenômeno UHI Extremo e Forte, com valores percentuais acima de 81%. Quanto ao Fenômeno UHI Muito Forte a acurácia balanceada se estabelece com média acima de 80% para a validação dos dados amostrais dos anos 2000, 2005, 2010, 2015 e 2020, conforme a Tabela 19.

Como forma de representar graficamente a aplicação do algoritmo para a classificação dos dados, temos na Figura 23 a modelagem da árvore resultante após o treinamento dos dados utilizando o algoritmo Floresta Aleatória (*Random Forest*) e nela pode-se observar os 5 níveis de profundidade com os seus respectivos nós terminais em uma classificação dos dados utilizando a variável dependente UTFVI que foi redefinida como uma variável discreta e categórica e as variáveis independentes Umidade Relativa, Velocidade e Direção dos Ventos.

Esta árvore apresenta o primeiro nó RH, definido em uma escolha da melhor variável aleatória por meio de *bootstrap*, de onde se origina os demais nós. Os quais são derivados a partir do corte que delimita a Umidade Relativa do Ar abaixo de

0,5022 (50,22%) e acima deste valor, um cálculo de votação que ocorre pela média da árvore de decisão. Esta divisão seleciona os valores da previsão mais votado como o resultado final.

Nesta árvore, podemos observar uma maior incidência dos valores do Fenômeno UHI Forte relacionados à variável dependente UTFVI na superfície do solo. Aproximadamente 97% da frequência se refere ao Fenômeno UHI Forte e o restante é atribuído ao Fenômeno UHI Muito Forte. Isto se dá em função da variação dos comportamentos das variáveis independentes (Umidade Relativa do Ar, Velocidade dos Ventos e Direção dos Ventos). Exemplificando, a árvore acima citada expressa que quando a Velocidade dos Ventos está entre 6,49 km/h e 6,54 km/h e a Umidade Relativa do Ar com limites abaixo de 50,22% haverá ocorrências do Fenômeno UHI Muito Forte.

Dentro dos acontecimentos dos Fenômenos UHI, importa destacar a decorrência dos eventos para o Fenômeno UHI Extremo que se sobressai quando a variável Umidade Relativa do Ar está acima de 50,22%. De acordo com os registros, a possibilidade de ocorrência, nas condições evidenciadas na árvore resultante, aumena em 1,08% em relação as demais classes de Fenômeno UHI.

Na seção de Apêndices há a apresentação da Floresta Aleatória para os anos 2005, 2010, 2015 e 2020.

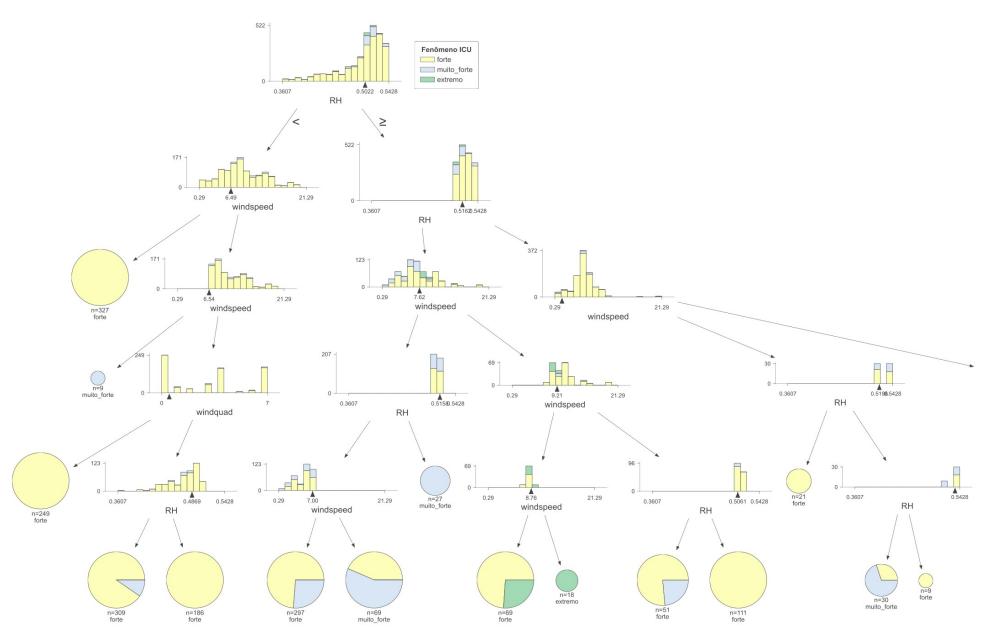


Figura 23: Floresta Randômica, com suas árvores de decisão, resultante para amostra do treinamento dos dados do ano 2000 com profundidade 6. As folhas (círculos) resultantes demonstram a classificação do Fenômeno UHI (Forte, Muito Forte e Extremo) para UTFVI na superfície do solo.

4.5.1.2. Caracterizando A Importância Das Variáveis Microclimáticas Na Floresta Aleatória

A importância das variáveis é dada pelo ajuste do resultado calculado como a média e o desvio padrão de acumulação da diminuição de impurezas dentro de cada árvore. Analisando o impacto das variáveis RH (Umidade Relativa do Ar), windspeed (Velocidade dos Ventos) e windquad (Direção dos Ventos) na variável UTFVI pode-se notar o protagonismo das variávels RH e windspeed. As variáveis RH e windspeed foram classificadas como as variáveis mais importantes para explicar o modelo. De acordo com a Figura 24 a variável RH tem 47,13% de importância relativa e a variável windspeed com 38,10% de importância relativa para o modelo de aprendizado de máquina para os dados referentes o ano 2000, no ano de 2005 tiveram respectivamente 53,82% e 37,21%, para o ano de 2010 a variável RH manteve-se como sendo a mais importância do que a variável windspeed sendo que a primeira ficou com 43,86% e esta última ficou com 39,93%.

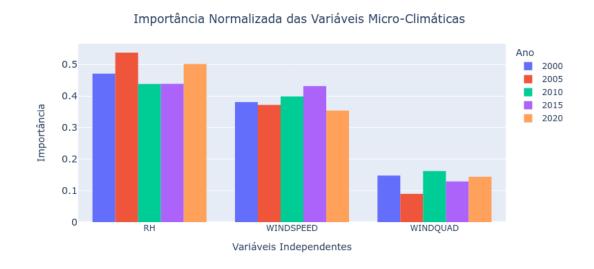


Figura 24: Classificação da importância das variáveis independentes para UTFVI na superfície do solo referentes aos dados amostrais dos anos 2000-2020.

No anos 2015 e 2020, a variável RH se tornou novamente a mais importante delas ficando com 43,89% para o ano 2015 e com 50,19% para o ano 2020. Já a

variável Direção dos Ventos (windquad) começou a demonstrar uma certa importância, apesar de baixa, 16% para o ano 2010, 12% para o ano 2015 e para o ano 2020 estabelecida em 14%, em 2005 houve uma baixa participação, 8% apenas. Adverte-se que para um bom modelo a confiabilidade destas importâncias depende de quão preciso é este modelo, assim quanto mais preciso o modelo mais confiáveis serão as importâncias calculadas.

Além de se avaliar a importância das variáveis independentes normalizadas é salutar destacar que a média do decréscimo da acurácia expressa quanta precisão o modelo irá perder quando da exclusão de cada variável. Neste caso, quanto mais a precisão é prejudicada, mais importante é a variável para o sucesso da classificação. Como pode-se observar nos gráficos anteriores, as variáveis são apresentadas em ordem decrescente de importância. Já a média do decréscimo no coeficiente Gini é uma medida de como cada variável contribui para a homogeneidade dos nós e folhas na floresta aleatória resultante. Então, quanto maior o valor da diminuição média da precisão ou da diminuição média do escore de Gini, maior será a importância da variável no modelo (MARTINEZ-TABOADA; REDONDO, 2020).

Nas Figuras 25 e 26 são apresentadas a maior média de decréscimo da precisão cujo objetivo é calcular a perda de precisão das previsões do modelo quando os valores de um variável são permutados entre as instâncias e a maior média de decréscimo de índice Gini que descreve o aumento de homegeneidade nos nós filhos (BÉNARD; DA VEIGA; SCORNET, 2022). O índice Gini considera as variáveis mais importantes aquelas com o maior valor da média de decréscimo e mede o ganho médio de pureza por divisões de uma variável e tem com finalidade calcular a probabilidade de um variável específica ser classificada incorretamente quando selecionada aleatoriamente, neste caso a variável Fenômeno UHI.

Ano 2000 2005 2010 2010 2020 100 RH WINDSPEED WINDQUAD

Média do Decréscimo da Acurácia

Figura 25: Médias do descréscimo da acurácia e índice Gini referentes ao período 2000-2020 para UTFVI em relação à Temperatura da Superfície do Solo.

Variáveis Micro-Climáticas

Nestas figuras estão evidenciadas que as variáveis com a maior participação para o decréscimo da acurácia do Fenômeno UHI no modelo são a Umidade Relativa do Ar (RH) e a Velocidade dos Ventos (windspeed), o mesmo se observa para o modelo da média do decréscimo Gini. Portanto, estas variáveis são as mais influentes para que se ocorra o Fenômeno UHI, sendo que em relação ao ano de 2015 pode-se destacar que a Velocidade dos Ventos e Umidade Relativa do Ar são bastantes parecidas nas suas importâncias, sendo esta última um pouco mais relevante.



Figura 26: Médias do descréscimo do índice Gini referentes ao período 2000-2020 para UTFVI em relação à Temperatura da Superfície do Solo.

4.5.2. Análise Dos Impactos Das Varíaveis Microclimáticas Sobre O Índice De Variância Do Campo Termal Urbano A 2 Metros De Altura

Nesta seção foram estudados dados amostrais com relação à UTFVI em função da Temperatura do Ar, a fim de estudar qual o impacto sofrido por esta variável pelos comportamentos dos ventos e a Umidade Relativa do Ar. E para tal análise foram aplicadas as mesmas técnicas para se analisar o UTFVI ao nível da superfície do solo. Ou seja, na etapa de treinamento dos dados, foram reservados 70% dos dados amostrais para uso no algoritmo das Floresta Aleatória, assim foram obtidos os resultados demonstrados na Tabela 20 que representam os resultados da matriz de confusão em relação à variável Fenômeno UHI a 2 metros, gerada a partir de uma Floresta Aleatória com um número de 500 árvores, para os dados amostrais referentes ao ano 2000.

Nesta matriz os valores preditivos do modelo também estão evidenciados em cinza e demonstram dados amostrais de treino e teste, onde são revelados valores sobre a performance da classificação do modelo. Dos 1593 registros que totalizam os dados amostrais de treino, 67,73% das classificações foram identificadas como corretas. Observando os resultados dos treinos aplicados na validação da amostra, seção Teste na Tabela 20, têm-se que para os dados amostrais de teste a taxa de classificações corretas foi de 92,54%. Ambos os valores Kappa dos dados amostrais para treino e teste sugerem um grau substancial de concordância nas avaliações, respectivamente 0,9681 e 0,8477, indicando uma boa confiabilidade para os dados de treinamento ou seja, entre 84% e 96% dos dados são confiáveis, de acordo com (MCHUGH, 2012).

Tabela 20: Matriz de Confusão e estatísticas da variável Fenômeno UHI a 2 metros de altura para os dados amostrais do ano 2000.

	Treino			Teste		
	Extremo	Forte	Muito Forte	Extremo	Forte	Muito Forte
Extremo	168	0	0	66	0	0
Forte	0	714	1	3	321	2
Muito Forte	0	0	230	0	0	88
Acurácia:	0,9991			0,9896	•	
Kappa:	0,9983			0,9788		

Os resultados apresentados pela matriz de confusão, no momento do treinamento, apontam que aproximadamente 98% dos dados relacionados ao Fenômeno UHI Extremo foram classificados corretamente, de um total de 161 registros 2 foram classificados como falsos negativos, ou seja, instâncias errôneamente preditas como negativas, neste caso 1 registro classificado como sendo Fenômeno UHI Forte e o mesmo número para Fenômeno UHI Muito Forte e 3 registros foram classificados como falsos positivos, ou preditos errôneamente como positvos. Nota-se que os demais 932 registros foram corretamente classificados como negativos para o Fenômeno UHI Extremo.

Para a categoria do Fenômeno UHI Forte, 706 registros (99,57%) do total de dados para este fenômeno foram classificados corretamente, 11 registros foram classificados como falsos positivos e 3 registros foram classificados como falsos negativos. Na categoria do Fenômeno Muito Forte 94,69% (214) de um total de 226 registros foram classificados corretamente, sendo que aproximadamente 1,32% foram classificados como falsos positivos (3 registros) e aproximadamente 6% foram classificados como falsos negativos (12 registros).

Quando da validação dos dados, o modelo retorna um aprendizado com base no treinamento do algoritmo que 84% dos registros da categoria Fenômeno UHI Extremo, 97,23% dos registros (317) da categoria Fenômeno UHI Forte e aproximadamente 83% (79) dos registros da categoria Fenômeno UHI Muito Forte foram classificados corretamente. No geral, temos uma acurácia de 92,54%. Detalhamento sobre a confiabilidade dos dados amostrais dos anos 2005, 2010, 2015

e 2020 com suas respectivas matrizes de confusão e análises estatísticas são apresentados no Apêndice.

Na Tabela 21, a qual se refere aos dados amostrais para ocorrência do Fenômeno UHI a 2 metros de altura de todos os períodos, o grau de concordância, dado pelo índice Kappa, é considerado uma confiação substancial (acima de 92%), de acordo com os níveis de concordância apresentados por McHugh (MCHUGH, 2012). No Apêndice há uma complementação das matrizes de confusão e suas respectivas análises estatísticas em relação à confiabilidade dos dados para os demais anos, no caso 2005, 2010, 2015 e 2020.

Tabela 21: Estatísticas da variável Fenômeno UHI para os dados amostrais de treinamento e teste para o Fenômeno UHI a 2 metros de altura dos anos de 2000-2020.

	Treinamento								
	2000	2005	2010	2015	2020				
Acurácia	0,9991	0,9975	0,9681	0,9948	0,9816				
Карра	0,9983	0,9950	0,9462	0,9908	0,9669				
			Teste						
	2000	2005	2010	2015	2020				
Acurácia	0,9896	0,9883	0,9549	0,9881	0,9510				
Карра	0,9788	0,9775	0,9211	0,9773	0,9112				

4.5.2.1. Indicadores De Robustez Do Modelo Do Fenômeno UHI A 2 Metros De Altura

Os indicadores de robustez dos modelos analisados apresentados na Tabela 22, a qual se referem aos dados dos anos 2000, 2005, 2010, 2015 e 2020, demonstram que as estatísticas resultantes para as amostras de teste do modelo revelam resultados suficientes para a sensibilidade o que expressa uma corretude na identificação das variáveis que correspondem aos Fenômenos UHI Extremo com valores médios acima de 95% e Forte com 99% de acerto e para ao Fenômeno UHI Muito Forte 97,14% dos dados foram classificados corretamente.

Quanto à especificidade, os resultados atingidos indicam que, na média, 98,48% da identificação das variáveis não se encontram nos níveis Extremo do Fenômeno UHI, 96,66% não se encontra no nível Forte e 99,86% não se encontra no

nível Muito Forte do Fenômeno UHI. Mensurando a força preditiva dos resultados observa-se que, conforme a Tabela 22, há uma acurácia balanceada mínima de 92,86% em relação aos dados amostrais de testagem e validação para o nível Muito Forte do Fenômeno UHI.

Tabela 22: Indicadores preditivos referentes aos dados amostrais de treino para os anos 2000-2020.

	In	dicadores do	Teste		
Extremo	2000	2005	2010	2015	2020
Sensibilidade	0,9565	1,0000	1,0000	1,0000	0,8469
Especificidade	1,0000	0,9931	0,9492	0,9928	0,9890
Precisão	1,0000	0,9625	0,8689	0,9908	0,9432
F1	0,9778	0,9809	0,9299	0,9831	0,8925
Acurácia Balanceada	0,9783	0,9966	0,9746	0,9964	0,9180
Forte	2000	2005	2010	2015	2020
Sensibilidade	1,0000	0,9910	0,9709	0,9908	0,9818
Especificidade	0,9686	0,9835	0,9832	0,9833	0,9144
Precisão	0,9847	0,9910	0,9877	0,9908	0,9444
F1	0,9923	0,9910	0,9792	0,9908	0,9627
Acurácia Balanceada	0,9843	0,9872	0,9771	0,9871	0,9481
Muito Forte	2000	2005	2010	2015	2020
Sensibilidade	0,9778	0,9714	0,8305	0,9677	0,9516
Especificidade	1,0000	1,0000	1,0000	1,0000	0,9930
Precisão	1,0000	1,0000	1,0000	1,0000	0,9752
F1	0,9888	0,9855	0,9074	0,9836	0,9633
Acurácia Balanceada	0,9889	0,9857	0,9153	0,9839	0,9723

Nota-se que os resultados obtidos, de acordo com a Tabela 22, para os indicadores de robustez que se referem aos períodos estudados, quais sejam 2005, 2010, 2015 e 2020, apresentam um comportamento bastante similar aos dados referenciados do ano 2000. A sensibilidade calculada para estes dados apontam que os níveis Extremo estão entre 84% e 100% e Forte em todos os anos estão na faixa de 96%, idem para o nível Muito Forte que também está na faixa de aproximadamente 98%. No caso do indicador especificidade os valores correspondentes são de aproximadamente 98%, em média para o nível Extremo, 96% para nível Forte e de 99% para o nível Muito Forte. Aplicando-se a média entre a sensibilidade e especificidade foi obtido uma acurácia balanceada mínima de 96%.

