



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÕES



sid.inpe.br/mtc-m21c/2021/06.08.17.14-TDI

**APRIMORAMENTO DO MÉTODO DE
CLASSIFICAÇÃO DIURNA DE NUVENS DO CPTEC
UTILIZANDO O SENSOR ABI/GOES-16**

Gilvani Gomes de Carvalho

Dissertação de Mestrado do
Curso de Pós-Graduação em
Meteorologia, orientada pelos Drs.
Daniel Alejandro Vila, e Renato
Galante Negri, aprovada em 27 de
maio de 2021.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/44RASEP>>

INPE
São José dos Campos
2021

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE
Coordenação de Ensino, Pesquisa e Extensão (COEPE)
Divisão de Biblioteca (DIBIB)
CEP 12.227-010
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/7348
E-mail: pubtc@inpe.br

CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELLECTUAL DO INPE - CEPPII (PORTARIA Nº 176/2018/SEI-INPE):

Presidente:

Dra. Marley Cavalcante de Lima Moscati - Coordenação-Geral de Ciências da Terra (CGCT)

Membros:

Dra. Ieda Del Arco Sanches - Conselho de Pós-Graduação (CPG)
Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia, Tecnologia e Ciência Espaciais (CGCE)
Dr. Rafael Duarte Coelho dos Santos - Coordenação-Geral de Infraestrutura e Pesquisas Aplicadas (CGIP)
Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon
Clayton Martins Pereira - Divisão de Biblioteca (DIBIB)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)
André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)

EDITORAÇÃO ELETRÔNICA:

Ivone Martins - Divisão de Biblioteca (DIBIB)
André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÕES



sid.inpe.br/mtc-m21c/2021/06.08.17.14-TDI

**APRIMORAMENTO DO MÉTODO DE
CLASSIFICAÇÃO DIURNA DE NUVENS DO CPTEC
UTILIZANDO O SENSOR ABI/GOES-16**

Gilvani Gomes de Carvalho

Dissertação de Mestrado do
Curso de Pós-Graduação em
Meteorologia, orientada pelos Drs.
Daniel Alejandro Vila, e Renato
Galante Negri, aprovada em 27 de
maio de 2021.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/44RASEP>>

INPE
São José dos Campos
2021

Dados Internacionais de Catalogação na Publicação (CIP)

Carvalho, Gilvani Gomes de.

C253a Aprimoramento do método de classificação diurna de nuvens do CPTEC utilizando o sensor ABI/GOES-16 / Gilvani Gomes de Carvalho. – São José dos Campos : INPE, 2021.
xxvi + 115 p. ; (sid.inpe.br/mtc-m21c/2021/06.08.17.14-TDI)

Dissertação (Mestrado em Meteorologia) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2021.

Orientadores : Drs. Daniel Alejandro Vila, e Renato Galante Negri.

1. Classificação de nuvens. 2. GOES-16/ABI. 3. K-Means.
4. Métodos de agrupamento. I.Título.

CDU 551.576.1



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).



INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

Pós Graduação
Meteorologia

ATA DE REUNIÃO

DEFESA FINAL DE DISSERTAÇÃO: GILVANI GOMES DE CARVALHO- REGISTRO 37414/2019

BANCA: 161/2021

No dia 27 de maio de 2021, às 10h00 por videoconferência, a aluna mencionada acima defendeu seu trabalho final, intitulado: **Método de Classificação de Nuvens em Tempo Real a partir do Sensor Goes-16/ABI** (apresentação oral seguida de arguição) perante uma Banca Examinadora, cujos membros estão listados abaixo. A aluna foi **APROVADA** pela Banca Examinadora, por unanimidade, em cumprimento ao requisito exigido para obtenção do Título de Mestre em Meteorologia.

A banca sugere incorporar as observações dos membros na dissertação, bem como mudança do título para:

Título: "APRIMORAMENTO DO MÉTODO DE CLASSIFICAÇÃO DIURNA DE NUVENS DO CPTEC UTILIZANDO O SENSOR ABI/GOES-16."

O texto final deverá ser conferido pelos orientadores e ser entregue no prazo estipulado pelo regimento do curso.

Eu, Simone Marilene Sievert da Costa Coelho como Presidente da Banca Examinadora, assino esta ATA em nome de todos os membros.

Membros da Banca

Simone Marilene Sievert da Costa Coelho - **Presidente** (CPTEC/INPE)

Daniel Alejandro Vila - **Orientador** (CPTEC/INPE)

Renato Galante Negri - **Orientador** (CPTEC/INPE)

Juan Ceballos - **Membro da banca** (CPTEC/INPE)

Anthony Carlos Silva Porfirio - **Membro da banca externo** (FUNCEME)



Documento assinado eletronicamente por **Simone Marilene Sievert da Costa Coelho, Pesquisadora**, em 31/05/2021, às 13:41 (horário oficial de Brasília), com fundamento no art. 6º do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site <http://sei.mctic.gov.br/verifica.html>, informando o código verificador



7497511 e o código CRC **E9740511**.

Referência: Processo nº 01340.003000/2021-14

SEI nº 7497511

*“To find yourself in the infinite,
You must distinguish and then combine;
Therefore my winged song thanks
The man who distinguished cloud from cloud.”*

GOETHE

*Aos meus pais **Dalvani Gomes e Gilmar de
Carvalho** e ao meu irmão **Gildeni Gomes**.*

AGRADECIMENTOS

A meus pais Dalvani Gomes e Gilmar de Carvalho, por todo apoio e incentivo e pelo esforço dedicado e por serem o meu porto seguro.

Ao meu irmão Gildeni Gomes pelo apoio e companheirismo.

Aos meus amigos Erykson Markony, Isabella Talamoni, Lorena Martina, Lívia, João Pedro, Yan, Marie, Dayanis, Geovanine Araújo, Cláudio Moisés, Eduardo Cho-Luck, Agostinho, Glícia, Ivo, Amanda Teotônio, Tarsila Martins, Manuel, Nelson, Eliseu, João, Laurízio, Tiago, Gabriel, Bianca, Luciano pela amizade.

Agradeço de coração aos meus orientadores Dr. Daniel Alejandro Vila e ao Dr. Renato Galante Negri, pela orientação, amizade, paciência, disponibilidade, suporte e todo conhecimento que obtive durante todo o período do mestrado.

Ao meu amigo Jussier Dantas pela amizade e pela ajuda com o texto.

A Taísa Pereira e Jaime Fernando que foram essenciais nessa jornada, dando apoio moral e pelas conversas e risadas que ajudaram a esquecer momentaneamente alguns problemas.

Gostaria de agradecer em especial ao Jaime Fernando, por ser um amigo sempre presente e disposto a ajudar sempre que precisei.

Aos professores Dra. Simone M. S. da Costa Coelho, Dr. Juan C. Ceballos, Dr. Anthony C. S. Porfírio pelas críticas e sugestões.

A todos que de forma direta ou indireta tenham contribuído para que este trabalho fosse realizado.

Ao CNPq pela concessão de bolsa de estudos, sob o processo número 130558/2019-2.

RESUMO

As nuvens desempenham um papel importante nos processos atmosféricos envolvendo o espalhamento e a absorção da radiação. Neste contexto, a classificação adequada das nuvens é fundamental para o entendimento de tais processos e os satélites geoestacionários são ferramentas valiosas. O objetivo deste estudo foi avaliar e aprimorar o classificador de nuvens mantido em operação pelo Centro de Previsão de Tempo e Estudos Climáticos do Instituto Nacional de Pesquisas Espaciais (CPTEC/INPE) e desenvolver um classificador novo explorando os novos canais presentes no sensor ABI/GOES-16. Imagens do sensor ABI/GOES-16 do período entre janeiro de 2019 a dezembro de 2020 foram utilizadas. O classificador operacional foi avaliado e aprimorado nos três horários disponíveis, 12:00, 15:00 e 18:00 UTC utilizando imagens dos canais em 0,64 e 10,3 μm do sensor ABI/GOES-16. Somado a isso, o classificador foi comparado à três produtos de nuvens disponibilizados pela NOAA: temperatura do topo da nuvem, fase da nuvem e máscara de nuvem visando observar o comportamento dos classificadores às características termodinâmicas das nuvens. Foi observado que alguns alvos tiveram uma melhora na identificação de aproximadamente 3%, como é o caso das nuvens cirrus. Outros alvos, como as nuvens cumulus, tiveram identificação semelhantes em ambas as versões do classificador. Devido à grande quantidade de parâmetros possíveis para o desenvolvimento do novo classificador, foi aplicada a técnica PCA para redução dos parâmetros, resultando em 12 (canais 0,64, 1,37, 6,2 8,4, 10,3 μm , texturas dos canais 0,64, 6,2, 10,3 μm e as diferenças espectrais 6,2-10,3, 11,2-12,3, 11,2-3,9 e 3,9-12,3 μm). Na comparação com os produtos da NOAA, foi verificada uma alta coerência das características termodinâmicas das nuvens, observando uma melhora de até 15% na categorização de nuvens cirrus, stratus e multicamadas. Por outro lado, a identificação das superfícies continentais e oceânicas foi semelhante àquelas dos classificadores operacional e atualizado nos três horários. O novo classificador baseado em 12 parâmetros, apresentou melhoras em relação ao original baseado em 4 parâmetros, porém ajustes ainda são necessários.

Palavras-chave: Classificação de nuvens. GOES-16/ABI. K-Means. Métodos de agrupamento.

IMPROVEMENT OF THE CPTEC DIURNAL CLOUD CLASSIFICATION METHOD USING THE ABI/GOES-16 SENSOR

ABSTRACT

Clouds play an important role in atmospheric processes involving the scattering and absorption of radiation. In this context, the proper cloud classification is essential for understanding such processes and geostationary satellites are valuable tools. The objective of this study was to evaluate and improve the cloud classifier kept in operation by the Center for Weather Forecasting and Climate Studies of National Institute for Space Research (CPTEC/INPE) and develop a new classifier exploring new channels in ABI/GOES-16 sensor. Images from ABI/GOES-16 sensor were used for the period between January 2019 and December 2020. The operational classifier was evaluated and improved at the three available times, 12:00, 15:00, and 18:00 UTC images of channels 0.64 and 10.3 μm from ABI/GOES-16 sensor was used. In addition, the classifiers were compared to three cloud products provided by NOAA: cloud top temperature, cloud top phase, and cloud mask in order to observe the classifiers' behavior to the thermodynamic characteristics of clouds. It was observed that some targets had an improvement in the identification of approximately 3%, as is the case of cirrus clouds. Other targets, such as cumulus clouds, had similar identification in both classifiers. Due to many possible parameters to develop the new classifier, the PCA technique was applied to reduce the parameters, resulting in 12 (channels 0.64, 1.37, 6.2 8.4, 10.3 μm , texture of the channels 0.64, 6.2, 10.3 μm , and the spectral differences 6.2-10.3, 11.2-12.3, 11.2-3.9 and 3.9-12.3 μm). In comparison with NOAA products, it was verified a high consistency of the clouds' thermodynamic characteristics, observing an improvement up to 15% in the categorization of cirrus, stratus, and multilayer clouds. On the other hand, the surface continental and oceanic identification was similar to the operational and update classifiers in the three times available. The new classifier based on 12 parameters, presented an improvement to the original classifiers, based on 4 parameters, however, adjustments are still needed.

Keywords: Cloud Classification. GOES-16/ABI. K-Means. Clustering methods.

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Classificação das nuvens nos três níveis atmosféricos.	6
2.2 Índices de refração do gelo e da água através da janela atmosférica: a) parte real; b) parte imaginária.	8
2.3 Ilustração da técnica <i>k-means</i> para 3 grupos utilizando dados em 2 dimensões.	13
3.1 Área escolhida para o desenvolvimento dos classificadores de nuvens. . .	15
3.2 Classificação de nuvens operacional do CPTEC/INPE para o dia 25/03/2020 às 15 UTC.	16
3.3 Funções resposta espectrais para as 16 bandas do sensor ABI/GOES-16.	20
3.4 Produto oficial da altura do topo da nuvem (m) para o dia 19/01/2019 em 15:00 UTC.	22
3.5 Produto oficial de profundidade óptica para o dia 19/01/2019 em 15:00 UTC.	23
3.6 Produto oficial da pressão do topo da nuvem (mb) para o dia 19/01/2019 em 15:00 UTC.	24
3.7 Produto oficial da temperatura do topo da nuvem (K) para o dia 19/01/2019 em 15:00 UTC.	25
3.8 Produto oficial da fase da nuvem para o dia 19/01/2019 em 15:00 UTC. .	26
3.9 Produto oficial da máscara de nuvem para o dia 19/01/2019 em 15:00 UTC.	27
4.1 Classificação nuvens de topo muito frio, 21/03/2019, 15:00 UTC.	36
4.2 Classificação de nuvens cirrus e multicamadas, 21/03/2019, 15:00 UTC. .	37
4.3 Classificação de nuvens cumulus e stratus, 28/04/2019, 15:00 UTC. . . .	38
4.4 Carta sinótica do dia 03/03/2020 no horário 12:00 UTC.	44
4.5 Carta de níveis altos para o dia 03/03/2020.	45
4.6 Carta de níveis médios para o dia 03/03/2020.	46
4.7 Imagem do visível e infravermelho no dia 03/03/2020 no horário 15:00 UTC.	47
4.8 Tipos de nuvens entre o classificador operacional (baseado no GOES-8) e o classificador atualizado (baseado no GOES-16) de todas as classes (cima) e classes agrupadas (baixo) no horário das 12:00 UTC do dia 03/03/2020.	49

4.9	Tipos de nuvens entre o classificador operacional (baseado no GOES-8) e o classificador atualizado (baseado no GOES-16) e o classificador atualizado de todas as classes (cima) e classes agrupadas (baixo) no horário das 15:00 UTC do dia 03/03/2020.	50
4.10	Tipos de nuvens entre o classificador operacional (baseado no GOES-8) e o classificador atualizado (baseado no GOES-16) e o classificador atualizado de todas as classes (cima) e classes agrupadas (baixo) no horário das 18:00 UTC do dia 03/03/2020.	51
4.11	Fração de pixels com classificação discordante entre os classificadores original (GOES-8) e atualizado (GOES-16) para o intervalo centrado em 12:00 UTC.	53
4.12	Diferença entre a classe cumulus do classificador operacional (GOES-8) e a classe superfície do classificador atualizado (GOES-16) no horário das 12:00 UTC do dia 03/03/2020.	54
4.13	Exemplo de um ponto onde houve diferença entre as classes do classificador operacional (baseado no GOES-8) e atualizado (baseado no GOES-16) no horário das 12:00 UTC do dia 03/03/2020.	56
4.14	Fração de pixels com classificação discordante entre os classificadores operacional (baseado no GOES-8) e atualizado (baseado no GOES-16) para o intervalo centrado em 15:00 UTC.	57
4.15	Diferença entre a classe cumulus do classificador operacional (GOES-8) e a classe stratus do classificador atualizado (GOES-16), no horário das 15:00 UTC do dia 03/03/2020.	58
4.16	Ponto de latitude e longitude onde houve diferença entre as classes do classificador operacional (baseado no GOES-8) e atualizado (baseado no GOES-16) no horário das 15:00 UTC do dia 03/03/2020.	60
4.17	Fração de pixels com classificação discordante entre os classificadores operacional (baseado no GOES-8) e atualizado (baseado no GOES-16) para o intervalo centrado em 18:00 UTC.	61
4.18	Diferença entre a classe superfície do classificador operacional (baseado no GOES-8) e a classe cirrus do classificador atualizado (baseado no GOES-16), no horário das 18:00 UTC do dia 03/03/2020.	62
4.19	Ponto de latitude e longitude onde houve diferença entre as classes do classificador operacional (baseado no GOES-8) e atualizado (baseado no GOES-16) no horário das 18:00 UTC do dia 03/03/2020.	64
4.20	Pontos selecionados para serem os novos centroides iniciais a partir das imagens do dia 21/03/2019 15:00 UTC.	76
4.21	Matriz de correlação dos canais ABI/GOES-16.	78

4.22	Matriz de correlação das texturas dos respectivos canais ABI/GOES-16.	79
4.23	Matriz de correlação dos canais, texturas e as diferenças entre as temperaturas de brilho.	81
4.24	Imagens classificadas pelo classificador operacional (cima) e atualizado (baixo) para o dia 03/03/2020 no horário 15:00 UTC.	91
4.25	Imagem do visível e infravermelho e imagem classificada para o dia 03/03/2020 no horário 12:00 UTC.	92
4.26	Imagem do visível e infravermelho e imagem classificada para o dia 03/03/2020 no horário 15:00 UTC.	93
4.27	Imagem do visível e infravermelho e imagem classificada para o dia 03/03/2020 no horário 18:00 UTC.	94
4.28	Imagens classificadas das classes agrupadas do classificador operacional (esquerda), classificador atualizado (direita) e novo classificador (baixo) para o dia 03/03/2020 no horário 15:00 UTC.	95

LISTA DE TABELAS

	<u>Pág.</u>
1.1 Classificação das nuvens de acordo com a altura da base.	1
3.1 Atributos espectrais do ABI, como canal, comprimento de onda central, resolução espacial e a descrição.	18
3.2 Parâmetros multiespectrais de entrada para o classificador, utilizando canais do GOES-16.	33
4.1 Centroides calculados com o conjunto de treinamento para o classificador operacional (baseado no GOES-8) e classificador atualizado (baseado no GOES-16) no horário das 12:00 UTC.	41
4.2 Centroides calculados com o conjunto de treinamento para o classificador operacional (baseado no GOES-8) e classificador atualizado (baseado no GOES-16) no horário das 15:00 UTC.	42
4.3 Centroides calculados com o conjunto de treinamento para o classificador operacional (baseado no GOES-8) e classificador atualizado (baseado no GOES-16) no horário das 18:00 UTC.	43
4.4 Distribuição média entre as classes de nuvens e o produto temperatura (K) do topo da nuvem da NOAA para os classificadores operacional (baseado no GOES-8) e o atualizado (baseado no GOES-16) para os intervalos de horário centrados em 12:00, 15:00 e 18:00 UTC.	68
4.5 Distribuição média entre as classes de nuvens e o produto fase termodinâmica da nuvem da NOAA para os classificadores operacional (baseado no GOES-8) e o atualizado (baseado no GOES-16) para os intervalos de horário centrados em 12:00, 15:00 e 18:00 UTC.	71
4.6 Distribuição média entre as classes de nuvens e o produto máscara da nuvem da NOAA para os classificadores operacional (baseado no GOES-8) e o atualizado (baseado no GOES-16) para os intervalos de horário centrados em 12:00, 15:00 e 18:00 UTC.	73
4.7 Diferença entre as temperaturas de brilho utilizadas na aplicação da técnica PCA.	80
4.8 Tabela resultante do treinamento das novas classes e parâmetros para 12:00 UTC.	84
4.9 Tabela resultante do treinamento das novas classes e parâmetros para 15:00 UTC.	86

4.10	Tabela resultante do treinamento das novas classes e parâmetros para 18:00 UTC.	88
4.11	Distribuição da temperatura do topo da nuvem em K em relação às classes de nuvens para as variantes 12:00, 15:00 e 18:00 UTC do novo classificador baseado em 12 parâmetros.	98
4.12	Distribuição da fase termodinâmica do topo da nuvem em relação às classes de nuvens para os horários 12:00, 15:00 e 18:00 UTC do novo classificador baseado em 12 parâmetros.	101
4.13	Distribuição da cobertura de nuvens em relação às classes de nuvens nos horários 12:00, 15:00 e 18:00 UTC do novo classificador baseado em 12 parâmetros	102

LISTA DE ABREVIATURAS E SIGLAS

AB	–	Alta da Bolívia
ABI	–	Advanced Baseline Imager
ACHA	–	ABI Cloud Height Algorithm
AVIRIS	–	Airborne Visible/Infrared Imaging Spectrometer
CMIP	–	Cloud Moisture Imagery Product
CONUS	–	Contínuo US
CPTEC	–	Centro de Previsão de Tempo e Estudos Climáticos
DISSM	–	Divisão de Satélites e Sensores Meteorológicos
DSA	–	Divisão de Satélites e Sistemas Ambientais
dTB	–	Diferença entre as temperaturas de brilho
ER2	–	Earth Resources 2
FY2D	–	Fengyun-2D
GIBS	–	Global Imagery Browse Service
GOES	–	Geostationary Operational Environmental Satellite
HIS	–	High Spectral Resolution Sounder
INPE	–	Instituto Nacional de Pesquisas Espaciais
JP	–	Jato Polar
JST	–	Jato Subtropical
MODIS	–	Moderate Resolution Imaging Spectroradiometer
MSG	–	Meteosat Second Generation
NASA	–	National Aeronautics and Space Administration
NDVI	–	Índice de Vegetação da Diferença Normalizada
NOAA	–	National Oceanic and Atmospheric Administration
PCA	–	Análise de Componentes Principais
SEVERI	–	Spinning Enhanced Visible and Infrared Imager
SST	–	Temperatura da Superfície do Mar
UTC	–	Universal Time Coordinated
VCAN	–	Vórtice Ciclônico de Altos Níveis
ZCAS	–	Zona de Convergência do Atlântico Sul
ZCIT	–	Zona de Convergência Intertropical
WMO	–	Organização Meteorológica Mundial

LISTA DE SÍMBOLOS

$^{\circ}\text{C}$	– grau Celcius
cb	– cumulonimbus
ci	– cirrus
ci-st	– cirrus semitransparentes
cu	– cumulus
d_{ik}	– distância Euclidiana
H_0	– hipótese nula
H_1	– hipótese alternativa
I	– total de variáveis
i	– número da i -ésima variável
C	– quantidade total de centroides
K	– temperatura em Kelvin
km	– quilômetro
M	– dimensão dos dados
mc	– multicamadas
μm	– micrômetro
n	– número de repetições
nm	– nanômetro
st	– stratus
x_{ik}	– variável no k -ésimo cluster
x_{jk}	– centroide do k -ésimo cluster

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
1.1 Objetivos	4
1.1.1 Objetivo geral	4
1.1.2 Objetivos específicos	4
2 FUNDAMENTAÇÃO TEÓRICA E REVISÃO DE LITERATURA	5
2.1 Nuvens	5
2.2 Classificação de nuvens	7
2.3 O método de agrupamento <i>K-Means</i>	10
3 DADOS E METODOLOGIA	15
3.1 Área de estudo	15
3.2 Satélite GOES-16	16
3.3 Produtos de topo de nuvem GOES-16/NOAA	21
3.4 Método <i>K-Means</i>	27
3.5 Classificadores	28
3.5.1 Modelo operacional	29
3.5.2 Atualizando o classificador a partir dos dados ABI/GOES-16	30
3.5.3 O novo classificador	30
4 RESULTADOS	35
4.1 Análise do classificador original	35
4.2 Revisão do classificador baseado em quatro parâmetros	39
4.3 Comparação entre os produtos da NOAA e os classificadores operacional (baseado no GOES-8) e atualizado (baseado no GOES-16)	64
4.3.1 Temperatura do topo da nuvem	65
4.3.2 Fase da nuvem	69
4.3.3 Máscara de nuvem	72
4.4 Criação de um novo classificador	74
4.4.1 Definição das sementes iniciais	74
4.4.2 Seleção dos parâmetros mais representativos	76
4.4.3 Treinando o classificador a partir do método K-Means	82
4.4.4 Avaliação do novo classificador	95

4.4.4.1	Temperatura do topo da nuvem	96
4.4.4.2	Fase da nuvem	98
4.4.4.3	Máscara de nuvem	101
5	CONCLUSÕES	105
5.1	Sugestões de trabalhos futuros	108
	REFERÊNCIAS BIBLIOGRÁFICAS	109

1 INTRODUÇÃO

As nuvens possuem um importante papel nos processos atmosféricos, que compõem o sistema climático. Por meio da análise da fração de cobertura e classificação de nuvens, podemos estudar o clima e fazer análises meteorológicas. A porção de nuvens que cobre o céu tem diferentes influências no sistema superfície-atmosfera, pois são grandes absorvedores de onda longa e refletores da radiação solar direta (STRABALA et al., 1994; XIA et al., 2015), contribuindo no saldo de energia e com o albedo planetário. Informações sobre o perfil vertical de nuvens também são essenciais na avaliação de modelos de circulação geral. Além disso, há também o impacto direto na vida humana, pois a quantidade e o tipo de nuvens exercem influência, por exemplo, na agricultura e geração de energias renováveis (BOTTINO; CEBALLOS, 2015).

As nuvens são classificadas de acordo com a altura da base, podendo ser baixa, média e alta. Além disso possui 10 subdivisões, como mostra a Tabela 1.1, são elas: cumulus, stratus, stratocumulus, nimbostratus, altostratus, altocumulus, cirrus, cirrostratus, cirrocumulus. Embora a cumulonimbus esteja classificada como sendo uma nuvem baixa, devido à altura da base, ela é mais conhecida como uma nuvem do tipo multicamada, pois se estende por toda a troposfera (HOUZE JUNIOR, 1993; JIN et al., 2014).

Tabela 1.1 - Classificação das nuvens de acordo com a altura da base.

Altura da Base	Nuvem
Baixa	Cumulus
	Cumulonimbus
	Stratus
	Stratocumulus
	Nimbostratus
Média	Altostratus
	Altocumulus
Alta	Cirrus
	Cirrostratus
	Cirrocumulus

Fonte: Produção da autora.

A identificação do tipo de nuvem ocorre basicamente de duas maneiras: a primeira, por meio de um observador e a outra mediante instrumentos, notadamente aqueles associados às técnicas de sensoriamento remoto, a exemplo dos satélites ou radares. O observador classifica as nuvens a partir da interpretação visual da característica e formato, dessa forma, a classificação correta está diretamente ligada à experiência do observador (ESCRIG et al., 2013; JIN et al., 2014). Contudo, se há uma porção do céu que está coberta por nuvens baixas e espessas, o observador é incapaz de identificar a existência de nuvens que estejam acima. Outro problema é a identificação de nuvens no período noturno. Nesse sentido, a utilização de instrumentos a bordo de satélites é de grande ajuda na identificação das nuvens principalmente com relação aos problemas citados acima.

A identificação das nuvens a partir de técnicas de sensoriamento remoto pode ser realizada por meio de instrumentos em superfície (radares) ou por instrumentos a bordo de satélites (imageadores). Os satélites se tornaram uma ferramenta de grande importância na estimativa da cobertura do céu e dos tipos de nuvem por possuírem várias bandas espectrais (canais) que possibilitam a identificação das nuvens de forma muito precisa. A partir da informação de um único canal desses instrumentos, bem como a partir de combinações de dois ou mais canais, é possível realizar a classificação das nuvens de acordo com as características a serem estudadas (FREY et al., 2008; BOTTINO; CEBALLOS, 2015).

Nesse contexto, os satélites geoestacionários são uma ótima ferramenta para acompanhar o desenvolvimento de sistemas meteorológicos, bem como seu trajeto e sua dissipação. Isso é possível devido à alta resolução temporal desses instrumentos, dessa maneira, a frequência entre imagens sucessivas é inferior ao tempo de vida médio das tempestades convectivas. Um exemplo é o GOES-16 (*Geostationary Operational Environmental Satellite - R Series*), que gera, atualmente, uma imagem *full disk* a cada 10 minutos (PAVOLONIS, 2010).

A tendência das novas gerações de satélites meteorológicos em órbita geoestacionária é reduzir ainda mais o intervalo entre as observações. Satélites em órbita polar ou equatorial possuem uma resolução temporal mais baixa que os geoestacionários. Contudo, há a possibilidade de serem equipados com sensores na banda de micro-ondas, permitindo inferir propriedades internas das nuvens como a presença de cristais de gelo, granizo ou gotas de água suficientemente grandes para produzir precipitação. Atualmente, a disposição dos satélites orbitando a Terra nos permite

uma cobertura muito ampla possibilitando estudos em diversas áreas, tais como uso de solo, aerossóis, classificação de nuvens, dentre outras (BOTTINO; CEBALLOS, 2015; LI et al., 2019).

As nuvens possuem um papel importante no aquecimento e resfriamento da temperatura do planeta e algumas de suas propriedades são utilizadas nos estudos de mudanças climáticas, monitoramento por sensoriamento remoto, previsão numérica de tempo, e assim por diante. Alguns parâmetros, tais como temperatura do topo, desenvolvimento vertical, espessura óptica, conteúdo de água líquida (JIN et al., 2014; BOTTINO; CEBALLOS, 2015; LI et al., 2019; OUKALI et al., 2019), são importantes na identificação de nuvens por imagens de satélite e determinados por meio dos canais que compõem o sensor. As imagens de satélite possuem um papel muito importante no que se refere à previsão do tempo, desastres naturais, incêndios florestais, além do monitoramento de eventos severos como tempestades (JIN et al., 2014).

De acordo com Jin et al. (2014) os algoritmos de reconhecimento do pixel são baseados em contrastes, ou seja, nuvem-superfície, para uma dada área de interesse. Como as nuvens possuem diferentes respostas em diferentes bandas espectrais, muitos estudos se utilizam da combinação de diferentes comprimentos de onda para a detecção de nuvens com diferentes propriedades, pois cada canal representa uma determinada característica. Por exemplo, Strabala et al. (1994) mostraram que a combinação tri-espectral pode diferenciar nuvens que contêm água de nuvens com gelo. Na época do referido estudo foi possível somente demonstrar que era possível utilizar esses comprimentos de onda para diferenciar esses tipos de nuvens, devido às diferentes propriedades de absorção e refração da água e do gelo. Atualmente é possível realizar esse tipo de separação, pois há satélites geoestacionários em órbita que possuem instrumentos acoplados com esses comprimentos de onda, um deles é o sensor *Advanced Baseline Imager* acoplado ao satélite GOES-16.

No contexto de métodos para a classificação de imagens de satélite, podemos destacar alguns trabalhos realizados que mostram a importância da classificação de nuvens, como é o caso de Bottino e Ceballos (2015), trabalho que resultou no produto que está em operação no Centro de Previsão de Tempo e Estudos Climáticos (CPTEC)/ Instituto Nacional de Pesquisas Espaciais (INPE). Há estudos que utilizam diferentes comprimentos de onda de sensores a bordo de satélites (geoestacionários ou de órbita polar), a fim de identificar alguns tipos de nuvens por meio das diferentes características que cada nuvem apresenta. A identificação dos tipos de nuvem, ajuda

na identificação e classificação de sistemas precipitantes (ACKERMAN et al., 1998; FREY et al., 2008; THIES et al., 2008). Esses estudos evidenciam a grande importância de utilizar algoritmos associados às imagens de satélite na detecção de regiões com ocorrência de chuvas convectivas.

1.1 Objetivos

1.1.1 Objetivo geral

Diante do exposto, o objetivo geral deste trabalho é avaliar e aprimorar o classificador de nuvens atualmente mantido pela Divisão de Satélites e Sensores Meteorológicos (DISSM), antiga Divisão de Satélites e Sistemas Ambientais (DSA), do Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), para toda a região da América do Sul a partir do sensor *Advanced Baseline Imager* ABI/GOES-16.