Um comportamento similar aos resultados dos indicadores para os dados amostrais se observa nos dados de treinamento para os referidos anos de 2005, 2010, 2015 e 2020. O índice de sensibilidade dos dados para os Fenômenos UHI Extremo e Forte se estabilizam na faixa de 98% e 100%, respectivamente. Para a classe Muito Forte os resultados ficam na faixa de 85%, aproximadamente. A acurácia balanceada média para estes períodos se refere ao ano de 2020 com um valor calculado de 95%.

Na Figura 27 é demonstrada a árvore que representa os resultados da votação após o treinamento dos dados utilizando o algoritmo Floresta Aleatória (*Random Forest*) com configuração para apresentar 5 níveis de profundidade com os seus respectivos nós terminais em uma classificação dos dados utilizando a variável dependente UTFVI a 2 metros de altura (UTFVI2m) sendo redefinida como uma variável discreta e categórica e as variáveis independentes umidade relativa, velocidade e Direção dos Ventos.

Esta árvore apresenta o primeiro nó identificado como windspeed, definido em uma escolha, da melhor variável, aleatória por meio de *bootstrap*, de onde se origina os demais nós, os quais são derivados a partir do corte que delimita a Velocidade dos Ventos abaixo de 4,20 km/h e acima deste valor, um cálculo de votação que ocorre pela média da árvore de decisão. Esta divisão seleciona os valores da previsão mais votado no resultado final.

Nesta árvore, podemos observar que a maior incidência dos valores do Fenômeno UHI Forte relacionados à variável dependente. Aproximadamente 64% da frequência se refere ao Fenômeno UHI Forte e o restante são atribuídos aos Fenômeno UHI Extremo e Muito Forte. Ao avaliar os eventos do Fenômeno UHI Extremo, na árvore resultante, o mesmo só acontece quando a velocidade do vento está abaixo de 4,2 km/h, denotando que há 51,19% de probabilidade de se aumentar os casos para este fenômeno, nesta condição.

Na seção de Apêndices há a apresentação da Floresta Aleatória para o UTFVI a 2 metros de altura dos anos 2005, 2010, 2015 e 2020.

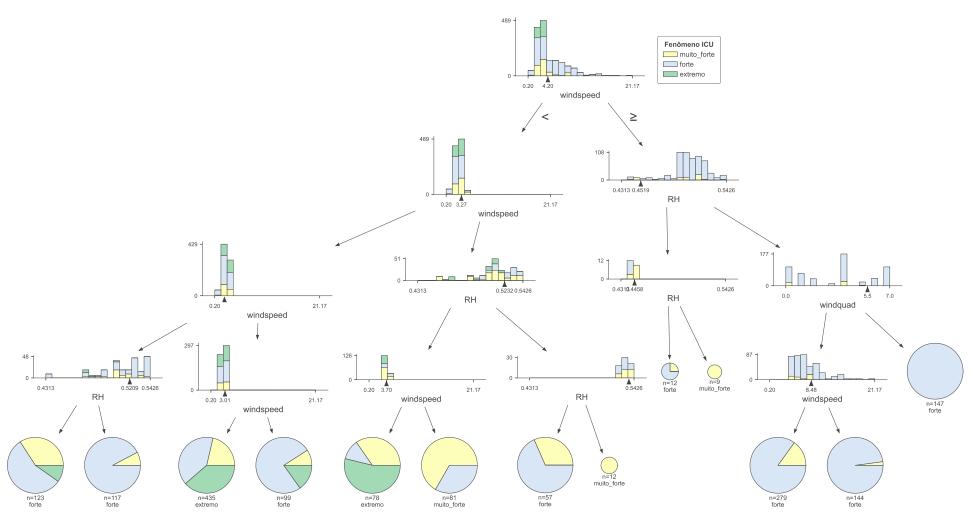


Figura 27: Floresta Randômica, com suas árvores de decisão, resultante para amostra do treinamento dos dados do ano 2000 com profundidade 5. As folhas (círculos) resultantes demonstram a classificação do Fenômeno UHI (Forte, Muito Forte e Extremo) para a UTFVI a 2 metros de altura.

4.5.2.2. Caracterizando A Importância Das Variáveis Microclimáticas Na Floresta Aleatória

Na Figura 28 são apresentadas o quão importante é uma variável no contexto da classificação do Índice de Variância do Campo Termal Urbano a 2 metros de altura. Nelas, pode-se observar que a variável windspeed é a mais importante para todos os períodos dos dados amostrais (2000-2020) e tem 54,78% de importância relativa e a variável RH com 35,58% de importância relativa para o modelo dos dados referentes o ano 2000, no ano de 2005 tiveram respectivamente 52,21% e 35,19%, para o ano de 2010 a variável windspeed manteve-se como sendo a mais importante ficando com 61,75% e a variável RH com 27,72%. No anos 2015 e 2020, a variável windspeed se manteve novamente como sendo a mais importante delas ficando com 50,42% para o ano 2015 e com 62,26% para o ano 2020. Já a variável Direção dos Ventos (windquad) ficou com valores abaixo de 12% no geral.

Novamente, assim como realizado para a análise da formação do UTFVI em relação à superfície do solo, a determinação da importância das variáveis independentes normalizadas é de grande auxílio para se destacar qual das variáveis tem uma maior influência na variável dependente em relação à formação do UTFVI a 2 metros. Na seção de Apêndices há a apresentação da Floresta Aleatória para os anos 2000-2020 referente à Fenômeno UHI a 2 metros de altura.

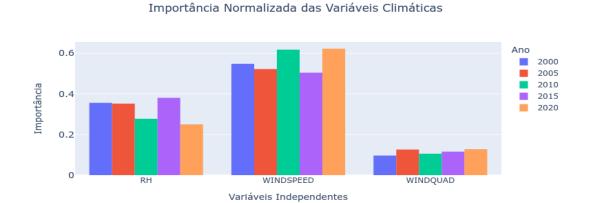


Figura 28: Classificação da importância das variáveis independentes para UTFVI a 2 metros de altura referentes aos dados amostrais do ano 2000-2020.

Nas Figuras 29 e 30 são apresentados a maiore média de Decréscimo da Acurácia (*MeanDecreaseAccuracy*) e a maior média de decréscimo de Índice *Gini* (*MeanDecreaseGini*) onde se envidencia o cálculo da perda de precisão das previsões do modelo quando os valores de um variável são permutados entre as instâncias.

Estas figuras suscintam que as variáveis com a maior participação para o decréscimo da acurácia do Fenômeno UHI no modelo *MeanDecreaseAccuracy* são a Velocidade dos Ventos (windspeed) e a Umidade Relativa do Ar (RH), idem para o modelo *MeanDecreaseGini*. Portanto, estas variáveis são as mais influentes para que se ocorra o Fenômeno UHI, sendo que em relação ao ano de 2005 pode-se destacar que a Velocidade dos Ventos em se tratando da média do decréscimo da acurácia tem mais importância em relação à Umidade Relativa do Ar e a Direção dos Ventos, sendo que para média do decréscimo do índice Gini a variável mais importante é a Umidade Relativa do Ar.

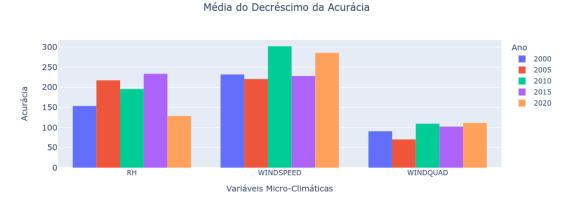


Figura 29: Médias do descréscimo da acurácia referentes ao período de 2000-2020 para UTFVI em relação à Temperatura do Ar (2m).

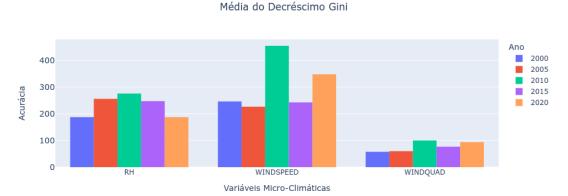


Figura 30: Médias do descréscimo do Índice Gini referentes ao período de 2000-2020 para UTFVI em relação à Temperatura do Ar (2m).

4.5.3. Análise Sintética Dos Impactos Das Variáveis Microclimáticas Sobre O Índice De Variância Do Campo Termal Urbano

Nas seções 4.5.1 e 4.5.2 foram analisados os impactos sobre o Índice de Variância do Campo Termal Urbano especificamente para os anos 2000, 2005, 2010, 2015 e 2020 sendo contextualizado os comportamentos das variáveis microclimáticas Umidade Relativa do Ar, Velocidade e Direção dos Ventos sobre este índice. Nesta seção será apresentada uma síntese dos dados destes períodos para formar um único período tanto em relação à UTFVI na superfície do solo quanto a 2 metros de altura a fim de observar de maneira mais ampla como se deu a influência do UTFVI em relação aos padrões dos ventos locais.

Nesta síntese, os períodos foram concatenados e as classes do UTFVI Nenhum, Fraco e Médio redenominadas como Normal e as classes Forte, Muito Forte e Extremo renomeadas como Anormal, o que nos permite obter uma visão geral sobre os Fenômenos UHI com uma distinção em apenas duas classes.

Na Figura 31 há uma demonstração parcial da árvore obtida da Floresta Aleatória calculada com uma profundidade 6 e 500 árvores, utilizando o critério Gini

que representa a classificação do Fenômeno IHU na superfície do solo. Nela, é apresentada uma predição para a classe Anormal dada uma Umidade Relativa do Ar em 40%, a Direção dos Ventos para o sudeste (quadrante 4) e sua velocidade acima de 9 km/h. Nestas condições, o UTFVI certamente será Anormal.

A Figura 32 externa uma visualização parcial da árvore obtida e nela observa-se a apresentação da classe Anormal (síntese das classes Forte, Muito Forte e Extremo) da uma condição de Velocidade dos Ventos a 3.91 km/h e Umidade Relativa do Ar abaixo de 39% independentemente da Direção dos Ventos. E de acordo com a Figura 33 as variáveis que mais influenciam no Fenômeno UHI são RH e windspeed, com respectivos 49,67% e 49,60% em relação à superfície do solorespectivamente. Quanto ao Fenômeno UHI a 2 metros de altura as variáveis mais importantes são windspeed e RH, com respectivos valores percentuais 49,81% e 49,37%, como se pode observar na Figura 34.

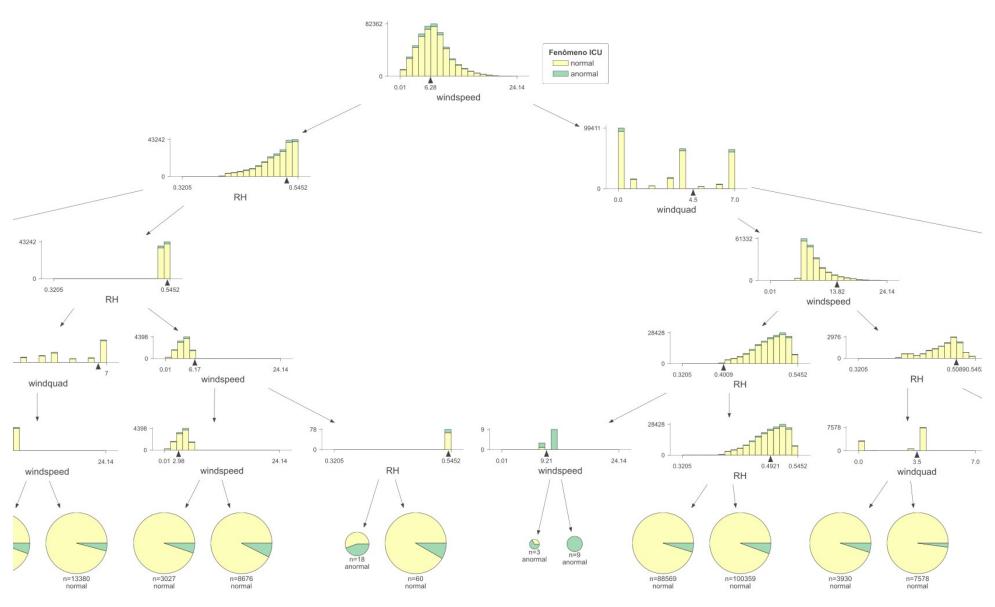


Figura 31: Visão parcial da Floresta Aleatória para o Fenômeno UHI na superfície do solo da síntese de todos os períodos entre 2000-2020). As categorias normal e anormal se referem aos níveis UTFVI onde anormal é a união dos níveis Extremo, Forte e Muito Forte e a categoria normal é a união dos níveis nenhum, fraco e médio.

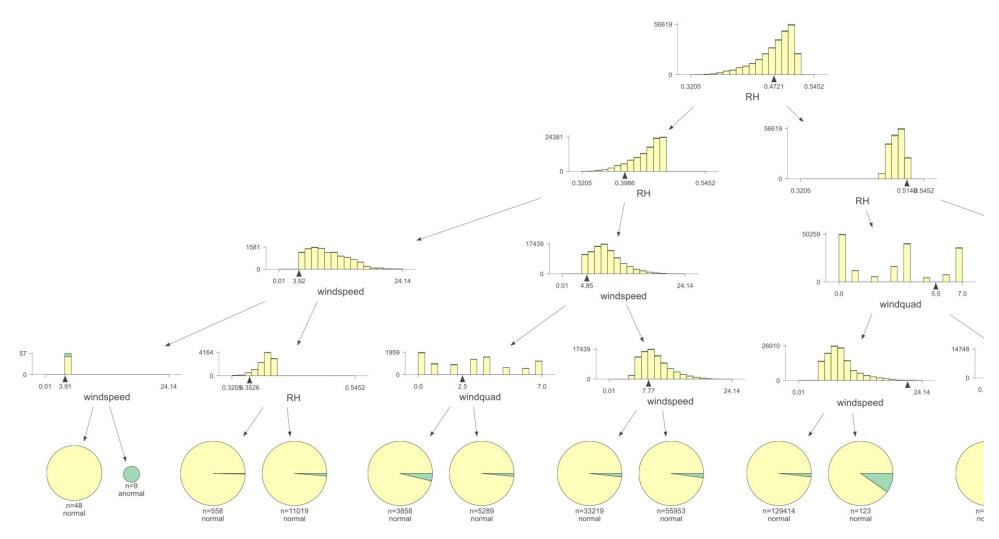


Figura 32: Visão parcial da Floresta Aleatória para o Fenômeno UHI a 2 metros de altura da síntese de todos os períodos (2000-2020). As categorias normal e anormal se referem aos níveis UTFVI onde anormal é a união dos níveis Extremo, Forte e Muito Forte e a categoria normal é a união dos níveis nenhum, fraço e médio.

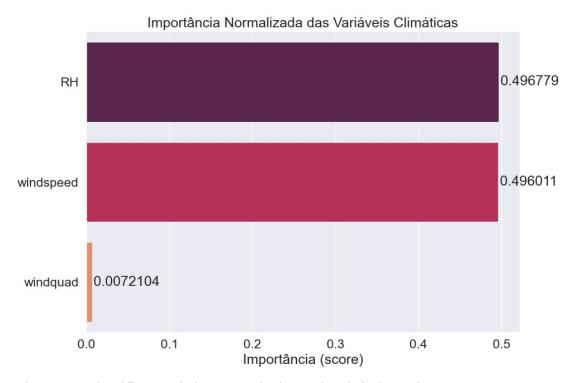


Figura 33: Classificação da importância das variáveis independentes para UTFVI na superfície do solo referente à síntese dos dados amostrais dos períodos 2000-2020.

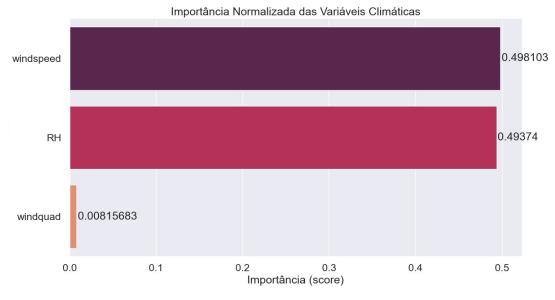


Figura 34: Classificação da importância das variáveis independentes para UTFVI a 2 metros de altura referente à síntese dos dados amostrais dos períodos 2000-2020.

4.5.4. Análise Dos Impactos Das Variáveis Microclimáticas Sobre O Índice De Variância Do Campo Termal Urbano Em Função Do NDVI

Com o objetivo de compreender qual a influência do NDVI, em relação ao UTFVI, em conjunto com a demais variáveis microclimáticas Umidade Relativa do AR, Velocidade dos Ventos e Direção dos Ventos, esta seção apresenta os resultados da classificação dos dados amostrais com o uso do algoritmo de Floresta Randômica no contexto do Fenômeno UHI na superfície do solo em relação ao NDVI.

(ULLAH et al., 2022) define, por meio da correlação de Pearson, que o NDVI tem uma relação negativa e insignificante com o Índice de Temperatura da Superfície do Solo. Desta forma, analisar a relação do NDVI com o UTFVI pode ser um bom indicador no desenvolvimento urbano sustentável, pois este índice tem uma relação direta com o respectivo índice LST.

A matriz de confusão apresentada na Tabela 23 demonstra os resultados em relação à variável Fenômeno UHI em função das variáveis microclimáticas Direção dos Ventos, Velocidade dos Ventos, Umidade Relativa do Ar e o NDVI para os dados amostrais do ano 2000. De acordo com os resultados desta Tabela, o treinamento determina uma acurácia de 99,24% nas classificações e quando aplicadas à validação dos dados (Teste) e obtido uma taxa de classificações corretas de 98,21%. Os índices Kappa aplicados aos dados amostrais para treinamento e validação sugerem um alto grau de concordância nas avaliações, entre 82% e 100%, neste caso, 90,88% e 96,21%.

Tabela 23: Matriz de Confusão e estatísticias da variável Fenômeno UHI para os dados amostrais do ano 2000.