1.1.2 Objetivos específicos

- Revisar o método desenvolvido por (BOTTINO; CEBALLOS, 2015);
- Aprimorar o método, adequando-o aos padrões do ABI/GOES-16;
- Avaliação do método desenvolvido por (BOTTINO; CEBALLOS, 2015) a três produtos de nuvens fornecidos pela NOAA;
- Criação de um novo classificador inteiramente baseado nas informações do sensor ABI/GOES-16, incorporando maiores detalhes espectrais;
- Validação do novo classificador desenvolvido.

2 FUNDAMENTAÇÃO TEÓRICA E REVISÃO DE LITERATURA

No capítulo anterior, introduziu-se de forma breve o conceito de nuvens, como identificá-las a partir de satélites e da importância dessas informações para identificar regiões com recorrência de precipitação. Neste capítulo serão expostos trabalhos que utilizaram diferentes técnicas a fim de identificar alguns tipos de nuvens.

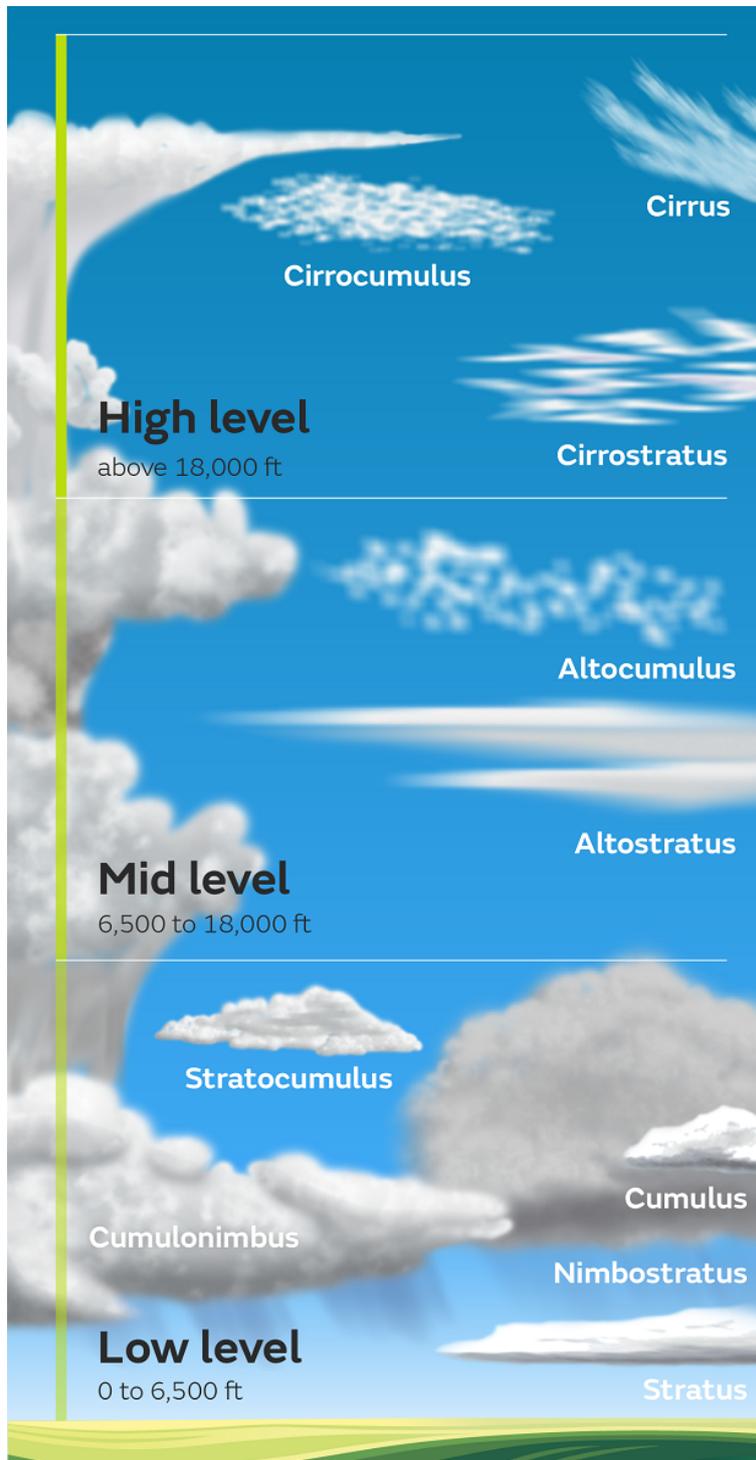
2.1 Nuvens

De acordo com a WMO (1975), nuvens são definidas como sendo um aglomerado de partículas de água e/ou gelo em suspensão na atmosfera, além disso, outras partículas maiores também podem fazer parte desse aglomerado, como partículas líquidas ou sólidas não aquosas presentes em poeira ou fumaça.

As nuvens foram primeiramente classificadas por Howard (1803), que descreveu seus movimentos e formas e, além disso, nomeou algumas estruturas utilizadas até os dias atuais. Posteriormente, a Organização Meteorológica Mundial (WMO, sigla em inglês para *World Meteorological Organization*) aumentou a classificação para 10 grupos conhecidos como genera, que se baseia na estrutura, formato e transparência das nuvens. Esses grupos estão ainda subdivididos em 3 níveis diferentes, como mostra a Tabela 1.1. As nuvens possuem um papel importante no sistema climático terrestre, seja regulando o balanço de energia, a precipitação, além de dispersar poluentes em suspensão na atmosfera (RASCHKE, 1988; GOODMAN; HENDERSON-SELLERS, 1988; STRABALA et al., 1994; XIA et al., 2015).

A Figura 2.1 apresenta a classificação das nuvens de acordo com os diferentes níveis atmosféricos, na qual as nuvens baixas têm sua base não mais que 2 km de distância da superfície terrestre, as nuvens médias possuem bases entre 2 e 6 km de distância da superfície e as nuvens altas possuem bases acima de 6 km de distância da superfície (HOUZE JUNIOR, 1993; JIN et al., 2014; WMO, 2017; METOFFICE, n.d.).

Figura 2.1 - Classificação das nuvens nos três níveis atmosféricos.



Fonte: Met Office (2020).

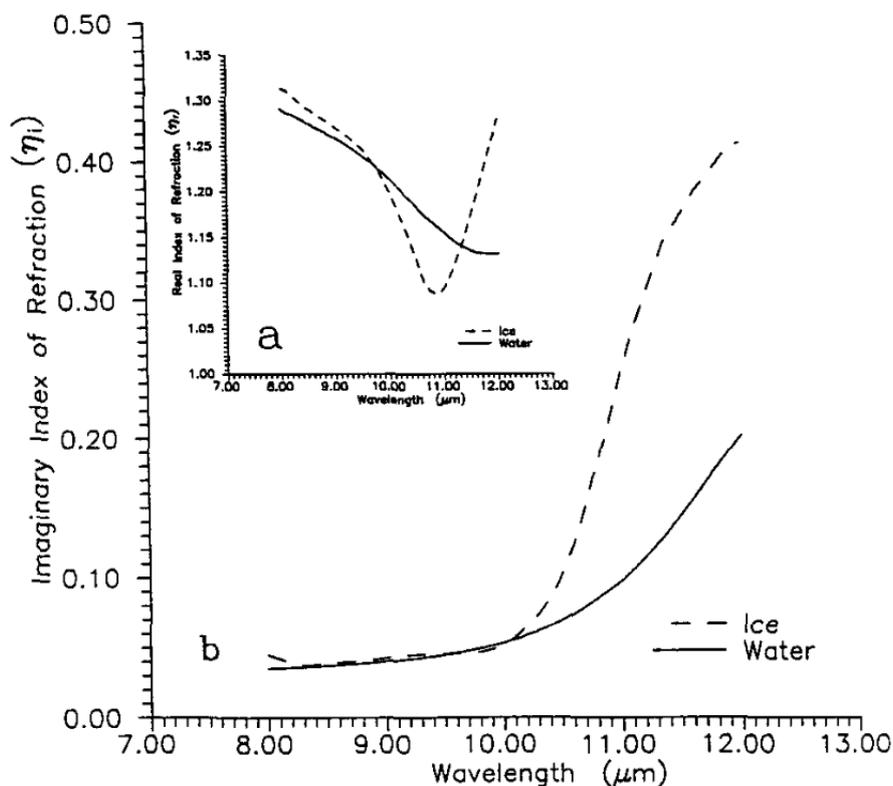
2.2 Classificação de nuvens

Diversos trabalhos utilizam dados de imagens de satélite para estudar propriedades e classificação de nuvens e intensidade da precipitação (STRABALA et al., 1994; THIES et al., 2008; ESCRIG et al., 2013; JIN et al., 2014; BOTTINO; CEBALLOS, 2015; OUKALI et al., 2019).

Strabala et al. (1994) utilizaram os dados do instrumento HIS (*High Spectral Resolution Sounder*) acoplado a uma aeronave operada pela NASA (*National Aeronautics and Space Administration*) chamada ER2 (*Earth Resources 2*) para estudar as propriedades de nuvens. Essa pesquisa baseia-se na técnica tri-espectral, que consiste na interpretação do diagrama de espalhamento das temperaturas de brilho na diferença resultante entre 8-11 μm e 11-12 μm na qual valores acima de 1 K representam nuvens de gelo e valores menores que 1 K representam nuvens de água. Dessa forma, foi possível identificar nuvens constituídas majoritariamente de água e nuvens majoritariamente de gelo, além das nuvens com mistura de fase.

O estudo de Strabala et al. (1994) é baseado na diferença entre as propriedades das nuvens de gelo e nuvens de água, já que sua microfísica se difere entre as diferenças dos canais 8-11 μm e 11-12 μm . Como pode ser observado na Figura 2.2, a absorção e a emissão das nuvens dependem do índice de refração da partícula, que se divide em parte real e imaginária. A parte real representa a magnitude do espalhamento e a parte imaginária representa a absorção, a linha pontilhada corresponde ao gelo e a linha contínua corresponde a água. É possível notar a diferença no comportamento do gelo e da água, principalmente na parte imaginária, na qual essa diferença é bem mais perceptível entre os comprimentos de onda 10 μm e 12 μm .

Figura 2.2 - Índices de refração do gelo e da água através da janela atmosférica: a) parte real; b) parte imaginária.



Fonte: Strabala et al. (1994).

No trabalho de Thies et al. (2008) foram utilizados dados do satélite geostacionário MSG/SEVIRI (*Meteosat Second Generation/Spinning Enhanced Visible and Infrared Imager*) com a finalidade de estudar a intensidade da precipitação, além de dados de radar para diferenciar a intensidade da precipitação durante o período de janeiro a agosto de 2004. Foram propostos 5 cenários, com intensidades que variam desde precipitação muito intensa até precipitação pouco intensa.

Isso foi possível devido à utilização de algumas combinações dos canais do *Meteosat Second Generation* (MSG), permitindo realizar o estudo durante o período diurno e noturno. Dessa forma, os autores puderam analisar a precipitação proveniente de sistemas convectivos, bem como a precipitação proveniente de nuvens estratiformes.

Escrig et al. (2013) utilizaram imagens de cinco canais do satélite geostacionário MSG, além de uma câmera que fotografa o céu, com a finalidade de checar a cobertura e a direção do movimento das nuvens. Os autores concluíram que, embora a comparação entre os dados de satélite e a utilização da câmera tenha se mostrado satisfatória, os resultados obtidos não foram os ideais. Em situações com ausência de nuvens ou cobertura total do céu, essa concordância foi de 90%. Contudo, em situações com cobertura parcial do céu essa concordância caiu para 75%. Para estimar o movimento das nuvens foi utilizado o método de correlação binária cruzada e os resultados mostraram 86% de acerto e as nuvens altas contribuíram significativamente nos erros encontrados.

A técnica de representação espaçada, é um tipo de inteligência artificial utilizada no reconhecimento de padrões em imagens para a classificação de nuvens, foi utilizada por Jin et al. (2014). Os autores propuseram 6 diferentes cenários: céu claro (superfície e oceano), nuvens baixas (nimbostratus e cumulus), nuvens médias (altostratus e altocumulus), nuvens altas (cirrostratus e cirrus densa) e nuvens multicamadas (cumulonimbus). As imagens utilizadas pelos autores foram referentes ao período diurno de 10 dias de observações do satélite geostacionário chinês FY2D (*Feng Yun-2D*). Foi concluído que o método empregado obteve 88,7% de acerto no geral, com destaque para a classificação das nuvens altostratus e altocumulus, nimbostratus e cumulus e cumulonimbus, com 90,6%, 93,3% e 97,8% respectivamente. Contudo, as nuvens cirrostratus e cirrus densa não obtiveram o mesmo resultado satisfatório, em razão dessas nuvens coexistirem com nuvens baixas e médias.

Frey et al. (2008) utilizaram algumas diferenças espectrais do *Moderate Resolution Imaging Spectrometer* (MODIS) para detectar alguns tipos de nuvens na região polar, por exemplo: nuvens altas, nuvens baixas e cirrus semi-transparentes. As diferenças espectrais utilizadas na detecção dessas nuvens foram: 3,9-12 μm para nuvens altas, 11-3,9 μm para nuvens baixas e 11-12 μm para cirrus semi-transparentes.

O estudo desenvolvido por Gao et al. (1993) mostra que nuvens cirrus poderiam ser detectadas utilizando um comprimento de onda no qual o comprimento central estivesse em torno de 1,38 μm . Os autores utilizaram imagens espectrais do AVIRIS (sigla em inglês para *Airborne Visible/Infrared Imaging Spectrometer*) a uma altura de 20 km. O método de detecção de nuvens cirrus deve-se à característica que este comprimento de onda possui, que é a forte absorção do vapor d'água. Quando há presença de nuvens cirrus o espalhamento recebido pelo sensor é muito maior do que

na ausência destas.

Como visto até o momento, métodos para classificação das nuvens são utilizados em diversos estudos, como estimar precipitação e a fração de cobertura do céu.

O trabalho de Bottino e Ceballos (2015) teve como objetivo analisar os aspectos das informações fornecidas pelo imageador a bordo do GOES-8, tendo seu foco direcionado para a identificação do tipo de nuvem no sul da América do Sul (43° S - 20° S e 76° W - 35° W). O método utilizado foi o agrupamento de *cluster* não supervisionado (*k-means*), objetivando identificar 6 cenários. Para determinar essas classes foram utilizadas 13 variáveis, possibilitando diferenciar nuvens com grande desenvolvimento vertical das mais finas, bem como separar nuvens estratiformes de cumulus. Dois grupos de imagens de setembro de 2002 foram utilizados. Um grupo correspondente ao treinamento e outro para validar o método. Um total de 32 centroides foram classificados a partir de uma análise visual do conjunto de treinamento. Os autores concluíram que a classificação automática de nuvens, por meio da aplicação da técnica, obteve um desempenho superior a 90%.

Nessa perspectiva, o presente estudo fundamenta-se na metodologia utilizada por Bottino e Ceballos (2015) para a classificação de nuvens, além de trabalhos que empregam alguns comprimentos de onda, bem como algumas diferenças das temperaturas de brilho de alguns sensores a bordo em satélites, podendo ser citados (STRABALA et al., 1994; ACKERMAN et al., 1998; THIES et al., 2008; FREY et al., 2008; OUKALI et al., 2019).

Como foi exposto na maior parte dos trabalhos citados acima, a utilização dos satélites geostacionários é de suma importância, pois a alta frequência na disponibilização dos dados permite a observação contínua da cobertura e tipo de nuvem, abrangendo toda a área de cobertura do satélite.

2.3 O método de agrupamento *K-Means*

As técnicas de agrupamento vem sendo bastante utilizada em diferentes áreas de estudo, como por exemplo, na medicina, psicologia, linguística, engenharia (GONG; RICHMAN, 1995; WEBB, 2003; JAIN, 2010) e ciências atmosféricas (CHU et al., 2012; BOTTINO; CEBALLOS, 2015), aplicando-se principalmente para o reconhecimento de padrões, processamento de imagens e aprendizado de máquina (JAIN et al., 1999;

XIE; JIANG, 2010). MacQueen (1967) descreve que o método *k-means* constitui-se na escolha de k grupos no começo do processo, no qual cada um deles consiste em um único membro escolhido aleatoriamente. Ao longo do processo novos membros são adicionados ou excluídos, ocasionando o seu ajustamento. Essa metodologia agrupa um conjunto de dados (indivíduos) que possuem características similares em diferentes grupos. Sendo assim, o objetivo desta técnica consiste em identificar e separar os grupos a partir de características distintas (JAIN et al., 1999; WEBB, 2003; GOVENDER; SIVAKUMAR, 2020).

Há duas categorias principais para a classificação de grupos que são bastante utilizadas: método hierárquico e método não hierárquico. O método hierárquico, baseado na similaridade dos grupos, produz uma classificação em formato de árvore, no qual os grupos formados não podem ser reajustados. Essa classificação se divide em duas técnicas, a aglomerativa (que une vários grupos em um) e a divisiva (que divide o conjunto em vários grupos). Essas técnicas, por serem baseadas em árvores, tornam o método computacionalmente dispendioso. Por outro lado, os métodos não hierárquicos executam a classificação com base em um número de grupos previamente definido a partir dos dados de entrada. Como esse método não possui a estrutura em forma de árvore, isto o torna computacionalmente mais leve, além disso, baseia-se na não semelhança entre os grupos (DOWNS; BARNARD, 1995; GONG; RICHMAN, 1995; JAIN, 2010; HAN et al., 2016; GOVENDER; SIVAKUMAR, 2020), o que representa uma vantagem. Por essa razão, no presente estudo utilizou-se o método não hierárquico *k-means*.

O procedimento computacional da técnica *k-means* tem como ponto inicial os dados de entrada e a escolha da quantidade de k grupos finais. Em seguida, há a separação simultânea de todos os grupos com centroides escolhidos aleatoriamente, além de escolher aleatoriamente os membros para cada um. A cada iteração, novos membros são adicionados ou excluídos a partir do cálculo da distância Euclidiana entre os membros e o centroide. Em seguida, sempre que há uma mudança, novos cálculos serão realizados, repetindo-se n vezes até que não haja mais mudanças na posição do centroide, chegando-se ao resultado final (GONG; RICHMAN, 1995; JAIN, 2010; CHU et al., 2012; BOTTINO; CEBALLOS, 2015; GOVENDER; SIVAKUMAR, 2020).

Um esquema gráfico representando todo o procedimento computacional é ilustrado na Figura 2.3. O procedimento padrão para calcular o número de grupos a partir da técnica *k-means* pode ser descrita de forma simples seguindo o passo a passo

apresentado por [Chu et al. \(2012\)](#) e [Jain \(2010\)](#).

Passo 1 – Escolher k grupos como sendo a quantidade de centroides.

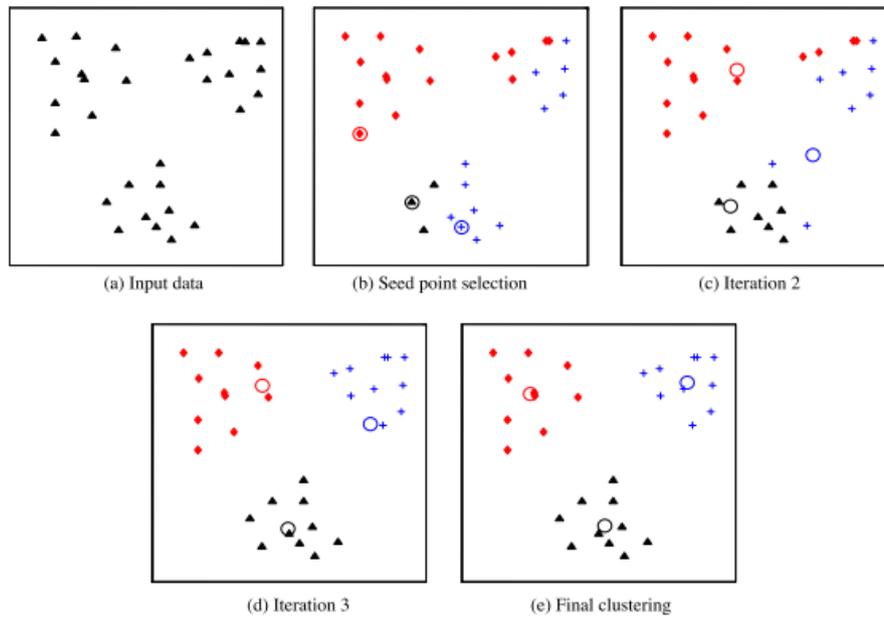
Passo 2 – Os membros do grupos são alocados de acordo com sua posição que deve ser mais próxima possível do centroide.

Passo 3 – Determinar novos centroides para os k grupos definidos. São determinados a partir do cálculo da média entre cada membro do grupo e seu centroide.

Passo 4 – Repetir os passos 2 e 3 até que haja estabilidade total dos membros e dos centroides.

A [Figura 2.3](#) mostra a ilustração do *k-means* para 3 grupos utilizando dados em 2 dimensões, na qual (a) representa os dados de entrada sem qualquer separação prévia; (b) identificação dos centroides e classificação inicial dos membros para cada grupo; (c) e (d) representam as iterações intermediárias, onde ainda há mudança na posição dos centroides e dos membros nos grupos e; (e) os 3 grupos finais com o resultado do *k-means*, após todas as iterações sem mudança do centroide.

Figura 2.3 - Ilustração da técnica *k-means* para 3 grupos utilizando dados em 2 dimensões.



Fonte: Jain (2010).

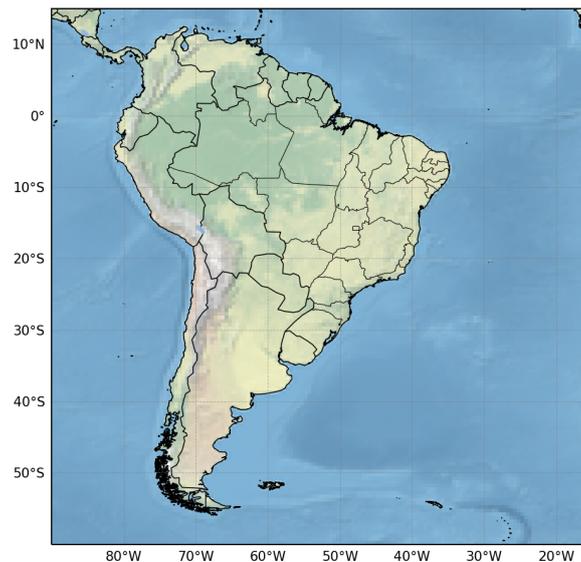
3 DADOS E METODOLOGIA

Neste capítulo serão apresentados a região de estudo, uma descrição dos dados, características gerais do satélite geoestacionário GOES-16 e os métodos a serem utilizados, além de abordar de forma breve os critérios para o desenvolvimento do algoritmo de classificação de nuvens.

3.1 Área de estudo

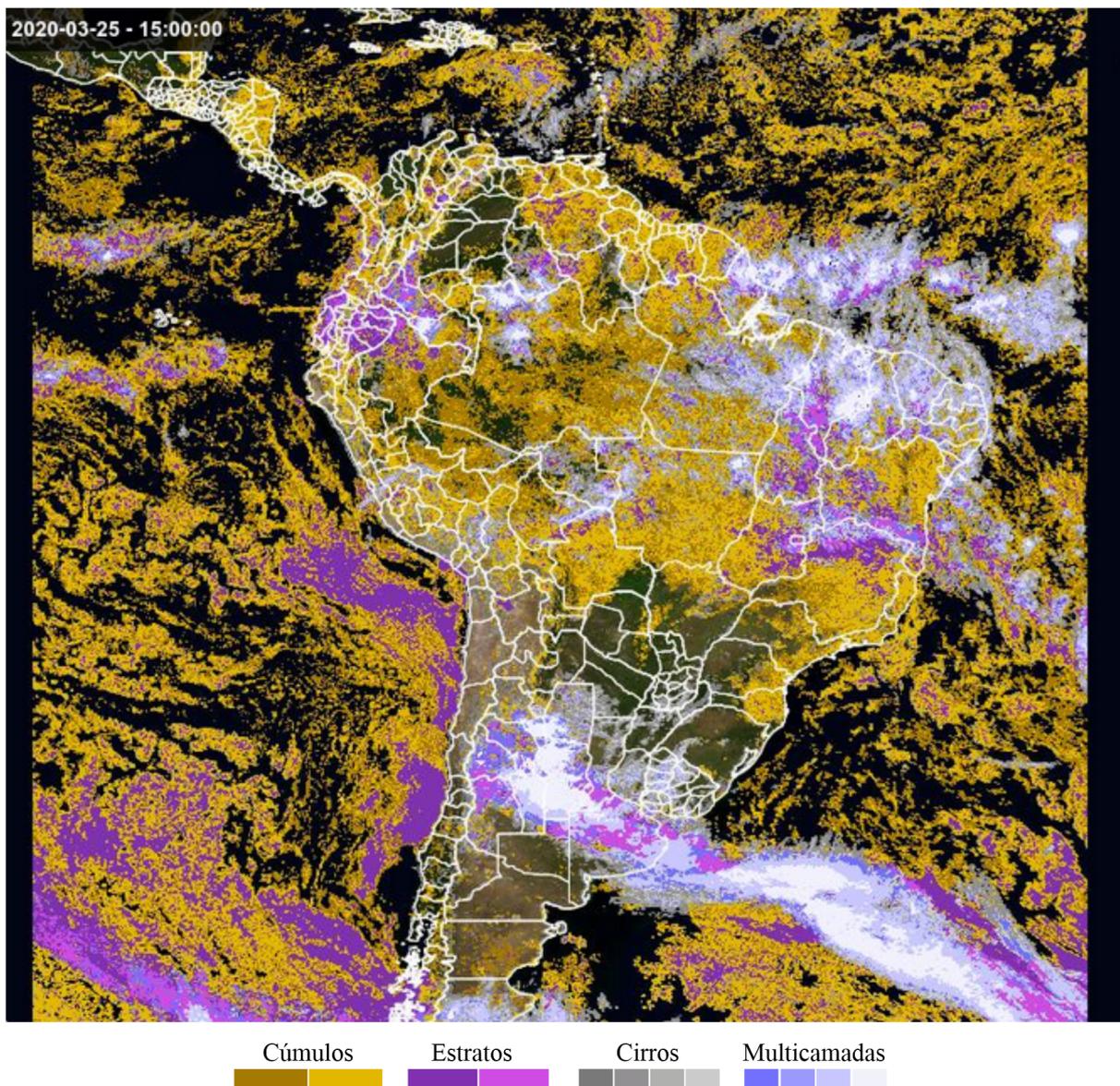
A área de estudo abrange toda a América do Sul, parte dos oceanos Pacífico e Atlântico (90° W - 15° W e 60° S - 15° N), conforme exibido na [Figura 3.1](#), que está dentro da região de cobertura do GOES-16. A [Figura 3.2](#), disponível no *website* do CPTEC/INPE, foi gerada utilizando o produto operacional de classificação de nuvens do Centro (BOTTINO; CEBALLOS, 2015), onde o resultado da classificação de nuvens é sobreposto a uma imagem média "blue marble" que representa o oceano e o continente, a cor amarela representa as nuvens cúmulos, a lilás representa as nuvens estratos, a cinza representa as nuvens cirros e a branca representa as nuvens multicamadas. Esse classificador será utilizado como ponto de partida para o desenvolvimento dos novos classificadores desse estudo.

Figura 3.1 - Área escolhida para o desenvolvimento dos classificadores de nuvens.



Fonte: Produção da autora.

Figura 3.2 - Classificação de nuvens operacional do CPTEC/INPE para o dia 25/03/2020 às 15 UTC.



Fonte: CPTEC/INPE (2020).

3.2 Satélite GOES-16

Serão utilizados dados do satélite geostacionário GOES-16 da NOAA (*National Oceanic and Atmospheric Administration*), em colaboração com a NASA, correspon-

dente ao período de 01/01/2019 a 31/12/2020. O sensor ABI (*Advanced Baseline Imager*), a bordo do GOES-16, permite monitorar continuamente as nuvens com maior detalhamento quando comparado aos seus antecessores, pois possui 11 bandas espectrais (canais) a mais. Além disso, houve também uma melhora na resolução espacial e temporal das observações (SCHMIT et al., 2017).

As 16 bandas espectrais do ABI abrangem regiões que vão desde o visível ($0,47 \mu m$ e $0,64 \mu m$), passando pelo infravermelho próximo ($0,86 \mu m$, $1,37 \mu m$, $1,6 \mu m$ e $2,2 \mu m$) até o infravermelho ($3,9 \mu m$, $6,2 \mu m$, $6,9 \mu m$, $7,3 \mu m$, $8,4 \mu m$, $9,6 \mu m$, $10,3 \mu m$, $12,3 \mu m$ e $13,3 \mu m$). Cada canal possui características que permitem estimar vários parâmetros atmosféricos e de superfície. Um exemplo é o canal 4 ($1,37 \mu m$), que ocupa uma região de forte absorção do vapor d'água e a sua principal aplicação é a detecção de nuvens altas. Outro exemplo é o canal 5 ($1,6 \mu m$), que permite diferenciar nuvens que contêm água de nuvens que contêm gelo. Mais detalhes podem ser encontrados em (<https://www.goes-r.gov/mission/ABI-bands-quick-info.html>) (PAVOLONIS, 2010; SCHMIT et al., 2017). Na Tabela 3.1 é observado os diferentes canais, resolução espacial e as respectivas aplicações principais.

Adicionalmente, é apresentada na Figura 3.3 a função resposta para as 16 bandas espectrais (canais) do ABI (área em azul), a transmitância (linha preta) para os canais do visível e infravermelho próximo e a temperatura de brilho terrestre (linha preta) para os canais na região do infravermelho. Com a adição desses novos canais, o universo de aplicações é maior, além disso, pode-se melhorar alguns dos produtos já existentes em determinados centros como, por exemplo, a classificação de nuvens, objeto central deste trabalho.

Tabela 3.1 - Atributos espectrais do ABI, como canal, comprimento de onda central, resolução espacial e a descrição.