		0	Teste			
	Extremo	Forte Muito Forte Ex		Extremo	Forte	Muito Forte
Extremo	26	2	0	10	0	0
Forte	2	1274	7	1	543	7
Muito Forte	0	2	129	0	3	51
Acurácia:	0,9924			0,9821		
Kappa:	0,9621			0,9088		

A matriz de confusão, apresentada na Tabela 23, demonstram que no momento do treinamento os resultados para o Fenômeno UHI Extremo indicam 26 registros classificados como verdadeiros positivos, ou seja, o valor real e o valor previsto são iguais. Dos 28 registros para o Fenômeno UHI Extremo 2 são apontados ou como falsos positivos (referindo-se a uma amostra pertencente à classe negativa, mas sendo classificada erroneamente como pertencente à classe positiva) ou falsos negativos (neste caso, referindo-se a uma amostra pertencente à classe positiva, mas sendo classificada erroneamente como pertencente à classe negativa). Quanto ao Fenômeno UHI Forte, que contém um total de 1276 registros, há 9 registros considerados falsos negativos e 4 considerados falsos positivos.

Após o treinamento do algoritmo da Floresta Aleatória, com um tamanho de 500 árvores e profundidade de nível 5, foi realizado a validação que nada mais é o uso de hiperparâmetros adequados e assim poder gerar resultados substanciais através da escolha da melhor variável para compor o nó raiz. Como pode-se observar na seção Teste da Tabela 23 que com apenas 30% dos dados amostrais o desempenho foi bastante similar aos resultados do treinamento, com 98,21% de acertos e uma confiança substancial de 90,88%.

Tabela 24: Estatísticas da variável Fenômeno UHI para os dados amostrais de treinamento e teste para o Fenômeno UHI na superfície do solo dos anos de 2000-2020

2000-2020.									
	Treinamento								
	2000	2005	2010	2015	2020				
Acurácia	0,9924	0,9737	0,9926	0,9887	0,9816				
Kappa	0,9621	0,9351	0,9783	0,9469	0,9578				
			•	•					
	Teste								
	2000	2005	2010	2015	2020				
Acurácia	0,9821	0,8995	0,9828	0,9796	0,9176				

0.7416

Kappa

0.9088

Na Tabela 24, a qual se refere aos dados amostrais para ocorrência do Fenômeno UHI na superfície do solo de todos os períodos, o grau de concordância obtido pelo índice Kappa, é considerado substancial (acima de 80%) para a confiabilidade da informação, de acordo com os níveis de concordância apresentados

0.9482

0.9011

0.8031

em (MCHUGH, 2012), com exceção para o ano de 2015 que é considerado moderado, determinando que de 35 a 63% dos dados são confiáveis. No Apêndice é apresentado as demais matrizes de confusão e suas respectivas análises estatísticas em relação à confiabilidade dos dados para os demais anos, no caso 2005, 2010, 2015 e 2020.

4.5.4.1. Indicadores De Robustez Do Modelo Para O Fenômeno UHI Na Superfície Do Solo Em Função Do NDVI E Demais Variáveis MicroClimáticas

Os indicadores de robustez dos modelos analisados e apresentados na Tabela 25, revelam que as estatísticas das amostras de teste do modelo são suficientes para a sensibilidade, o que expressa uma corretude na identificação das variáveis que correspondem aos Fenômenos UHI Forte com valores médios acima de 96% e Muito Forte com média acima de 73% de acerto e para os dados amostrais do ano 2015 e 2020 temos uma exceção em relação ao Fenômeno UHI Extremo que demonstrou uma precisão explicando que entre 62% e 64% são verdadeiros positivos classificados corretamente.

Tabela 25: Indicadores preditivos dos dados amostrais do período 2000-2020.

Indicadores do Teste Extremo 2000 2005 2010 2015 2020 Sensibilidade 0.9375 0,9285 0,6406 0,9090 0,6250 Especificidade 1,0000 0,9947 1,0000 1,0000 0,9871 Precisão 1,0000 0,8490 1,0000 0,9285 0,8118 F1 0,9523 0,7200 0,9677 1,0000 0,7161 Acurácia Balanceada 0,9545 0,8098 0,9687 0,9642 0,8138 2000 2005 2010 2015 2020 Forte Sensibilidade 0,9931 0,9790 0,9945 0,9673 0,9986 0,7978 Especificidade 0,8841 0,7436 0,9206 0,8814 0,9232 Precisão 0,9855 0,9109 0,9801 0,9839 0,9900 0,9383 0,9893 0,9885 0,9503 0,9393 0,8555 0,9596 0,9372 0,8884 Acurácia Balanceada 2020 **Muito Forte** 2000 2005 2010 2015 Sensibilidade 0,8793 0,7339 0,9120 0,9286 0,8133 0,9843 Especificidade 0,9946 0,9660 0,9988 0,9931 Precisão 0,8618 0,9913 0,9310 0,9444 0,9455 0,8682 0,9107 0,7927 0,9500 0,9369 Acurácia Balanceada 0,9369 0.8499 0,9554 0,9608 0,8988 A árvore relacionada aos dados amostrais, Figura 35, estruturada com um quantidade de 500 árvores e profundidade de 6 níveis, apresenta os seus respectivos nós terminais em uma classificação dos dados utilizando a variável dependente UTFVI, derivada da Temperatura da Superfície do Solo, sendo redefinida como uma variável discreta e categórica e as variáveis independentes Umidade Relativa do Ar, Velocidade e Direção dos Ventos Locais e NDVI.

Tem-se nesta árvore a apresentação do primeiro nó identificado como a variável microclimática RH (Umidade Relativa do Ar), definido em uma escolha, da melhor variável, aleatória por meio de bootstrap, de onde se origina os demais nós, os quais são derivados a partir do corte que delimita a umidade relativa o valor abaixo de 0,5022 e acima deste valor, um cálculo de votação que ocorre pela média da árvore de decisão.

Nesta árvore, podemos observar que a maior incidência dos valores do Fenômeno UHI Forte relacionados à variável dependente. Aproximadamente 91% da frequência se refere ao Fenômeno UHI Forte e o restante são atribuídos aos Fenômeno UHI Extremo e Muito Forte. Frisa-se o resultado onde se demonstra o nó terminal para o Fenômeno Extremo que determina que para esta classificação aconteça deve haver uma Umidade Relativa do Ar entre 50% e 51% e Velocidade dos Ventos entre 7,82 km/h e 9,21 km/h, independente das demais variáveis microclimáticas. Na seção de Apêndices há a apresentação da Floresta Aleatória para o UTFVI calculado em relação à Temperatura da Superfície do Solo para os períodos 2005-2020.

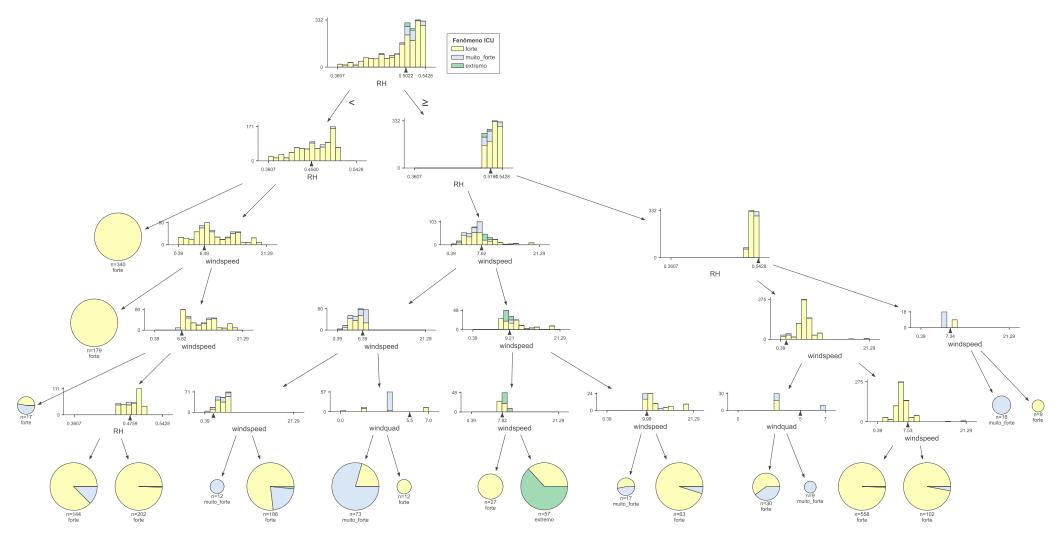


Figura 35: Floresta Randômica, com suas árvores de decisão, resultante para amostra do treinamento dos dados do ano 2000 com profundidade 5. As folhas (círculos) resultantes demonstram a classificação do Fenômeno UHI (Forte, Muito Forte e Extremo) em função das variáveis microclimáticas Umidade Relativa do Ar, NDVI, Velocidade e Direção dos Ventos locais para UTFVI em relação à superfície do solo.

4.5.4.2. Caracterizando A Importância Das Variáveis Climáticas Na Floresta Aleatória

Muito estudos descrevem que o NDVI é preponderante na atenuação da formação de ilhas de calor, neste estudo a classificação do Fenômeno UHI demonstra que, conforme a Figura 36, o Índice por Diferença Normalizada da Vegetação não contribui para a formação do mesmo, o que contribui para afirmar que o NDVI tem influência negativa com relação a formação do Fenômeno UHI. Nesta Figura, a Umidade do Relativa do Ar e a Velocidade dos Ventos são de maior importância, em especial a Umidade do Ar.

A distribuição indica que a Umidade do Ar tem um importância acima dos 35% formação dos Fenômeno UHI Forte, Muito Forte e Extremo. O NDVI contribui com uma média de apenas 11% e a Direção dos Ventos com média de 12%, se equivalendo na pouca influência no Índice de Variação do Campo Termal Urbano.

Importância Normalizada das Variáveis Climáticas

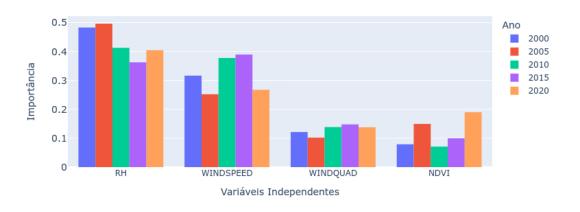


Figura 36: Classificação da importância das variáveis independentes para UTFVI em relação à superfície do solo referentes aos dados amostrais do período de 2000-2020.

Para verificar a precisão do modelo a Figura 37 externa a ordem da importância das variáveis e enseja que se as variáveis RH (Umidade Relativa do Ar), windspeed (Velocidade dos Ventos), windquad (Direção dos Ventos) e NDVI forem retiradas do modelo causarão imprecisão do mesmo. Quanto a medida da

importância baseada na impureza, a Figura 38 revela a contribuição de cada variável para homogeneidade dos nós e folhas na Floresta Aleatória resultante (Figura 35).

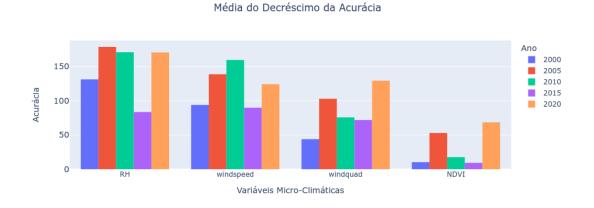


Figura 37: Médias do descréscimo da acurácia referentes ao período 2000-2020 para UTFVI em relação à Temperatura da Superfície do Solo.

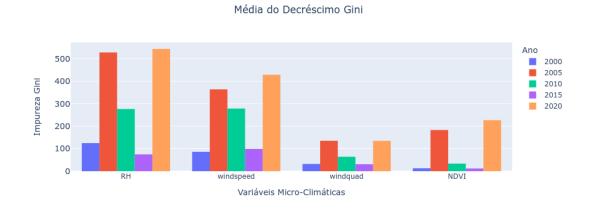


Figura 38: Médias do descréscimo Gini referentes ao período 2000-2020 para UTFVI em relação à Temperatura da Superfície do Solo.

É demonstrado na Figura 38 que a variável NDVI não causará grande impacto na homegeneidade do modelo caso seja removida, o mesmo acontecerá para a variável windquad (Direção dos Ventos) em especial para os dados amostrais dos anos 2000 e 2015.

4.5.5. Análise Dos Impactos Das Variáveis Microclimáticas Sobre O Índice De Variância Do Campo Termal Urbano A 2 Metros De Altura Em Função Do NDVI.

Na seção anterior foram estudados os efeitos do NDVI e as variáveis microclimáticas no UTFVI derivado da Temperatura da Superfície do Solo e nesta serão estudados a repercussão do NDVI e em consórcio com as demais variáveis microclimáticas (Umidade Relativa do Ar, Velocidade e Direção dos Ventos) na formação dos Fenômenos UHI Forte, Muito Forte e Extremo. Na Tabela 26 expressa-se os resultados da matriz de confusão em função da variável Fenômeno UHI a 2 metros (UTFVI2m) e respectivas variáveis microclimáticas Direção e Velocidade dos ventos locais, Umidade Relativa do Ar e o índice espectral NDVI.

A matriz de confusão delimita uma acurácia de 99,44% para o treinamento e 96% para a validação dos dados amostrais pela Floresta Aleatória, o que remete à uma taxa de acerto da predição com uma performance bastante alta e um alto grau de concordância nas avaliações dos dados amostrais. Percebe-se que há 1 registro detectado no momento do treinamento como falso-negativo, ou seja, este registro foi classificado erroneamente como pertencente à classe negativa (Fenômeno UHI Muito Forte) e que há 3 registros nesta mesma categoria classificado como falsos negativos.

Tabela 26: Matriz de Confusão e estatísticias da variável Fenômeno UHI a 2 metros de altura para os dados amostrais do 2000.

	Treino			Teste			
	Extremo	Forte	Muito Forte	Extremo	Forte	Muito Forte	
Extremo	15	0	1	6	0	3	
Forte	0	140	0	0	60	0	
Muito Forte	0	0	22	0	0	6	
Acurácia:	0,9944			0,9600	•		
Kappa:	0,9843			0,8826			

Os índices estatísticos visualizados na Tabela 27 indicam uma alta precisão da predição dos dados e uma confiança, no geral, substancial.

Tabela 27: Estatísticas da variável Fenômeno UHI a 2 metros de altura para os dados amostrais de treinamento e teste dos anos de 2000-2020.

dados amos	mais uc mema	annemo e teste	dos anos de	2000-2020.				
	Treinamento							
	2000	2005	2010	2015	2020			
Acurácia	0,9944	0,9654	0,9802	1,0000	0,9873			
Kappa	0,9843	0,9078	0,9624	1,0000	0,9734			
			•		•			
	Teste							
	2000 2005 2010 2015 2020							
Acurácia	0,9600	0,9104	0,9813	1,0000	0,9704			
Kappa	0,8826	0,7469	0,9543	1,0000	0,9346			

4.5.5.1. Indicadores De Robustez Do Modelo Para O Fenômeno UHI A 2 Metros De Altura Em Função Do NDVI E Demais Variáveis MicroClimáticas

A robustez dos indicadores preditivos para os modelos identificados na Tabela 28, expõe que as estatísticas das amostras de afirmação do treinamento dos dados são suficientes para todos os indicadores estatísticos, determinando assim uma corretude na identificação das variáveis que correspondem aos Fenômenos UHI Extremo e Forte. No que tange a Sensibilidade (taxa de valores reais previstos e relação ao número total de valores reais) para o Fenômeno UHI Muito Forte, os dados do ano 2000 explicita uma baixa correspondência na suficiência, porém no decorrer dos demais períodos os valores se reafirmaram, sendo suficientes.

Tabela 28: Indicadores preditivos referentes à validação (Teste) dos dados amostrais para o período 2000-2020.

Indicadores do Teste						
Extremo	2000	2005	2010	2015	2020	
Sensibilidade	1,0000	1,0000	1,0000	1,0000	1,0000	
Especificidade	0,9565	1,0000	1,0000	1,0000	1,0000	
Precisão	0,6667	1,0000	1,0000	1,0000	1,0000	
F1	0,8000	1,0000	1,0000	1,0000	1,0000	
Acurácia Balanceada	0,9783	1,0000	1,0000	1,0000	1,0000	
Forte	2000	2005	2010	2015	2020	
Sensibilidade	1,0000	0,9515	1,0000	1,0000	1,0000	
Especificidade	1,0000	0,7742	0,9298	1,0000	0,9024	
Precisão	1,0000	0,9333	0,9752	1,0000	0,9592	
F1	1,0000	0,9423	0,9874	1,0000	0,9792	
Acurácia Balanceada	1,0000	0,8629	0,9649	1,0000	0,9512	
Muito Forte	2000	2005	2010	2015	2020	
Sensibilidade	0,6667	0,7586	0,7777	1,0000	0,8261	
Especificidade	1,0000	0,9524	1,0000	1,0000	1,0000	
Precisão	1,0000	0,8148	1,0000	1,0000	1,0000	
F1	0,8000	0,7857	0,8750	1,0000	0,9048	
Acurácia Balanceada	0,8334	0,8555	0,8888	1,0000	0,9130	

A árvore resultante, exposta na Figura 42, expressa a classificação dos dados utilizando a variável dependente UTFVI a 2 metros de altura e as variáveis independentes Umidade Relativa do Ar, Velocidade e Direção dos Ventos Locais e NDVI.

Nesta árvore, o primeiro nó é identificado como a variável microclimática windspeed (Velocidade dos Ventos), classificada como a mais importante consoante a Figura 39, de onde se origina os demais nós. A árvore, a partir desta escolha, delimita os valores preditivos para a velocidade do vento onde a classificação determina os nós terminais a partir de valores acima e abaixo de 1,46 km/h (uma média).

Percebe-se que quando a velocidade do vento está abaixo de 1,46 km/h e independente dos níveis do NDVI haverá ocorrências restritas do Fenômeno UHI Extremo. Já numa combinação de vento com velocidade mais alta e umidade do ar acima de 51,29% resulta na formação do Fenômeno UHI Forte e quando a umidade do ar está abaixo de 51,29% haverá a formação do Fenômeno UHI Muito Forte.