Canal	Comprimento de Onda Central (μm)	Resolução Espacial (km)	Descrição
1	0,47	1	Azul
2	0,64	0,5	Vermelho
3	0,86	1	Vegetação
4	1,37	2	Cirrus
5	1,6	1	Neve/Gelo
6	2,2	2	Tamanho da partícula da nuvem
7	3,9	2	Janela atmosférica de onda curta
8	6,2	2	Vapor d'água em altos níveis
9	6,9	2	Vapor d'água em médios níveis
10	7,3	2	Vapor d'água em baixos/médios níveis
11	8,4	2	Fase do topo da nuvem
12	9,6	2	Ozônio
13	10,3	2	Janela atmosférica limpa
14	11,2	2	Janela atmosférica de onda longa
15	12,3	2	Janela atmosférica suja
16	13,3	2	CO ₂

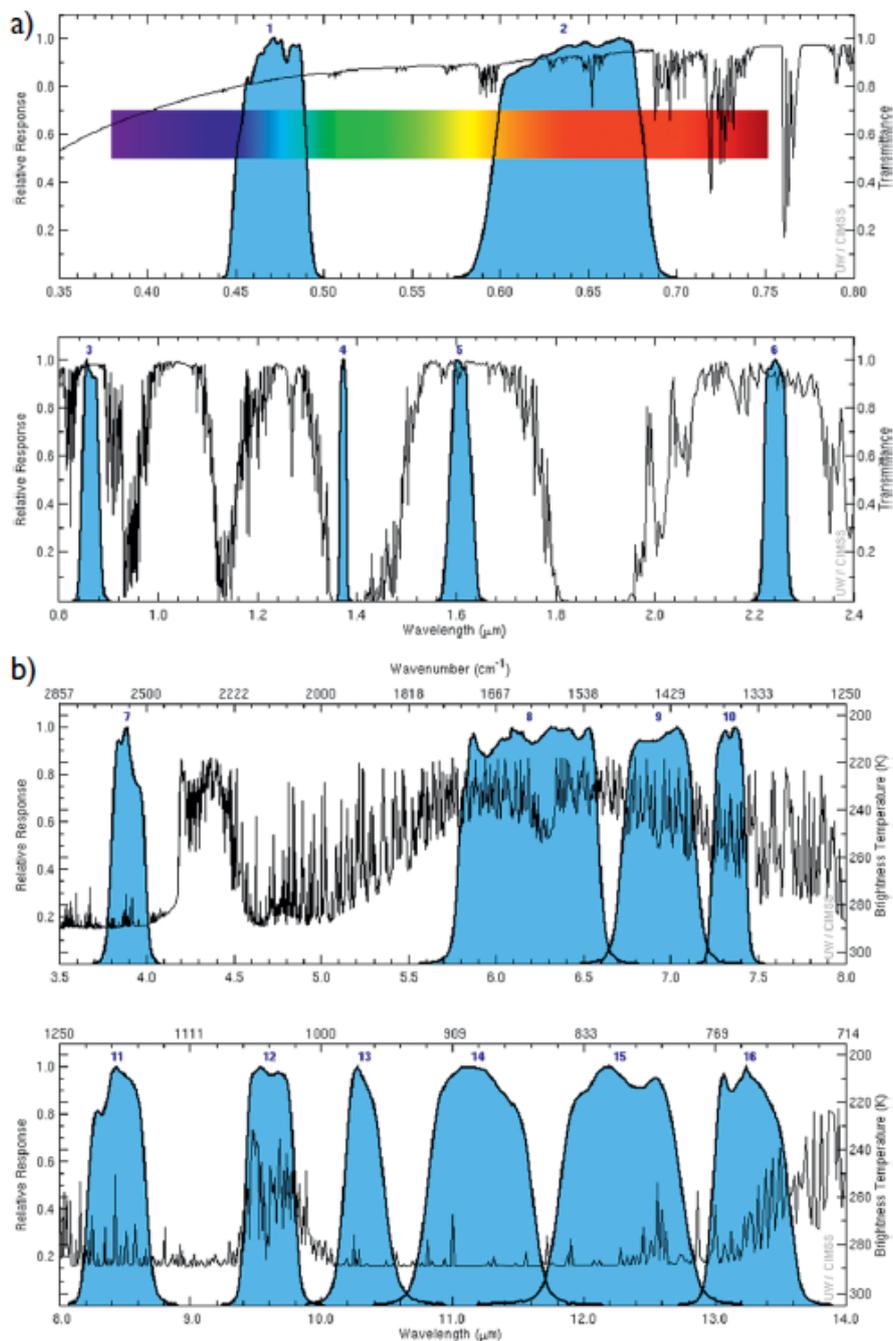
Fonte: Adaptado de Schmit et al. (2017).

A localização operacional do satélite GOES-16 é 75,2° W no Equador (MCCORKEL et al., 2020). Além disso, possui três tipos de varreduras com resoluções espaciais diferentes: mesoescala, contínuo US (CONUS) e *full disk*. O modo mais utilizado atualmente gera uma imagem *full disk* a cada 10 minutos, CONUS a cada 5 minutos e mesoescala a cada 1 minuto, cobrindo uma área de 1000 km x 1000 km (SCHMIT; GUNSHOR, 2020; SCHMIT et al., 2017).

Há um conjunto abrangente de produtos oficiais disponíveis que podem ser adquiridos, a exemplo de detecção de aerossol, cobertura de neve, profundidade óptica

de aerossóis, nuvens, dentre outros. As documentações técnicas de todos os produtos disponíveis podem ser encontradas em <https://www.goes-r.gov/products/overview.html#ABI>. Nesse estudo foi utilizado o produto de imagens de nuvens e umidade (CMIP, sigla em inglês para *Cloud Moisture Imagery Product*), o qual é constituído pelos 16 canais cobrindo as três áreas de varredura (mesoescala, CONUS e *full disk*). Esse produto pode ser adquirido separadamente para cada canal, com resolução espacial de acordo com a [Tabela 3.1](#) ou na forma de um único arquivo (*Cloud Moisture Imagery Product - Multi-Band Format*) que contém todos os 16 canais, porém na resolução espacial de 2 x 2 km [Schmit et al. \(2010\)](#).

Figura 3.3 - Funções resposta espectrais para as 16 bandas do sensor ABI/GOES-16.



As áreas preenchidas são as funções resposta de cada canal que compõe o ABI; a) canais do visível e infravermelho próximo, em que a linha preta representa a transmitância e; b) canais do infravermelho, em que a linha preta representa a temperatura de brilho emitida pela superfície.

Fonte: Schmit et al. (2017).

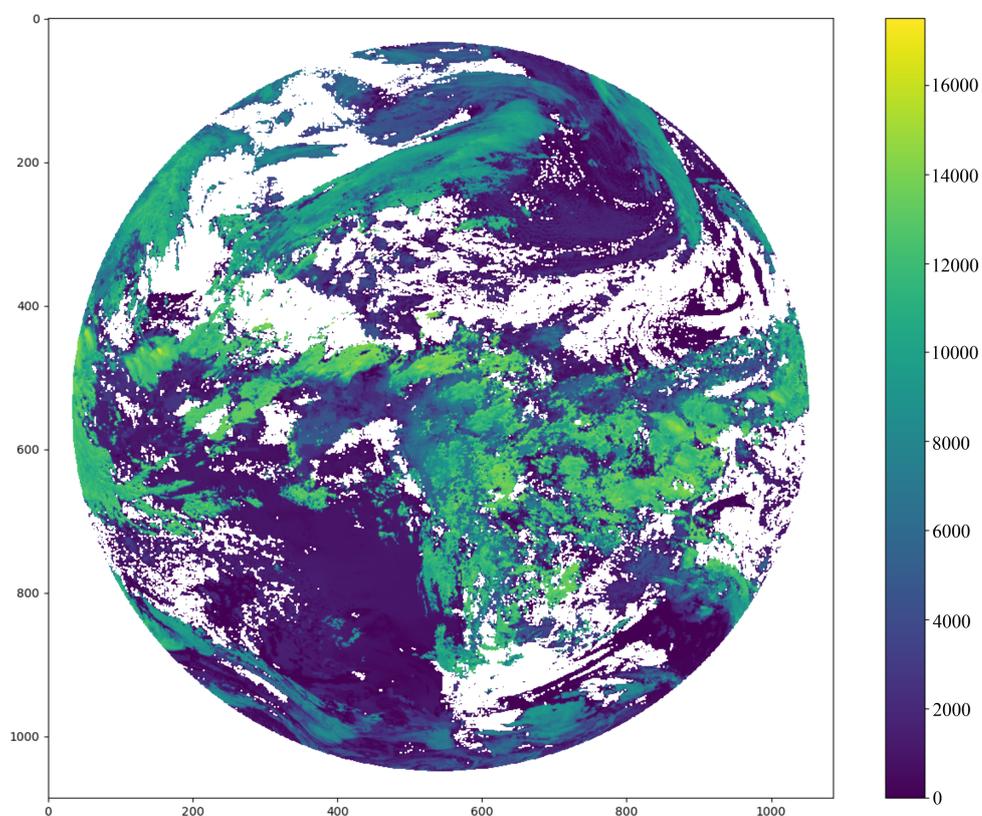
3.3 Produtos de topo de nuvem GOES-16/NOAA

Além do produto CMIP, outros produtos derivados do sensor ABI proveem informações acerca de nuvens, sendo eles: máscara de nuvens (*Clear Sky Mask*), altura de nuvens (*Cloud Layers/Heights*), profundidade óptica da nuvem (*Cloud Optical Depth*), distribuição do tamanho da partícula da nuvem (*Cloud Particle Size Distribution*), altura do topo da nuvem (*Cloud Top Height*), fase da nuvem (*Cloud Top Phase*), pressão do topo da nuvem (*Cloud Top Pressure*) e temperatura do topo da nuvem (*Cloud Top Temperature*).

O produto fase do topo da nuvem, fornece informações sobre quatro categorias de fase da nuvem: gelo, água, gotículas super-resfriadas e mistura de fase. Além de fornecer informações sobre seis diferentes propriedades físicas das nuvens: nuvens quentes, com gotículas super-resfriadas, com mistura de fase, de gelo opticamente finas, de gelo opticamente densas e multicamadas, as quais permitem descrever a fase da nuvem utilizando quatro canais do ABI, são eles: canal 10 ($7,3 \mu m$), 11 ($8,4 \mu m$), 14 ($11,2 \mu m$) e 15 ($12,3 \mu m$) (PAVOLONIS, 2010).

Outro exemplo é a altura do topo da nuvem, responsável por estimar a extensão vertical da nuvem para todos os pixels classificados como nublado. Além disso, também é utilizado para gerar um sinalizador de camada de nuvem, classificando-as como altas, médias ou baixas. Como utiliza 3 canais do infravermelho, é possível utilizar esse produto durante o dia e durante a noite. Na [Figura 3.4](#) podem ser observadas as diferentes alturas dos topos das nuvens, onde as cores mais claras representam nuvens mais altas, enquanto cores mais escuras representam nuvens mais baixas e a cor branca refere-se às regiões com ausência de nuvens. Esse produto é muito importante, pois fornece informações necessárias para outros produtos de nuvem, como temperatura do topo, pressão no topo, altura do topo e camada de nuvens (HEIDINGER, 2012).

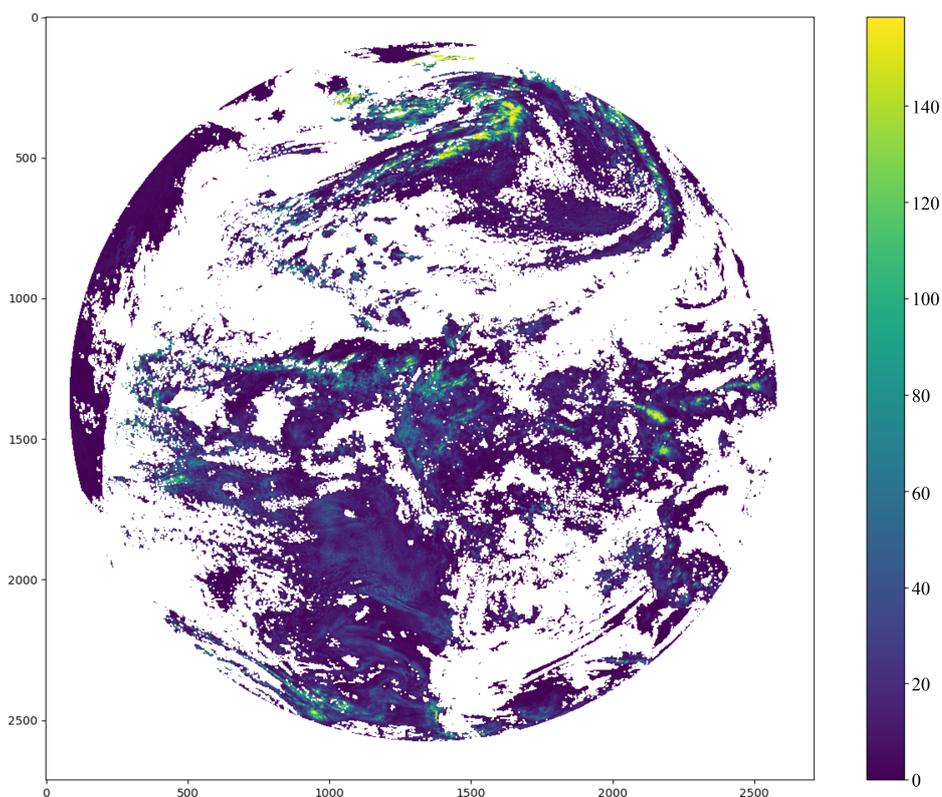
Figura 3.4 - Produto oficial da altura do topo da nuvem (m) para o dia 19/01/2019 em 15:00 UTC.



Fonte: Produção da autora.

O produto de profundidade óptica utiliza dois conjuntos diferentes de canais, um para o período diurno e outro para o período noturno, e possui resolução espacial de 4 km para o *full disk* nos dois períodos. A partir dos pixels marcados como nublado ou provavelmente nublado, representa-se a profundidade óptica (Figura 3.5) entre o topo e a base da nuvem. Para o período diurno, utiliza-se o canal 2 ($0,64 \mu m$) e o canal 6 ($2,2 \mu m$). Para o período noturno, os canais utilizados são os canais 7 ($3,9 \mu m$), 14 ($11,2 \mu m$) e 15 ($12,3 \mu m$) (MINNIS; HECK, 2012; WALTHER et al., 2013).

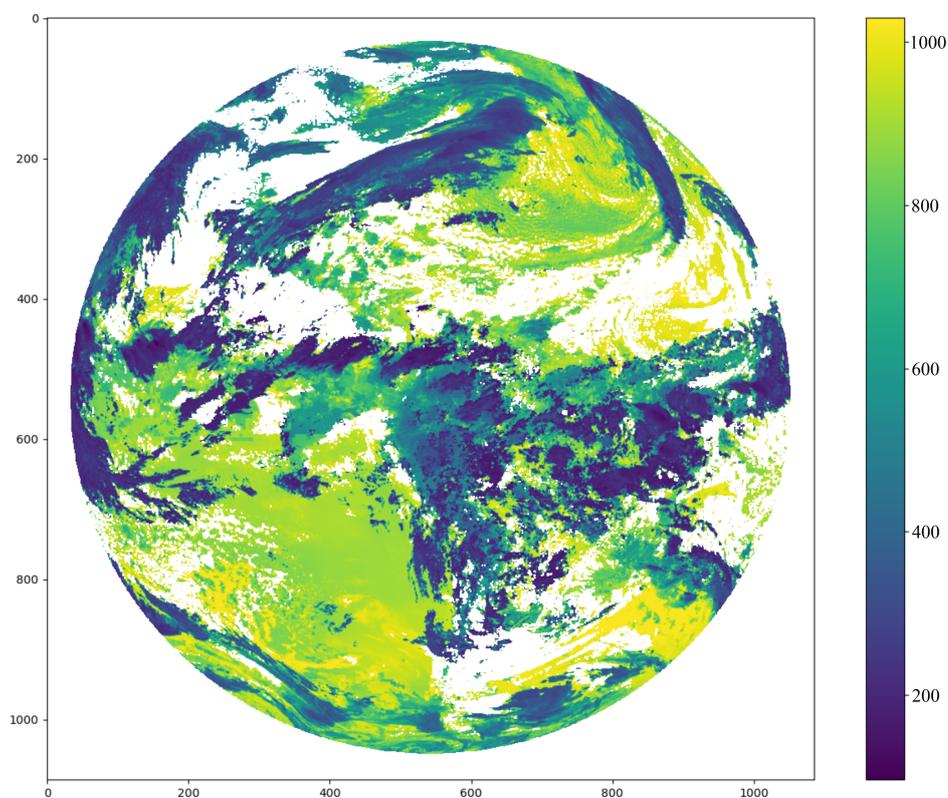
Figura 3.5 - Produto oficial de profundidade óptica para o dia 19/01/2019 em 15:00 UTC.



Fonte: Produção da autora.

Os produtos de pressão no topo da nuvem, bem como altura do topo da nuvem possuem resolução espacial de 10 km, enquanto o produto temperatura no topo da nuvem possui resolução espacial de 2 km (HEIDINGER, 2012). A Figura 3.6 mostra diferentes pressões no topo da nuvem, na qual as cores mais claras indicam pressões maiores, enquanto cores mais escuras indicam pressões no topo das nuvens menores e áreas em branco indicam ausência de nuvens.

Figura 3.6 - Produto oficial da pressão do topo da nuvem (mb) para o dia 19/01/2019 em 15:00 UTC.

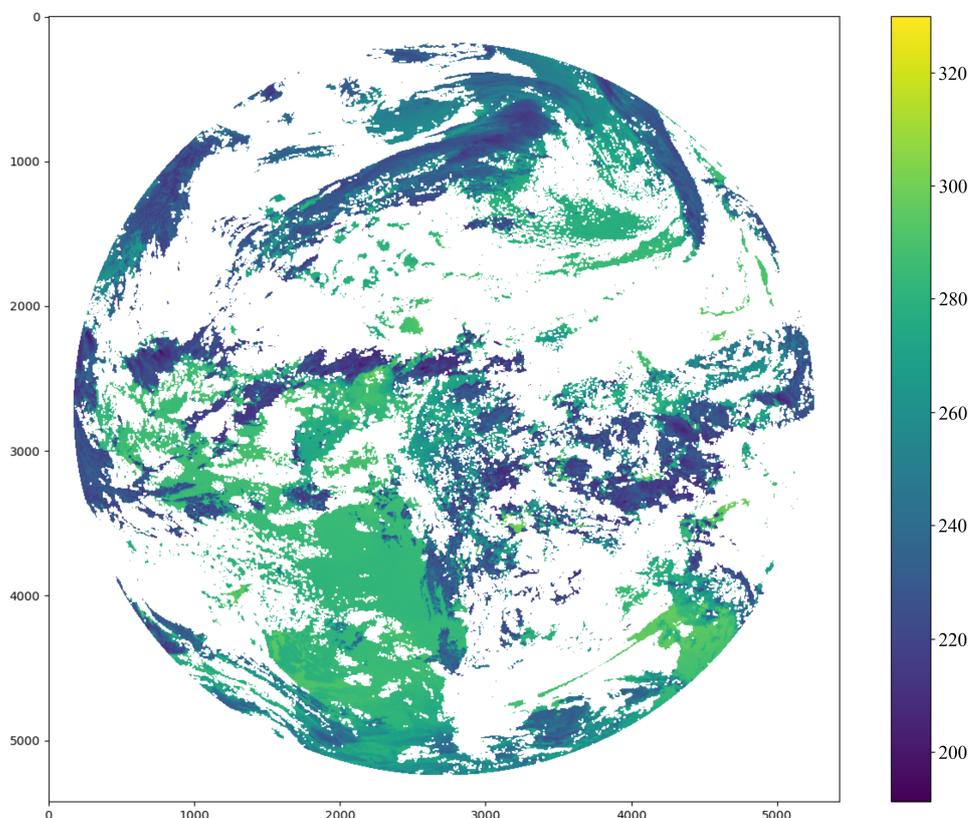


Fonte: Produção da autora.

Alguns dos produtos de nuvens serão utilizados a fim de compará-los ao classificador proposto no presente trabalho. Esses produtos serão descritos a seguir.

Os produtos temperatura no topo da nuvem, altura do topo da nuvem e pressão no topo da nuvem, seguem o *ABI Cloud Height Algorithm (ACHA)*, em que a [Figura 3.7](#) representa a temperatura do topo da nuvem. Esses produtos utilizam alguns canais na região do infravermelho gerando, assim, informações durante o período diurno e o período noturno. São responsáveis por estimar a extensão vertical de todos os tipos de nuvens, utilizando os canais 14 ($11,2 \mu m$), 15 ($12,3 \mu m$) e 16 ($13,3 \mu m$) ([HEIDINGER, 2012](#)).

Figura 3.7 - Produto oficial da temperatura do topo da nuvem (K) para o dia 19/01/2019 em 15:00 UTC.

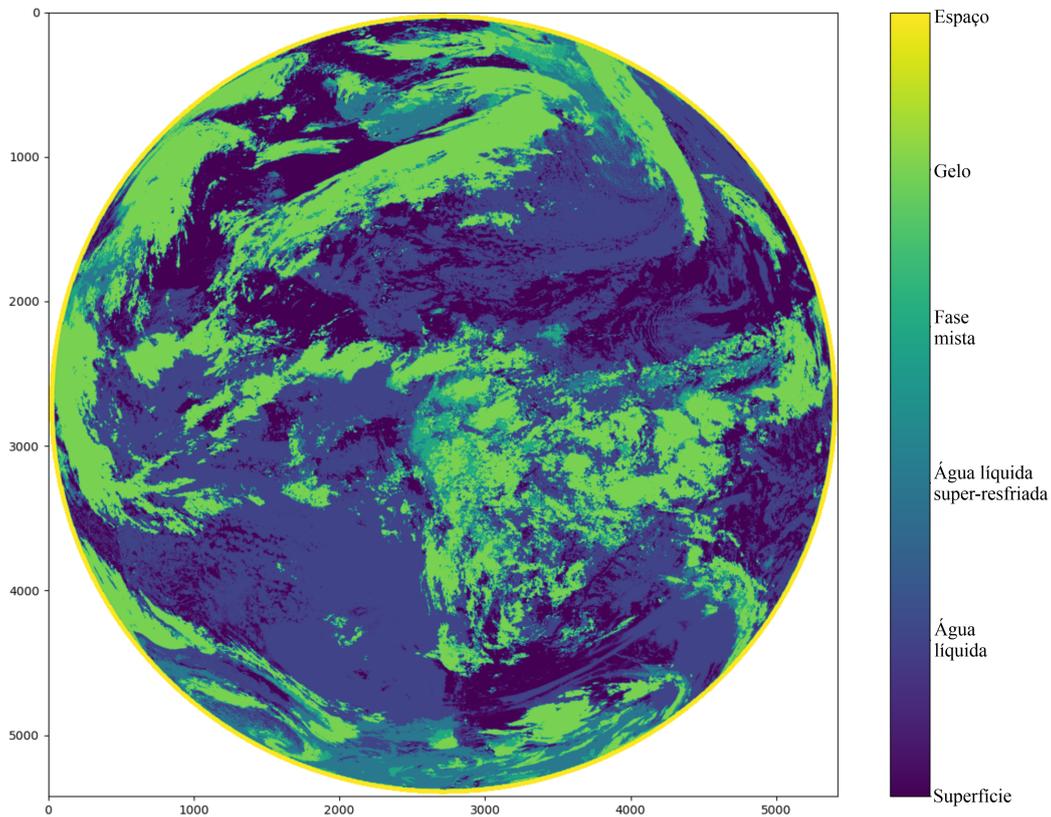


Fonte: Produção da autora.

O produto fase do topo da nuvem (Figura 3.8) nos dá informações sobre a fase da nuvem, bem como o tipo de nuvem, para cada pixel classificado como "nublado". Para a fase da nuvem há quatro categorias, são elas: fase de água líquida quente, fase de água líquida super-resfriada, fase mista e fase gelo (PAVOLONIS, 2010). A fase de água líquida quente refere-se às nuvens de água líquida com temperaturas de brilho superior a 273 K; a fase de água líquida super-resfriada é referente à nuvem que possui topo de água líquida com temperatura de brilho inferior a 273 K; a fase mista possui uma probabilidade alta de conter água líquida, bem como gelo próximo ao topo da nuvem e; a fase gelo refere-se a todas as nuvens que contém gelo em seu topo (PAVOLONIS, 2010). Este produto possui 2 km de resolução espacial e utiliza alguns canais na região do infravermelho: canal 10 ($7,3 \mu m$), 11 ($8,4 \mu m$), 14 ($11,2$

μm) e 15 ($12,3 \mu m$). Ao empregar esses canais, as informações são ininterruptas, ou seja, são mensuradas durante o período diurno e noturno (PAVOLONIS, 2010).

Figura 3.8 - Produto oficial da fase da nuvem para o dia 19/01/2019 em 15:00 UTC.

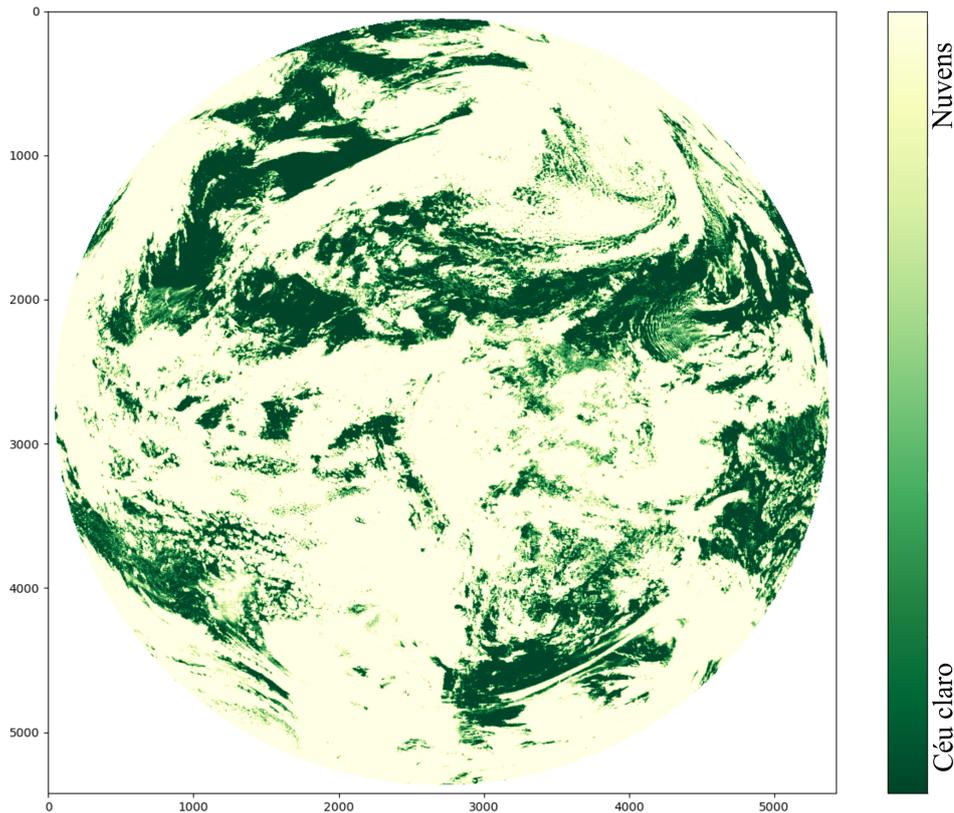


Fonte: Produção da autora.

O produto máscara de nuvens possui 2 km de resolução espacial e produz quatro tipos de máscara de nuvens: céu claro, céu provavelmente claro, céu provavelmente nublado e céu nublado (Figura 3.9). Este produto utiliza mais canais do que os demais descritos anteriormente, canais estes que estão tanto na região do infravermelho como também na região do visível. O produto emprega os canais 2 ($0,64 \mu m$), 4 ($1,37 \mu m$), 5 ($1,6 \mu m$), 7 ($3,9 \mu m$), 9 ($6,9 \mu m$), 10 ($7,3 \mu m$), 11 ($8,4 \mu m$), 14 ($11,2 \mu m$) e 15 ($12,3 \mu m$), onde o canal 9 só é utilizado quando o canal 10 não está disponível (HEIDINGER; STRAKA III, 2012). Este produto é fundamental para várias aplicações que necessitam que o pixel seja céu claro como, por exemplo, o Índice de Vegetação

da Diferença Normalizada (NDVI, sigla em inglês para *Normalized Difference Vegetation Index*) e a Temperatura da Superfície do Mar (SST, sigla em inglês para *Sea Surface Temperature*) (HEIDINGER; STRAKA III, 2012).

Figura 3.9 - Produto oficial da máscara de nuvem para o dia 19/01/2019 em 15:00 UTC.



Fonte: Produção da autora.

3.4 Método *K-Means*

O método *k-means* baseia-se na não similaridade (dissimilaridade) dos grupos. Esse método, apesar de ser necessário inicialmente especificar a quantidade de grupos de saída (final) e ser sensível a valores extremos, possui vantagens que viabilizam o seu uso, tais como: baixa complexidade, pode ser aplicado a uma grande quantidade de dados, e baixo custo computacional (JAIN et al., 1999; LIKAS et al., 2003; JAIN, 2010; BAGIROV et al., 2011).

A similaridade que cada grupo possui é definida pela distância entre seus membros e o centroide. A distância Euclidiana é uma das métricas mais utilizadas na associação entre as amostras e os centroides, ou seja, ela define em qual grupo (centróide) cada amostra será associada em cada iteração do método (JAIN et al., 1999; BOTTINO; CEBALLOS, 2015; GOVENDER; SIVAKUMAR, 2020). A distância Euclidiana é determinada pela Equação 3.1:

$$d_{ik} = \left(\sum_{i=1}^I (x_{i,k} - x_{j,k})^2 \right)^{1/2} \quad (3.1)$$

Em que, d_{ik} representa a distância Euclidiana, I representa o total de variáveis (parâmetros), i representa o número da i -ésima variável ($i = 1, 2, \dots, I$), $x_{i,k}$ representa a variável no k -ésimo grupo e $x_{j,k}$ representa o centroide do k -ésimo grupo.

Os centroides são calculados a partir da média do conjunto de variáveis ($x_{i,k}$) e é dado por:

$$x_{j,k} = \frac{1}{C} \sum_{k=1}^C x_{i,k} \quad (3.2)$$

Em que, C representa o número total de grupos (ou centroide).

Por ser um método de classificação não supervisionado, as classes definidas (*clusters*) deverão ser interpretadas para definir o que cada classe representa, ou seja, quais os tipos de nuvens ou quaisquer outros objetos (superfícies, nuvens de fuligem, cinzas vulcânicas, etc). Essa associação será feita a partir da análise visual em que cenas escolhidas aleatoriamente são interpretadas.

3.5 Classificadores

Nesta seção serão discutidas as características e procedimentos realizados para a atualização do classificador atualmente em operação no CPTEC/INPE, bem como o desenvolvimento do novo classificador utilizando mais canais do sensor ABI/GOES-16.

3.5.1 Modelo operacional

Nesta seção serão descritas as características do modelo em operação no DISSM que é a base do presente estudo.

Bottino e Ceballos (2015) utilizaram dados provenientes do satélite GOES-8, o qual possuía 5 canais espectrais: um no visível, outro na região do vapor d'água e 3 na região do infravermelho. Esse satélite possuía uma resolução espacial de 4 km e resolução temporal de 30 minutos.

Primeiramente, para desenvolvimento do classificador atualmente em operação, os autores realizaram um estudo em que utilizaram dois conjuntos de dados: um para o treinamento e outro para a validação, correspondente a 10 dias do mês de setembro de 2002. As variáveis utilizadas foram os cinco canais espectrais disponíveis, bem como oito novas variáveis, três delas sendo diferenças entre os canais termais em relação ao canal 4 e outras cinco referentes às texturas dos cinco canais, totalizando 13 variáveis. As texturas foram calculadas tendo como base uma grade de 3x3 pixels, onde essa grade “varre” toda a imagem em questão calculando a textura do pixel central da grade.

Foram definidos seis cenários, que variam desde situações de céu nublado (nuvens cumulus, cirrus, stratus, multicamadas e cumulonimbus) até situações de céu claro (superfícies continentais e oceânicas). Para classificar os seis diferentes cenários, os autores aplicaram a técnica *cluster* dinâmico, também conhecida como *k-means*. Na aplicação da técnica é necessário escolher previamente a quantidade de grupos, para tal os autores definiram a quantidade de grupos igual a 35, entretanto 3 deles apresentaram valores espúrios, resultando em 32 grupos contendo as características das 13 variáveis.

Outra característica desse método é definir o critério de convergência dos grupos, e o critério utilizado pelos autores foi a distância Euclidiana. A distância Euclidiana utilizada pelos autores foi igual a 0,04.

A aplicação da técnica de análise de componentes principais foi utilizada a fim de verificar possíveis variáveis redundantes, chegando a um total de cinco variáveis com relevância: dois canais (um na região do visível e outro no infravermelho), uma diferença entre temperaturas de brilho e duas texturas (relativas aos canais do visível

e do infravermelho).

Contudo, o modelo em operação atualmente possui quatro variáveis: dois canais (visível e infravermelho janela) e suas respectivas texturas. Além disso, o número total de grupos (centroides) é igual a 30, os quais identificam superfície (céu claro), nuvens cumulus, cirrus, stratus e multicamadas, esta última representando nuvens que apresentam intenso desenvolvimento vertical (cumulonimbus).

3.5.2 Atualizando o classificador a partir dos dados ABI/GOES-16

De posse dos dados, primeiramente o classificador desenvolvido por (BOTTINO; CEBALLOS, 2015) foi atualizado a partir do uso das observações ABI/GOES-16 utilizando os mesmos parâmetros do classificador original, ou seja, o canal do visível, infravermelho e suas respectivas texturas. A técnica *k-means* é sensível aos centroides iniciais, portanto os centroides obtidos por (BOTTINO; CEBALLOS, 2015) foram utilizados como ponto de partida para a definição das novas tabelas de centroides correspondentes aos dados ABI/GOES-16.

Após recriar os classificadores, os mesmos foram comparados entre si, bem como a três dos produtos oficiais de nuvem da NOAA (temperatura do topo da nuvem, a fase da nuvem e a máscara de nuvem). A comparação teve como finalidade analisar o comportamento dos dois classificadores baseados no GOES-8 (o classificador em operação e o classificador atualizado).

3.5.3 O novo classificador

Estudos de caso foram realizados com a finalidade de verificar quais foram as melhorias na classificação de nuvens em relação ao classificador operacional (<http://sigma.cptec.inpe.br/nuvens/>). Imagens de satélite na região do visível e do infravermelho janela (canal 2 e 13) foram utilizadas auxiliando na nefanálise. Após essa etapa foi possível seguir com a definição de novos parâmetros (variáveis) para o desenvolvimento do novo classificador e a comparação das classes de nuvens aos produtos de topo de nuvens derivados dos dados do sensor ABI/GOES-16 pela NOAA.

A técnica estatística Análise de Componentes Principais (PCA, sigla em inglês para *Principal Components Analysis*) permite reduzir o número grande de variáveis de um determinado conjunto de dados para um número menor, perdendo o mínimo de

informações possível (MAĆKIEWICZ; RATAJCZAK, 1993; WEBB, 2003; WILKS, 2011; HONGYU et al., 2015). Portanto, essa técnica foi utilizada para definir quais canais ou combinações tinham potencial para o desenvolvimento do novo classificador de nuvens.

A técnica PCA transforma um conjunto de variáveis que se correlacionam entre si em um conjunto de variáveis ortogonais, em outras palavras, não possuem correlação entre si. As variáveis são ordenadas em termos da variância original dos dados, onde a primeira componente é responsável pela maior variação total dos dados, a segunda componente descreve a segunda maior e assim por diante. Isso por si só, já leva a uma redução na dimensão das variáveis (WEBB, 2003; WILKS, 2011; HONGYU et al., 2015). Há alguns passos a serem seguidos para calcular as componentes principais:

- Passo 1 - Obter os dados de dimensão M ;
- passo 2 - Calcular a média ou o vetor médio dos dados;
- Passo 3 - Subtrair a média de todos os itens dos dados;
- Passo 4 - Calcular a matriz de covariância utilizando todas as subtrações, que terá dimensão $M \times M$;
- Passo 5 - Calcular os autovalores e autovetores da matriz de covariância.

Os autovetores representam o peso de cada variável no conjunto original sobre cada uma das componentes. Enquanto o autovalor representa a contribuição que cada uma das componentes possui na explicação da variância total dos dados, ou seja, indica a importância que cada variável possui (WEBB, 2003; WILKS, 2011; HONGYU et al., 2015). A técnica PCA foi utilizada objetivando a redução na dimensão dos parâmetros, que ao total foi correspondente a 40: os 16 canais do ABI/GOES-16, suas respectivas texturas e 8 diferenças entre temperaturas de brilho. Essas 8 diferenças entre as temperaturas de brilho podem ser visualizadas na Tabela 4.7. Um total de 12 parâmetros foram obtidos ao final do processo.

A escolha dos parâmetros é uma etapa de suma importância, sendo assim, após o cálculo da técnica PCA chegou-se aos parâmetros finais a serem utilizados no desenvolvimento do novo classificador. Estes parâmetros fornecem informações importantes a respeito das características das nuvens. Estas propriedades, são descritas a seguir:

- Conteúdo de água líquida: esse parâmetro depende tanto do tamanho da gotícula de água ou formação de gelo na nuvem, além da espessura óptica da nuvem;
- Temperatura do topo da nuvem: a informação obtida com esse parâmetro é especificamente para nuvens convectivas. Quanto mais frio, maior é a convecção;
- Fase da nuvem: informa se a nuvem em questão possui gelo em seu topo, corroborando com a temperatura do topo da nuvem, ou seja, se o topo for mais frio e há presença de gelo, então essa nuvem é convectiva;
- Altura do topo da nuvem: importante na identificação de nuvens com grande desenvolvimento vertical.

Uma breve descrição dos parâmetros espectrais utilizados para compor o novo classificador será feita a seguir. Na [Tabela 3.2](#) são mostrados todos os parâmetros multiespectrais utilizados no desenvolvimento do novo classificador. Além dos canais do ABI/GOES-16 (0,64 μm , 1,37 μm , 6,2 μm , 8,4 μm , 10,33 μm , 11,2 μm e 12,3 μm), também foram utilizadas as texturas do canal do visível, do vapor d'água e do infravermelho correspondente à janela atmosférica (0,64, 6,2 e 10,3 μm), além de 4 diferenças entre as temperaturas de brilho, totalizando 12 parâmetros diurnos. A [Tabela 3.2](#) mostra os parâmetros, tal como a característica e algumas das referências relacionadas aos parâmetros utilizados no desenvolvimento do novo classificador. Além das referências citadas na tabela, maiores detalhes podem ser encontrados em (INOUE, 1985; STRABALA et al., 1994; DESSLER; YANG, 2003; LIU et al., 2004; LAZRI et al., 2014a; LAZRI et al., 2014b).

Tabela 3.2 - Parâmetros multiespectrais de entrada para o classificador, utilizando canais do GOES-16.

Informações	Canais/Combinações (μm)	Referência
Conteúdo de água líquida	0,64	(THIES et al., 2008)
	11,2-3,9	(OUKALI et al., 2019) (FREY et al., 2008)
Temperatura do topo da nuvem	10,3	(FREY et al., 2008)
	12,3-3,9	(OUKALI et al., 2019)
Fase da nuvem	1,37	(GAO et al., 1993)
	6,2	(FREY et al., 2008)
	6,2-10,3	(THIES et al., 2008)
	11,2-12,3	(ACKERMAN et al., 1998)
Altura do topo da nuvem	10,3	(BOTTINO; CEBALLOS, 2015)
	6,2-10,3	(THIES et al., 2008)
	12,3-3,9	(FREY et al., 2008)

Fonte: Produção da autora.

Além de desenvolver o novo classificador, o mesmo foi comparado à três produtos oficiais da NOAA, a fim de verificar o seu comportamento quanto à classificação de nuvens. Os produtos utilizados foram: temperatura do topo da nuvem, fase da nuvem e máscara de nuvem.

4 RESULTADOS

4.1 Análise do classificador original

Antes de propor alguma modificação ou a criação de um novo classificador, foi verificada a qualidade do classificador operacional.

Primeiramente, foram realizadas análises de casos com a finalidade de verificar o comportamento do classificador gerado a partir de imagens do classificador atualmente em operação na DISSM. Para isso, 24 dias foram escolhidos aleatoriamente, dois dias de cada mês, durante o ano de 2019, no horário de 15:00 UTC (sigla em inglês para *Coordinate Universal Time*), o qual corresponde aproximadamente ao meio-dia local na maior parte do Brasil. Para cada dia escolhido, foram obtidas imagens de satélite na região do visível e do infravermelho, análogas aos canais do satélite GOES-8, assim como, cartas sinóticas com o objetivo de identificar a presença ou não de sistemas meteorológicos atuantes, tal qual padrões referentes à pressão em superfície.

Após aquisição de todos os dados, realizou-se o diagnóstico através de reanálise para identificar possíveis erros na classificação. Dentre os 24 dias selecionados, houveram ocorrência de diferentes sistemas meteorológicos, como por exemplo, passagem de frentes frias, ocorrência da Zona de Convergência do Atlântico Sul (ZCAS), bem como a Zona de Convergência Intertropical (ZCIT). Essa variedade possibilita uma análise mais criteriosa, já que a presença de diferentes sistemas, significa que há presença de diferentes tipos de nuvens em uma mesma imagem. A seguir, serão discutidos apenas 2 dos 24 dias selecionados, os quais observou-se algumas discordâncias pertinentes na classificação.

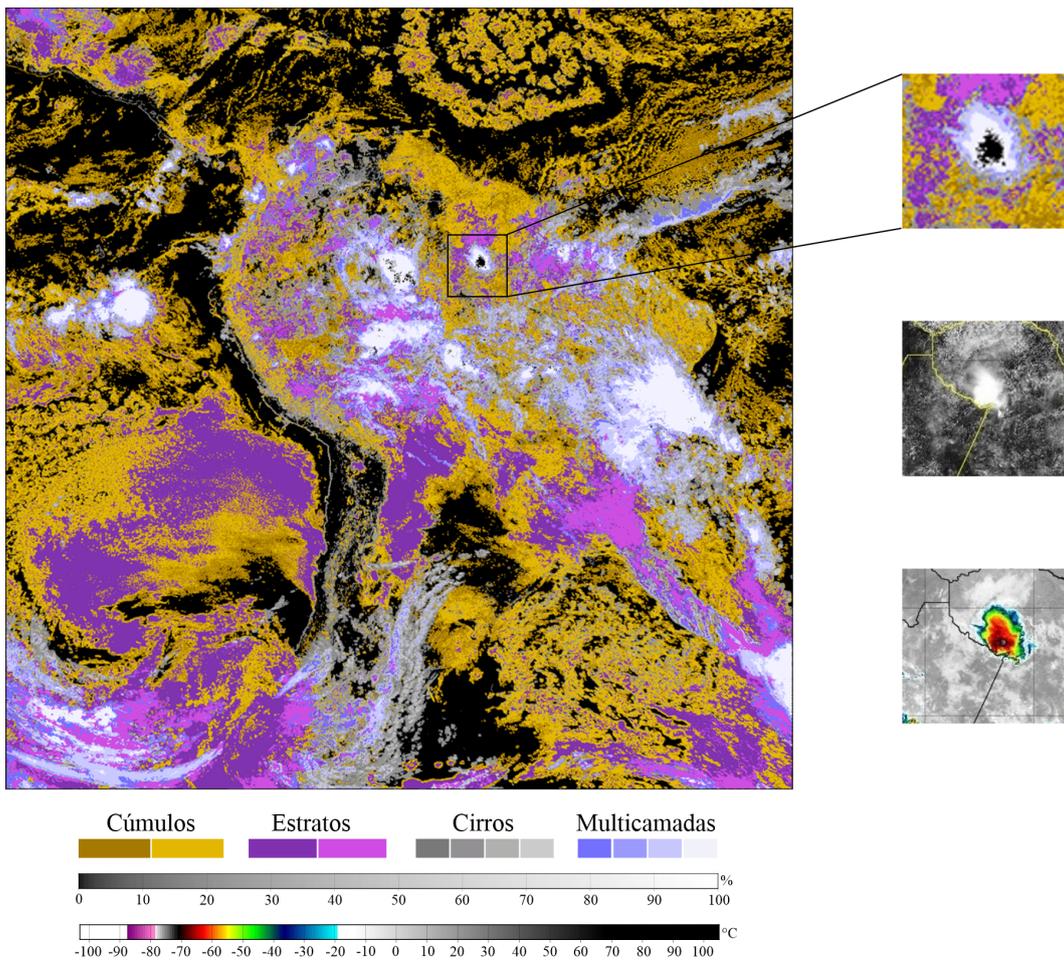
O primeiro dia a ser discutido é 21/03/2019, neste dia os sistemas meteorológicos atuantes foram a ZCIT posicionada mais ao sul do equador, além da presença de uma frente fria estacionária passando por São Paulo e adentrando o continente até a fronteira com a Bolívia.

O classificador mostra, em maior parte, uma coerência na identificação de nuvens para este dia, como por exemplo, as nuvens stratus, cirrus e multicamadas associadas à frente fria, visualizada na imagem da esquerda representada pelas cores lilás, cinza e branca, respectivamente. Contudo, na região norte do continente sul-americano,

houve uma discordância na classificação de nuvens com topos muito frios, em torno de $-70\text{ }^{\circ}\text{C}$ (Figura 4.1).

Na região da Figura 4.1 demarcado com uma caixa em preto, é possível observar presença de nuvens com topos muito frios, porém nota-se que há uma pequena região classificada como céu claro. Esse tipo de erro pode ser devido à forma como o método de classificação atual está implementado. Supondo que o limiar da temperatura de brilho for igual a 220 K ($-53\text{ }^{\circ}\text{C}$), então o pixel é ignorado, mostrando região de céu claro quando é nuvem. À esquerda da Figura 4.1 mostra a imagem do classificador em operação e à direita, o zoom da área do classificador, da imagem do visível e infravermelho, respectivamente.

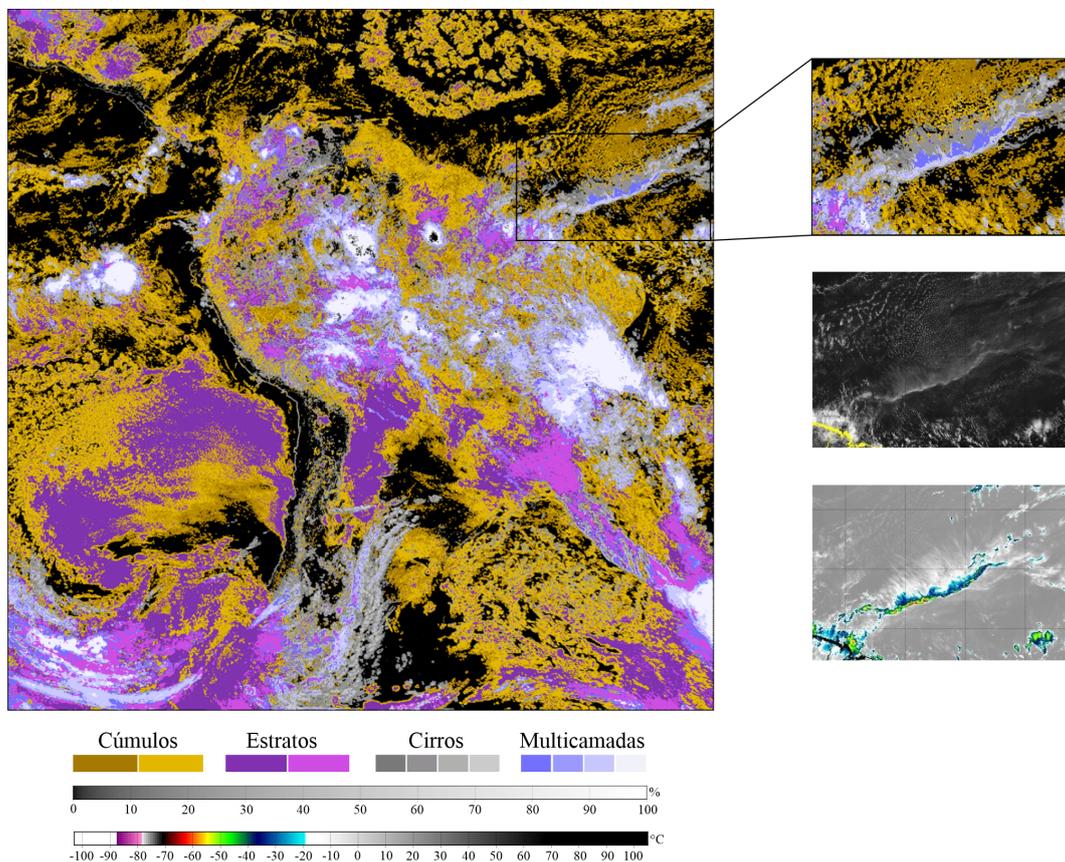
Figura 4.1 - Classificação nuvens de topo muito frio, 21/03/2019, 15:00 UTC.



Fonte: Adaptado de CPTEC / DISSM (2020).

No mesmo dia 21/03/2019, houve outro local com ocorrência de classificação errônea de nuvem. Localizado na região do Atlântico Norte, próximo à costa norte do Nordeste. Nesta região, o classificador indicou a presença de nuvens multicamadas e nuvens cirrus, na qual a porção de nuvens multicamadas representa maior quantidade. Contudo, ao analisar as imagens do visível e do infravermelho, é possível observar que a extensão de nuvens multicamadas é inferior àquela apresentada pelo classificador, observando nuvens cirrus em predominância. A Figura 4.2 mostra o classificador operacional à esquerda e à direita, o zoom da área destacada, a imagem do visível e do infravermelho, respectivamente.

Figura 4.2 - Classificação de nuvens cirrus e multicamadas, 21/03/2019, 15:00 UTC.

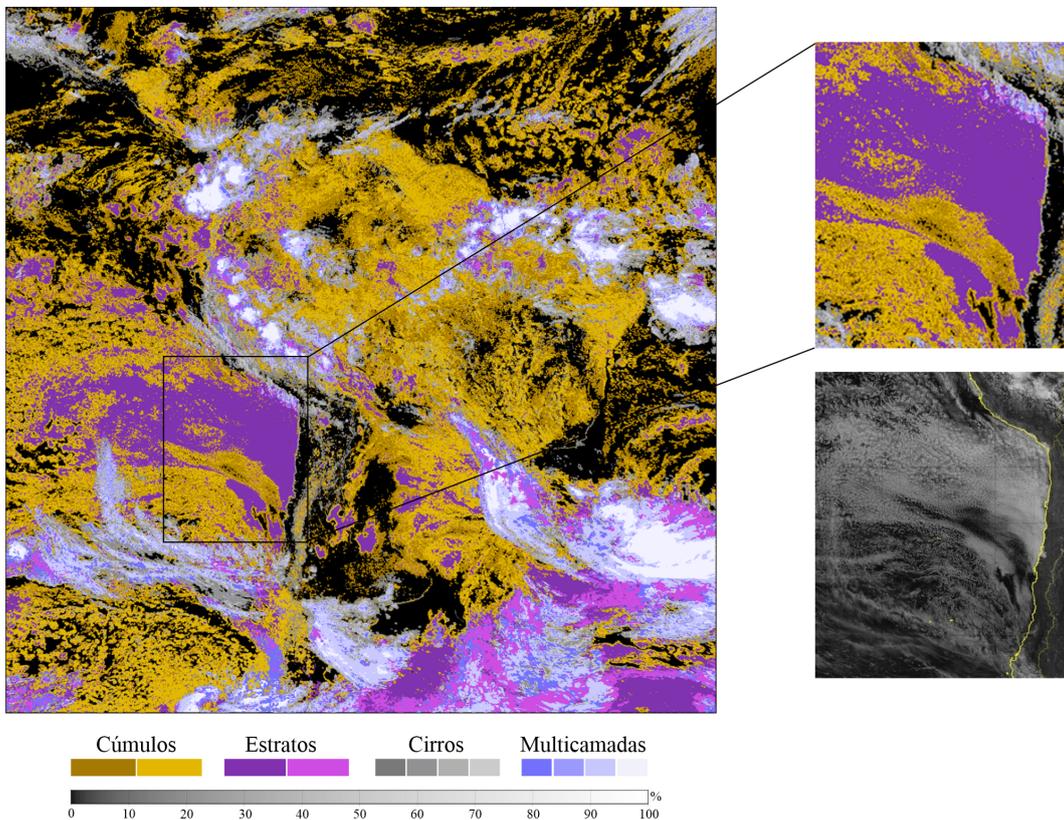


Fonte: Adaptado de CPTEC/DISSM (2020).

No dia 28/04/2019, a ZCIT estava posicionada mais ao sul, houve também presença

de duas frentes frias, uma passando por Santa Catarina e Paraná, e a outra localizada na Argentina. De forma geral, a classificação de nuvens foi de acordo em toda a área de cobertura, entretanto, no Oceano Pacífico, na região próximo à costa do Chile, há uma grande área classificada como sendo nuvens stratus, porém ao analisar a imagem do visível, percebe-se que há também nuvens cumulus na região, como pode ser visualizado na Figura 4.3. Como a resolução espacial do GOES-16 é maior, isso pode ajudar na identificação mais precisa de nuvens cumulus, por exemplo, na região do infravermelho, a resolução espacial do GOES-16 é de 2 km, enquanto o GOES-8 possuía uma resolução espacial de 4 km.

Figura 4.3 - Classificação de nuvens cumulus e stratus, 28/04/2019, 15:00 UTC.



Fonte: Adaptado de CPTEC / DISSM (2020).

4.2 Revisão do classificador baseado em quatro parâmetros

Essa seção apresenta os resultados da atualização do classificador original baseado nas imagens do sensor *Imager*, que equipou o satélite GOES-8, para uso das imagens do sensor ABI que equipa o satélite GOES-16, atualmente em operação. Apesar do sensor ABI possuir canais semelhantes àqueles do sensor *Imager*, estes apresentam resolução espacial maiores. Por exemplo, o canal referente ao vermelho do ABI (*Imager*) apresenta resolução 0,5 km (1 km), enquanto no infravermelho janela 2 km (4 km).

Vale ressaltar que as funções resposta desses canais são diferentes podendo alterar levemente os valores de temperatura de brilho no canal do infravermelho, usado como parâmetro na classificação. Portanto, é importante verificar quais são os impactos dessas diferenças ao utilizar um classificador criado a partir do sensor *Imager* ou um criado utilizando dados do sensor ABI.

Para atualizar o classificador, foram empregados os mesmos parâmetros finais determinados em Bottino e Ceballos (2015), onde os canais empregados pelos autores foram o visível ($0,65 \mu m$) e o infravermelho janela ($10,3 \mu m$), além das texturas dos respectivos canais, calculados a partir das matriz 3×3 centrada no pixel em questão, tal qual são calculadas no classificador atualmente em operação.

Para determinar os novos centroides, foi utilizado um conjunto de treinamento, correspondente a um conjunto de 120 imagens equivalente a 10 dias de cada mês do ano de 2019, escolhidos aleatoriamente. Durante o treinamento do classificador, 50.000 (cinquenta mil) pontos (amostras), foram extraídas aleatoriamente em cada uma das imagens. Cada amostra consiste em um vetor de atributos composto pela posição geográfica da amostra (latitude e longitude), as informações a respeito da reflectância, temperatura de brilho no infravermelho e as texturas de cada um desses dois parâmetros.

A partir das informações extraídas em cada um dos pontos de latitude e longitude, calculou-se os centroides a partir do método de agrupamento *k-means*. O cálculo dos novos centroides foi realizado levando em consideração as listas de centroides utilizadas atualmente, as quais são baseadas no GOES-8 e representam os encontrados por (BOTTINO; CEBALLOS, 2015). Esse procedimento foi realizado para cada um dos três horários atualmente em operação na DISSM: 12:00, 15:00 e 18:00 UTC.

As [Tabela 4.1](#), [Tabela 4.2](#) e [Tabela 4.3](#) representam os dois grupos de centroides nos horários 12:00, 15:00 e 18:00 UTC, respectivamente. A esquerda em cada uma das tabelas, encontram-se os centroides referentes ao classificador operacional (baseado no GOES-8) e a direita os centroides referentes ao classificador atualizado (baseado no GOES-16).

O número total de amostras utilizadas no treinamento em cada um dos horários, foi de 6×10^6 (seis milhões), correspondentes aos 120 dias de 2019. O método *k-means* foi executado aplicando os quatro parâmetros utilizados pelo classificador atualmente em operação, e a distância Euclidiana foi a medida de semelhança entre as amostras e centroides utilizados.

De acordo com as tabelas, [Tabela 4.1](#), [Tabela 4.2](#) e [Tabela 4.3](#), cada classe representada por cada centroide do classificador operacional possui um equivalente no conjunto estimado a partir do sensor ABI/GOES-16 nos três horários analisados. Dentre os três horários, foi observado que houve maior divergência entre os centroides dos classificadores operacional e atualizado na tabela das 18:00 UTC, em que alguns pares de centroides apresentaram diferenças nos valores de reflectância de até 10 unidades. Portanto ao utilizar as imagens do novo sensor, ABI, partindo dos centroides obtidos utilizando o sensor GOES-8/*Imager*, o método não convergiu para pontos muito próximos. As maiores diferenças encontradas no parâmetro "reflectância" pode ser explicado pela maior resolução espacial do sensor ABI, as variações de contraste nas bordas das nuvens ou nas próprias nuvens cumulus ou cumulonimbuns.

Para verificar a equivalência entre os dois conjuntos de centroides, ou seja, se realmente cada classe do classificador operacional equivale à mesma classe no classificador atualizado para o sensor ABI/GOES-16 foi realizado o teste-t, que é utilizado para comparar dois grupos. Para tal assumimos a hipótese nula (H_0) e a hipótese alternativa (H_1) a fim de verificar qual destas é verdadeira, avalia-se o p-valor que nos ajuda na tomada de decisão na rejeição ou não de H_0 (WILKS, 2011; PETER, 2008). Como o teste-t é calculado a partir da média, assume-se duas hipóteses (H_0 e H_1), onde H_0 equivale a média do conjunto de centroides do classificador operacional ser igual a média do conjunto do classificador atualizado e H_1 é equivalente a média dos dois conjuntos de centroides serem diferentes. Dessa forma, considerando H_0 com p-valor $> 0,05$ e H_1 com p-valor $\leq 0,05$, ao calcular o teste-t para as três tabelas, obteve-se os p-valores iguais a 0,98; 0,92 e; 0,88 respectivamente para 12:00, 15:00 e

18:00 UTC, portanto, a hipótese nula (H_0) é aceita. Isso significa que as médias dos dois conjuntos são iguais, então por consequência é possível afirmar que os mesmos não possuem diferença estatística significativa.

Tabela 4.1 - Centroides calculados com o conjunto de treinamento para o classificador operacional (baseado no GOES-8) e classificador atualizado (baseado no GOES-16) no horário das 12:00 UTC.

Classes	Centroides Atuais (GOES-8)				Centroides Novos (GOES-16)			
	vis (%)	ir (K)	texvis	texir	vis (%)	ir (K)	texvis	texir
sup	9.1	291.8	1.0	0.3	7.7	293.3	0.5	0.2
cu1	14.5	289.5	3.9	0.5	13.3	291.1	3.0	0.6
cu1	15.0	286.6	8.3	1.1	16.1	289.5	7.0	1.1
ci1	14.7	281.2	3.9	3.0	14.4	280.9	3.0	3.2
cu2	21.3	278.2	6.0	1.8	20.8	282.8	4.5	1.6
cu2	19.1	278.7	9.3	3.1	21.5	281.4	7.9	3.0
ci3	25.1	267.6	9.7	8.0	26.2	266.2	7.2	8.3
ci2	17.5	277.9	6.4	5.1	18.9	276.7	5.2	5.4
ci3	31.4	267.1	4.4	1.3	22.7	266.9	2.3	1.6
cu2	18.8	279.5	14.4	1.9	24.4	285.8	13.0	1.8
sup	18.0	273.4	3.0	0.9	11.5	281.3	0.9	0.4
st1	35.3	283.4	5.3	0.5	34.3	281.5	2.9	0.3
cu2	37.7	272.8	24.2	2.6	49.3	277.3	19.5	2.5
cu2	36.4	271.0	13.9	4.3	37.3	271.8	12.1	4.9
mc1	45.8	252.6	4.3	1.6	46.0	257.1	3.1	1.3
sup	10.8	287.0	2.3	1.5	9.2	286.9	1.5	1.6
ci4	35.0	259.3	6.0	4.4	35.9	255.7	4.4	4.4
cu2	33.4	277.6	10.3	1.3	35.5	283.1	8.3	1.0
ci4	34.4	252.3	13.8	13.7	39.2	257.9	13.5	12.6
cu2	109.2	260.3	43.2	3.9	60.3	264.6	31.5	4.3
mc1	87.1	255.9	20.3	2.9	75.6	232.8	11.7	2.3
st2	38.1	267.9	7.5	2.6	43.3	268.1	6.2	2.5
cu2	52.3	273.6	14.7	1.6	59.2	275.3	11.9	1.4
st1	50.4	272.0	8.3	0.9	52.8	277.8	6.2	0.6
mc3	53.6	239.5	6.8	3.9	65.1	242.0	5.2	3.7
st1	53.2	271.4	3.7	0.4	56.3	279.1	2.4	0.2
mc2	59.5	248.5	13.9	7.6	58.5	246.8	9.9	7.1
mc3	58.5	232.3	3.7	1.4	56.1	232.2	2.6	1.4
st2	70.3	256.6	9.8	2.0	72.1	258.9	4.8	1.2
mc4	71.8	211.0	5.0	1.1	83.1	221.0	2.9	1.0

Fonte: Produção da autora.

Tabela 4.2 - Centroides calculados com o conjunto de treinamento para o classificador operacional (baseado no GOES-8) e classificador atualizado (baseado no GOES-16) no horário das 15:00 UTC.

Classes	Centroides Atuais (GOES-8)				Centroides Novos (GOES-16)			
	vis (%)	ir (K)	texvis	texir	vis (%)	ir (K)	texvis	texir
sup	6.3	294.8	0.7	0.2	5.8	293.8	0.5	0.2
cu1	10.3	292.7	3.6	0.6	11.5	292.3	3.7	0.7
cu1	12.2	291.2	7.2	0.9	13.2	291.4	7.4	1.3
ci1	9.8	289.2	2.2	2.2	11.5	292.6	2.3	2.4
cu2	17.4	283.4	5.3	1.4	21.9	286.8	5.2	1.5
cu2	15.3	284.5	9.2	2.9	18.8	288.4	8.2	3.3
ci3	18.9	268.5	6.7	8.7	21.4	270.5	5.9	8.6
ci2	13.2	279.7	3.9	4.9	15.5	282.0	3.8	5.0
ci2	18.4	270.2	2.9	2.5	21.5	273.2	2.9	2.8
cu1	13.7	287.7	11.9	1.3	19.6	288.8	12.1	1.9
sup	8.1	287.3	1.1	0.6	10.5	288.7	1.4	0.8
st1	34.7	280.4	4.9	0.5	39.2	280.2	4.1	0.5
cu2	18.6	283.6	17.4	2.3	23.8	287.6	17.2	2.7
ci2	28.0	277.5	15.8	5.1	32.5	280.5	13.5	5.6
mc1	35.6	250.3	3.4	2.3	39.6	253.3	3.1	2.3
sup	10.3	301.7	1.0	0.7	12.6	304.5	1.2	0.9
ci4	31.8	255.9	5.3	5.2	34.1	256.8	4.9	5.4
cu2	31.5	284.9	9.9	0.9	35.2	284.6	9.0	1.1
ci4	34.2	254.1	13.6	14.3	34.7	259.6	11.3	13.5
cu2	63.8	269.9	26.8	4.3	53.5	281.1	21.5	3.8
mc1	70.4	254.8	12.2	3.0	70.7	266.3	10.8	2.9
st1	36.7	272.9	9.1	2.5	41.5	273.5	8.0	2.9
cu2	48.8	280.5	15.2	1.7	52.9	282.4	14.4	1.7
st1	51.1	279.8	8.6	0.6	59.1	278.9	8.2	0.6
mc3	54.8	235.0	5.5	4.0	65.9	242.5	4.4	4.1
st1	54.6	279.3	3.6	0.3	63.6	279.7	3.4	0.2
mc2	51.5	244.0	10.2	8.1	55.2	247.6	8.3	8.1
mc3	53.6	231.7	2.9	1.4	59.3	234.7	2.3	1.4
st2	56.7	259.9	4.7	1.3	74.4	263.9	4.4	1.3
mc4	74.7	217.5	3.7	1.2	84.4	221.0	2.1	1.1

Fonte: Produção da autora.