4.5.5.2. Caracterizando A Importância Das Variáveis Microclimáticas Na Floresta Aleatória

A importância das variáveis microclimáticas, no tocante à classificação do Fenômeno UHI a 2 metros é visualizada na Figura 39 a qual demonstra que a variável Velocidade dos Ventos (windspeed) tem uma importância definida entre aproximadamente 40% e 56%. Direção dos Ventos e NDVI não tem tanta importância quando se trata de predição do Fenômeno UHI em Ilha de Calor Urbana.

Importância Normalizada das Variáveis Climáticas em Função do UTFVI a 2 metros

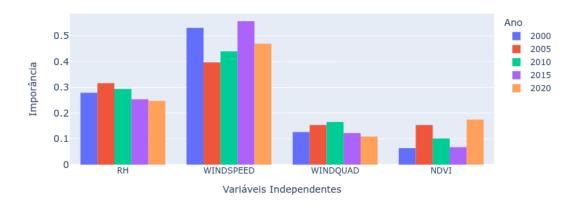


Figura 39: Classificação da importância das variáveis independentes para UTFVI a 2 metros de altura resultantes da classificação dos dados amostrais do período de 2000-2020.

É nítido que a manutenção da precisão do modelo recai sobre variável microclimática Velocidade dos Ventos (windspeed) no contexto do Fenômeno UHI em função do NDVI e demais variáveis. Nota-se também que a variável NDVI não influencia na formação do UTFVI a 2 metros de altura, dada a baixa Média do Decréscimo da Acurácia, conforme é apresentado na Figura 40. A média do Decréscimo da Acurária para a variável NDVI ficou estabelecida entre aproximados - 3 e 15.



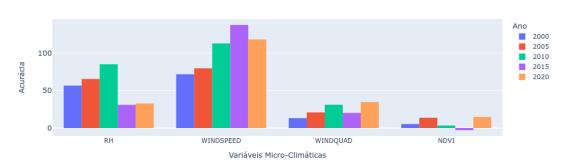


Figura 40: Médias do descréscimo da acurácia referentes ao período 2000-2020 para UTFVI a 2 metros de altura.

Quanto à Média do Decréscimo Gini, que mensura a homogeneidade dos dados, a Figura 41 externa a importância das variáveis microclimáticas Velocidade dos Ventos e a Umidade Relativa do Ar.

Média do Decréscimo Gini (UTFVI a 2 metros) Ano 2000 2015 2010 2015 2020 Variávels Micro-Climáticas

Figura 41: Médias do descréscimo da acurácia e índice Gini referentes aos anos 2000-2020 para UTFVI em relação da Temperatura do Ar (2 metros).

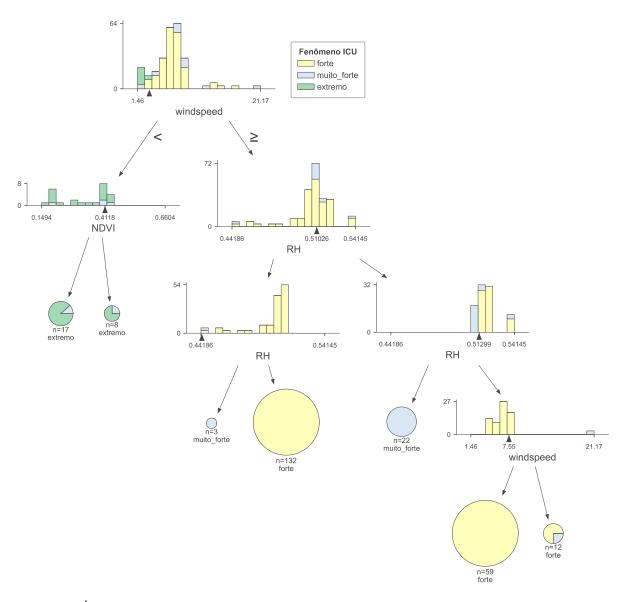


Figura 42: Árvores de decisão, resultante para amostra do treinamento dos dados do ano 2000 com profundidade de 5 níveis. As folhas resultantes demonstram a classificação do Fenômeno UHI a 2 metros de altura (Forte, Muito Forte e Extremo) em função das variáveis microclimáticas Umidade Relativa do Ar, NDVI, Velocidade e Direção dos Ventos locais.

4.6. DISCUSSÃO

A aplicação, neste estudo, da técnica de classificação de dados utilizando Floresta Aleatória proporcionou resultados analíticos que permitiu observar como se deu a evolução da expansão do fenômeno de ilhas de calor na cidade de Cuiabá.

4.6.1. Expansão Da Ilha De Calor Em Cuiabá

Inicialmente, com o uso da ferramenta *Google Earth Engine*, foram obtidas imagens que demonstraram os níveis atingidos pela Temperatura da Superfície do Solo, o índice de vegetação por diferença normalizada (NDVI) e também o índice de área construída por diferença normalizada (NDBI) e dessa forma pôde-se obter uma análise visual da expansão das áreas construídas e a consequente ocupação desordenada do solo, como pode ser observado na Figura 20. Nesta imagem, é possível observar que, proporcionalmente, a área de vegetação diminui na medida que se aumenta a área construída, isto não necessariamente implica em construções prediais mas pode ter ocorrido a diminuição da área vegetada pela ação humana, como desmatamento e queimadas.

Mas, de modo geral, na Figura 20 houve apenas representação dos índices espectrais LST, NDVI e NDBI que são obtidos através de cálculos que analisa o espectro da vegetação, albedo dos materiais e também do solo. Considerando apenas os limites dos índices espectrais já foi possível visualizar a dimensão das áreas de vegetação e das áreas construídas e na finalidade de confirmar e até complementar os resultados destes índices aplicou-se o algoritmo de classificação de Floresta Aleatória (smileRandomForest) que demonstrou uma clara e evidente diferença na distinção de áreas vegetadas das áreas construídas, inclusive corpos d'água e solo nu.

Na imagem da Figura 21 observou-se uma nítida diferença quando da aplicação do algoritmo de classificação resultando em uma acurácia de 86,47%, o que a caracteriza como satisfatória na classificação da cobertura do solo para a imagem do ano 2000 onde foi utilizada o satélite Landsat 7. Percebe-se na imagem

uma tímida representação da cobertura de solo nu, em contraste a área vegetada e a área urbana estão bem caracterizadas. Por outro lado, com o uso das imagens do satélite Landsat 8 já foi possível haver uma melhor classificação, através do algoritmo, onde notou-se uma evolução na apresentação da imagem relacionado ao ano 2020, com resultado da acurácia ficando em torno de 95%, o que é bem representativo.

Em resumo, quando obtidos bons dados e imagens de satélites e quando bem treinado o algoritmo de Floresta Aleatória apresenta respostas quase perfeitas na classificação da cobertura do solo.

4.6.2. Análise Estátistica Da Classificação Do Fenômeno UHI

Na Tabela 29 são exibidos a avaliação estatística onde é demonstrado o desempenho obtido na classificação dos dados em relação ao UTFVI na Superfície do Solo. Nesta tabela, o coeficiente de determinação (R²) revela que 97% da variável categorizada derivada de UTFVI é explicada pelas variáveis independentes (RH, windspeed e windquad) para o treinamento e 96% para a validação dos dados amostrais na classificação das variáveis referentes ao ano 2000. Em 2005, o coeficiente de determinação foi estabelecido em 86% para o treinamento dos dados e 82% para a testagem dos dados.

Para 2010, R² foi de 92% para o treinamento e 94% como resultado da classificação na validação dos dados. 2015 o resultado obtidos evidenciam que os coeficientes da classificação para treinamento e teste dos dados ficaram em 97%. Na Tabela 29 são exibidos ainda, para a classificação dos dados do ano de 2020, os resultados do coeficiente de determinação em 86% para o treinamento e 85% quando da testagem dos dados, indicando o quanto a variação observada pode ser explicada pelos dados do modelo.

Apesar da avaliação estatística revelar que o desempenho do coeficiente de determinação para a classificação dos dados dos anos 2005 e 2020 ficarem abaixo em relação aos demais períodos avaliados, no geral os resultados acima 80% do R²

indicam que mais variabilidade é explicada pelo modelo. Quanto a análise da diferença absoluta entre os valores previstos pelo modelo e os valores observados expressada pelo erro médio absoluto (MAE), têm-se que o acerto das predições são mais acentuados para o ano 2000 e 2015, para os demais anos o MAE indica um certo afastamento das previsões do modelo real, em média.

Para os modelos relacionados à UTFVI a 2 metros de altura, percebe-se, conforme a Tabela 30, que apenas para o ano 2000 houve uma predição para a validação dos dados amostrais que se aproxima muito da realidade do modelo onde o Erro Médio Absoluto está em 0,02. Nos demais anos os erros médios absolutos tendem a erros na previsão e demonstram uma certo distanciamento das previsões dos modelos reais.

Tabela 29: Avaliação estatística sumarizada da classificação dos dados para os anos 2000-2020 em relação à UTFVI a 2 metros de altura.

Treino							
Indicador	2000	2005	2010	2015	2020		
MAE	0,17	0,19	0,28	0,28	0,24		
MSE	0,18	0,33	0,41	0,48	0,40		
RMSE	0,42	0,57	0,64	0,69	0,63		
R ²	0,84	0,89	0,78	0,81	0,84		
	Teste						
Indicador	2000	2005	2010	2015	2020		
MAE	0,02	0,23	0,33	0,31	0,32		
MSE	0,03	0,42	0,47	0,54	0,54		
RMSE	0,18	0,65	0,69	0,73	0,74		
R ²	0,78	0,87	0,75	0,81	0,79		

Ao avaliar o modelo com a inclusão do índice espectral NDVI na classificação do treinamento dos dados amostrais em relação ao UTFVI na superfície do solo temos os resultados indicando que a maioria das variáveis explicativas (RH, windspeed, windquad e NDVI) são estatisticamente significativas, com o coeficiente de determinação (R²) acima de 0,80 tanto para o treinamento quanto para a avaliação dos dados amostrais, o que implica em dizer que a relação entre as variáveis é considerada forte, conforme podemos observar na Tabela 31.

Tabela 30: Avaliação estatística sumarizada da classificação dos dados para os anos 2000-2020 em relação à UTFVI a 2 metros de altura.

Treino						
Indicador	2000	2005	2010	2015	2020	
MAE	0,17	0,19	0,28	0,28	0,24	
MSE	0,18	0,33	0,41	0,48	0,40	
RMSE	0,42	0,57	0,64	0,69	0,63	
R ²	0,84	0,89	0,78	0,81	0,84	
		Te	este			
Indicador	2000	2005	2010	2015	2020	
MAE	0,02	0,23	0,33	0,31	0,32	
MSE	0,03	0,42	0,47	0,54	0,54	
RMSE	0,18	0,65	0,69	0,73	0,74	
R ²	0,78	0,87	0,75	0,81	0,79	

Equiparadamente, o modelo que utiliza o NDVI na classificação tem um desempenho semelhante ao modelo classificado sem o NDVI como se pode observar nos valores dos seus respectivos Erros Médios Absolutos. A predição para a validação dos dados amostrais para ambos os modelos, nos períodos de 2000 e 2015, se aproximam muito da veracidade do modelo, com o MAE abaixo de 0,03.

Tabela 31: Avaliação estatística sumarizada da classificação dos dados para os anos 2000-2020 em relação à UTFVI na Superfície do Solo em função do NDVI e demais variáveis microclimáticas.

Treino							
Indicador	2000	2005	2010	2015	2020		
MAE	0,03	0,26	0,10	0,00	0,19		
MSE	0,39	0,48	0,16	0,00	0,31		
RMSE	0,62	0,70	0,41	0,03	0,55		
R ²	0,99	0,85	0,93	1,00	0,86		
		Te	este				
Indicador 2000 2005 2010 2015 2020							
MAE	0,02	0,36	0,11	0,02	0,20		
MSE	0,02	0,67	0,17	0,03	0,32		
RMSE	0,13	0,82	0,41	0,17	0,56		
R ²	0,98	0,80	0,92	0,98	0,86		

A sumarização da avaliação estatística apresentada na Tabela 32 expressa a magnitude absoluta média (Erro Médido Absoluto) entre os valores reais e os valores

previstos pelo modelo de classificação do treinamento dos dados amostrais como sendo perfeito, pois o mesmo é zero. Os resultados para R² diz que proporção da informação dos dados explicada pelo modelo é de 100% para o treinamento dos dados. Na validação dos dados (Teste), os valores para R² são bem próximos dos obtidos no treinamento. Muito embora, para os anos 2005 e 2020, na aplicação do algoritmo Floresta Aleatória para validação dos dados os Erros Médios Absolutos tendem a não se aproximar da realidade do modelo, ou seja a eficácia da predição do modelo não se confirma.

Tabela 32: Avaliação estatística sumarizada da classificação dos dados para os anos 2000-2020 em relação à UTFVI a 2 metros de altura em função do NDVI e demais variáveis microclimáticas

e demais varia	aveis illicioci	imaticas.			
		Tre	eino		
Indicador	2000	2005	2010	2015	2020
MAE	0,00	0,00	0,00	0,00	0,00
MSE	0,00	0,00	0,00	0,00	0,00
RMSE	0,00	0,00	0,00	0,00	0.06
R ²	1,00	1,00	1,00	1,00	1,00
		Te	este		
Indicador	2000	2005	2010	2015	2020
MAE	0,05	0,12	0,00	0,00	0,04
MSE	0,05	0,24	0,00	0,00	0,04
RMSE	0,23	0,49	0,07	0,00	0,21
R ²	0,95	0,94	1,00	1,00	0,96

De modo geral, o Erro Quadrático Médio (MSE) para os modelos classificados com base no UTFVI em relação à superfície do solo e a 2 metros de altura, sem a influencia do NDVI, não é impactante. Quando o modelo é submetido à influencia do NDVI, no contexto do UTFVI na superfície do solo, a avaliação da diferença quadrática média entre os valores observados e previstos torna-se importante. E quando se trata da avaliação do modelo segundo o UTFVI a 2 metros de altura, sob a influencia do NDVI, no geral é insignificante. Neste caso, o modelo não apresenta erros.

4.6.3. Análise Da Correlação Entre UTFVI Na Superfície Do Solo E A 2 Metros

As informações da Temperatura do Ar e da superfície do solo foram provinientes do Era5 e são referentes aos 12 pontos geográficos listados na Tabela 4 aos quais foram disponibilizadas medições a cada 1 hora, conforme mostram as Figuras 43, Erro: Origem da referência não encontrada, 44, 45 e 46. Nos períodos estudados, a maior parte dos valores estimados para a Temperatura da Superfície do Solo são superiores à tempetura do ar, exceto entre os meses de julho e agosto que em algumas medições as tempeturas do ar ficaram com medições acima dos valores estimados para a Temperatura da Superfície do Solo.

Já em relação ao Índice de Variância do Campo Termal Urbano, os cálculos apontam que este índice a 2 metros de altura é na maioria da suas medições superior ao seu equivalente da superfície do solo, em todos os períodos observados especificamente para as categorias dos Fenômenso UHI Forte, Muito Forte e Extremo. Isto implica em deduzir que: (i) a Temperatura do Ar tem maior impacto na área urbana, em relação ao Fenômeno UHI, no Índice de Variância do Campo Termal Urbano que a Temperatura da Superfície do Solo; (ii) a amplitude da média da Temperatura do Ar para a região estudada (Cuiabá) é superior a amplitude da média da Temperatura da Superfície do Solo da região em estudo.

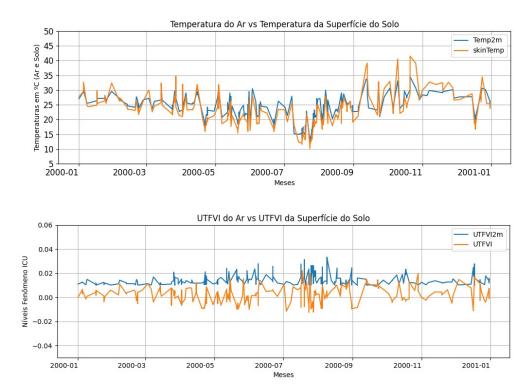
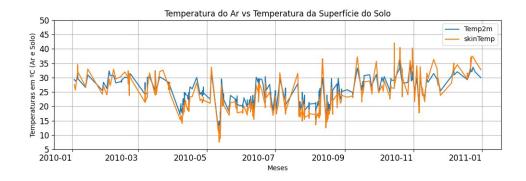


Figura 43: Séries das Temperaturas diárias da superfície do solo e das temperaturas do ar (2 metros) e UTFVI da superfície do solo e a 2 metros para ano 2000.



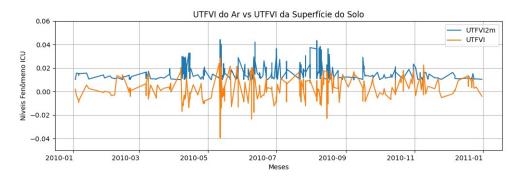
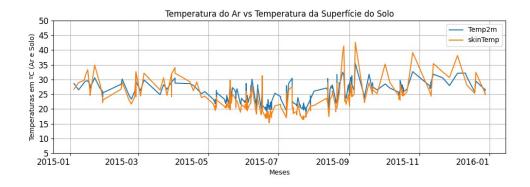


Figura 44: Séries das Temperaturas diárias da superfície do solo e das temperaturas do ar (2 metros) e UTFVI da superfície do solo e a 2 metros para ano 2010.



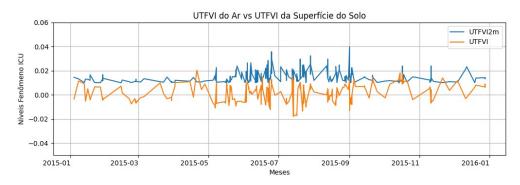


Figura 45: Séries das Temperaturas diárias da superfície do solo e das temperaturas do ar (2 metros) e UTFVI da superfície do solo e a 2 metros para ano 2015.