Tabela 4.3 - Centroides calculados com o conjunto de treinamento para o classificador operacional (baseado no GOES-8) e classificador atualizado (baseado no GOES-16) no horário das 18:00 UTC.

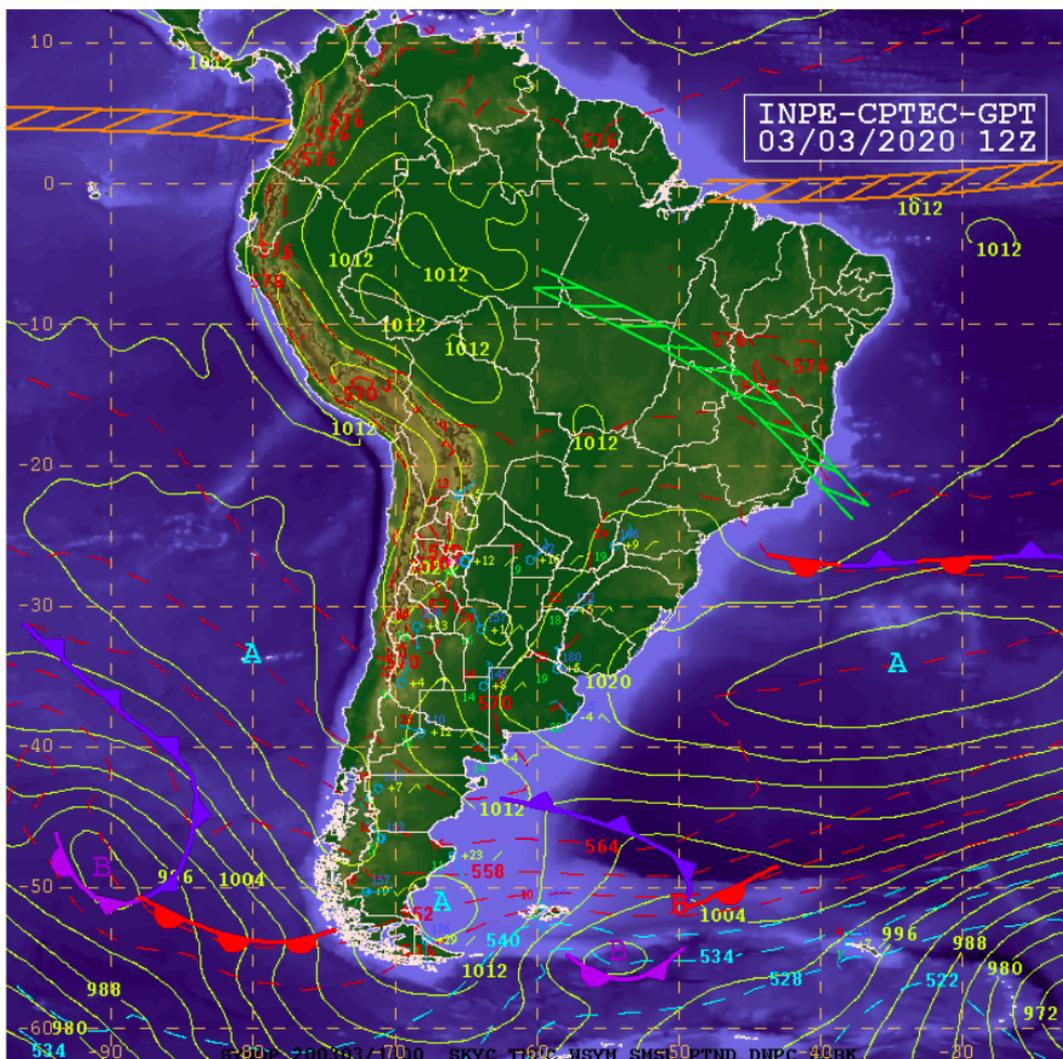
Classes	Centroides Atuais (GOES-8)				Centroides Novos (GOES-16)			
	vis (%)	ir (K)	texvis	texir	vis (%)	ir (K)	texvis	texir
sup	5.8	293.5	0.7	0.3	7.6	293.7	0.6	0.2
cu1	9.8	293.1	4.3	0.8	10.9	293.9	3.8	0.9
cu1	10.7	290.6	8.3	1.3	13.8	293.5	8.0	1.9
ci1	10.5	289.2	2.8	2.7	13.5	291.0	3.0	2.9
cu2	22.0	284.9	5.7	0.9	24.9	286.8	5.6	1.0
cu1	15.4	287.5	10.1	3.6	20.2	291.6	11.1	4.6
ci3	18.4	268.5	6.5	9.7	24.2	271.1	7.8	10.1
ci2	13.2	279.4	3.9	5.6	17.8	280.7	4.5	5.9
ci3	18.4	268.0	2.8	2.8	21.1	270.0	2.5	2.8
cu1	13.1	288.6	13.3	1.6	18.8	290.8	13.9	2.3
sup	10.2	286.5	1.6	0.9	16.5	287.4	1.5	0.6
st1	35.3	276.6	3.9	0.6	40.2	280.8	3.5	0.4
cu1	16.8	287.8	18.3	3.1	26.3	289.1	20.3	3.6
cu2	22.3	284.6	17.2	6.6	33.9	283.6	20.2	8.0
mc2	35.8	248.9	3.2	2.4	40.4	246.7	3.0	2.6
sup	12.0	302.4	1.2	0.7	15.5	306.1	1.5	1.0
ci4	29.7	252.8	5.0	5.8	36.3	252.9	5.2	6.1
cu2	28.9	284.7	10.2	1.1	33.6	286.3	10.2	1.6
ci4	30.9	259.2	14.0	15.87	40.3	257.6	15.6	16.3
cu2	61.2	278.3	20.3	3.9	68.4	279.4	23.4	4.2
ci4	46.4	270.9	12.3	4.8	51.6	272.7	13.3	5.1
cu2	31.8	273.0	7.7	2.8	38.5	272.0	6.9	3.2
mc1	44.9	282.2	14.4	1.8	57.9	281.9	15.0	1.9
st1	44.1	280.6	8.1	0.6	52.5	280.6	8.0	0.7
cu2	54.8	236.2	5.2	4.2	68.0	234.4	6.7	4.5
st1	53.5	274.1	3.4	0.5	65.6	277.5	3.6	0.4
mc2	51.7	245.2	10.4	9.3	60.2	242.8	11.0	9.7
mc3	49.6	233.4	2.3	1.3	62.4	233.0	2.6	1.5
st2	59.8	261.8	7.7	1.8	76.3	264.6	7.3	1.8
mc4	68.2	219.3	2.5	1.2	84.6	215.7	3.1	1.2

Fonte: Produção da autora.

Antes de comparar os resultados obtidos pelos dois classificadores, é necessário uma avaliação do estado da atmosfera para o respectivo dia em diferentes níveis atmosféricos.

Em superfície (Figura 4.4), é possível verificar na carta sinótica, a presença de sistemas meteorológicos atuantes na América do Sul, como a Zona de Convergência do Atlântico Sul, e a Zona de Convergência Intertropical, além de sistemas frontais tanto no Oceano Pacífico quanto no Oceano Atlântico. É possível notar que há presença de sistemas de alta pressão no próximo à costa do Rio Grande do Sul e Paraná, além de outro sistema de alta pressão no Oceano Pacífico, próximo à costa do Chile.

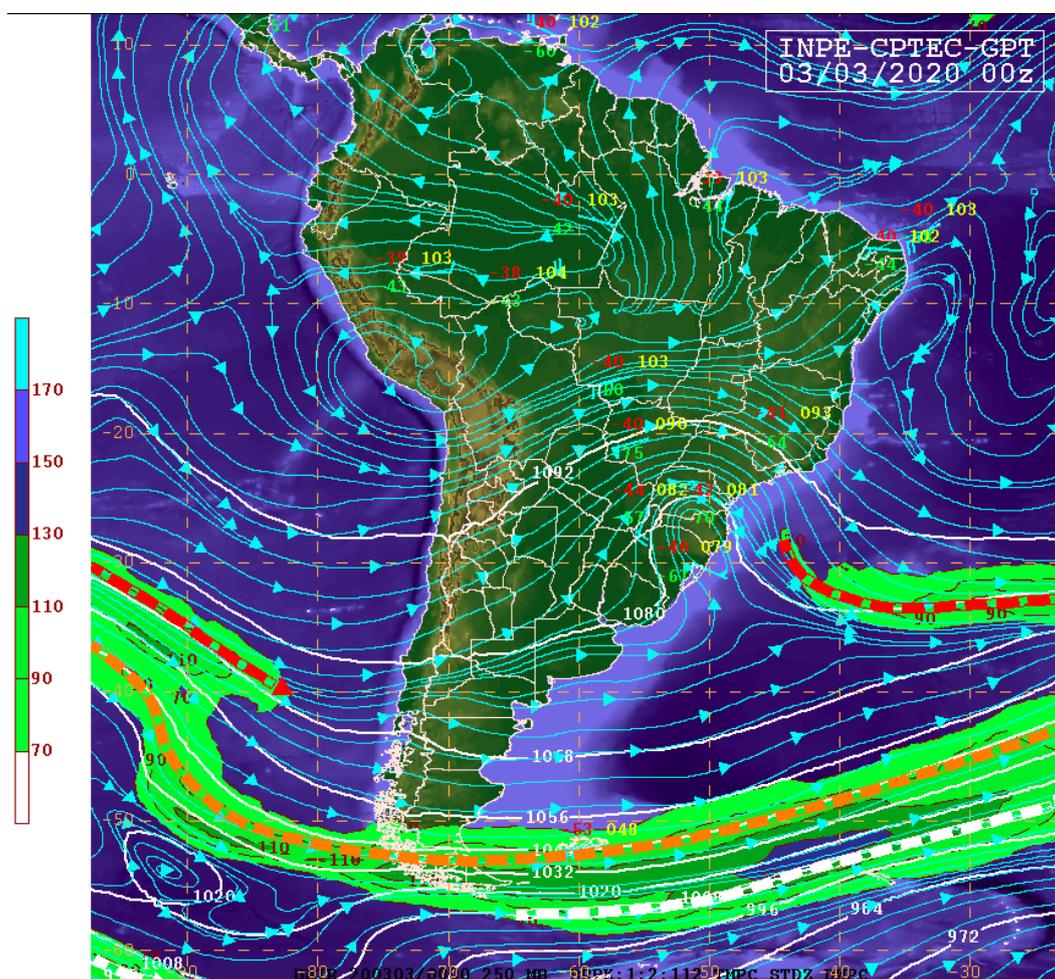
Figura 4.4 - Carta sinótica do dia 03/03/2020 no horário 12:00 UTC.



Fonte: CPTEC (2021).

Em altos níveis (Figura 4.5) é possível perceber um escoamento anticiclônico associado à Alta da Bolívia (AB), com a presença de Vórtices Ciclônicos de Altos Níveis (VCAN) no Oceano Atlântico, próximo à costa do Nordeste, assim como sobre a região Sul do Brasil, com o Jato Subtropical (JST) a leste deste último VCAN mencionado, já no extremo sul do continente sul americano o Jato Polar (JP) é observado. É possível perceber uma forte convergência sobre a parte norte da região Centro-Oeste e parte do Sudeste do Brasil.

Figura 4.5 - Carta de níveis altos para o dia 03/03/2020.

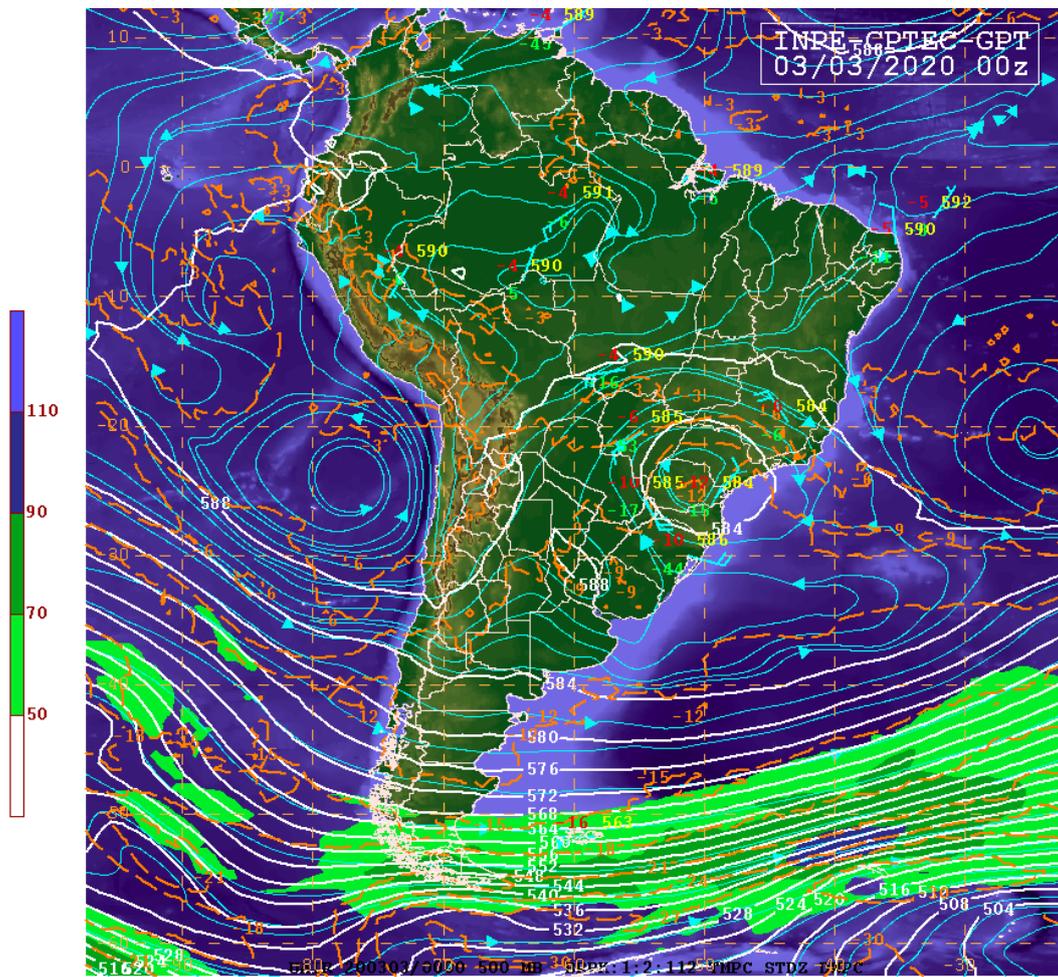


Fonte: CPTEC (2021).

Em médios níveis (Figura 4.6) é possível notar a presença de um vórtice ciclônico

sobre parte da região Centro-Oeste, da região Sul e parte da região Sudeste do Brasil referente ao VCAN em altos níveis, assim como é possível observar no Oceano Pacífico um vórtice anticiclônico referente à Alta da Bolívia, isto se deve a intensidade destes sistemas, que se estendem até níveis mais baixos. Enquanto no extremo sul do continente sul americano e oceanos nota-se uma região com baroclinia mais intensa, a qual está associada ao escoamento dos jatos em altos níveis.

Figura 4.6 - Carta de níveis médios para o dia 03/03/2020.

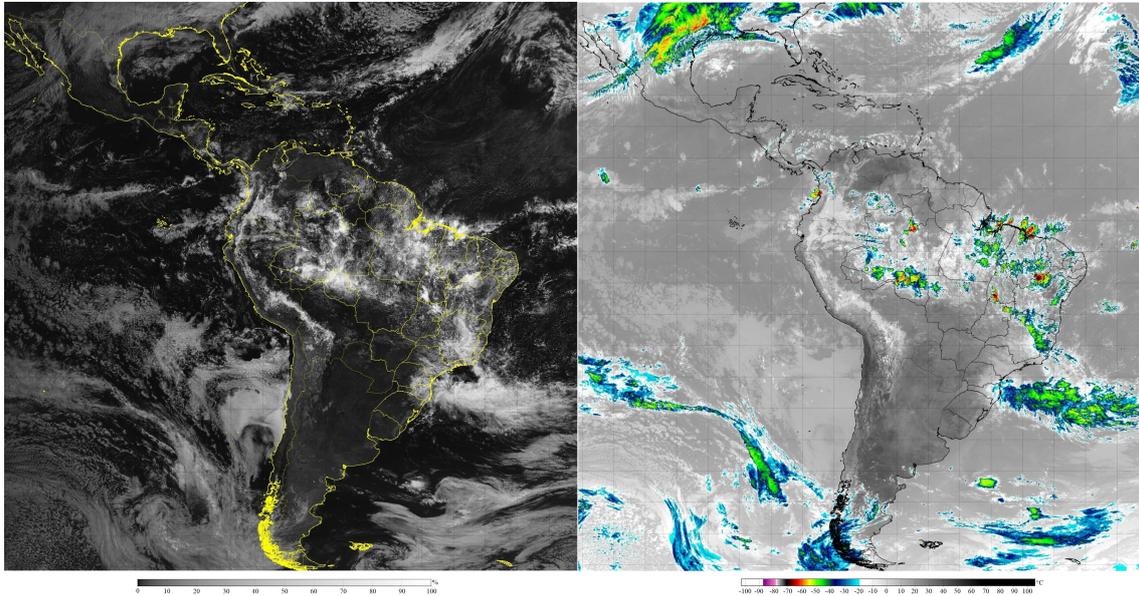


Fonte: CPTEC (2021).

A Figura 4.7 mostra imagens na região do visível (esquerda) e infravermelho (direita), exibindo o padrão de nuvens no horário 15:00 UTC, complementando as

informações observadas nas cartas sinóticas (Figura 4.4, Figura 4.6 e Figura 4.5).

Figura 4.7 - Imagem do visível e infravermelho no dia 03/03/2020 no horário 15:00 UTC.



Fonte: CPTEC (2021).

Buscando comparar os dois classificadores de forma mais detalhada, o dia 03/03/2020 foi escolhido por não estar inserido no conjunto de treinamento. As imagens classificadas para este dia foram geradas utilizando os dois classificadores (operacional e atualizado) nos três horários. Dessa forma foi possível verificar visualmente as situações em que os classificadores divergiram em relação a cada classe.

Outra forma de verificar a ocorrência dessa discrepância foi através de histogramas tridimensionais. Os valores localizados na diagonal principal representam as situações em que os dois classificadores (operacional e atualizado) indicaram a mesma classe, enquanto os pontos fora dessa diagonal representam os pontos onde houveram discordância entre os dois classificadores (Figuras 4.8, 4.9 e 4.10).

As Figuras 4.8, 4.9 e 4.10 apresentam os histogramas tridimensionais da distribuição das classes de nuvens definidas pelo classificador operacional (chamado de GOES-8 na figura) e o classificador atualizado (chamado de GOES-16 na figura). São apre-

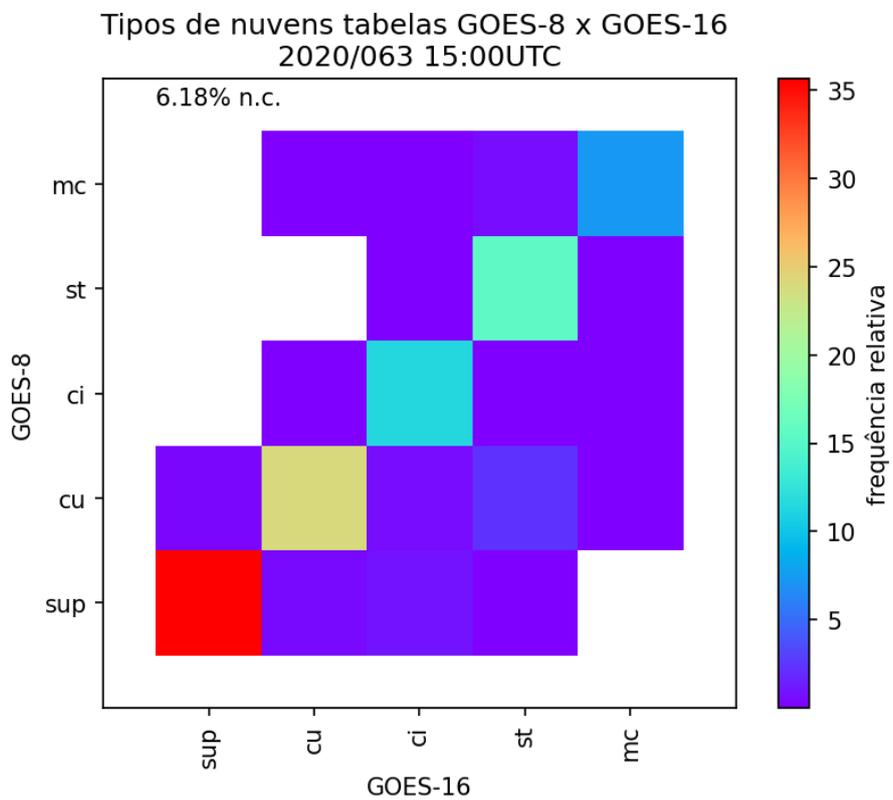
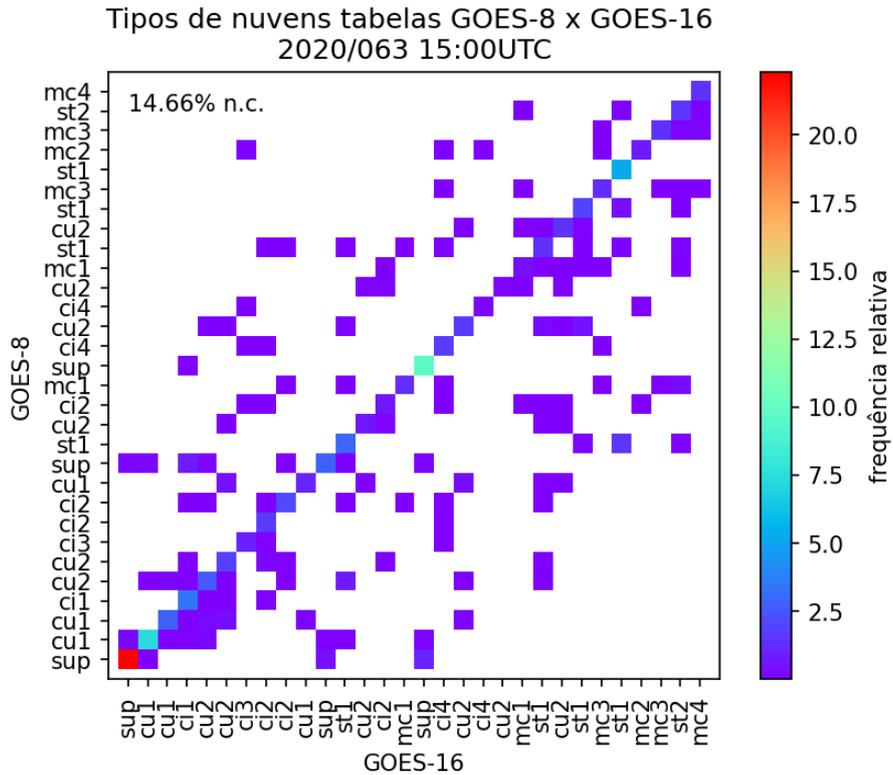
sentadas as distribuições considerando todas as 30 classes (topo) e as distribuições em que as variações de um mesmo tipo de nuvem, por exemplo, ci, agrupadas sob uma mesma classe geral representando todas as subclasses ci (baixo). Nos histogramas são mostrados também os valores (em %) correspondentes às discordâncias entre os dois classificadores indicados por "n.c."(não concordam).

Pode-se notar que dentre os três horários, 12:00 UTC apresenta maior discordância entre os dois classificadores, com um total de 15,10% (considerando as subclasses agrupadas). Nesse horário, os tipos de alvos que mais apresentaram discordâncias foram a classe 2 do classificador operacional (GOES-8), correspondente à cu1 e a classe 1 do classificador atualizado (GOES-16), correspondente à sup.

No horário 15:00 UTC, a discordância entre os dois classificadores foi de 6,18%, a menor dentre os três horários. O tipo de alvo que mais apresentou discordância entre os dois classificadores é a classe 2 do classificador operacional (cu1) e a classe 12 do classificador atualizado (st1). E assim como na figura anterior, as outras diferenças entre as classes é próximo de 1. As classes que mais apresentaram discordância entre os dois classificadores foram a classe 2 (cu1) para o operacional e as classes 4 (ci1) e a 12 (st1) no atualizado.

No horário 18:00 UTC, a discordância entre os dois classificadores foi de 7,51%. Os tipos de alvos que mais apresentaram discordâncias entre os dois classificadores foram as classes 2 (cu1) do classificador operacional e a classe 4 (ci1), como também a classe 12 (st1) do classificador atualizado, sendo a primeira diferença mais sutil do que a segunda. E assim como nas figuras anteriores, as outras diferenças entre os dois classificadores ficam em torno de 1. Entre os classificadores, as classes mais discordantes foram a classe 2 (cu1) e a 4 (ci1) no operacional e a classe 2 (cu1), a 4 (ci1) e a 12 (st1) no atualizado.

Figura 4.9 - Tipos de nuvens entre o classificador operacional (baseado no GOES-8) e o classificador atualizado (baseado no GOES-16) e o classificador atualizado de todas as classes (cima) e classes agrupadas (baixo) no horário das 15:00 UTC do dia 03/03/2020.



Fonte: Produção da autora.

Como foi visto nas Figuras 4.8, 4.9 e 4.10, os dois classificadores, apresentam diferentes identificações para alguns tipos de alvos. Como pode-se perceber, no horário 12:00 UTC, a maior discordância foi entre as classes 2 (cu1) e 1 (sup) e a 2 (cu1) e 12 (st1).

Antes de analisar a imagem de satélite a fim de investigar mais detalhadamente os pontos onde ocorreram os erros, foi confeccionado um gráfico de linha referente aos 3 horários, entretanto, foram adicionados 4 outros horários próximos, com 30 minutos de intervalo entre eles, para ser possível perceber tanto a variabilidade dos horários, quanto a amplitude na porcentagem dos erros nos respectivos horários. O erro entre os dois classificadores, foram calculados para 24 dias distribuídos ao longo do ano de 2019, os dias utilizados foram iguais aos da [Seção 4.1](#).

Na [Figura 4.11](#) o horário 12:00 UTC é representado pela linha azul, nesse horário o erro entre as classes agrupadas oscila em torno de 14,4% durante todo o período do ano, isso corrobora com o mostrado na [Figura 4.8](#), calculada para um dia, onde o erro foi de 15,10%. Entretanto, ao analisar a variabilidade pode ser notado que a partir de agosto a porcentagem dos erros diminuem consideravelmente, principalmente no horário 11:00 UTC, representado pela linha vermelha.

Figura 4.11 - Fração de pixels com classificação discordante entre os classificadores original (GOES-8) e atualizado (GOES-16) para o intervalo centrado em 12:00 UTC.



Fonte: Produção da autora.

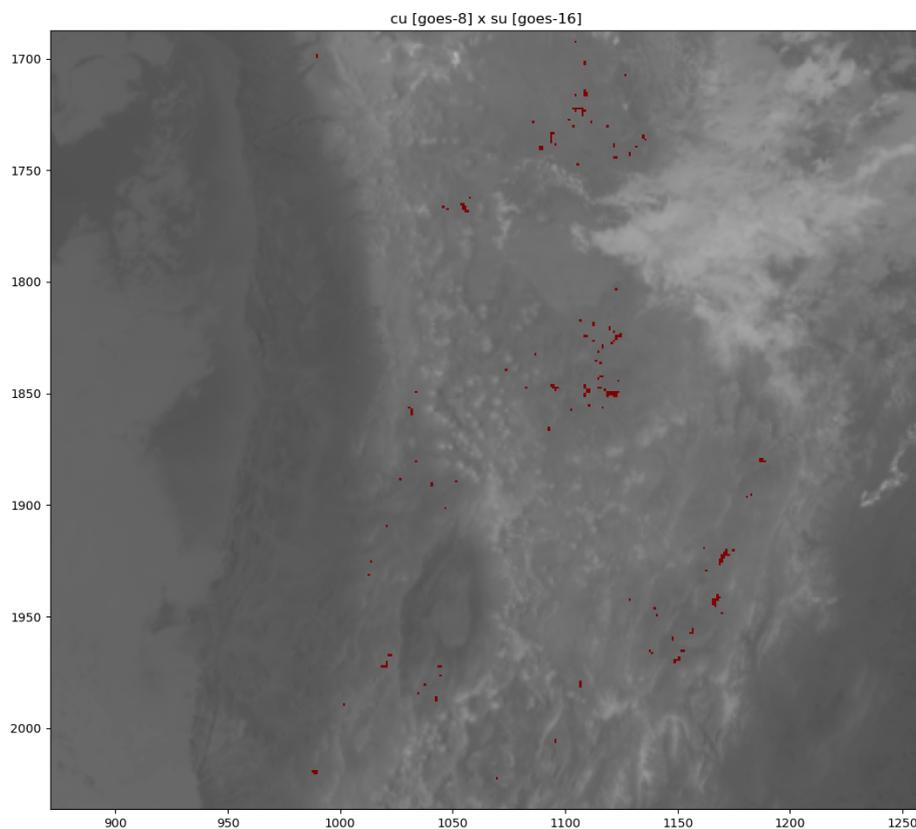
Ao plotar imagens do infravermelho, a fim de verificar os pixels onde houveram diferenças nas classificações entre os dois classificadores, pode-se perceber que de maneira geral, essas diferenças não são facilmente percebidas na imagem, pois apesar da porcentagem ser relativamente alta para todas as classes, em torno de 15% no horário de 12:00 UTC como foi visto. Contudo, a diferença entre os pares de classes nos dois classificadores possuem percentagens bem menores, isto torna a identificação mais difícil, sendo necessário uma análise mais detalhada.

Portanto, algumas classes que apresentaram diferenças entre os dois classificadores foram averiguadas, mas mostrada e discutida apenas uma dessas diferenças em cada um desses horários. Apesar de serem mostrados apenas uma região, essas diferenças ocorrem em toda a área de estudo, mas mostrada unicamente um parte.

A Figura 4.12 mostra a região dos Andes e pode-se notar os pontos em vermelho, os

quais representam essa diferença entre os dois classificadores. Nota-se que todos os pontos, ou quase todos os pontos, mostram que a classificação correta foi realizada pelo classificador atualizado, mostrando superfície. Mesmo ao analisar uma área bem menor que toda a região de estudo, é possível perceber que os pontos ainda são bem poucos, no entanto ainda gera dúvidas em definir a precisão dos dois classificadores.

Figura 4.12 - Diferença entre a classe cumulus do classificador operacional (GOES-8) e a classe superfície do classificador atualizado (GOES-16) no horário das 12:00 UTC do dia 03/03/2020.



Fonte: Produção da autora.

Nuvens do tipo cumulus, no geral, possuem a temperatura do topo quente, podem ser confundidas com algumas superfícies (céu claro) ligeiramente mais frias. Portanto, é necessário explorar alguns desses pontos mais detalhadamente. Mesmo tendo uma imagem relativamente mais detalhada a definição dos pontos ainda geram dúvidas,

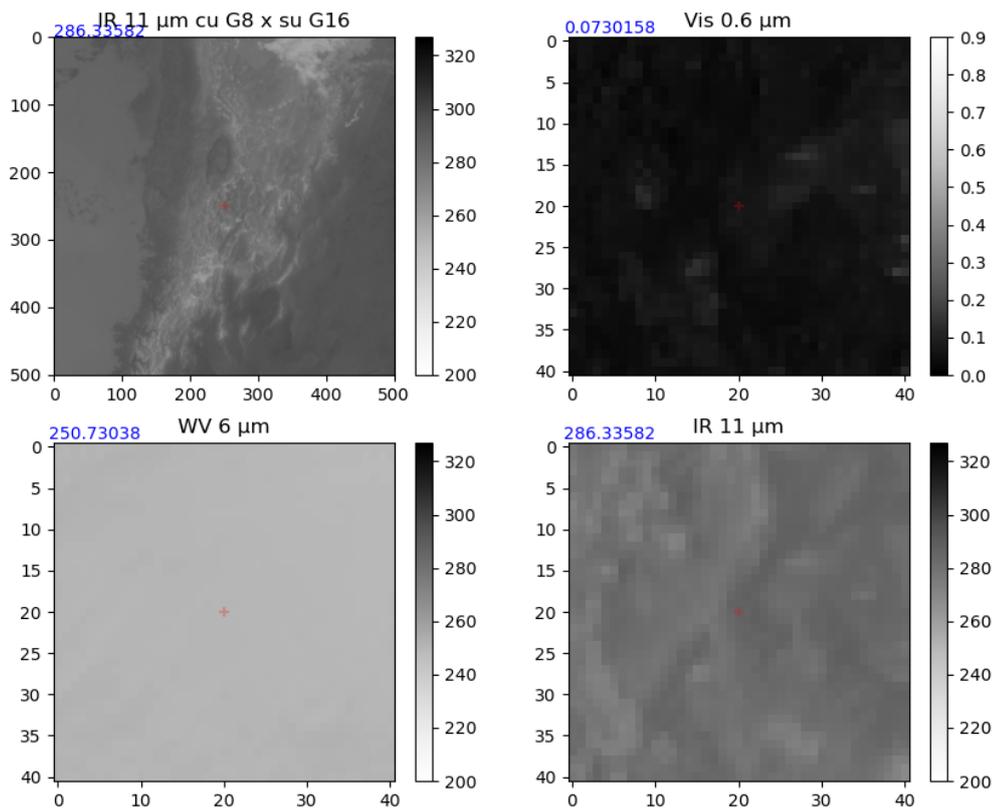
por isso, uma avaliação somente qualitativa não é suficiente. Portanto, para definir qual dos dois classificadores obteve mais êxito em classificar corretamente, foram isolados alguns pontos, os quais forneceram informações sobre a temperatura de brilho e reflectância.

A [Figura 4.13](#) mostra o ponto de latitude e longitude referente à diferença nas classificações entre o classificador operacional (indicado como G8) e o classificador atualizado (indicado como G16), representado pelo ponto vermelho. No canto superior de cada imagem há uma indicação do valor referente à cada uma das características, infravermelho e visível.

A imagem esquerda superior, mostra a localização do ponto, com um zoom de 250 pontos em torno do ponto central (ponto de interesse), indicando a localização do mesmo. Nas outras três imagens, há 20 pontos de distância em torno do ponto central, que mostra detalhes da região, permitindo analisar a homogeneidade (ou não) da superfície em seu entorno. A imagem direita superior, mostra a reflectância com o valor 7, valores dessa magnitude indicam céu claro, pois na ausência de nuvens as superfícies apresentam baixa refletividade. A imagem inferior esquerda indica a região do vapor d'água em altos níveis ($6,2 \mu m$) e a imagem inferior direita indica a região do infravermelho com valor na temperatura de brilho igual a 286,3 K, onde essa temperatura pode indicar superfície (céu claro) ou o topo de nuvens baixas, como a nuvem cumulus, por exemplo.

Entretanto, como o ponto está localizado na região dos Andes, essa temperatura de brilho que poderia ser confundido com nuvens cumulus. Por ser uma região de altitude, a temperatura de brilho é mais baixa do que regiões mais baixa, como por exemplo, regiões localizadas ao nível médio do mar. Nesta região em específico estava atuando um sistema de alta pressão, corroborando com a ausência de nuvens, já que um sistema de alta pressão impossibilita convecção. Ao analisar outras imagens (não mostradas), mostrando também pontos isolados de latitude e longitude, apresentaram as mesmas características, reflectância e temperaturas de brilho similares à mostrada na [Figura 4.13](#), sendo assim, é possível afirmar que dentre os dois classificadores, o que mais obteve êxito na classificação foi o classificador atualizado.

Figura 4.13 - Exemplo de um ponto onde houve diferença entre as classes do classificador operacional (baseado no GOES-8) e atualizado (baseado no GOES-16) no horário das 12:00 UTC do dia 03/03/2020.

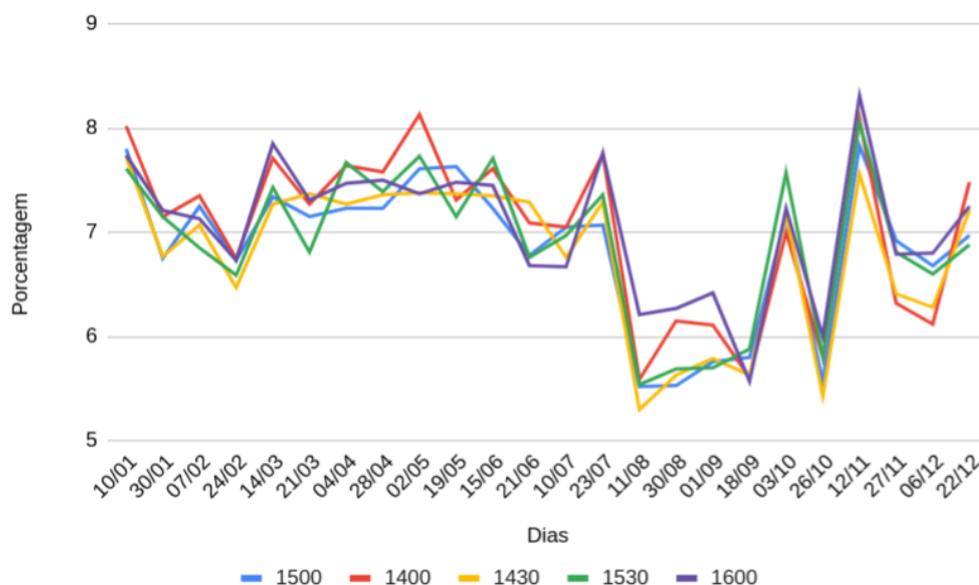


Fonte: Produção da autora.

A Figura 4.14 mostra o gráfico de linha para o erro entre as classes agrupadas dos dois classificadores no horário das 15:00 UTC, onde a linha azul representa este horário. Assim como na Figura 4.11, foram utilizados horários adjacentes a fim de observar a variabilidade. Na média para quase todo o período do ano, o erro fica em torno de 7%, porém a partir do mês de agosto há uma queda nos valores, assim como no gráfico anterior, mas diferentemente, aqui todos os horários plotados apresentam uma amplitude mais homogênea. A média do erro durante o período analisado foi

de 6,18%, similar à Figura 4.9 que foi calculada apenas para um dia.

Figura 4.14 - Fração de pixels com classificação discordante entre os classificadores operacional (baseado no GOES-8) e atualizado (baseado no GOES-16) para o intervalo centrado em 15:00 UTC.



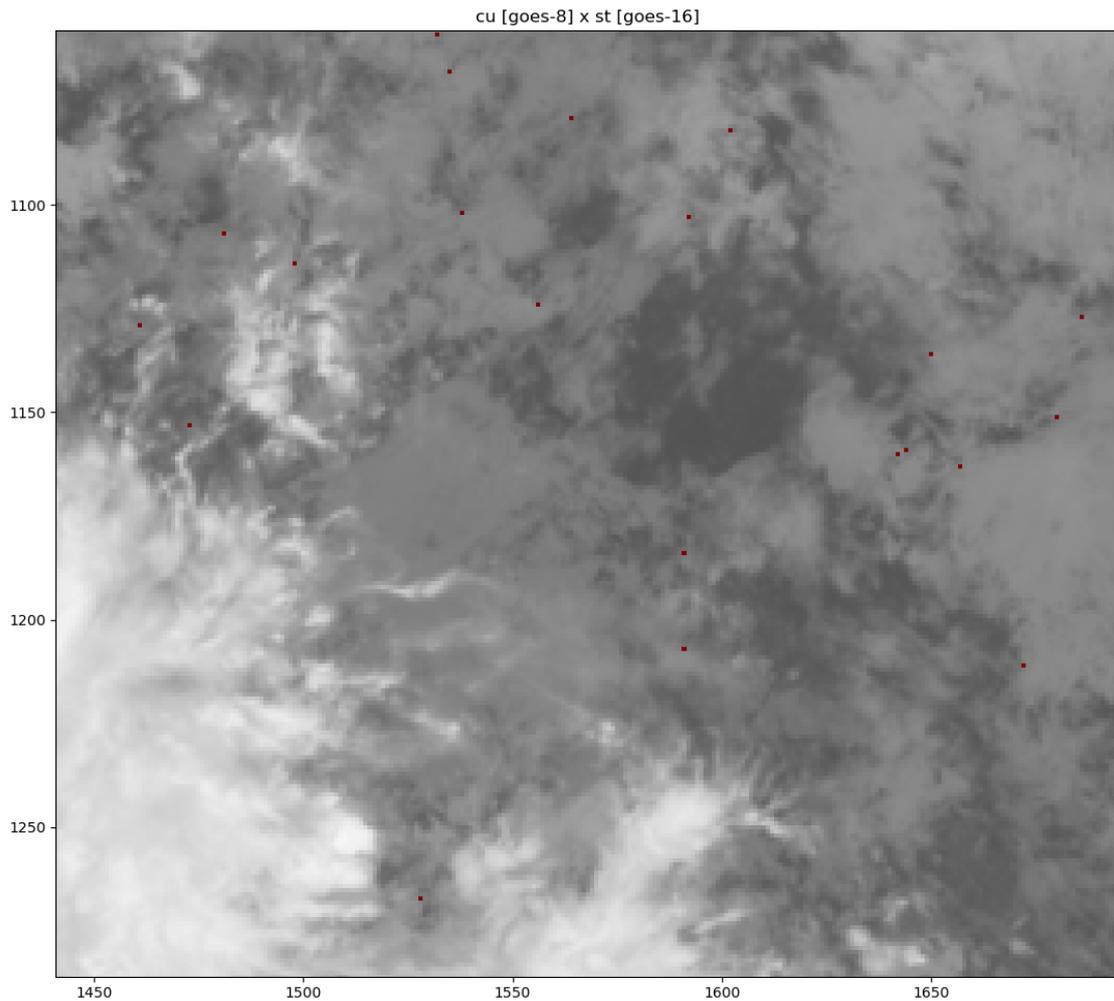
Fonte: Produção da autora.

No horário 15:00 UTC os erros mais significativos foram entre a classe 2 (cu1) do classificador operacional e a classe 12 (st1) do classificador atualizado. Para essa diferença específica, a frequência do erro foi de aproximadamente 3%. Essa porcentagem dificulta ainda mais a busca e análise desses pixels. A imagem abaixo mostra a região correspondente a porção Norte e Centro-Oeste do Brasil, onde há ocorrência de diversos tipos de nuvens.

Observando a Figura 4.15 é possível notar que os pontos onde se encontram as diferenças entre os dois classificadores estão sobre uma região onde há uma grande ocorrência de nebulosidade. Na região especificada, há predominância de nuvens do tipo stratus, diagnosticada através da nefanálise, sendo possível inferir que a

classificação correta foi realizada pelo classificador atualizado.

Figura 4.15 - Diferença entre a classe cumulus do classificador operacional (GOES-8) e a classe stratus do classificador atualizado (GOES-16), no horário das 15:00 UTC do dia 03/03/2020.



Fonte: Produção da autora.

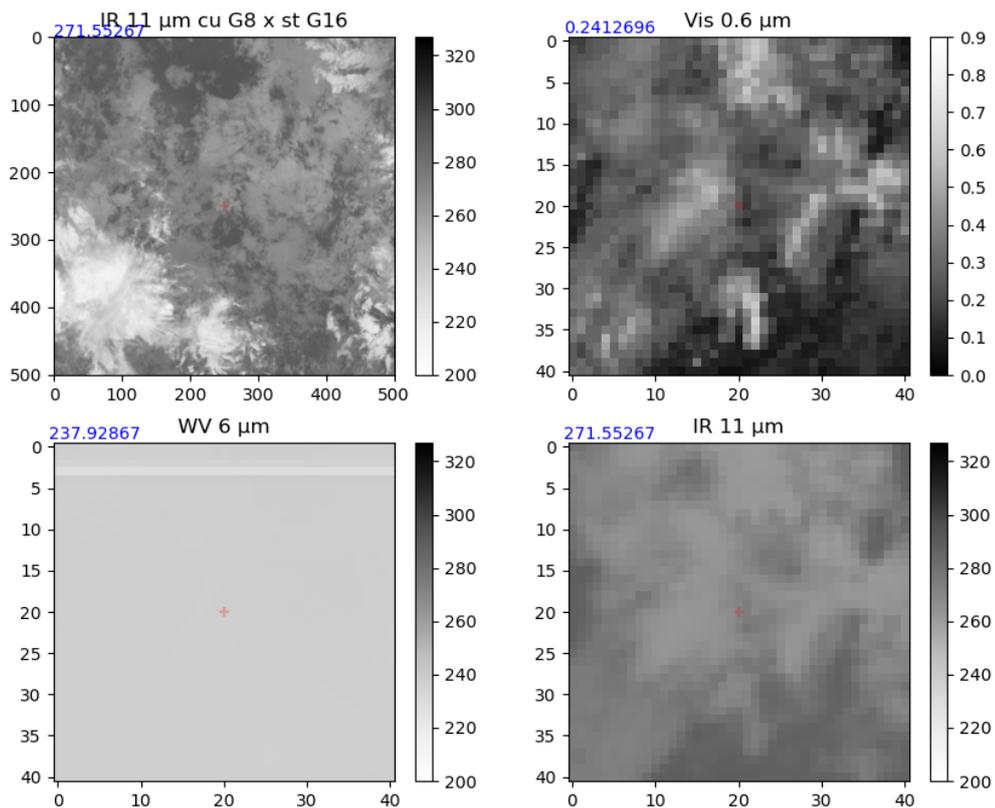
Para uma maior certeza na afirmação feita através da nefanálise, foram geradas diversas figuras de vários pontos isolados (não mostrados) onde foi possível verificar suas características, tanto de reflectância quanto de temperatura de brilho.

A [Figura 4.16](#) mostra um destes pontos, o qual está inserido na imagem anterior, a reflectância referente à este ponto é de 24 e a temperatura de brilho é igual a 271,5 K. Estes valores indicam a nuvem classificada como sendo stratus.

Apesar de uma nuvem cumulus ter um valor de reflectância baixo, os valores de temperatura de brilho correspondentes à essa reflectância são mais elevados, portanto, para que uma nuvem cumulus tenha uma temperatura de brilho mais baixa, é necessário que haja desenvolvimento vertical, conseqüentemente esse desenvolvimento resulta em uma maior reflectância. Por outro lado, as nuvens stratus são mais homogêneas, podendo apresentar temperaturas de brilho mais baixas e ter uma reflectância mais baixa, devido sua homogeneidade e por se estendem horizontalmente por uma área maior.

Ao analisar os conjuntos de centroides ([Tabela 4.2](#)) referentes ao mesmo horário, alguns centroides relacionados às nuvens do tipo cumulus possuem temperatura de brilho próximo ao mostrado na [Figura 4.16](#). Porém a reflectância é significativamente mais elevada, enquanto alguns centroides referentes às nuvens do tipo stratus possuem valores mais baixo de reflectância e temperaturas de brilho. Portanto, pode-se concluir que o classificador que obteve melhor resultado foi o atualizado, esta afirmação pôde ser confirmada após análise de várias figuras de pontos isolados que mostraram as mesmas características apresentadas na figura a seguir, além disso os detalhes mostrados tanto na imagem do visível quanto na imagem do infravermelho indicam uma região onde há nuvens mais uniformes, representando um padrão de nuvens do tipo stratus.

Figura 4.16 - Ponto de latitude e longitude onde houve diferença entre as classes do classificador operacional (baseado no GOES-8) e atualizado (baseado no GOES-16) no horário das 15:00 UTC do dia 03/03/2020.



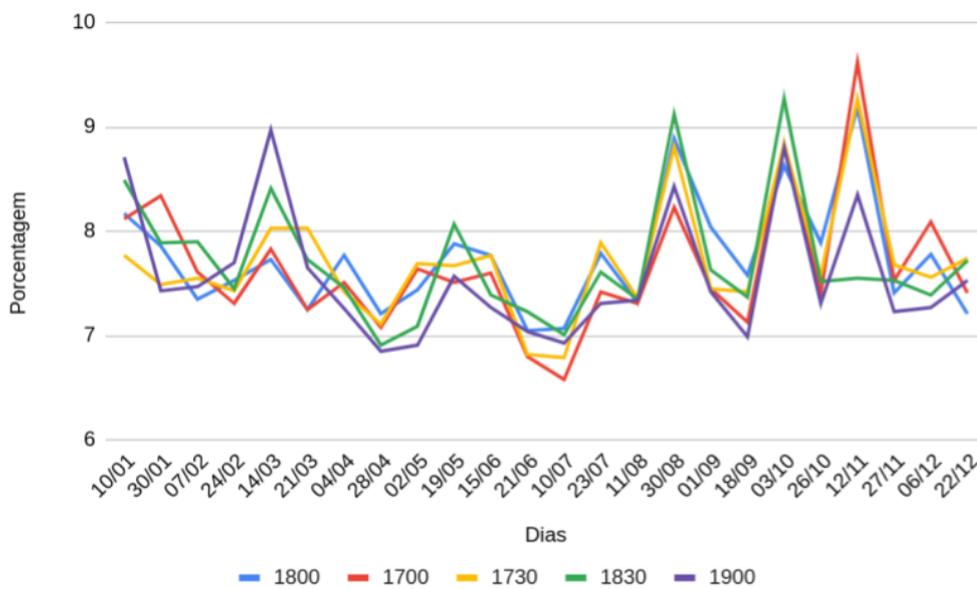
Fonte: Produção da autora.

O horário 18:00 UTC, representado pela linha azul (Figura 4.17), a média o erro entre as classes agrupadas calculado para todo o período entre os dois classificadores foi de aproximadamente 7,5%, enquanto na Figura 4.10, calculado apenas para um dia, foi de 7,51%.

Todos os horários apresentaram a mesma variabilidade, entretanto, diferente dos outros dois horários, onde a partir do mês de agosto houve uma queda nos valores

dos erros, aqui o comportamento foi o oposto, apresentando um aumento a partir deste mês, com valores chegando próximo à 10% no mês de novembro. Outro ponto a ser notado é que o período onde houve um decréscimo nos valores foi durante o período de abril a julho, ficando entre 7 e 8%.

Figura 4.17 - Fração de pixels com classificação discordante entre os classificadores operacional (baseado no GOES-8) e atualizado (baseado no GOES-16) para o intervalo centrado em 18:00 UTC.

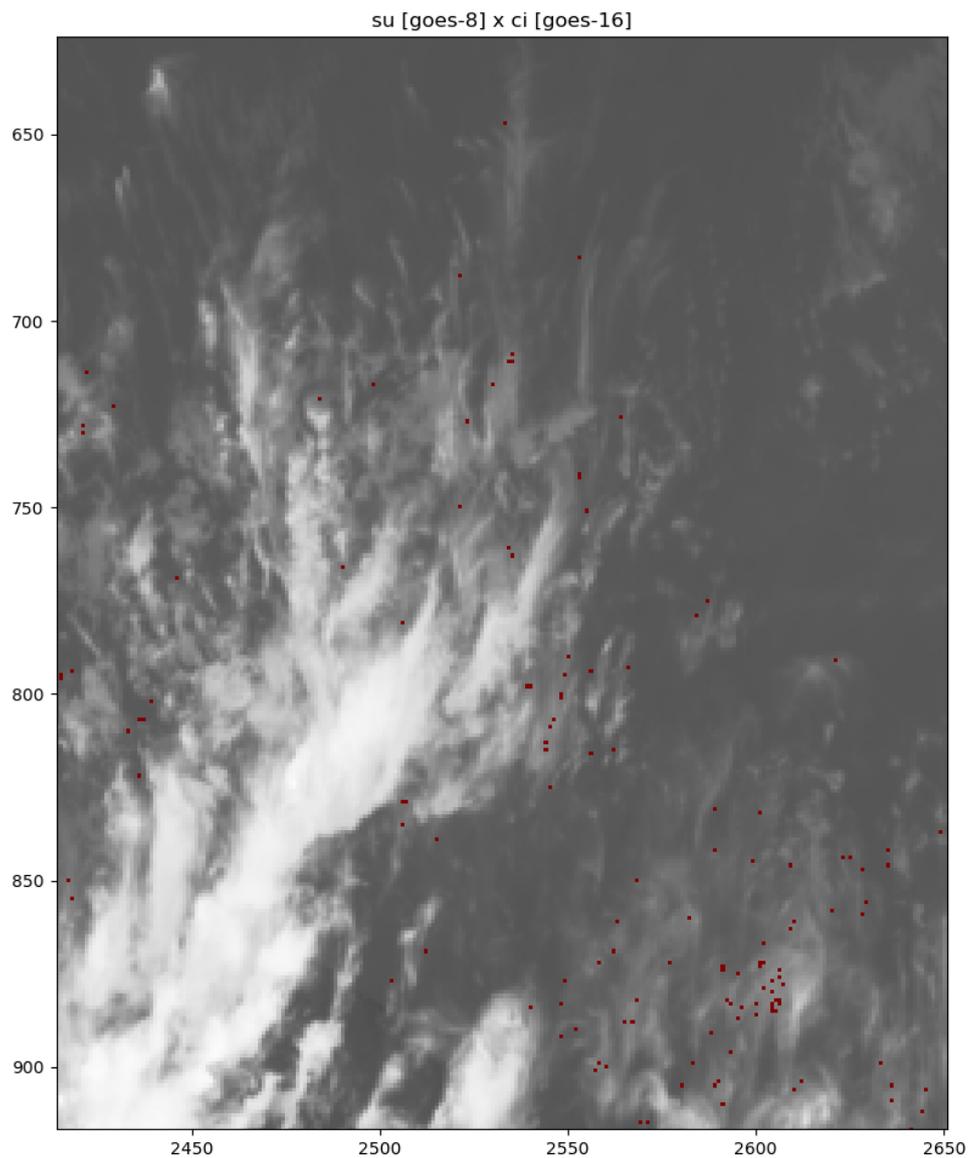


Fonte: Produção da autora.

Como dito anteriormente, a porcentagem entre as classes agrupadas, apesar de ser relativamente alta, a diferença entre classes específicas, como é o exemplo aqui utilizado, superfície e cirrus é bem menor. Ao observar a imagem do infravermelho com as duas classes, superfície para o classificador operacional e cirrus para o classificador atualizado, no horário 18:00 UTC, é possível notar que a maior parte dos pontos estão classificados como cirrus, sendo possível inferir que a categorização correta foi realizada pelo classificador atualizado. Essa diferença entre os classificadores pode ser parcialmente explicada devido aos pontos aparentarem ser nuvens cirrus semi-

transparentes, que eventualmente pode ser confundida com superfícies. A área em destaque na Figura 4.18, está sobre o Oceano Atlântico próximo à costa do estado do Ceará.

Figura 4.18 - Diferença entre a classe superfície do classificador operacional (baseado no GOES-8) e a classe cirrus do classificador atualizado (baseado no GOES-16), no horário das 18:00 UTC do dia 03/03/2020.



Fonte: Produção da autora.

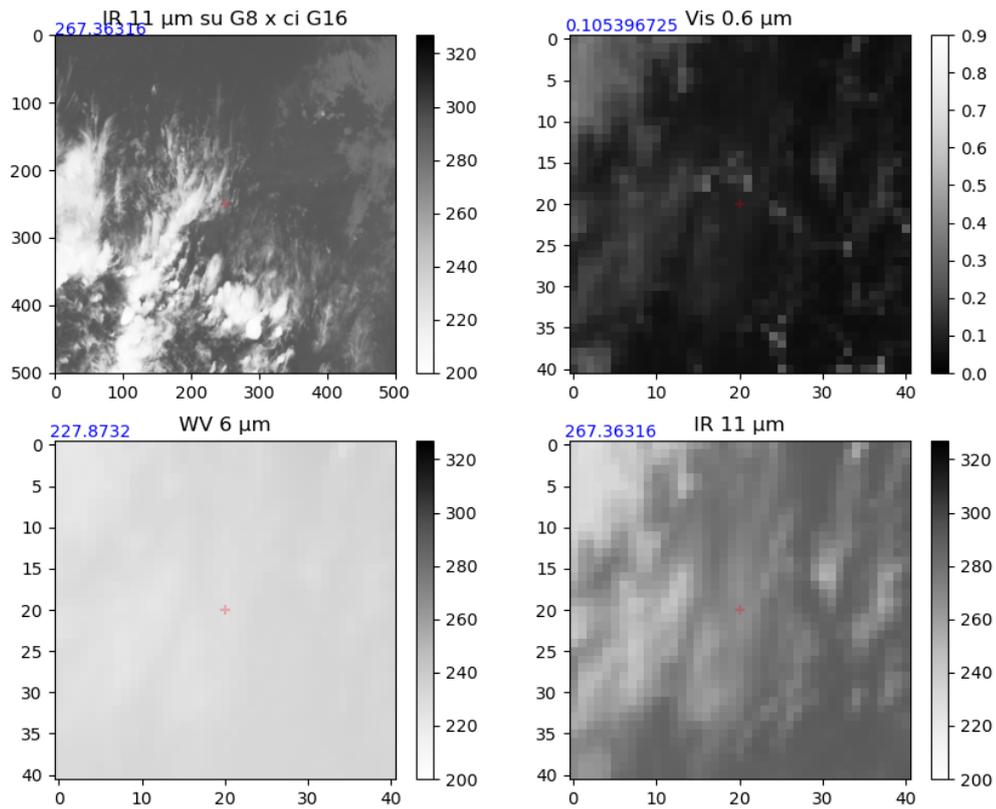
Sob a perspectiva de uma análise quantitativa, para assim ter mais certeza na classificação, assim como nos dois horários anteriores, foram avaliadas diversas figuras (não mostradas) de pontos isolados, com a finalidade de verificar as características destes pontos de latitude e longitude.

A [Figura 4.19](#) mostra um dos pontos inserido na imagem anterior, a reflectância referente a este ponto é de 10 e temperatura de brilho igual a 267,4 K. Estes valores, principalmente a temperatura de brilho indicam que o ponto foi classificado como uma nuvem do tipo cirrus.

Apesar da reflectância ser baixa, principalmente tratando-se de nuvens altas compostas de gelo, se deve ao fato da mesma ser uma nuvem cirrus semi-transparente, em decorrência disto, parte da radiação proveniente do solo ultrapassa a nuvem, chegando ao sensor a bordo do satélite. Afirmar que o ponto analisado, assim como a maior parte dos pontos da imagem anterior, refere-se a uma nuvem cirrus, pois estão localizados sobre o oceano que é uma superfície sem rugosidade, e além de analisar somente o valor da temperatura de brilho, é possível verificar as características rugosas, sendo plausível afirmar que de fato os pontos são correspondentes à nuvens.

Essas mesmas características discutidas puderam ser verificadas ao analisar outros pontos isolados (não mostrados), apresentando características semelhantes.

Figura 4.19 - Ponto de latitude e longitude onde houve diferença entre as classes do classificador operacional (baseado no GOES-8) e atualizado (baseado no GOES-16) no horário das 18:00 UTC do dia 03/03/2020.



Fonte: Produção da autora.

4.3 Comparação entre os produtos da NOAA e os classificadores operacional (baseado no GOES-8) e atualizado (baseado no GOES-16)

Nesta seção serão discutidos como os classificadores operacional e atualizado se comportaram em relação aos produtos da NOAA. Os produtos utilizados para elaborar uma comparação junto aos dois classificadores foram:

- a) temperatura do topo da nuvem (*Cloud Top Temperature*), com o intuito de verificar se os tipos de nuvens classificadas através dos dois classificadores eram coerentes com a temperatura do topo da nuvem;
- b) fase da nuvem (*Cloud Top Phase*), a fim de verificar se existiram algumas situações discrepantes, como por exemplo, nuvens cumulus formadas por gelo ou nuvens cirrus formadas de gotículas de água e;
- c) máscara de nuvem (*Clear Sky Mask*) com a finalidade de verificar se as áreas classificadas como superfícies nos dois classificadores estão adequadas ao produto oficial, além disso, esse produto foi utilizado com o objetivo de identificar situações onde houveram ocorrência de nuvens do tipo cirrus semitransparentes e foram classificadas como superfície.

A comparação entre os classificadores operacional e atualizado com os produtos de nuvem da NOAA permite melhor compreender as deficiências desses métodos. Como os produtos utilizados aqui possuem a mesma resolução espacial que os dados do ABI/GOES-16 empregados (2 km), por este motivo não foi necessário realizar nenhuma adequação na resolução espacial dos produtos.

Assim como na comparação entre os dois classificadores operacional e atualizado, foi realizada uma análise para vários dias, como os utilizados na [Seção 4.1](#). Entretanto, de todos os produtos utilizados, apenas o produto fase de nuvem está disponível para todo o ano de 2019, os outros produtos estão disponíveis apenas a partir do início de dezembro do mesmo ano, isto impossibilitou a utilização dos mesmos dias analisados em [Seção 4.1](#).

Para analisar o comportamento dos classificadores operacional e atualizado com relação aos produtos de nuvens, foram utilizados dados referente ao ano de 2020. Com relação aos dados referentes ao ano de 2020, foram selecionados 20 dias. Vale ressaltar que este é um período diferente dos utilizados para o treinamento, já que os centroides gerados a partir da aplicação do método *k-means* são utilizados aqui para poder comparar os classificadores com os produtos oficiais da NOAA.