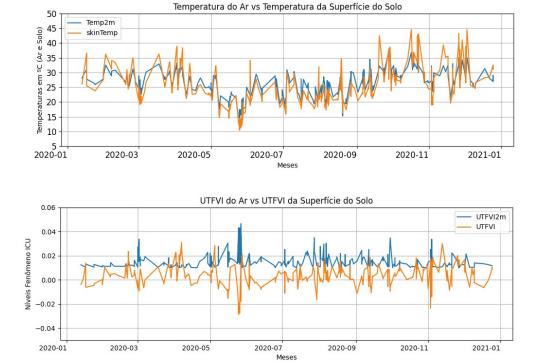


Figura 46: Séries das Temperaturas diárias da superfície do solo e das temperaturas do ar (2 metros) e UTFVI da superfície do solo e a 2 metros para ano 2020.

4.6.4. Análise Dos Resultados Da Classificação Do Fenômeno UHI

4.6.4.1. UTFVI Na Superficie Do Solo

Retratar a incidência do Fenômeno UHI torna-se necessário para averiguar quais níveis deste fenômeno têm maior significância em relação às variáveis climáticas Umidade Relativa do Ar e direção e Velocidade dos Ventos. E para tal, as árvores de decisões construídas nas respectivas Floresta Aleatória dos períodos relacionados aos dados amostrais apresentaram resultados que indicaram, para o ano 2000, uma frequência de 1% do Fenômeno UHI Extremo quando a Velocidade dos Ventos é superior a 7,8 km/h, a direção é no sentido sul para norte e ainda a umidade relativa está em 50%, aproximadamente.

E quando o vento está com sua velocidade acima de 2,2 km/h e a Umidade Relativa do Ar abaixo de 50% a maior ocorrência, chegando a 42%, é para o

Fenômeno UHI Forte e quando a umidade do ar está entre 52% e 54% com Velocidade dos Ventos acima de 2,2 km/h houve uma ocorrência de 31%. Para o Fenômeno UHI Muito Forte a maior incidência acontece quando a Umidade Relativa do Ar variou entre 51% e 54% com ventos vindos da direção oeste.

Para o ano 2005, as maiores incidências foram para o Fenômeno UHI Forte quando a Umidade Relativa do Ar está entre 43% e 49% e Velocidade dos Ventos abaixo de 10 km/h onde este fenômeno atingiu 42% de cobertura e quando a umidade do ar atinge um patamar acima de 49% há uma ocorrência de 36% do fenômeno forte e na coincidência da Direção dos Ventos se originasse a nordeste de Cuiabá e a Umidade Relativa do Ar ficasse entre 41% e 43% aconteceria o Fenômeno UHI Muito Forte.

No ano de 2010, passou a ocorrer o Fenômeno UHI Extremo nas condições em que a Umidade Relativa do Ar estivesse entre 52% e 53%, com ventos vindo da direção nordeste com velocidade entre 2,4 km/h e 8,2 km/h, chegando a um total de 3% dos registros dos dados amostrais. A categoria Forte ocorreu em 49% dos registros quando a Umidade Relativa do Ar ficou abaixo de 51% e a Direção dos ventos vindo do nordeste e a ocorrência da categoria Muito Forte foi de 2% quando a umidade relativa se estabeleceu entre 51% e 53% e a Direção dos Vento vindo do sudoeste e noroeste, outros 2% acontenceu quando a umidade do ar ficou acima de 51% e a Direção dos Ventos vindo do sul.

O ano de 2015 se demonstrou atípico na incidência dos fenômenos categorizados do UTFVI, onde houve uma difusão para ambos, fenômenos Forte e Muito Forte. Destaca-se que a ocorrência do Fenômeno UHI Forte se deu em 14% quando a Umidade Relativa do Ar ficou abaixo de 51% com Velocidade dos Ventos acima de 8 km/h e com direção sul e noroeste.

Interessante notar o comportamento dos fenômenos para o ano de 2020 que expressou uma incidência de 96% do Fenômeno UHI Forte quando a Umidade Relativa do Ar estivesse acima dos 40%, 1% da sua ocorrência se estabeleceu quando a Umidade Relativa do Ar ficou abaixo dos 40% com ventos vindos do nordeste e velocidade acima de 5,7 km/h o mesmo acontecendo para o Fenômeno UHI Muito

Forte. Outros 3% foram classificados na categoria de Muito Forte quando a Umidade Relativa do Ar estive abaixo de 40% e ventos vindos da direção nordeste.

4.6.4.2. UTFVI Na Superficie Do Solo E O NDVI

Na classificação dos dados amostrais do ano 2000 e sob o aspecto do NDVI, quando esta variável é incluída como variável dependente na Floresta Aleatória, haverá a ocorrência de 1% de frequência do Fenômeno ICU Forte sob a condição de que a Umidade Relativa do Ar tenha de estar em 45% e o NDVI acima de 0.21.

O NDVI influencia diretamente na Direção dos Ventos, para nordeste, quando esta variável estiver com o índice com valor maior ou igual a 0,22, Velocidade dos Ventos maior ou igual a 6,37 km/h e Umidade Relativa do Ar menor que 43%. Já quando o NDVI estiver menor que 0.22, a Velocidade dos Ventos maior ou igual a 6,37 km/h com a Umidade Relativa do Ar menor que 43%, a Direção dos Ventos ocorrerá no sentido sudoeste. A inclusão da variável NDVI causa influência indireta de modo que o Fenômeno ICU Forte ocorrerá em 32% dos casos quando a Umidade Relativa do Ar estiver entre 52% e 54% e a Velocidade dos Ventos for maior que 1,7 km/h.

Diferentemente do que ocorre quando não há a inclusão da variável NDVI, a incidência do Fenômeno UHI Extremo ocorrerá em 2% dos casos quando a Umidade Relativa do Ar estiver em 51% com a Velocidade dos Ventos maior ou igual a 7,8 km/h e a Direção dos Ventos for sul com tendência direcional para o noroeste.

Quando realizada a classificação dos dados amostrais referentes ao ano 2005, observou-se que 10% dos casos do Fenômeno ICU Muito Forte ocorreram quando o nível do NDVI ficou maior ou igual a 0.24, a Umidade Relativa do Ar ficou entre 41% e 45% e a Velocidade dos Ventos maior ou igual a 7,1 km/h. Em 1% dos casos ocorreu o Fenômeno ICU Forte quando a Umidade Relativa do Ar ficou entre 41% e 45%, com uma Velocidade dos Ventos em 7,1 km/h e o NDVI menor que 0,24.

Para os anos 2010 e 2015 não houveram ocorrências dos Fenômenos ICU que sofreram influência direta do NDVI. Com relação ao ano 2020 2% dos casos do

Fenômeno ICU Forte ocorreram quando a Umidade Relativa do Ar ficou abaixo de 40% e o NDVI abaixo de 0,33, houve uma ocorrência de 3% dos casos para o Fenômeno ICU Muito Forte quando a Umidade Relativa do Ar ficou menor que 40% e o NDVI maior ou igual a 0,33.

4.6.4.3. UTFVI A 2 Metros De Altura

Peculiarmente, os eventos para o Fenômeno UHI Extremo calculados a 2 metros de altura tiveram registros para todos os períodos em estudo neste trabalho, diferentemente dos registros do mesmo fenômeno ao nível da superfície do solo. Em 2000, 13% das classificações se deram nesta classe onde o pico fícou em 3% quando a Velocidade dos Ventos alternou entre 2,8 km/h e 3,1 km/h com a Umidade Relativa do Ar entre 46% e 52% e ventos vindos da direção sul e também nordeste. A maior incidência para o Fenômeno UHI Forte aconteceu quando a Velocidade dos Ventos foi maior que 4,2 km/h, independentemente da Direção dos Vento e da umidade do ar. 4% dos registros apontaram para a ocorrência da classe Muito Forte e isto foi devido a Umidade Relativa do Ar que ficou abaixo de 52% com a Velocidade dos Ventos ficando entre 3,8 km/h e 4,2 km/h.

Para 2005, a distribuição das classes dos fenômenos ocorreram de forma que 13% do total dos registros foram classificados como Extremo e destes 26% aconteram devido a Velocidade dos Ventos estar abaixo de 3,3 km/h e sua direção vindo do sul e noroeste e ainda a Umidade Relativa do Ar permanecendo entre 47% e 49%. Ressalta-se que o Fenômeno UHI Forte aconteceu em 28% do total das ocorrências quando a Velocidade dos Ventos atingiu um patamar acima de 6,5 km/h, independentemente da sua direção ou da Umidade Relativa do Ar. 17% do total dos registros foram para o Fenômeno UHI Muito Forte, onde foi observado que a Umidade Relativa do Ar se estabeleceu entre 45% e 52%.

Em 2010, 27% do total de ocorrências foi para o Fenômeno UHI Extremo e desta totalização 25% coincidiram quando a Velocidade dos Ventos se estabeleceu entre 2,3 e 4,7 km/h e Direção dos Ventos vindo do noroeste e Umidade Relativa do Ar entre 44% e 53%. 34% das ocorrências foram identificadas como Forte quando a Velocidade dos Ventos ficou acima de 5,1 km/h. Com relação à classe Muito Forte,

15% do total dos registros se vinculam a esta categoria onde 40% das ocorrências foram quando a Velocidade dos Ventos se alternou entre 1,2 e 2,3 km/h com a Umidade Relativa do Ar em torno de 51%.

Os resultados da classificação das classes do Fenômeno UHI para o ano 2015 demonstraram que 19% do total dos dados amostrais foram identificados como Extremo e destes 68% das ocorrências se deram quando a Velocidade dos Ventos intercalou 1,5 km/h e 3,5 km/h com a Umidade Relativa do Ar entre 46% e 51%. 69% do total dos registros foram denominados para a categoria Forte, sendo que 36% desta totalização ocorreram quando a Direção dos Ventos não se originaram do nordeste e com sua velocidade acima de 4,4 km/h e a Umidade Relativa do Ar acima de 44%. A categoria Muito Forte registrou 16% do total das ocorrências e o seu acontecimento se deu, na sua maioria quando a Direção dos Ventos veio do noroeste da cidade de Cuiabá.

Finalmente, no ano de 2020 ocorreram também as categorias Extremo, Forte e Muito Forte do Fenômeno UHI. Onde 13%, 59% e 25% do total dos resgistros foram identificados, respectivamente, como Extremo, Forte e Muito Forte. Ressaltase que 52% dos registros da categoria Muito Forte aconteceram quando a Velocidade dos Ventos variou entre 2,0 km/h e 3 km/h e a umidade relativa entre 47% e 52%. Dos registros da categoria Forte, pouco mais de 69% foram detectados quando a Velocidade dos Ventos ficou acima de 4,6 km/h, independente da sua direção e Umidade Relativa do Ar. Já em relação à categoria Extremo, 38% foram registrados quando a Velocidade dos Ventos se intercalou 3,1 km/h e 4 km/h com a Direção dos Ventos vindo do sudeste e nordeste e ainda a umidade do acima se estabelecendo acima de 50%.

4.6.4.4. UTFVI A 2 Metros De Altura E O NDVI

A investigação dos resultados da análise da classificação do Fenômeno UHI a 2 metros de altura demonstram não ocorrer influência do NDVI em sua formação para o período de 2000, existe apenas influência quando se trata de avaliar a Direção dos Ventos que se direciona à nordeste quando a Umidade Relativa do Ar está em

51%, o NDVI maior ou igual 0,37 e Velocidade dos Ventos em 4,5 km/h. Quando o NDVI é menor que 0,37 e mantendo as mesmas características para a Velocidade dos Ventos e Umidade Relativa do Ar a Direção dos Ventos passa ser sudeste. Ressalta-se que 2005, 2010 e 2015 tiveram um comportamento similar ao período de 2000 onde foi constatado influencia indireta.

No período de 2020, o NDVI passou a ter impacto na formação dos Fenômenos UHI Forte e Muito Forte. Em 10% dos casos o NDVI contribui para a formação do Fenômeno UHI Forte, sendo que em 6% dos casos este fenômeno ocorreu nas condições em que o NDVI foi maior ou igual a 0,25 com a Direção dos Ventos ficando a sudeste e nordeste e ainda a Velocidade dos Ventos em 8,0 km/h. Em 5% dos casos, o NDVI em conjunto com as variáveis Umidade Relativa do Ar, Direção dos Ventos e Velocidade dos Ventos ocasionou o Fenômeno UHI Muito Forte.

4.6.5. Distribuição Espacial Da Temperatura Da Superfície Do Solo E Do Ar

Os gráficos de distribuição para Temperatura do Ar e da Temperatura da Superfície do Solo em relação à Direção dos Ventos (quadrantes) são elenvados nas Figuras 47 e 48. Os dados dispersos para Temperatura do Ar (Figura 48) foram mais homogêneos do que aqueles para Temperatura da Superfície do Solo (Figura 47) e a partir destes boxplots pode-se observar que as médias das temperaturas do ar se estabelece por volta de 30°C e da superfície do solo, em torno de 32°C estão vinculadas ao quadrante 7 para ambas variáveis microclimáticas, porém não se pode afirmar que que as temperaturas do ar e do solo ficam mais elevadas quando o vento surge da direção do quadrante 7, ou seja direção noroeste.

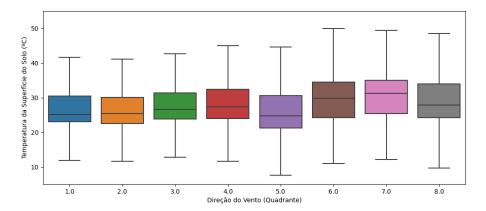


Figura 47: Temperatura da Superfície do Solo em relação à Direção dos Ventos denotada pelos quadrantes de 1 a 8. O traço dentro da caixa indica o valor mediano dos dados das temperaturas da superfície do solo para os anos de 2000-2020.

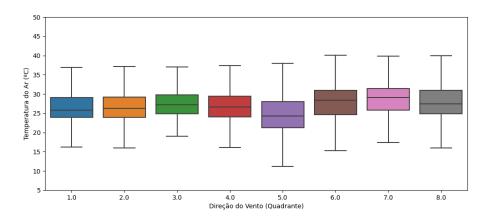


Figura 48: Temperatura do Ar (2m) em relação à Direção dos Ventos denotada pelos quadrantes de 1 a 8. O traço dentro da caixa indica o valor mediano dos dados das temperaturas do ar para a cidade de Cuiabá para os anos de 2000-2020.

5. CONCLUSÕES

O propósito deste trabalho foi analisar como os padrões de temperaturas da superfície do solo e da Temperatura do Ar das áreas urbanas, através do uso do Índice de Variância do Campo Termal Urbano, podem impactar nos comportamentos dos ventos locais e assim favorecer indiretamente para a formação de ilhas de calor. As ilhas de calor tem por si só um arranjo espacial que é intrisicamente relacionado com a classificação termal em função do uso e cobertura do solo, assim foi feito, inicialmente, uma análise da ocupação do solo onde se verificou que houve uma crescente variação das temperaturas, tanto da superfície do solo como do ar, muito em função da expansão das áreas construídas.

O comportamento dos ventos locais na cidade de Cuiabá-MT, segundo os dados de reanálise do ERA5, tem como característica uma combinação das direções nor-nordeste e norte, com ocorrências em torno de 30% e com velocidade máxima média de aproximadamente de 22 km/h para a estação quente e seca. Muito embora acontecesse esta predominância na direção norte, houve também ocorrências de ventos na direção sul. Para a estação quente e úmida houve uma predominância no sentido nor-nordeste com cerca de 60% das ocorrências e com a velocidade dos ventos em torno de 20 km/h.

Importa ressaltar que de acordo com o aumento da temperatura do ar, consoante as Figuras 10 a 19, houve um espalhamento dos ventos locais, o que indica a influência da mesma no comportamento dos ventos locais. Assim, investir na mitigação da formação de ilha de calor pode colaborar para um maior equilíbrio do fluxo dos ventos locais, o que pode levar a melhorias de fatores como saúde, economia e etc.

Ainda nesse mesmo sentido, as Figura 20 e 21 demonstram a relação proporcional da cobertura do solo com a Temperatura da Superfície do Solo. Nestas figuras, é possível se ater na evolução das áreas construídas que se expandiram para a região norte e sudeste com consequente diminuição da área verde. Numa classificação específica da cobertura do solo, retratada na Tabela 10, percebe-se uma

nítida característica da formação de uma ilha de calor, aumento de áreas impermeáveis (áreas construídas e solo nu) e a diminuição da área vegetada.

Como consequência da correlação positiva entre os índices de áreas construídas e as temperaturas da superfície do solo e do ar, procurou-se avaliar os fatores que mais contribuiem na formação dos fenômenos UHI em Cuiabá, o que resultou na identificação de três níveis ou graus que podem impactar negativamente no desenvolvimento da vegetação, natural ou não, no cotidiano das pessoas quanto aos aspectos econômicos e de saúde, dentre outros impactos negativos. Cuiabá apresenta, conforme a Figura 22, áreas com piores condições térmicas com UTFVI > 0,010, categorizados neste trabalho como Fenômeno UHI Forte, Muito Forte e Extremo.

Na Figura 22, houve uma demonstração da dinâmica do UTFVI onde no ano 2000 a variância do campo termal, com piores condições, se concentrou na região norte e no ano de 2020 houve uma maior distribuição da variância do campo termal, se expandido para região sul e sudeste, o que se expressa um aumento da área de ilha de calor na região da cidade de Cuiabá-MT.

Como forma de avaliar quais os impactos causados por estas condições, foi então realizado uma classificação dos níveis do UTFVI em função da Velocidade e Direção dos Ventos e Umidade Relativa do Ar. Como resultado direto desta classificação, notou-se que ao nível da superfície do solo a Umidade Relativa do Ar tem mais importância, ou seja, esta variável contribui efetivamente para a formação do UTFVI em relação à superfície do solo.