4.3.1 Temperatura do topo da nuvem

A comparação permitiu construir três tabelas (1 para cada produto NOAA) para os mesmos três horários de referência (12:00, 15:00 e 18:00 UTC). No produto de

temperatura do topo da nuvem (*Cloud Top Temperature*), foram destacados oito diferentes temperaturas de brilho para os cinco diferentes tipos de alvos (Tabela 4.4), onde para cada uma dessas temperaturas e superfícies, temos os valores (em porcentagem) de acertos ou erros, a depender do tipo de par alvo-temperatura. Além das colunas referentes a cada uma das temperaturas de brilho, a última coluna (-1), é referente à porcentagem que não foi classificada em nenhuma das outras temperaturas, isto é, consideradas como classificação correta.

Ao avaliar o alvo do tipo superfície no horário das 12:00 UTC, é possível perceber que o classificador operacional identificou em cerca de 63,1% de forma correta. Contudo, aproximadamente 19% foram classificados com temperaturas mais altas (297 e 282 K) e aproximadamente 18% foram classificados com temperaturas mais baixas (267 a 207 K).

Enquanto o classificador atualizado, obteve um resultado cerca de 3% melhor ao identificar a superfície (céu claro). Ao analisar os outros dois horários, é possível notar que a categorização da superfície foi melhor (acima de 70%). Entretanto, dentre os três horários, em 15:00 UTC, o classificador operacional teve um melhor desempenho, já no horário das 18:00 UTC os dois classificadores foram equivalentes.

Cumulus são nuvens geralmente mais baixas, compostas majoritariamente de gotículas de água, a não ser que possuam um desenvolvimento vertical intenso (WALLACE; HOBBS, 2006; COTTON et al., 2011). É possível notar que as maiores porcentagens, variam em torno de 54%, nos três horários, encontram-se entre as temperaturas de brilho 297 e 282 K, nos dando um bom indício de uma classificação correta deste alvo. Cerca de 30% da categorização das nuvens cumulus no horário 12:00 UTC e aproximadamente 20% nos dois outros horários (15:00 e 18:00 UTC), essas altas porcentagens referem-se a uma identificação errônea deste alvo (última coluna) pelo classificador operacional.

O classificador atualizado, apresentou cerca de 52% de acerto nos três horários ao identificar nuvens cumulus (temperaturas entre 297 e 282 K), por outro lado, a categorização errônea deste tipo de nuvem, representado na última coluna, foi igual ao classificador operacional, apresentando maior porcentagem no horário de 12:00 UTC (27%).

Ao analisar as nuvens do tipo cirrus, no classificador operacional pode-se notar que

esta classe foi majoritariamente classificada entre as temperaturas 252 e 207 K, correspondendo a uma porcentagem de aproximadamente 74% nos horários 12:00 e 18:00 UTC. No horário 15:00 UTC essa porcentagem foi ligeiramente menor 69,3%.

Mesmo apresentando porcentagens altas na identificação correta, cerca de 20% das nuvens cirrus foram categorizadas entre as temperaturas de brilhos mais altas (297 a 267 K), essas classificações podem indicar a presença de nuvens cirrus semi-transparentes. As nuvens cirrus semi-transparente por não serem tão espessas são bastantes difíceis de detectar, além disso parte da radiação emitida pela superfície pode atravessar a nuvem e alcançar o sensor do satélite, podendo ser confundido com superfície (céu claro) (GAO et al., 1993; DESSLER; YANG, 2003).

O classificador atualizado, no geral, teve melhor êxito principalmente nos horários 12:00 e 18:00 UTC (79,1%) para temperaturas de topo mais frios, no horário 15:00 UTC o resultado foi levemente inferior, com 71,5%. Assim como o classificador operacional, o atualizado houve classificações com topos mais quentes, em torno de 17,6%.

As nuvens stratus, são nuvens uniformes e espessas, ou seja, cobrem grandes extensões de área, entretanto, a depender do quão espessa seja esta nuvem, a temperatura do seu topo pode variar, desde temperaturas um pouco mais elevadas, quanto temperaturas mais baixas (HOUZE JUNIOR, 1993; TAPAKIS; CHARALAMBIDES, 2013). O classificador operacional, identificou este tipo de nuvem em sua maioria com temperaturas de topo entre 282 a 222 K, em torno de 91% nos horários 15:00 e 18:00 UTC. No horário 12:00 UTC foi aproximadamente 88%. O classificador atualizado, no geral teve uma melhora de aproximadamente 3% ao identificar este tipo de nuvens.

As nuvens cumulonimbus, aqui denominadas de multicamadas (mc), possui um desenvolvimento vertical bastante acentuado, e como visto no [Capítulo 1](#), a altura da sua base se encontra em baixos níveis da atmosfera e seu topo em altos níveis (HOUZE JUNIOR, 1993; WALLACE; HOBBS, 2006). O classificador operacional teve uma média na identificação deste alvo em aproximadamente 92% nos três horários, variando a temperatura do topo entre 252 e 207 K. O classificador atualizado teve uma média na identificação em aproximadamente 90% das nuvens multicamadas.

Tabela 4.4 - Distribuição média entre as classes de nuvens e o produto temperatura (K) do topo da nuvem da NOAA para os classificadores operacional (baseado no GOES-8) e o atualizado (baseado no GOES-16) para os intervalos de horário centrados em 12:00, 15:00 e 18:00 UTC.

	312	297	282	267	252	237	222	207	-1
12:00 UTC									
Classificador operacional									
su	0,0	7,3	11,3	1,2	0,4	1,3	4,9	10,5	63,1
cu	0,0	18,8	32,1	4,4	0,9	1,9	6,2	8,7	27,0
ci	0,0	1,9	8,3	13,0	5,3	13,9	34,8	20,4	2,4
st	0,0	4,5	49,5	15,3	5,0	5,5	12,8	4,9	2,5
mc	0,0	0,0	0,2	1,1	6,3	25,4	54,3	12,6	0,1
Classificador atualizado									
su	0,0	6,1	10,7	0,7	0,4	0,9	4,0	11,2	66,0
cu	0,0	17,6	28,3	5,4	1,2	2,2	7,3	8,9	29,1
ci	0,0	1,7	7,0	9,1	5,1	14,5	36,8	22,7	3,1
st	0,0	2,8	48,8	15,9	4,6	6,5	13,2	5,6	2,6
mc	0,0	0,0	0,0	3,7	7,1	25,8	49,4	13,5	0,5
15:00 UTC									
Classificador operacional									
su	1,2	6,6	2,5	0,1	0,1	0,6	3,2	7,1	78,6
cu	0,5	25,4	29,5	1,9	0,5	1,4	6,2	9,9	24,7
ci	0,8	6,9	8,3	5,8	3,2	11,2	34,5	20,4	8,9
st	0,0	2,9	56,3	13,9	5,3	6,3	10,5	4,4	0,4
mc	0,0	0,0	0,5	2,9	3,9	27,8	50,7	13,8	0,4
Classificador atualizado									
su	1,0	7,3	3,0	0,3	0,2	0,8	4,0	8,5	74,9
cu	0,6	23,8	31,7	2,4	0,5	1,7	6,9	9,4	23,0
ci	0,8	5,6	7,6	5,9	3,4	11,9	35,9	20,3	8,6
st	0,0	1,8	56,8	15,7	4,9	5,8	11,0	3,6	0,4
mc	0,0	0,0	0,7	3,1	5,0	26,9	49,9	14,0	0,4
18:00 UTC									
Classificador operacional									
su	2,5	8,5	5,1	0,2	0,2	0,7	3,7	7,1	72,0
cu	1,6	28,6	28,8	2,3	0,6	1,4	6,0	8,4	22,3
ci	1,4	7,6	7,9	3,9	2,5	10,9	34,9	25,3	5,6
st	0,0	2,8	52,3	15,6	5,9	6,5	10,8	4,2	1,9
mc	0,0	0,3	2,2	2,6	2,8	23,8	49,3	18,2	0,8
Classificador atualizado									
su	1,5	8,5	4,8	0,3	0,2	0,8	3,6	8,4	71,9
cu	1,8	25,8	28,9	2,6	0,6	1,6	6,3	9,7	22,7
ci	1,2	5,3	6,7	3,9	2,5	11,5	37,8	27,3	3,8
st	0,0	1,9	51,6	17,7	5,2	6,4	11,8	4,3	1,1
mc	0,0	0,3	2,5	2,3	3,7	23,3	48,4	18,6	0,9

Fonte: Produção da autora.

Nesta comparação entre o classificador operacional e o classificador atualizado com o produto oficial da temperatura do topo da nuvem, percebeu-se que a diferença entre os dois classificadores e o produto é muito pequena.

Algumas dessas diferenças foram de aproximadamente 3%, como é o caso da superfície no horário das 12:00 UTC. Contudo, a nuvem do tipo cumulus, foi a que apresentou maiores erros na identificação dentre todos os tipos de superfícies analisados, e as nuvens do tipo stratus e multicamadas foram as melhores categorizadas.

4.3.2 Fase da nuvem

Foram utilizados os mesmos 20 dias para a comparação com o produto temperatura do topo da nuvem, onde foi calculado a média dos 20 dias para todas as fases da nuvem: céu claro, água líquida, água líquida super-resfriada, fase mista e gelo (Tabela 4.5).

O classificador operacional categorizou a superfície como céu claro em 55,2% e 58,2% nos horários 12:00 e 18:00 UTC, respectivamente e 67,3% em 15:00 UTC. O classificador atualizado categorizou a superfície como céu claro em 61,4%, 65,5% e 59%, nos três horários respectivamente.

Os dois classificadores, em média categorizaram a superfície corretamente em aproximadamente 60%, para todo o período avaliado, entretanto, o classificador atualizado, identificou 2% mais que o classificador operacional, em média nos três horários. Contudo houve identificações errôneas com porcentagens altas nos dois classificadores, próximo a 45% no operacional e próximo a 41% no atualizado.

As nuvens do tipo cumulus, nos três horários foram categorizadas em sua maioria como sendo nuvens de água líquida, nos dois classificadores, havendo então, uma correspondência, já que as nuvens cumulus geralmente são nuvens quentes, ou seja, nuvens que estão abaixo da isoterma de 0° (WALLACE; HOBBS, 2006; COTTON et al., 2011). Nos dois classificadores houveram identificações errôneas em outras fases, como por exemplo, nuvens cumulus formadas por gelo, foram similares, cerca de 17%.

As nuvens cirrus por outro lado, foram identificadas majoritariamente como nuvens de gelo nos dois classificadores. Esta informação está coerente, já que essas nuvens

são classificadas como nuvens altas, como descrito na [Tabela 1.1](#). Contudo, houve categorização errônea deste alvo relativamente alta nos dois classificadores, chegando a 23% no classificador operacional e 20% no classificador atualizado na categoria água líquida, o que é improvável já que a base destas nuvens estão acima de 6 km da superfície.

As nuvens multicamadas, foram em sua maioria categorizadas como nuvens de gelo nos dois classificadores, cerca de 93% nos três horários. Também houveram classificações em fase mista, água líquida e água líquida super-resfriada, nos dando indício que este tipo de nuvem realmente seja a mc, já que se trata de um tipo de nuvem que possui um desenvolvimento vertical bastante intenso.

Na comparação do classificador operacional (baseado no GOES-8) e a versão atualizada (baseada no GOES-16) com o produto oficial fase da nuvem, verificou-se que a diferença entre os dois classificadores é bastante similar em média para as três variantes de horários para alguns tipos de nuvens, como por exemplo nuvens multicamadas e nuvens stratus.

Contudo, houveram diferenças onde a versão do classificador atualizado teve uma pequena melhora em relação ao classificador operacional, como por exemplo as nuvem do tipo cirrus que foi de aproximadamente 4% e a superfície foi cerca de 2%.

Tabela 4.5 - Distribuição média entre as classes de nuvens e o produto fase termodinâmica da nuvem da NOAA para os classificadores operacional (baseado no GOES-8) e o atualizado (baseado no GOES-16) para os intervalos de horário centrados em 12:00, 15:00 e 18:00 UTC.

	Céu Claro	Água Líquida	Água Líquida Super-resfriada	Fase Mista	Gelo
12:00 UTC					
Classificador operacional					
su	55,2	25,7	0,4	0,3	18,4
cu	8,8	69,7	1,5	1,3	18,7
ci	1,1	13,9	3,8	4,5	76,7
st	1,6	57,5	8,3	4,9	27,7
mc	0,0	0,4	1,0	4,8	93,8
Classificador atualizado					
su	61,4	22,3	0,1	0,1	17,1
cu	11,1	64,5	1,8	1,7	20,9
ci	1,9	11,7	2,6	3,9	79,9
st	2,1	55,8	8,6	6,0	27,5
mc	0,0	0,0	2,2	5,1	92,7
15:00 UTC					
Classificador operacional					
su	67,3	20,0	0,0	0,0	12,7
cu	4,0	73,5	0,4	0,2	21,9
ci	4,0	23,5	0,8	1,6	70,1
st	0,2	61,3	6,2	5,9	26,4
mc	0,0	1,4	1,3	3,4	93,9
Classificador atualizado					
su	65,5	20,4	0,0	0,0	14,1
cu	4,2	73,8	0,5	0,4	21,1
ci	3,6	20,5	1,0	2,1	72,8
st	0,3	61,0	7,1	5,5	26,1
mc	0,0	1,2	1,5	3,8	93,5
18:00 UTC					
Classificador operacional					
su	58,5	26,9	0,0	0,0	14,6
cu	2,8	76,4	0,7	0,7	19,4
ci	2,1	22,3	0,4	1,1	74,1
st	0,9	58,5	6,9	6,9	26,8
mc	0,0	3,2	0,9	2,7	93,2
Classificador atualizado					
su	59,0	26,5	0,0	0,0	14,5
cu	3,0	74,7	0,6	0,5	21,2
ci	1,5	18,6	0,4	1,2	78,3
st	0,6	60,1	7,9	6,5	24,9
mc	0,0	3,5	1,0	2,5	93,0

Fonte: Produção da autora.

4.3.3 Máscara de nuvem

O produto de máscara de nuvem foi utilizado para verificar e reforçar se os tipos de alvos analisados estão coerentes e identificar possíveis situações as quais envolvam nuvens semi-transparentes que possivelmente foram classificadas como superfície.

Assim como nos dois outros produtos, foi utilizado o mesmo conjunto de 20 dias, onde as tabelas geradas também foram resultado da média dos 20 dias. O produto de máscara de nuvem tem duas classificações, céu limpo e céu com nuvens (Tabela 4.6).

Os dois classificadores em comparação ao produto máscara de nuvem em média são bastantes similares nos três horários. A superfície e a nuvem cumulus foram erroneamente identificadas nos dois classificadores, cerca de 30% e 25% em média nos três horários, respectivamente. Enquanto as outras nuvens foram categorizadas corretamente mais de 90%. Destas três categorias de nuvens, cirrus foi, em média, classificada como céu claro em aproximadamente 5% nos dois classificadores, esse erro pode indicar cirrus semi-transparentes categorizada como superfície (céu claro).

Tabela 4.6 - Distribuição média entre as classes de nuvens e o produto máscara da nuvem da NOAA para os classificadores operacional (baseado no GOES-8) e o atualizado (baseado no GOES-16) para os intervalos de horário centrados em 12:00, 15:00 e 18:00 UTC.

	Céu Limpo	Céu com Nuvens
12:00 UTC		
Classificador operacional		
su	62,5	37,5
cu	27,1	72,9
ci	2,7	97,3
st	2,1	97,9
mc	0,0	100,0
Classificador atualizado		
su	65,0	35,0
cu	28,0	72,0
ci	2,6	97,4
st	3,0	97,0
mc	0,0	100,0
15:00 UTC		
Classificador operacional		
su	77,8	22,2
cu	24,8	75,2
ci	8,9	91,1
st	0,5	99,5
mc	0,0	100,0
Classificador atualizado		
su	74,9	25,1
cu	24,7	75,3
ci	8,2	91,8
st	0,4	99,6
mc	0,0	100,0
18:00 UTC		
Classificador operacional		
su	70,6	29,4
cu	22,1	77,9
ci	6,4	93,6
st	1,5	98,5
mc	0,0	100,0
Classificador atualizado		
su	70,8	29,2
cu	22,3	77,7
ci	4,4	95,6
st	1,1	98,9
mc	0,0	100,0

Fonte: Produção da autora.

4.4 Criação de um novo classificador

Nesta seção serão vistos e discutidos todos os passos realizados para a definição do novo classificador baseado no *Advanced Baseline Imager* a bordo do satélite GOES-16. Esse classificador utiliza outros parâmetros além dos quatro discutidos anteriormente e utilizados pelo classificador operacional e a sua versão atualizada desenvolvida nesse estudo.

4.4.1 Definição das sementes iniciais

O primeiro passo foi escolher um conjunto de sementes para iniciar o procedimento de agrupamento pelo método *k-means*. Optou-se por escolher estas amostras manualmente pois assim foi possível evitar que o método convergisse para vários centroides idênticos. A partir da análise do estudo de caso descrito na Seção 4.1, foi escolhida a imagem das 15:00 UTC do dia 21/03/2019, para que fossem então definidos pontos de latitude e longitude, marcados em rosa (Figura 4.20), de onde foram extraídos os parâmetros para serem utilizados como sementes iniciais na geração do novo classificador.

Para a escolha destes pontos foram levados em consideração os vários tipos de alvos. Os alvos selecionados foram: cumulonimbus (cb), cumulus (cu), stratus (st), cirrus (ci), cirrus semi-transparentes (ci-st), multicamadas (mc), além de superfícies oceânicas e continentais (su), desde semiárido, floresta, rio, cordilheira dos Andes, dentre outros. No total, foram escolhidos 60 pontos mas posteriormente reduzidos para 41 pontos baseado na similaridade dos centroides obtidos ao final do processo de agrupamento. Para tornar o texto mais claro apenas os pontos utilizados serão apresentados.

Os diferentes alvos foram escolhidos utilizando uma imagem na região do visível e outra do infravermelho janela em $11 \mu m$ do dia escolhido para tal, criando uma amostra semelhante a lista de tipos de alvos utilizados em (BOTTINO; CEBALLOS, 2015).

As nuvens do tipo cb foram escolhidas devido ao seu grande desenvolvimento vertical, chegando até o topo da troposfera. Pelo seu grande desenvolvimento vertical, estas nuvens possuem a temperatura do topo da nuvem muito fria e alta refletância (WALLACE; HOBBS, 2006; COTTON et al., 2011).

Esse tipo de nuvem, em particular, apresentou algumas falhas na classificação, como pôde ser visto na [Seção 4.1](#). Nuvens do tipo mc podem se assemelhar com nuvens st e cb em valores de reflectância e temperatura de brilho do topo. Por definição a nuvem mc, é classificada de forma, quando há presença de nuvens em duas ou nas três camadas atmosféricas simultaneamente. Esse tipo de nuvem é mais comum de ser identificada quando associada à sistemas meteorológicos, assim como é mais fácil identificar nuvens mc próximos a ocorrência de nuvens cb ([WALLACE; HOBBS, 2006](#); [BOTTINO; CEBALLOS, 2015](#)). Devido ao grande desenvolvimento vertical, uma nuvem do tipo cb estará presente desde os baixos níveis até níveis altos. Por este motivo as nuvens mc e cb serão consideradas como sendo somente mc.

As nuvens cu foram escolhidas tendo sua ocorrência tanto sobre o continente, como sobre o oceano. Este último apresentou algumas inconsistências na classificação na região do Pacífico, próximo à costa do Chile. Foram escolhidos nuvens cumulus com diferentes estágios de desenvolvimento, os quais apresentam diferentes reflectâncias, bem como temperatura de brilho, tornando possível classificá-los em diferentes estágios.

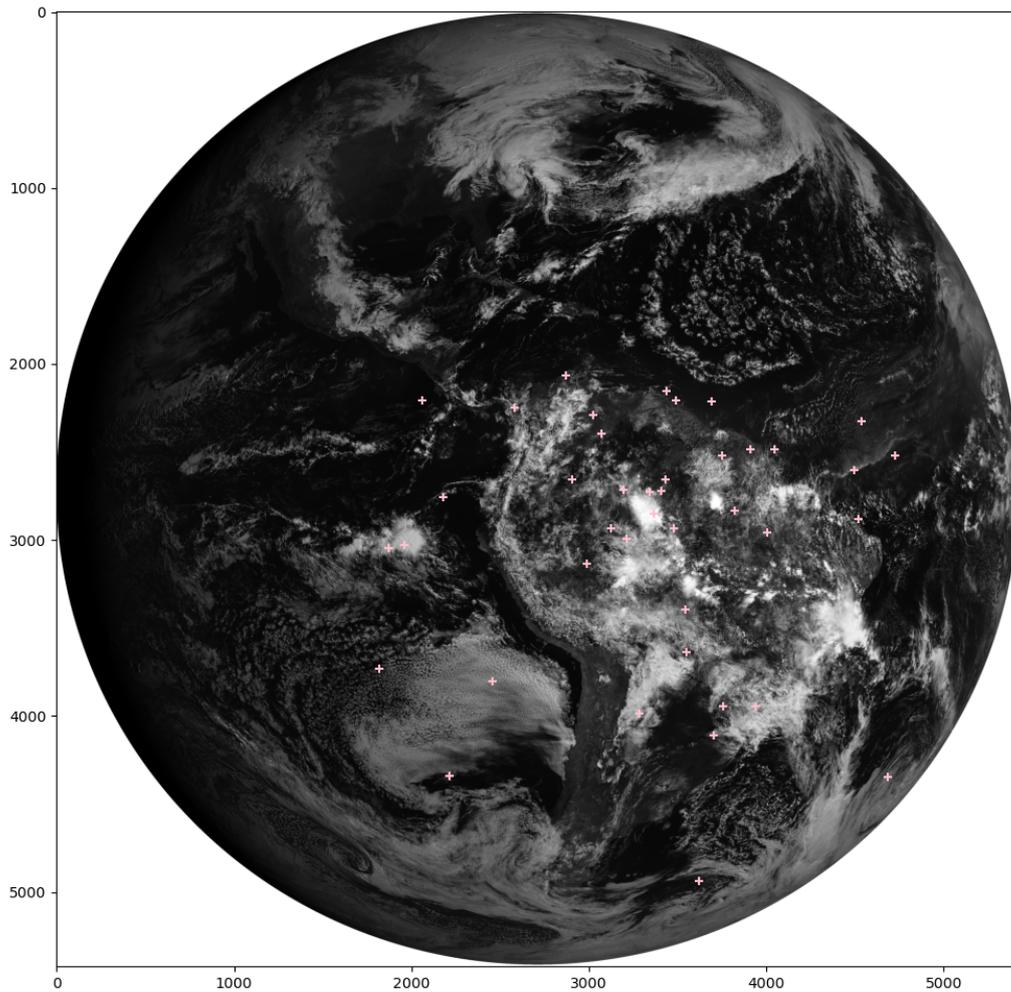
Foram selecionadas nuvens st em diferentes estágios. Esse tipo de nuvem pode variar desde nuvens menos densa até mais densa, implicando em várias combinações entre valores de reflectância e temperatura de brilho.

As nuvens cirrus foram subdividas em 2 classes, as nuvens cirrus mais espessas, denominadas apenas de cirrus (ci) e as mais finas, denominadas de cirrus semi-transparentes (ci-st). Essa subdivisão visa separar essas duas classes, já que do ponto de vista da reflectância as ci (mais densas), por consequência possuem valores de reflectância mais elevados que as ci-st. No estudo de caso, houveram algumas classificações errôneas quanto às nuvens ci, que acabaram sendo categorizadas como sendo nuvens st. As duas denominações de ci e ci-st foram utilizadas para uma melhor escolha dos pontos, portanto, daqui em diante serão chamadas apenas de ci.

Além dos tipos de nuvens também foram selecionados alguns pontos de céu claro, ou seja, sem a presença de nuvens. Estes pontos possuem características diferentes, como por exemplo, superfícies oceânicas, como Oceano Atlântico mais ao sul e mais ao norte, o mesmo se aplica ao Oceano Pacífico. Diferentes superfícies continentais também foram escolhidas levando em consideração diferentes tipos de rugosidade e cobertura do solo, bem como regiões de lagos e rios. Todos compuseram a classe su.

A Figura 4.20 apresenta todos os pontos escolhidos como sementes iniciais indicados pelas cruzes em rosa.

Figura 4.20 - Pontos selecionados para serem os novos centroides iniciais a partir das imagens do dia 21/03/2019 15:00 UTC.



Fonte: Produção da autora.

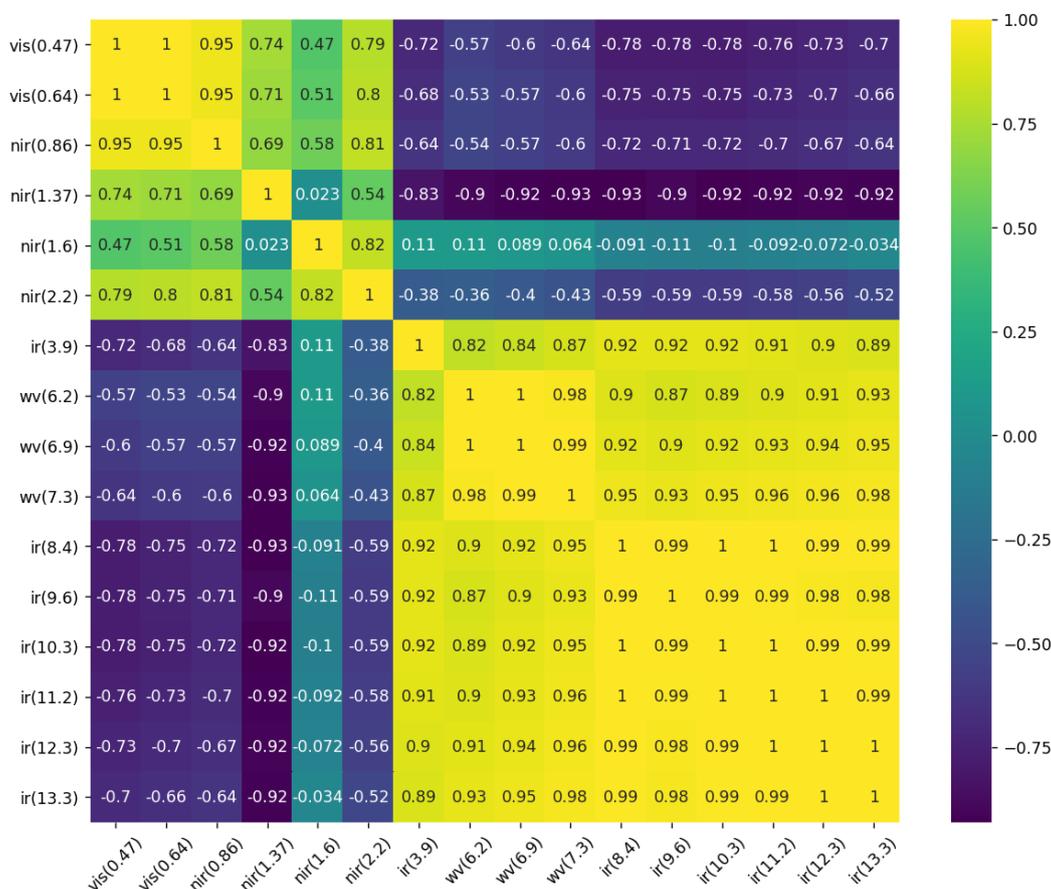
4.4.2 Seleção dos parâmetros mais representativos

Como a princípio a quantidade de variáveis selecionadas para utilizar no novo classificador totalizaram 40 variáveis, somando todos os 16 canais do ABI/GOES-16,

suas correspondentes texturas, além de 8 diferenças entre temperaturas de brilho (Tabela 4.7), foi necessário verificar a existência de variáveis redundantes. Para isso, a análise de componentes principais, a qual visa a redução na dimensão das variáveis sem que haja perda significativa na variação total dos dados, foi aplicada a esse conjunto de 40 variáveis.

A análise de componentes principais foi realizada em 3 etapas distintas para que assim se chegasse ao conjunto de variáveis finais. Na primeira etapa foram selecionados apenas os 16 canais do ABI/GOES-16 e ao aplicar a técnica PCA encontrou-se que 5 das 16 variáveis explicam mais de 98% da variância total dos dados. As 5 variáveis foram os canais de número 2 ($0,64 \mu m$), 4 ($1,37 \mu m$), 8 ($6,2 \mu m$), 11 ($8,4 \mu m$) e 13 ($10,3 \mu m$). A figura Figura 4.21 apresenta a matriz de correlação que embasou essa conclusão.

Figura 4.21 - Matriz de correlação dos canais ABI/GOES-16.



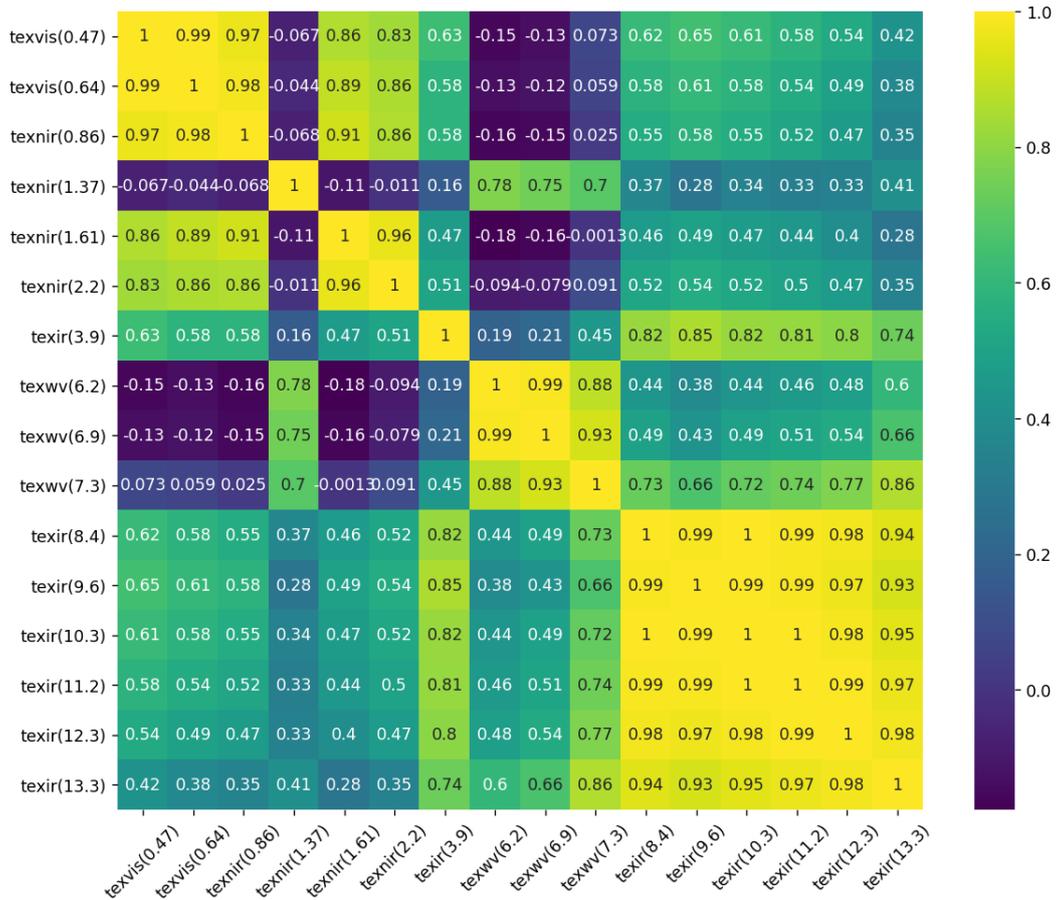
Fonte: Produção da autora.