E quando se trata da classificação do UTFVI a 2 metros de altura a variável Velocidade dos Ventos tem mais importância. Assim, para classificação do UTFVI ao nível da superfície do solo a Floresta Aleatória determinou que deve-se utilizar a Umidade Relativa do Ar como fator de definição do Fenômeno UHI resultante. Já para a classificação do UTFVI a 2 metros a Floresta Aleatória resultante determinou que a Velocidade dos Ventos é o fator principal na definição do Fenômeno UHI resultante.

Quando adicionada na classificação, a variável NDVI se mostrou, de fato, ser fator de atenuação da formação de ilha de calor, consoante a baixa importância,

retratadas nas Figuras 36, 37 e 38, a partir dos resultados da classificação da Floresta Aleatória quando avaliada na perspectiva da UTFVI em relação à superfície do solo.

Ainda de acordo com os resultados desta classificação, o NDVI não se repercute no desempenho e nem na homogeneidade da formação dos Fenômenos UHI Forte, Muito Forte e Extremo. Ao avaliar na perspectiva do UTFVI a 2 metros de altura, o mesmo se repetiu, o Índice de Vegetação por Diferença Normalizada se revelou não impactar na formação dos Fenômenos UHI Forte, Muito Forte e Extremo. Presume-se, a partir destes resultados, que o NDVI é um fator de grande importância na mitigação na formação de ilha de calor.

Um aspecto a salientar é que a chance de ocorrências do Fenômeno UHI Extremo aumenta consideravelmente quando se compara a variável Índice de Variância do Campo Termal Urbano a 2 metros de altura com esta mesma variável em relação à superficie do solo, o que nos leva a considerar que as variações entre a temperatura da superfície do solo do ponto geográfico e a temperatura média da superfície solo de toda área em estudo (Cuiabá-MT) é menor quando comparadas com as variações temperatura a 2 metros de altura do ponto geográfico em relação a toda área em estudo. O que implica em observar que os dados das variáveis microclimáticas a 2 metros de altura são mais significativos dos os respectivos dados em relação à superfície do solo e, portanto, tem mais influência na formação do Fenômeno UHI.

Finalmente, o uso do algoritmo de aprendizado de máquina supervisionado Floresta Aleatória se mostrou bastante versátil na classificação tanto na cobertura do solo, com análise de imagens, onde demonstrou a sua capacidade de detecção com grande acurácia de pontos (pixels) distintos como água, vegetação e solo nu (Figura 21) quanto na classificação de dados amostrais realizando predições com acurácia dentro das expectativas, como pode ser visualizado nas árvores de decições, tais como as Figuras 23 e 42.

Dentre as suas características no processamento das informações, a mais importante foi o fato de a Floresta Aleatória lidar muito bem com a alta dimensionalidade dos dados, haja vista que dados microclimáticas são propensos a serem vultosos tanto na quantidade de variáveis quanto no volume dos dados em si.

Grande quantidade de dados podem causar necessidade de ajustes excessivos e o algoritmo da Floresta Aleatória lida muito bem com estes ajustes, além de combinar vários modelos, subdivisões dos dados em árvores menores, para se fazer previsões.

Um outro aspecto não menos importante é a capacidade de análises estatísticas inerente ao próprio algoritmo que podem sem aplicadas tanto na classificação de dados bem como na regressão de problemas.

5.1. CONTRIBUIÇÕES

Como contribuição deste trabalho buscou-se: (i) uma representação espaçotemporal das temperaturas da superfície do solo e do ar a fim de avaliar a extensão dos efeitos negativos da expansão da área construída na cidade de Cuiabá através da classificação do uso e cobertura do solo; (ii) uma análise comparativa espaçotemporal e sua respectiva representação gráfica com objetivo de facilitar a compreensão da relação negativa entre índice da diferença normalizada da vegetação e o Índice de Variância do Campo Termal Urbano e por fim a classificação dos Fenômenos UHI em função das variáveis microclimáticas como uma forma preditiva para analisar quais as tendências do acontecimento dos níveis de UTFVI, tanto em relação à superfície do solo quando a 2 metros de altura.

Um outro aspecto de relevância da contribuição deste trabalho é quanto ao planejamento urbano. Gestores públicos de cidades médias e grandes podem fazer uso dos conhecimentos científicos em favor de melhorias no conforto térmico dos habitantes. Neste sentido, os gestores podem identificar as áreas com piores índices do Fenômeno UHI e assim realizar trabalhos para mitigar os efeitos nocivos da intensidade das ilhas de calor urbano seja controlando a umidade relativa do ar por meios de artificialização de lagos, seja pela diminuição de áreas impermeáveis com a substituição destas áreas por vegetação, de preferência nativas, ou mesmo com a criação de espaços verdes ou reflorestamento de áreas nativas. Em outra frente de mitigação do Fenômeno UHI, os gestores podem tirar proveito dos ventos locais para proporcionar arrefecimento da sensação de calor com o direcionamento dos ventos

para áreas mais quentes (criação de cânions), ou mesmo projetar a construção de novas áreas havitadas com aproveitamento da ventilação proporcionada pelos ventos.

5.2. TRABALHOS FUTUROS

A partir dos resultados deste trabalho é possível expandir análises classificatórias das condições climáticas que contribuem para a formação de ilhas de calor urbano. Dessa forma, seria importante que:

- A metodologia empregada neste trabalho possa ser utilizada em estudos que venham realizar análise comparativa entre cidades que sofrem com os efeitos da formação de ilhas de calor;
- Sejam realizadas classificações de zonas climáticas locais e verificar a influência das suas características variáveis microclimáticas de interesse;
- Se empregue o uso de classificação em conjunto com tecnologias geoespaciais para analisar as tendências das principais características de ilhas de calor urbano;
- Sejam feitos estudos que proponham, através de classificação de dados pertinentes, melhorias na morfologia de cidades e assim obter ganhos com a redução dos efeitos nocivos de ilhas de calor urbano.
- Buscar o uso de tecnologias alternativas à Floresta Aleatória como análise geoespacial para prover compreensão da relação entre pontos geográficos e assim fornecer relações entre as variáveis microclimáticas e revelar padrões e tendências.
- Aplicar regressões, com o uso de modelos KNN, para se analisar tendências de crescimento das áreas de ilhas de calor, bem como análise das tendências da intensidade do Fenômeno UHI.

Códigos Fontes

Os códigos fontes utilizados neste projeto estão disponíveis na plataforma Github no endereço https://github.com/albertosales/pgfa

6. REFERÊNCIAS BIBLIOGRÁFICAS

ALBONWAS, R. K.; AL-KHAKANI, E. T. Using the multiple regression model to predict the land surface temperature in Al Najaf province, Iraq. **International Journal of Special Education**, v. 37, n. 3, p. 9970–9981, 31 mar. 2022.

ALCARDE ALVARES, C. et al. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, 1 dez. 2013.

ALI, J. et al. Random Forests and Decision Trees. **International Journal of Computer Science Issues(IJCSI)**, v. 9, 1 set. 2012.

BENALI, A. et al. Estimating air surface temperature in Portugal using MODIS LST data. **Remote Sensing of Environment**, v. 124, p. 108–121, 4 ago. 2012.

BÉNARD, C.; DA VEIGA, S.; SCORNET, E. Mean decrease accuracy for random forests: inconsistency, and a practical solution via the Sobol-MDA. **Biometrika**, v. 109, n. 4, p. 881–900, 1 dez. 2022.

BREIMAN, L. Random Forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 1 out. 2001.

CALLAHAN, C. W.; ELANSARI, A. M.; FENTON, D. L. Chapter 8 - Psychrometrics. Em: YAHIA, E. M. (Ed.). **Postharvest Technology of Perishable Horticultural Commodities**. [s.l.] Woodhead Publishing, 2019. p. 271–310.

CARLSON, T. N.; RIPLEY, D. A. On the relation between NDVI, fractional vegetation cover, and leaf area index. **Remote Sensing of Environment**, v. 62, n. 3, p. 241–252, 1 dez. 1997.

CHANDER, G.; MARKHAM, B. L.; HELDER, D. L. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. **Remote Sensing of Environment**, v. 113, n. 5, p. 893–903, maio 2009.

CORREIA, A. H. C.; PEHARZ, R.; DE CAMPOS, C. **Joints in Random Forests**. arXiv, , 19 nov. 2020. Disponível em: http://arxiv.org/abs/2006.14937. Acesso em: 1 jun. 2023

COSTA, R. L. Use of remote sensing in the identification of Urban Heat Islands and in the evaluation of Human Thermal Comfort. **Journal of Hyperspectral Remote Sensing**, v. 7, n. 7, p. 408–422, 2017.

DOS SANTOS, R. S. Estimating spatio-temporal air temperature in London (UK) using machine learning and earth observation satellite data. International Journal of Applied Earth Observation and Geoinformation, v. 88, p. 102066, 1 jun. 2020. DUARTE, D. H. S. O impacto da vegetação no microclima em cidades adensadas e seu papel na adaptação aos fenômenos de aquecimento urbano. Contribuições a uma abordagem intersdisciplinar. text—[s.l.] Universidade de São Paulo, 7 dez. 2015.

ENTERIA, N.; SANTAMOURIS, M.; EICKER, U. Urban Heat Island (UHI) Mitigation Hot and Humid Regions: Hot and Humid Regions. [s.l: s.n.].

ERELL, E.; PEARLMUTTER, D.; WILLIAMSON, T. **Urban Microclimate**. Routledge, 25 jun. 2012. Disponível em: https://www.taylorfrancis.com/books/9781136539435>. Acesso em: 5 jun. 2023

FERREIRA, L. DE F.; CARRILHO, S. T.; MENDES, P. C. áreas verdes urbanas: uma contribuição aos estudos das ilhas de frescor. **Brazilian Geographical Journal**, v. 6, n. 2, p. 101–120, 2015.

FILHO, W. L. et al. Impacts, strategies and tools to mitigate UHI. **A+BE** | **Architecture and the Built Environment**, n. 20, p. 67–102, 2017.

FRANCO, F. M. et al. TRAÇADO URBANO E SUA INFLUÊNCIA NO MICROCLIMA: UM ESTUDO DE CASO EM CENTRO HISTÓRICO. Revista Eletrônica em Gestão, Educação e Tecnologia Ambiental, p. 1916–1931, 18 fev. 2013.

GALL, J.; RAZAVI, N.; VAN GOOL, L. An Introduction to Random Forests for Multi-class Object Detection. Em: DELLAERT, F. et al. (Eds.). **Outdoor and Large-Scale Real-World Scene Analysis**. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. v. 7474p. 243–263.

GENUER, R. et al. **Random Forests for Big Data**. arXiv, , 22 mar. 2017. Disponível em: http://arxiv.org/abs/1511.08327. Acesso em: 1 jun. 2023

GUHA, S. et al. Analytical study of land surface temperature with NDVI and NDBI using Landsat 8 OLI and TIRS data in Florence and Naples city, Italy. **European Journal of Remote Sensing**, v. 51, n. 1, p. 667–678, 1 jan. 2018.

HERSBACH, H. et al. The ERA5 global reanalysis. **Quarterly Journal of the Royal Meteorological Society**, v. 146, n. 730, p. 1999–2049, jul. 2020.

HOWARD, L. The Climate of London: Deduced from Meteorological Observations. Cambridge: Cambridge University Press, 2012. v. 1

HOWELL, T.; DUSEK, D. Comparison of Vapor-Pressure-Deficit Calculation Methods—Southern High Plains. **Journal of Irrigation and Drainage Engineering**, v. 121, 1 mar. 1995.

HUSSEIN ABDULZHRAA AL-SAGHEER, R.; ALHARAN, A.; AL-HABOOBI, A. Popular Decision Tree Algorithms of Data Mining Techniques: A Review. p. 133–142, 1 jan. 2017.

KAFY, A.-A. Estimation of Urban Heat Islands Effect and Its Impact on Climate Change: A Remote Sensing and GIS-Based Approach in Rajshahi District.

Rochester, NY, 1 out. 2019. Disponível em: https://papers.ssrn.com/abstract=3681227. Acesso em: 25 jun. 2023

KAFY, A.-A. et al. Prediction of seasonal urban thermal field variance index using machine learning algorithms in Cumilla, Bangladesh. **Sustainable Cities and Society**, v. 64, p. 102542, 1 jan. 2021.

KANG, L. et al. Use of Geographically Weighted Regression Model for Exploring Spatial Patterns and Local Factors Behind NDVI-Precipitation Correlation. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 7, n. 11, p. 4530–4538, nov. 2014.

KAUR, L.; MITTAL, R. Variables in Social Science Research. 24 abr. 2021.

KONG, F.; SINGH, R. P. 2 - Chemical Deterioration and Physical Instability of Foods and Beverages. Em: SUBRAMANIAM, P. (Ed.). **The Stability and Shelf Life of Food (Second Edition)**. Woodhead Publishing Series in Food Science, Technology and Nutrition. Second Edition ed. [s.l.] Woodhead Publishing, 2016. p. 43–76.

KUMAR, D.; SHEKHAR, S. Statistical analysis of land surface temperature–vegetation indexes relationship through thermal remote sensing. **Ecotoxicology and Environmental Safety**, Green Technologies for Environmental Pollution Control and Prevention (Part 1). v. 121, p. 39–44, 1 nov. 2015.

LAN, T. et al. A Comparative Study of Decision Tree, Random Forest, and Convolutional Neural Network for Spread-F Identification. **Advances in Space Research**, v. 65, 1 fev. 2020.

LAWRENCE, M. G. The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications. **Bulletin of the American Meteorological Society**, v. 86, n. 2, p. 225–234, 1 fev. 2005.

LEHOCZKY, A. et al. The Urban Heat Island Effect in the City of Valencia: A Case Study for Hot Summer Days. **Urban Science**, v. 1, n. 1, p. 9, mar. 2017.

LIU, L.; ZHANG, Y. Urban heat island analysis using the landsat TM data and ASTER Data: A case study in Hong Kong. **Remote Sensing**, v. 3, p. 1535–1552, 1 dez. 2011.

LIU, Z.; XIONG, H. **Object Detection and Localization Using Random Forest**. 2012 Second International Conference on Intelligent System Design and Engineering Application. **Anais**... Em: 2012 SECOND INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEM DESIGN AND ENGINEERING APPLICATION. jan. 2012.

LUZ, L. et al. Utilização do NDVI e NDBI para avaliação da expansão urbana no município do Rio das Ostras - RJ, utilizando a plataforma Google Earth Engine. **ANAIS DO XIX SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO**, 2019.

MARTINEZ-TABOADA, F.; REDONDO, J. I. The SIESTA (SEAAV Integrated evaluation sedation tool for anaesthesia) project: Initial development of a multifactorial sedation assessment tool for dogs. **PLOS ONE**, v. 15, n. 4, p. e0230799, 1 abr. 2020.

MCHUGH, M. L. Interrater reliability: the kappa statistic. **Biochemia Medica**, v. 22, n. 3, p. 276–282, 2012.

MILLS, G. Luke Howard and The Climate of London. **Weather**, v. 63, p. 153–157, 1 jun. 2008.

NAIM, MD. N. H.; KAFY, A.-A. Assessment of urban thermal field variance index and defining the relationship between land cover and surface temperature in Chattogram city: A remote sensing and statistical approach. **Environmental Challenges**, v. 4, p. 100107, 1 ago. 2021.

NAJAFZADEH, F. et al. Spatial and Temporal Analysis of Surface Urban Heat Island and Thermal Comfort Using Landsat Satellite Images between 1989 and 2019: A Case Study in Tehran. **Remote Sensing**, v. 13, n. 21, p. 4469, 7 nov. 2021.

NERY, C. V. M.; MOREIRA, A. A.; FERNANDES, F. H. S. Análise do Comportamento Espectral da Floresta Estacional Decidual no Parque Estadual Lapa Grande (Behavior Analysis of Spectral Deciduous Forest in Lapa Grande State Park). **Revista Brasileira de Geografia Física**, v. 7, n. 2, p. 417–433, 16 set. 2014.

NGUYEN, T. T. Landsat Time-series Images-based Urban Heat Island Analysis: The Effects of Changes in Vegetation and Built-up Land on Land Surface Temperature in Summer in the Hanoi Metropolitan Area, Vietnam. **Environment and Natural Resources Journal**, v. 18, n. 2, 2020.

NUGRAHA, A. S. A.; GUNAWAN, T.; KAMAL, M. Comparison of Land Surface Temperature Derived From Landsat 7 ETM+ and Landsat 8 OLI/TIRS for Drought Monitoring. **IOP Conference Series: Earth and Environmental Science**, v. 313, n. 1, p. 012041, ago. 2019.

O'MALLEY, C. et al. An Investigation into Minimizing Urban Heat Island (UHI) Effects: A UK Perspective. **Energy Procedia**, 6th International Conference on Sustainability in Energy and Buildings, SEB-14. v. 62, p. 72–80, 1 jan. 2014.

PEREIRA, R. H. M.; GONÇALVES, C. N. **Geobr: Loads Shapefiles of Official Spatial Data Sets of Brazil.** https://github.com/ipeaGIT/geobr, 22 nov. 2019. Disponível em: https://github.com/ipeaGIT/geobr>. Acesso em: 22 nov. 2023

RANAGALAGE, M.; ESTOQUE, R.; MURAYAMA, Y. An Urban Heat Island Study of the Colombo Metropolitan Area, Sri Lanka, Based on Landsat Data (1997–2017). **International Journal of Geo-Information**, v. 6, p. 17, 23 jun. 2017.

RAVANELLI, R. et al. MONITORING URBAN HEAT ISLAND THROUGH GOOGLE EARTH ENGINE: POTENTIALITIES AND DIFFICULTIES IN DIFFERENT CITIES OF THE UNITED STATES. . Em: ISPRS - INTERNATIONAL ARCHIVES OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES. 30 abr. 2018.

ROUSE, J. W. et al. **Monitoring vegetation systems in the Great Plains with ERTS**. 1 jan. 1974. Disponível em: https://ntrs.nasa.gov/citations/19740022614>. Acesso em: 6 jun. 2023

SEJATI, A. W.; BUCHORI, I.; RUDIARTO, I. The spatio-temporal trends of urban growth and surface urban heat islands over two decades in the Semarang Metropolitan Region. **Sustainable Cities and Society**, v. 46, p. 101432, 1 abr. 2019. SHAHMOHAMADI, P. et al. The Impact of Anthropogenic Heat on Formation of Urban Heat Island and Energy Consumption Balance. **Urban Studies Research**, v. 2011, p. e497524, 2 jul. 2011.

SOBRINO, J.; JIMENEZ, J.-C.; PAOLINI, L. Land surface temperature retrieval from LANDSAT TM 5. **Remote Sensing of Environment**, v. 90, p. 434–440, 1 abr. 2004.

SOUTO, J. I. DE O.; COHEN, J. C. P. Spatiotemporal variability of urban heat island: Influence of urbanization on seasonal pattern of land surface temperature in the Metropolitan Region of Belém, Brazil. **urbe. Revista Brasileira de Gestão Urbana**, v. 13, p. e20200260, 17 maio 2021.

STEWART, I. D.; OKE, T. Local Climate Zones for Urban Temperature Studies. **Bulletin of the American Meteorological Society**, v. 93, p. 1879–1900, 1 dez. 2012. TESFAMARIAM, S.; GOVINDU, V.; UNCHA, A. Spatio-temporal analysis of urban heat island (UHI) and its effect on urban ecology: The case of Mekelle city, Northern Ethiopia. **Heliyon**, v. 9, n. 2, p. e13098, 1 fev. 2023.

TOMLINSON, C. J. et al. Including the urban heat island in spatial heat health risk assessment strategies: a case study for Birmingham, UK. **International Journal of Health Geographics**, v. 10, n. 1, p. 42, 17 jun. 2011.

VENTER, Z. S. et al. Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. **Remote Sensing of Environment**, v. 242, p. 111791, 1 jun. 2020.

VILANOVA, S. R. F.; MAITELLI, G. T. A importância da conservação de áreas verdes remanescentes no centro político administrativo de Cuiabá-MT. **UNICIÊNCIAS**, v. 13, n. 1, 2009.

VILLANUEVA-SOLIS, J. Urban Heat Island Mitigation and Urban Planning: The Case of the Mexicali, B. C. Mexico. **American Journal of Climate Change**, v. 06, p. 22–39, 1 jan. 2017.

WENG, Q.; LU, D.; SCHUBRING, J. Estimation of land surface temperature-vegetation abundance relationship for urban heat island studies. **Remote Sensing of Environment**, v. 89, n. 4, p. 467–483, [s.d.].

WERNECK, D. R. Variabilidade da temperatura de superfície diurna entre as zonas climáticas locais (LCZ): um estudo para a área urbana do Distrito Federal. 28 set. 2022.

WICKI, A.; PARLOW, E.; FEIGENWINTER, C. Evaluation and Modeling of Urban Heat Island Intensity in Basel, Switzerland. **Climate**, v. 6, n. 3, p. 55, set. 2018. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, 1 dez. 2013.

BENALI, A. et al. Estimating air surface temperature in Portugal using MODIS LST data. **Remote Sensing of Environment**, v. 124, p. 108–121, 4 ago. 2012.

CARLSON, T. N.; RIPLEY, D. A. On the relation between NDVI, fractional vegetation cover, and leaf area index. **Remote Sensing of Environment**, v. 62, n. 3, p. 241–252, 1 dez. 1997.

CORREIA, A. H. C.; PEHARZ, R.; DE CAMPOS, C. **Joints in Random Forests**. arXiv, , 19 nov. 2020. Disponível em: http://arxiv.org/abs/2006.14937. Acesso em: 1 jun. 2023

COSTA, R. L. Use of remote sensing in the identification of Urban Heat Islands and in the evaluation of Human Thermal Comfort. **Journal of Hyperspectral Remote Sensing**, v. 7, n. 7, p. 408–422, 2017.

DOS SANTOS, R. S. Estimating spatio-temporal air temperature in London (UK) using machine learning and earth observation satellite data. International Journal of Applied Earth Observation and Geoinformation, v. 88, p. 102066, 1 jun. 2020. DUARTE, D. H. S. O impacto da vegetação no microclima em cidades adensadas e seu papel na adaptação aos fenômenos de aquecimento urbano. Contribuições a uma abordagem intersdisciplinar. text—[s.l.] Universidade de São Paulo, 7 dez. 2015.

ERELL, E.; PEARLMUTTER, D.; WILLIAMSON, T. **Urban Microclimate**. Routledge, 25 jun. 2012. Disponível em: https://www.taylorfrancis.com/books/9781136539435. Acesso em: 5 jun. 2023 FERREIRA, H. V. L.; JÚNIOR, J. C. U. VARIAÇÃO DA TEMPERATURA DA SUPERFÍCIE ATRAVÉS DE IMAGENS ASTER EM ZONAS CLIMÁTICAS LOCAIS DA CIDADE DE CUIABÁ, BRASIL. **Revista Brasileira de Climatologia**, v. 26, n. 0, 26 mar. 2020.

GALL, J.; RAZAVI, N.; VAN GOOL, L. An Introduction to Random Forests for Multi-class Object Detection. Em: DELLAERT, F. et al. (Eds.). **Outdoor and Large-Scale Real-World Scene Analysis**. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. v. 7474p. 243–263.

GENUER, R. et al. **Random Forests for Big Data**. arXiv, , 22 mar. 2017. Disponível em: http://arxiv.org/abs/1511.08327>. Acesso em: 1 jun. 2023

GUHA, S. et al. Analytical study of land surface temperature with NDVI and NDBI using Landsat 8 OLI and TIRS data in Florence and Naples city, Italy. **European Journal of Remote Sensing**, v. 51, n. 1, p. 667–678, 1 jan. 2018.

HERSBACH, H. et al. The ERA5 global reanalysis. **Quarterly Journal of the Royal Meteorological Society**, v. 146, n. 730, p. 1999–2049, jul. 2020.

LAN, T. et al. A Comparative Study of Decision Tree, Random Forest, and Convolutional Neural Network for Spread-F Identification. **Advances in Space Research**, v. 65, 1 fev. 2020.

LEHOCZKY, A. et al. The Urban Heat Island Effect in the City of Valencia: A Case Study for Hot Summer Days. **Urban Science**, v. 1, n. 1, p. 9, mar. 2017.

LIU, Z.; XIONG, H. **Object Detection and Localization Using Random Forest**. 2012 Second International Conference on Intelligent System Design and Engineering Application. **Anais**... Em: 2012 SECOND INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEM DESIGN AND ENGINEERING APPLICATION. jan. 2012.

Livro Perfil 2011.indb. [s.d.].

MILLS, G. Luke Howard and The Climate of London. **Weather**, v. 63, p. 153–157, 1 jun. 2008.

NERY, C. V. M.; MOREIRA, A. A.; FERNANDES, F. H. S. Análise do Comportamento Espectral da Floresta Estacional Decidual no Parque Estadual Lapa Grande (Behavior Analysis of Spectral Deciduous Forest in Lapa Grande State Park). **Revista Brasileira de Geografia Física**, v. 7, n. 2, p. 417–433, 16 set. 2014.

NGUYEN, T. T. Landsat Time-series Images-based Urban Heat Island Analysis: The Effects of Changes in Vegetation and Built-up Land on Land Surface Temperature in Summer in the Hanoi Metropolitan Area, Vietnam. **Environment and Natural Resources Journal**, v. 18, n. 2, 2020.

NUGRAHA, A. S. A.; GUNAWAN, T.; KAMAL, M. Comparison of Land Surface Temperature Derived From Landsat 7 ETM+ and Landsat 8 OLI/TIRS for Drought Monitoring. **IOP Conference Series: Earth and Environmental Science**, v. 313, n. 1, p. 012041, ago. 2019.

O'MALLEY, C. et al. An Investigation into Minimizing Urban Heat Island (UHI) Effects: A UK Perspective. **Energy Procedia**, 6th International Conference on Sustainability in Energy and Buildings, SEB-14. v. 62, p. 72–80, 1 jan. 2014.

RAVANELLI, R. et al. MONITORING URBAN HEAT ISLAND THROUGH GOOGLE EARTH ENGINE: POTENTIALITIES AND DIFFICULTIES IN DIFFERENT CITIES OF THE UNITED STATES. . Em: ISPRS - INTERNATIONAL ARCHIVES OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES. 30 abr. 2018.

RENARD, F. et al. Evaluation of the Effect of Urban Redevelopment on Surface Urban Heat Islands. **Remote Sensing**, v. 11, p. 299, 1 fev. 2019.

ROUSE, J. W. et al. **Monitoring vegetation systems in the Great Plains with ERTS**. 1 jan. 1974. Disponível em: https://ntrs.nasa.gov/citations/19740022614>. Acesso em: 6 jun. 2023

SOBRINO, J.; JIMENEZ, J.-C.; PAOLINI, L. Land surface temperature retrieval from LANDSAT TM 5. **Remote Sensing of Environment**, v. 90, p. 434–440, 1 abr. 2004.

SOUTO, J. I. DE O.; COHEN, J. C. P. Spatiotemporal variability of urban heat island: Influence of urbanization on seasonal pattern of land surface temperature in the Metropolitan Region of Belém, Brazil. **urbe. Revista Brasileira de Gestão Urbana**, v. 13, p. e20200260, 17 maio 2021.

STEWART, I. D.; OKE, T. Local Climate Zones for Urban Temperature Studies. **Bulletin of the American Meteorological Society**, v. 93, p. 1879–1900, 1 dez. 2012. VENTER, Z. S. et al. Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. **Remote Sensing of Environment**, v. 242, p. 111791, 1 jun. 2020.

VILANOVA, S. R. F.; MAITELLI, G. T. A importância da conservação de áreas verdes remanescentes no centro político administrativo de Cuiabá-MT. **UNICIÊNCIAS**, v. 13, n. 1, 2009.

VILLANUEVA-SOLIS, J. Urban Heat Island Mitigation and Urban Planning: The Case of the Mexicali, B. C. Mexico. **American Journal of Climate Change**, v. 06, p. 22–39, 1 jan. 2017.

WENG, Q.; LU, D.; SCHUBRING, J. Estimation of land surface temperature-vegetation abundance relationship for urban heat island studies. **Remote Sensing of Environment**, v. 89, n. 4, p. 467–483, [s.d.].

WERNECK, D. R. Variabilidade da temperatura de superfície diurna entre as zonas climáticas locais (LCZ): um estudo para a área urbana do Distrito Federal. 28 set. 2022.

7. APÊNDICES

7.1. ANÁLISE DOS IMPACTOS DAS VARIÁVEIS MICROCLIMÁTICAS SOBRE O ÍNDICE DE VARIÂNCIA DO CAMPO TERMAL NA SUPERFÍCIE DO SOLO

Em relação ao 2005, 41,79% dos registros do Fenômeno UHI Extremo utilizados no treinamento (135 registros) foram classificados como falsos positivos, ou seja, instâncias errôneamente preditas como positivas e 58,20% (188 registros) com instâncias classificadas correta mente. Para o Fenômeno UHI Forte, houve apenas um registro considerado falso negativo de um universo de 3067 registros para essa categoria. Observa-se também que para a categoria de Fenômeno UHI Muito Forte, 28,09% dos registros foram classificados como falsos positivos, neste caso são instâncias sendo consideradas errôneamente preditas como positivas.

Quanto a validação (teste) dos dados, os resultados apresentam que 42,64% dos dados foram corretamente classificados para o Fenômeno UHI Extremo, 57,35% dos dados do Fenômeno UHI Forte foram classificados errôneamente como sendo Extremo indicando um valor falso negativo, na categoria Muito Forte houve 32,21% das instâncias sendo classificados como falsos positivos. Neste período o valor Kappa indica que há uma concordância moderada com 0,7294 para os dados de treinamento e 0,7198 para os dados de teste quanto a avaliação nominal das mesmas amostras, de acordo com a Tabela 33.

	Treino			Teste		
	Extremo	Forte	Muito_Forte	Extremo	Forte	Muito Forte
Extremo	96	0	0	39	0	0
Forte	134	3603	326	64	1535	160
Muito Forte	10	28	733	5	26	269
Acurácia:	0,9000	•		0,8785	•	
Kappa:	0,7188			0,6487		

Tabela 33: Matriz de Confusão e estatísticas da variável Fenômeno UHI para os dados amostrais do ano 2005.

A classificação dos dados para o ano 2010, expostos na Tabela 34, demonstra que os Fenômenos UHI Extremo, apenas 57% (85) dos dados foram classificados corretamente, bem como 16,87% dos dados foram classificados incorretamente, sendo os mesmos considerados falsos positivos. Na categoria Forte, 100% dos dados amostrais foram considerados corretamente preditos. Em relação à predição da validação dos dados amostrais classificados, 56,71% dos registros na categoria Extremo foram classificados corretamente, o restante como sendo falso negativo.

Assim como no treinamento os valores para a categoria Forte foram 100% classificados corretamente. Para a categoria Muito Forte houve uma classificação dos dados indicando que 17,56% dos dados foram classificados incorretamente como falsos positivos, incorrendo em um compreensão de que as instâncias foram erroneamente preditas como positivas. Os respectivos valores Kappa, para treino e teste, 0,8343 e 0,8294 indicam um nível de concordância forte, ou seja, entre 64 e 81% dos dados são confiáveis.

	Treino			Teste		
	Extremo	Forte	Muito Forte	Extremo	Forte	Muito Forte
Extremo	85	0	0	38	0	0
Forte	60	1929	44	33	822	25
Muito Forte	0	0	282	0	0	123
Acurácia:	0,9567			0,9443		
Kappa:	0,8561			0,8221		

Tabela 34: Matriz de Confusão e estatísticas da variável Fenômeno UHI para os dados amostrais do ano 2010.

Na Tabela 35, são mostrados os valores 0,9909 e 0,9863 referentes à acurácia dos dados treinados e testados, respectivamente e os valores Kappa 0,9532 para os dados treinamentos e 0,9354 para os dados de validação indicando que nível de concordância quase perfeito para o ano de 2015.

	Treino			Teste		
	Extremo	Forte	Muito Forte	Extremo	Forte	Muito Forte
Extremo	8	0	0	4	0	0
Forte	3	1930	20	0	851	7
Muito Forte	0	0	237	0	32	99
Acurácia:	0,9895			0,9927		
Kappa:	0,9479			0,8855		

Tabela 35: Matriz de Confusão e estatísticas da variável Fenômeno UHI para os dados amostrais do ano 2015.

Para o ano de 2020, conforme a Tabela 36, a taxa de acerto das predições se mantêm em torno de 90% com o indíce Kappa determinando uma confiança de 0,7841 para a treinagem dos dados e para validação 0,7964 indicando que, entre os avaliadosres, 64% e 81% dos dados são confiáveis. Dos registros avaliados, 42,63% dos valores dos dados amostrais para o Fenômeno UHI Extremo foram preditos errôneamente como negativos (falsos negativos), idem para outros 40% dos dados para a mesma categoria quando aplicados na validação dos dados. 25% dos dados relacionados ao evento Muito Forte foram classificados como falsos positivos.

	Treino			Teste		
	Extremo	Forte	Muito Forte	Extremo	Forte	Muito Forte
Extremo	186	0	0	51	0	0
Forte	138	3089	238	78	1321	116
Muito Forte	3	0	623	3	3	226
Acurácia:	0,9114			0,8888		
Kappa:	0,7222			0,7364		

Tabela 36: Matriz de Confusão e estatísticas da variável Fenômeno UHI para os dados amostrais do ano 2020.

7.2. ANÁLISE DOS IMPACTOS DA VARIÁVEIS MICROCLIMÁTICAS SOBRE O ÍNDICE DE VARIÂNCIA DO CAMPO TERMAL URBANO A 2 METROS DE ALTURA

Na realização da classificação dos dados amostrais para o ano de 2005, 41% dos registros do fenômeno UHI Extremo utilizados no treinamento (68 registros)

foram classificados como falsos negativos, ou seja, instâncias errôneamente preditas como negativas e 58,20% (148 registros) com instâncias classificadas corretamente e classificados como falsos positivos 15 registros. Para o Fenômeno UHI Forte, houve 60 registros demonstrados como falsos negativos de um montante de 733 registros para essa categoria. Nota-se também que na categoria de Fenômeno UHI Muito Forte, 15 registros, ou seja aproximadamente 7%, foram classificados como falsos negativos, neste caso são instâncias com significado positivo para o valor real.

Aplicando este aprendizado nos dados de testagem aos dados de validação dos dados, os resultados apresentam que 87,32% dados foram corretamente classificados para o Fenômeno UHI Extremo, 15,49% dos dados do Fenômeno UHI desta classe foram classificados errôneamente como sendo Fenômeno UHI Forte e outros 28,16% classificados como Fenômeno UHI Muito Forte, o que indica uma classificação de valores falsos negativos. Na classe Fenômeno UHI Forte houve 6% das instâncias sendo classificados como falsos positivos (18 registros), 36 registros classificados como falsos negativos e os demais registros, 273, classificados corretamente.

Quanto às instâncias relacionadas ao Fenômeno UHI Muito Forte, 44 (43,56%) foram classificados corretamente, significando que o valor real e também os valores previstos são iguais, classificados como falsos positivos 49 (48,51%) e 12,87% sendo classificados como sendo falsos negativos. O índice Kappa para este modelo indica que há uma concordância substancial com 0,7486 para os dados de treinamento e 0,6642 para os dados de teste quanto a avaliação nominal das mesmas amostras e a precisão calculada foi de 87,40% na predição dos dados amostrais e 82,75% para os dados de validação, conforme externado na Tabela 37.

	Treino			Teste		
	Extremo	Forte	Muito_Forte	Extremo	Forte	Muito Forte
Extremo	166	0	0	77	3	0
Forte	0	781	3	0	329	3
Muito Forte	0	0	240	0	0	102
Acurácia:	0,9975			0,9883	•	
Kappa:	0,9950			0,9775		

Tabela 37: Matriz de Confusão e estatísticas da variável Fenômeno UHI a 2 metros de altura para os dados amostrais do ano 2005.

As Tabelas 38, 39, 40 apresentam dados que indicam uma classificação bastante similar das variáveis em relação aos dados relacionados aos anos 2010, 2015 e 2020. Tanto as instâncias corretamente classificadas como as demais apresentam dados com valores muito próximos o que indica que um padrão na interação entre essas categorias, o que nos permite entender onde o modelo está se destacando ou mesmo onde está falhando. Os índices Kappa para os modelos dos respectivos anos indicam uma confiança substancial de concordância nas avaliações, tanto para os dados amostrais de treinamento do algoritmo como para os dados de validação do algoritmo. Indicando uma boa confiabilidade para os dados de treinamento ou seja, entre 61% e 80% dos dados são confiáveis.

	Treino			Teste		
	Extremo	Forte	Muito Forte	Extremo	Forte	Muito Forte
Extremo	427	18	33	179	12	15
Forte	0	883	1	0	401	5
Muito Forte	0	0	268	0	0	98
Acurácia:	0,9681			0,9549		
Kappa:	0,9462			0,9211		

Tabela 38: Matriz de Confusão e estatísticas da variável Fenômeno UHI a 2 metros de altura para os dados amostrais do ano 2010.

	Treino			Teste		
	Extremo	Forte	Muito Forte	Extremo	Forte	Muito Forte
Extremo	222	3	0	87	3	0
Forte	0	694	3	0	323	3
Muito Forte	0	0	221	0	0	90
Acurácia:	0,9948			0,9881		
Kappa:	0,9908			0,9773		

Tabela 39: Matriz de Confusão e estatísticas da variável Fenômeno UHI a 2 metros de altura para os dados amostrais do ano 2015.

	Treino			Teste		
	Extremo	Forte	Muito Forte	Extremo	Forte	Muito Forte
Extremo	196	0	1	83	3	2
Forte	18	769	5	15	323	4
Muito Forte	0	0	314	0	3	118
Acurácia:	0,9816	-		0,9510		
Kappa:	0,9669			0,9112		

Tabela 40: Matriz de Confusão e estatísticas da variável Fenômeno UHI a 2 metros de altura para os dados amostrais do ano 2020.

7.3. CLASSIFICAÇÃO DAS VARIÁVEIS MICROCLIMÁTICAS EM FUNÇÃO DO ÍNDICE DE VARIÂNCIA DO CAMPO TERMAL URBANO (FENÔMENO UHI) PARA OS ANOS 2005, 2010, 2015 E 2020

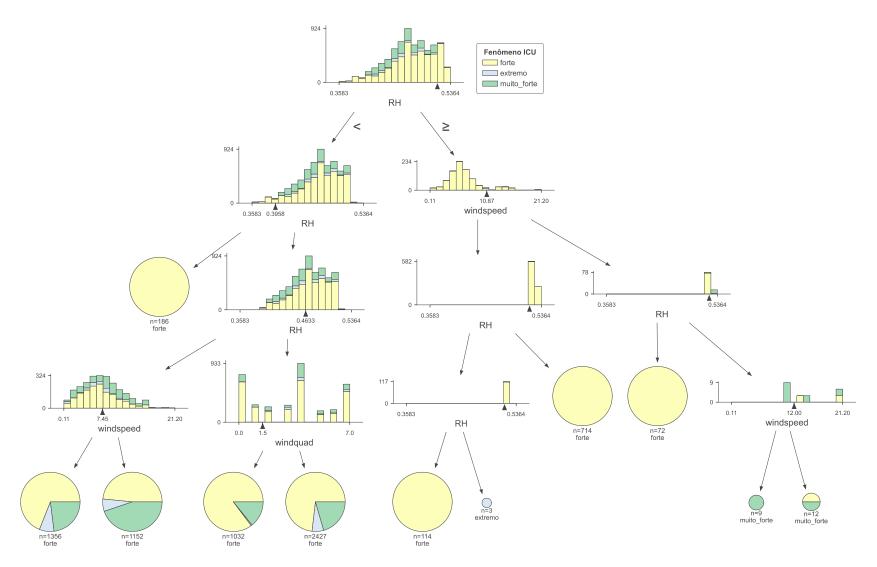


Figura 49: Floresta Aleatória para Fenômeno UHI - ano 2005.

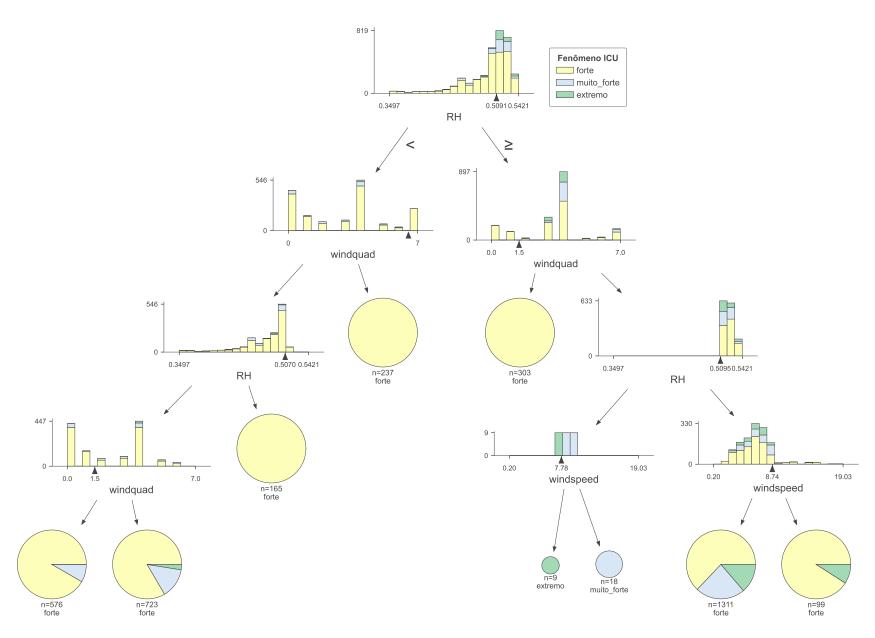


Figura 50: Floresta Aleatória para Fenômeno UHI - ano 2010.

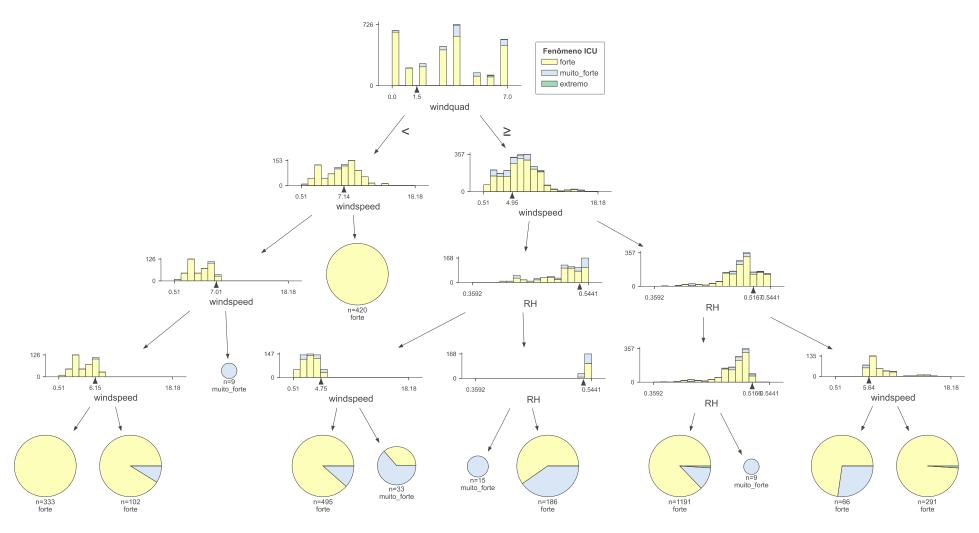


Figura 51: Floresta Aleatória para Fenômeno UHI - ano 2015.

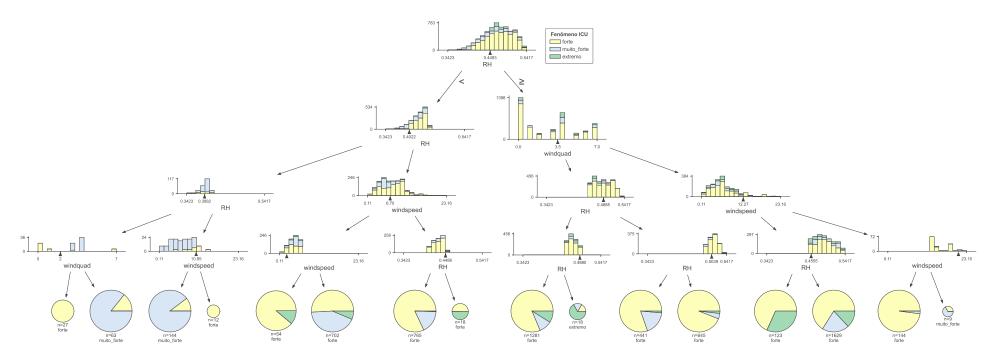


Figura 52: Floresta Aleatória para Fenômeno UHI - ano 2020.

7.4. CLASSIFICAÇÃO DAS VARIÁVEIS MICROCLIMÁTICAS EM FUNÇÃO DO ÍNDICE DE VARIÂNCIA DO CAMPO TERMAL URBANO A 2 METROS DE ALTURA PARA OS ANOS 2005, 2010, 2015 E 2020

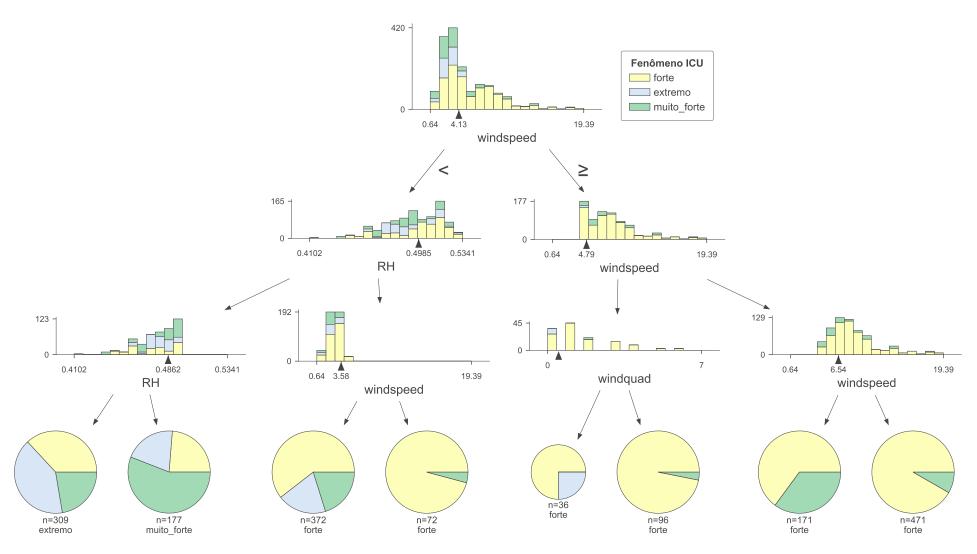


Figura 53: Floresta Aleatória para Fenômeno UHI a 2 metros de altura para o ano de 2005.

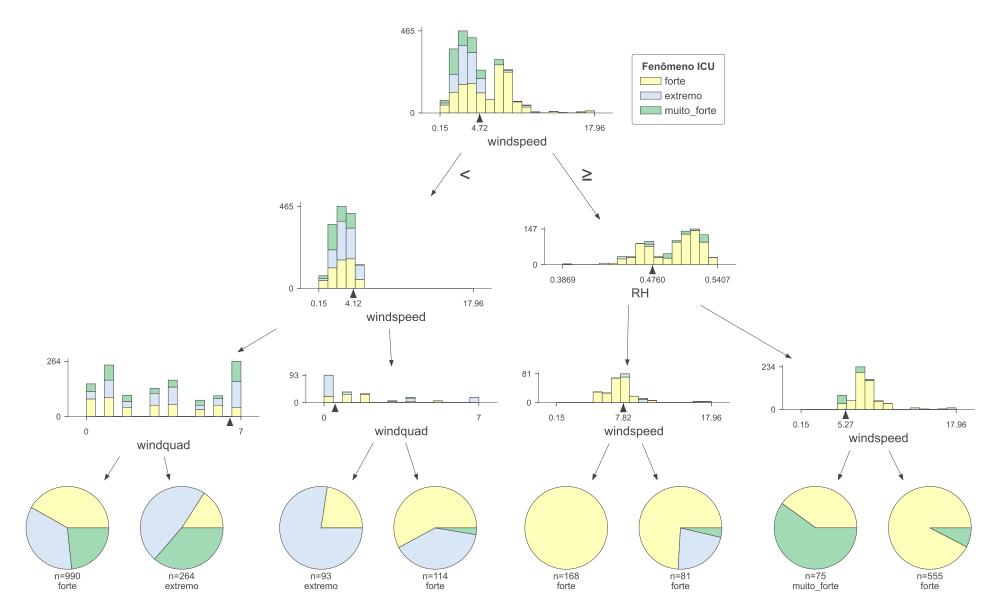


Figura 54: Floresta Aleatória para Fenômeno UHI a 2 metros de altura para o ano de 2010.

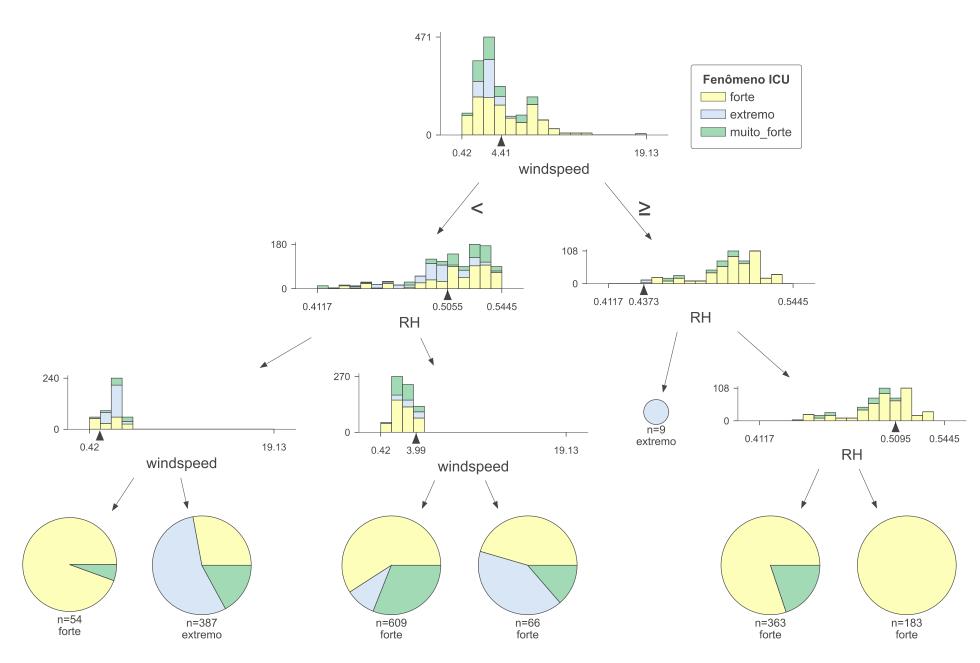


Figura 55: Floresta Aleatória para Fenômeno UHI a 2 metros de altura para o ano de 2015 com profundidade 5.

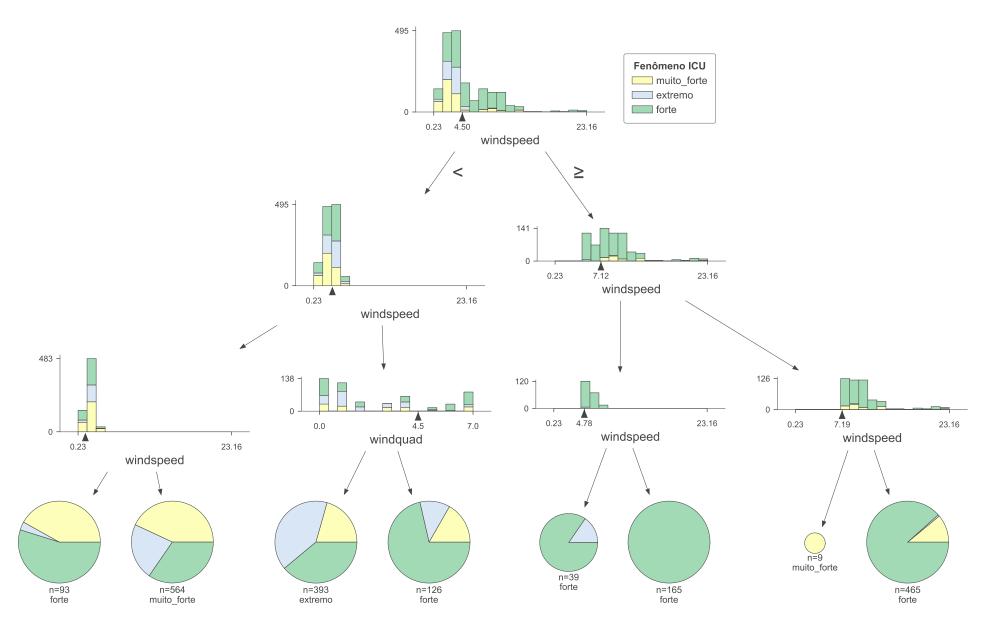


Figura 56: Floresta Aleatória para Fenômeno UHI a 2 metros de altura para o ano de 2020 com profundidade 5.