Na segunda etapa para a aplicação da técnica PCA foram selecionados as texturas respectivas a cada um dos 16 canais que compõem o ABI/GOES-16 e encontrou-se que 3 das 16 texturas explicam mais 98% da variância total dos dados foram: a textura na região do visível em $0,64 \mu m$, na região do vapor d'água em $6,2 \mu m$ e na região do infravermelho em $10,3 \mu m$, como pode ser visto na Figura 4.22.

Esse resultado é coerente pois as variações espaciais nos canais nas bandas de absorção são geralmente sempre baixas exceto para topos de nuvens localizados em camadas acima do pico da função peso desses canais, já que a radiação emitida por esses topos não é mais atenuado ao ponto de não chegar ao topo da atmosfera. Para os canais nas janelas atmosféricas na porção do infravermelho é esperado que a

variação espacial sejam similares, variando apenas na intensidade do sinal, ou seja, temperaturas de brilho maiores ou menores. O mesmo pode ser esperado para as imagens dos canais na porção do visível.

Figura 4.22 - Matriz de correlação das texturas dos respectivos canais ABI/GOES-16.



Fonte: Produção da autora.

Na terceira e última etapa, a técnica PCA foi aplicada para as 8 diferenças entre as temperaturas de brilho. A Figura 4.23 apresenta a matriz de correlação, com os canais e as texturas finais encontradas nas etapas anteriores inseridos, além das 8 combinações entre as temperaturas de brilho.

Na Figura 4.23 as diferenças entre as temperaturas de brilho (dTB) utilizadas estão

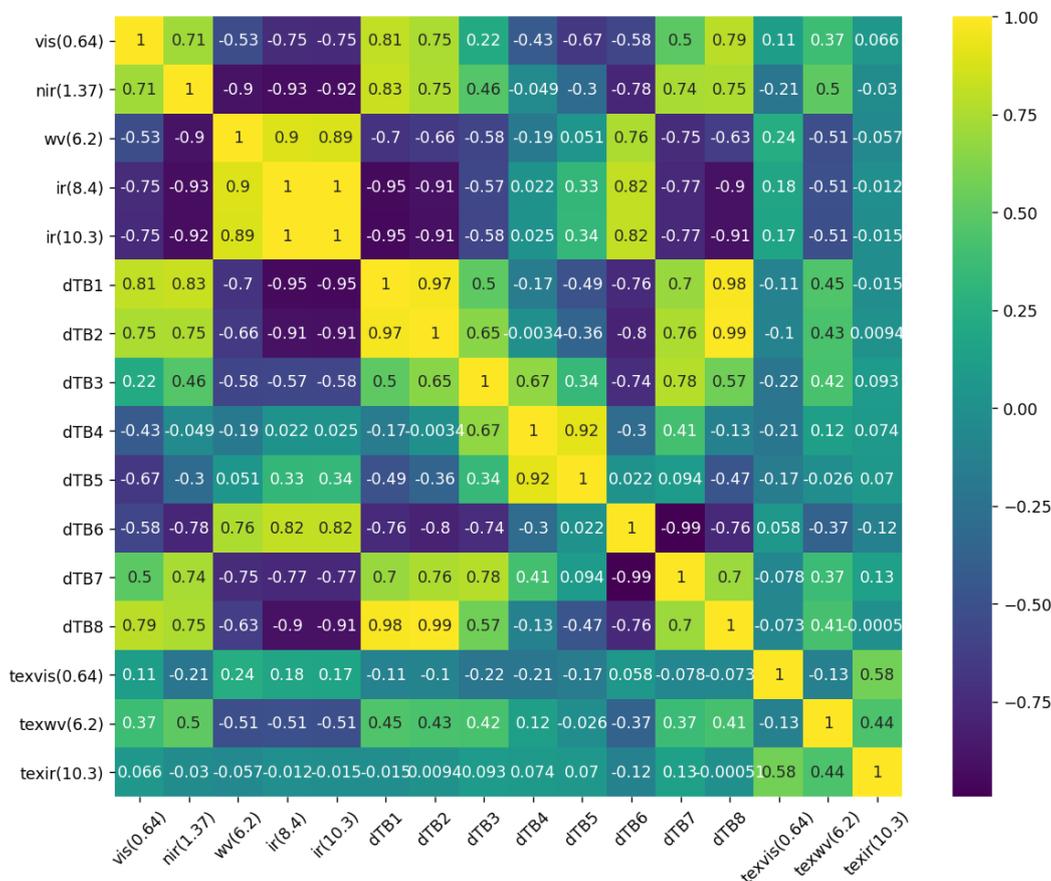
descritas como dTBx, onde na [Tabela 4.7](#) pode-se visualizar as diferenças utilizadas e seus respectivos índices. As 4 diferenças entre as temperaturas de brilho finais que explicam mais de 98% da variância total dos dados foram: dTB1, dTB5, dTB6 e dTB7.

Tabela 4.7 - Diferença entre as temperaturas de brilho utilizadas na aplicação da técnica PCA.

Diferenças entre as temperaturas de brilho (dTB)	
dTB1	6,2 μm - 10,3 μm
dTB2	7,3 μm - 12,3 μm
dTB3	8,4 μm - 11,2 μm
dTB4	10,3 μm - 12,2 μm
dTB5	11,2 μm - 12,3 μm
dTB6	11,2 μm - 3,9 μm
dTB7	3,9 μm - 12,3 μm
dTB8	7,3 μm - 11,2 μm

Fonte: Produção da autora.

Figura 4.23 - Matriz de correlação dos canais, texturas e as diferenças entre as temperaturas de brilho.



Fonte: Produção da autora.

Portanto, a partir da técnica de PCA obtivemos um total de 12 variáveis para o uso na geração do novo classificador, sendo elas: canal 2 ($0,64 \mu m$), canal 4 ($1,37 \mu m$), canal 8 ($6,2 \mu m$), canal 11 ($8,4 \mu m$), canal 13 ($10,3 \mu m$), textura do canal 2, textura do canal 8, textura do canal 13, dTB1, dTB5, dTB6 e dTB7.

A aplicação da técnica PCA foi separada em três etapas foi realizada desta forma simplesmente para se ter uma melhor visualização das matrizes de correlação mostradas nas figuras anteriores, já que ao executar a técnica nas 40 variáveis não foi possível se ter uma boa visualização, mas o resultado final foi igual ao obtido quando empregada em três etapas.

Nas matrizes de correlação, temos cada par de variáveis variando de +1 a -1. Isso indica o quão forte ou fraco é a correlação entre os pares: i) a correlação positiva indica que ambas as variáveis mudam na mesma direção; ii) a correlação negativa indica que as variáveis mudam em direções opostas e; iii) a correlação nula, não há nenhuma mudança entre as variáveis (FIGUEIREDO FILHO; SILVA JÚNIOR, 2009; WILKS, 2011).

Quando o objetivo é a redução das variáveis e há uma correlação positiva entre as analisadas, se faz necessário escolher apenas uma delas, já quando há uma correlação negativa, pode-se escolher as duas variáveis analisadas pois ambas são significativas. Entretanto, há ocasiões onde mesmo se duas variáveis possuem uma forte correlação negativa, apenas se escolhe uma delas, quando uma destas variáveis também possui uma forte correlação positiva com alguma outra variável.

4.4.3 Treinando o classificador a partir do método K-Means

Após eliminar os parâmetros redundantes ao aplicar a técnica PCA, foram utilizados os mesmos 120 dias de 2019 usados na atualização do classificador baseado em quatro atributos (vir , IR , tex_{vis} , tex_{IR}) como conjunto de treinamento.

Para a execução do método de agrupamento *k-means*, foram selecionados, de maneira aleatória, 50.000 (cinquenta mil) amostras de cada uma das imagens. Cada amostra consiste em um vetor formado pelos 12 atributos selecionados a partir da análise PCA. O parâmetro textura de cada ponto foi calculado a partir de uma matriz de 3 x 3 pixels centrada no pixel em questão.

O critério de semelhança entre as amostras e os centroides (sementes) utilizado na execução do *k-means* foi a distância Euclidiana, a qual é utilizada para definir a convergência dos pontos em k centroides, foi 0,05.

Os 41 pontos selecionados a partir da imagem do visível (Figura 4.20) foram utilizados como sementes iniciais. Após a obtenção do novo conjunto de centroides (Tabela 4.9), verificou-se a necessidade de um deles ser descartado por haver valores discrepantes em alguns parâmetros, totalizando assim, 40 centroides finais. A Tabela 4.9 apresenta os centroides obtidos para o horário 15:00 UTC.

Pela dependência de canais na região do visível esse classificador deve ser dividido em

variantes referentes à períodos específicos do dia. Portanto, tal qual os classificadores baseados em 4 parâmetros, o novo classificador que utiliza 12 parâmetros possui três variantes, 12:00, 15:00 e 18:00 UTC e que devem ser utilizados para classificar imagens nos intervalos de tempo 11:00 a 13:59, 14:00 a 16:59 e 17:00 a 19:59 UTC, respectivamente.

Tabela 4.8 - Tabela resultante do treinamento das novas classes e parâmetros para 12:00 UTC.

Classes	vis0,64	nir1,37	wv6,2	ir8,4	ir10,3	texvis0,64	texwv6,2	texir10,3	dTB1	dTB5	dTB6	dTB7
1	86.6	41.4	211.0	212.6	213.0	3.0	0.9	1.2	-2.0	0.8	-44.4	-45.3
2	50.0	0.1	250.5	278.3	279.6	3.7	0.1	0.2	-29.1	0.5	-3.5	-4.0
3	7.9	0.1	248.7	290.9	293.6	1.5	0.1	0.4	-44.8	2.6	-2.8	-5.4
4	38.9	9.2	226.4	252.2	252.8	4.2	1.0	4.2	-26.3	3.9	-35.0	-39.0
5	80.5	28.1	217.2	220.9	221.3	3.7	1.0	1.6	-4.1	0.9	-33.0	-33.8
6	71.5	11.1	226.2	235.9	236.2	5.2	0.9	2.1	-10.0	1.0	-17.5	-18.5
7	9.3	0.1	247.8	286.6	288.7	1.7	0.1	0.6	-40.9	1.7	-1.0	-2.7
8	56.8	18.3	218.3	229.3	229.3	3.9	1.2	3.3	-11.0	2.0	-39.8	-41.8
9	53.0	10.9	224.7	241.7	241.8	4.5	1.0	3.6	-17.0	2.2	-28.1	-30.2
10	63.1	0.3	241.3	277.1	279.0	6.5	0.1	0.7	-37.7	1.0	-6.9	-7.9
11	7.7	0.1	244.2	291.4	294.6	1.3	0.1	0.4	-50.3	3.4	-5.1	-8.5
12	62.3	3.8	234.2	255.7	257.3	5.2	0.3	1.9	-23.1	1.6	-28.2	-29.8
13	40.1	0.1	242.9	280.5	282.4	4.7	0.1	0.4	-39.5	1.1	-1.9	-3.0
14	24.2	3.3	233.0	271.8	273.8	3.9	0.5	3.4	-40.8	5.4	-25.3	-30.7
15	33.7	6.3	228.7	260.3	260.8	4.5	1.0	4.2	-32.1	3.6	-24.2	-27.9
16	16.9	0.3	245.0	267.8	268.8	3.2	0.2	1.3	-23.8	0.4	-3.8	-4.2
17	17.0	0.4	236.8	285.7	288.6	4.6	0.2	1.2	-51.9	3.1	-7.6	-10.7
18	45.8	0.2	238.2	282.6	285.2	7.9	0.2	1.0	-47.0	2.2	-7.4	-9.6
19	16.5	0.3	239.5	288.2	291.4	2.8	0.2	0.9	-51.9	3.5	-12.2	-15.7

Classes	vis0,64	nir1,37	wv6,2	ir8,4	ir10,3	texvis0,64	texwv6,2	texir10,3	dTB1	dTB5	dTB6	dTB7
20	64.5	4.3	232.7	249.6	250.1	7.0	0.6	2.1	-17.3	1.0	-10.5	-11.5
21	74.2	16.4	218.2	223.2	223.4	4.6	0.9	1.6	-5.2	0.9	-26.5	-27.5
22	48.9	5.3	229.8	252.5	252.9	5.1	0.7	3.2	-23.1	2.0	-18.9	-20.9
23	10.8	0.1	240.9	279.3	280.9	1.5	0.1	0.7	-39.9	0.9	-0.7	-1.6
24	7.5	0.1	239.4	287.8	290.6	1.3	0.1	0.5	-51.2	2.5	-2.0	-4.5
25	7.2	0.2	237.2	289.8	293.2	1.2	0.1	0.4	-56.0	3.5	-4.9	-8.4
26	8.9	0.4	237.5	289.4	292.9	1.6	0.2	0.8	-55.4	4.2	-8.7	-12.9
27	14.4	0.1	250.7	280.6	281.9	1.6	0.1	0.5	-31.2	0.6	0.0	-0.6
28	39.4	1.4	236.5	266.7	268.8	5.8	0.2	2.4	-32.3	2.4	-23.0	-25.4
29	58.4	1.3	235.6	273.8	275.8	7.1	0.4	2.0	-40.2	2.4	-14.8	-17.2
30	15.0	1.4	235.2	282.6	285.2	2.8	0.4	2.2	-50.0	4.6	-15.1	-19.8
31	29.4	0.2	247.1	280.6	282.6	6.3	0.1	1.2	-35.5	1.4	-7.9	-9.4
32	43.9	0.7	237.7	266.2	267.7	6.4	0.2	1.4	-30.0	1.2	-8.5	-9.6
33	28.6	3.0	221.3	228.2	228.5	4.0	1.1	2.2	-7.2	1.0	-14.0	-15.0
34	32.5	0.3	239.5	279.7	282.3	6.5	0.2	1.5	-42.8	2.1	-13.8	-15.9
35	16.0	0.2	245.7	287.3	290.0	4.5	0.1	1.2	-44.4	2.8	-8.5	-11.3
36	25.0	2.7	233.6	272.7	274.4	4.6	0.6	3.3	-40.8	3.8	-18.7	-22.6
37	19.1	0.3	237.2	279.2	281.5	5.1	0.2	1.5	-44.3	2.0	-5.3	-7.3
38	20.4	1.4	235.3	277.8	279.9	4.6	0.4	2.7	-44.6	3.4	-12.4	-15.8
39	62.3	1.4	237.7	263.8	265.3	6.3	0.3	1.4	-27.5	1.2	-15.9	-17.1
40	29.8	1.5	234.9	266.2	267.6	5.1	0.4	2.7	-32.7	2.2	-13.9	-16.0

Fonte: Produção da autora.

Tabela 4.9 - Tabela resultante do treinamento das novas classes e parâmetros para 15:00 UTC.

Classes	vis0,64	nir1,37	wv6,2	ir8,4	ir10,3	texvis0,64	texwv6,2	texir10,3	dTB1	dTB5	dTB6	dTB7
1	88.2	49.7	208.7	210.0	210.5	2.9	1.1	1.4	-1.7	0.9	-42.0	-42.9
2	56.7	0.9	248.9	277.7	279.2	5.1	0.1	0.4	-30.3	0.6	-19.2	-19.8
3	73.5	0.5	240.2	277.8	279.8	7.0	0.2	0.9	-39.5	1.2	-14.1	-15.4
4	9.4	0.4	237.3	289.9	293.2	3.3	0.2	1.2	-55.9	4.3	-10.2	-14.5
5	78.5	34.7	217.4	220.8	221.1	3.0	1.2	1.8	-3.8	0.9	-30.4	-31.3
6	63.2	8.1	233.9	252.7	254.2	4.9	0.4	2.2	-20.3	1.6	-37.5	-39.1
7	46.9	13.1	225.2	241.7	241.8	3.9	1.2	4.0	-16.6	2.0	-27.2	-29.2
8	31.9	11.0	226.4	254.0	254.4	3.9	1.3	4.9	-28.0	4.4	-35.6	-39.9
9	52.9	21.7	218.9	229.2	229.2	3.4	1.3	3.5	-10.3	2.0	-38.2	-40.2
10	41.6	3.0	236.4	266.1	268.2	5.7	0.3	2.9	-31.8	2.8	-31.7	-34.5
11	6.4	0.1	239.2	287.5	290.4	1.3	0.1	0.4	-51.2	2.6	-2.2	-4.8
12	71.9	3.2	237.1	262.7	264.0	5.8	0.4	1.7	-26.9	1.2	-25.3	-26.6
13	6.5	0.1	249.3	288.3	290.4	1.9	0.1	0.5	-41.1	1.9	-2.2	-4.1
14	63.5	10.9	230.0	244.1	244.5	4.1	0.8	2.6	-14.5	1.1	-15.7	-16.7
15	10.9	0.1	244.8	300.4	304.6	1.6	0.1	1.1	-59.8	4.5	-8.2	-12.8
16	18.9	1.4	246.3	302.4	305.1	1.7	0.2	1.5	-58.8	1.5	-7.5	-8.9
17	21.2	4.5	232.9	273.4	275.2	3.9	0.8	3.9	-42.3	5.4	-26.8	-32.1
18	15.3	1.0	238.1	290.6	294.0	4.1	0.3	2.4	-55.9	5.5	-16.8	-22.3
19	34.1	0.5	247.8	280.3	282.3	5.7	0.1	0.9	-34.5	1.2	-18.7	-19.9

Classes	vis0,64	nir1,37	wv6,2	ir8,4	ir10,3	texvis0,64	texwv6,2	texir10,3	dTB1	dTB5	dTB6	dTB7
20	17.0	0.6	235.2	281.2	283.9	5.6	0.3	1.8	-48.7	2.6	-7.9	-10.5
21	70.1	21.9	223.9	230.9	231.3	3.2	1.1	2.2	-7.4	1.0	-22.4	-23.4
22	42.8	1.5	238.0	269.3	271.2	6.5	0.3	2.1	-33.2	1.7	-24.2	-25.9
23	15.7	1.8	235.3	282.1	284.4	4.4	0.5	2.9	-49.1	4.4	-16.8	-21.2
24	6.3	0.1	247.5	292.2	295.2	1.5	0.1	0.4	-47.7	3.0	-4.0	-7.0
25	12.8	0.2	241.8	301.7	306.6	2.0	0.1	1.4	-64.7	6.1	-12.8	-18.9
26	6.4	0.2	237.9	290.7	294.2	1.4	0.1	0.5	-56.3	3.6	-5.2	-8.8
27	64.6	0.6	249.7	279.3	280.6	5.6	0.1	0.3	-31.0	0.6	-11.3	-11.9
28	35.3	0.6	249.9	281.5	283.1	7.3	0.1	0.9	-33.2	0.9	-10.9	-11.8
29	23.5	0.4	237.7	286.7	289.7	8.0	0.2	1.8	-52.0	3.3	-12.9	-16.2
30	31.9	7.6	228.7	257.4	257.8	4.5	1.1	4.8	-29.1	3.0	-24.2	-27.2
31	50.8	0.3	240.7	279.2	281.2	7.8	0.1	0.9	-40.5	1.2	-9.1	-10.3
32	9.2	0.2	243.8	291.4	294.5	3.3	0.1	0.8	-50.7	3.6	-7.7	-11.3
33	29.6	1.3	238.0	277.2	279.7	6.4	0.3	3.0	-41.7	3.5	-23.1	-26.5
34	48.1	0.3	239.2	283.5	286.2	11.0	0.1	1.4	-47.0	2.3	-12.5	-14.8
35	20.2	4.1	231.3	270.1	271.2	4.3	0.9	4.5	-39.9	3.5	-18.4	-21.9
36	57.2	0.9	237.2	276.9	279.1	8.3	0.3	1.8	-42.0	2.1	-19.4	-21.5
37	19.8	0.4	246.1	285.5	287.7	6.9	0.1	1.4	-41.6	1.8	-8.8	-10.6
38	29.1	0.6	238.9	279.5	281.9	7.1	0.2	1.9	-43.0	2.1	-16.2	-18.3
39	53.3	2.9	234.7	262.0	263.1	6.1	0.5	2.7	-28.3	1.3	-14.8	-16.1
40	19.6	0.4	246.1	288.5	291.4	6.7	0.2	1.9	-45.2	3.4	-13.1	-16.5

Fonte: Produção da autora.

Tabela 4.10 - Tabela resultante do treinamento das novas classes e parâmetros para 18:00 UTC.

Classes	vis0,64	nir1,37	wv6,2	ir8,4	ir10,3	texvis0,64	texwv6,2	texir10,3	dTB1	dTB5	dTB6	dTB7
1	56.2	0.6	249.7	280.0	281.4	6.5	0.1	0.4	-31.6	0.6	-14.3	-14.9
2	70.1	21.7	223.0	229.5	229.8	4.2	1.3	2.5	-6.8	1.0	-22.8	-23.8
3	7.0	0.1	249.5	288.7	290.9	1.7	0.1	0.5	-41.4	1.9	-2.6	-4.5
4	25.7	6.2	231.8	270.3	271.8	4.3	1.0	4.6	-39.9	5.7	-31.4	-37.1
5	87.1	49.7	205.8	207.2	207.6	3.8	1.2	1.5	-1.8	1.0	-45.8	-46.9
6	66.7	8.0	234.8	253.2	254.7	6.2	0.5	2.6	-19.9	1.5	-37.7	-39.2
7	71.0	0.3	239.4	279.6	282.0	10.5	0.2	1.4	-42.6	1.7	-13.0	-14.7
8	36.4	13.2	223.9	247.6	247.4	4.6	1.7	5.6	-23.5	3.9	-36.6	-40.5
9	57.4	23.2	216.6	224.6	224.5	4.4	1.7	3.8	-7.9	1.8	-37.0	-38.8
10	10.6	0.1	238.9	289.5	292.6	1.5	0.1	0.4	-53.7	2.9	-2.9	-5.8
11	11.1	0.2	247.1	295.0	298.4	4.1	0.1	1.3	-51.3	4.0	-8.2	-12.2
12	47.8	12.5	225.1	241.1	241.1	4.7	1.3	4.5	-16.0	2.1	-26.6	-28.7
13	11.0	0.5	238.1	290.9	294.2	4.6	0.2	1.7	-56.1	4.5	-11.2	-15.7
14	35.0	0.8	250.1	279.8	281.5	4.7	0.1	0.6	-31.3	0.8	-21.7	-22.6
15	49.8	0.2	241.6	278.7	280.7	7.8	0.1	1.0	-39.1	1.2	-7.0	-8.2
16	62.6	2.9	236.1	261.4	262.5	7.9	0.6	3.0	-26.4	1.1	-13.6	-14.8
17	31.8	2.0	238.4	275.2	277.5	7.1	0.4	3.7	-39.1	3.5	-26.9	-30.4
18	16.7	0.6	240.6	297.6	301.8	4.5	0.2	2.4	-61.2	6.1	-15.3	-21.3
19	20.3	1.4	246.2	304.8	307.6	1.9	0.3	1.5	-61.4	1.6	-7.3	-8.9

Classes	vis0,64	nir1,37	wv6,2	ir8,4	ir10,3	texvis0,64	texwv6,2	texir10,3	dTB1	dTB5	dTB6	dTB7
20	28.6	7.7	228.4	260.5	260.8	4.8	1.4	5.6	-32.3	3.8	-26.3	-30.1
21	80.3	37.6	215.9	219.1	219.6	4.6	1.5	2.2	-3.7	1.0	-31.7	-32.7
22	17.9	0.5	237.1	286.6	289.5	6.3	0.3	1.8	-52.4	3.4	-9.4	-12.7
23	50.3	2.8	237.5	265.3	267.0	7.4	0.4	2.8	-29.5	1.8	-28.8	-30.7
24	9.4	0.1	247.0	292.0	295.0	2.2	0.1	0.5	-48.0	3.0	-4.3	-7.2
25	14.2	0.2	243.5	302.2	306.7	2.6	0.1	1.4	-63.3	4.9	-8.8	-13.7
26	8.6	0.2	239.0	291.6	295.1	2.2	0.1	0.7	-56.1	3.8	-6.0	-9.8
27	75.7	2.1	240.0	268.2	269.8	8.3	0.4	2.1	-29.7	1.1	-23.5	-24.7
28	16.7	0.3	245.5	288.2	290.9	6.0	0.1	1.6	-45.4	2.8	-12.5	-15.3
29	60.7	10.2	228.5	240.8	241.2	4.6	0.9	2.7	-12.7	1.0	-15.2	-16.2
30	36.5	5.4	230.5	256.6	257.1	5.6	0.9	4.8	-26.7	2.2	-18.9	-21.1
31	42.1	0.5	241.2	277.9	279.9	7.3	0.2	1.3	-38.7	1.3	-16.4	-17.6
32	18.2	3.0	234.2	279.6	281.6	4.4	0.8	3.9	-47.3	5.0	-20.8	-25.8
33	37.9	0.5	239.3	284.9	287.9	12.4	0.2	2.7	-48.5	3.1	-15.5	-18.6
34	20.4	0.4	248.9	284.1	285.9	6.0	0.1	1.2	-37.1	1.4	-11.5	-12.9
35	57.7	1.0	238.7	278.0	280.3	12.1	0.3	2.9	-41.6	2.3	-20.4	-22.7
36	12.0	0.2	237.3	281.7	284.3	2.5	0.2	0.9	-47.0	2.1	-2.6	-4.6
37	26.1	0.6	242.3	280.5	282.7	6.2	0.2	1.9	-40.4	2.1	-20.2	-22.2
38	37.9	0.3	239.2	282.9	285.7	9.9	0.2	1.6	-46.5	2.2	-9.7	-11.9
39	19.1	0.9	239.4	287.2	290.0	6.7	0.3	2.7	-50.6	4.1	-16.1	-20.2
40	21.4	3.1	232.6	273.0	274.4	5.3	0.9	4.6	-41.7	3.4	-16.7	-20.1

Fonte: Produção da autora.

Após a obtenção do novo classificador baseado em 12 parâmetros iniciou-se a fase de validação. Uma imagem que não fez parte do conjunto utilizado inicialmente no treinamento foi classificada, o dia escolhido foi 03/03/2020, o mesmo utilizado para validar os classificadores operacional e atualizado.

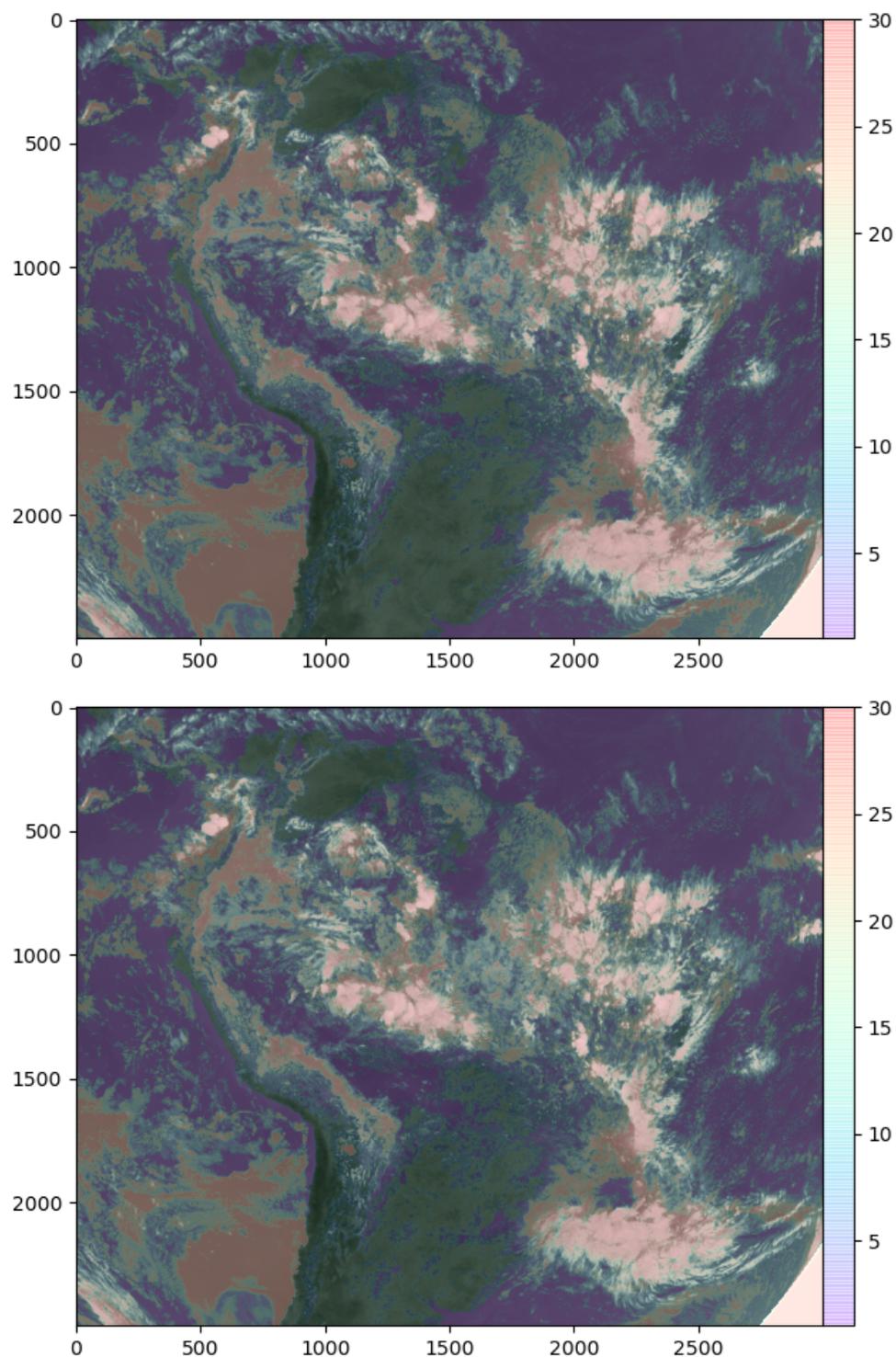
A partir da validação do método, foi possível identificar os alvos de interesse (superfície, cumulus, cirrus, stratus e multicamadas). Vale ressaltar que a classe multicamada aqui refere-se à nuvem cumulonimbus, já que a mesma possui sua base em baixos níveis e devido ao intenso desenvolvimento vertical, se estende até altos níveis (HOUBE JUNIOR, 1993; WALLACE; HOBBS, 2006).

Os mesmos alvos que compõem o classificador operacional, foram utilizados no novo classificador baseado em 12 parâmetros, a fim de obter uma melhor comparação entre os dois classificadores. Novas classes poderiam ter sido criadas, mas isso prejudicaria a comparação entre os dois métodos, o que não impede que esse trabalho seja feito futuramente.

A Figura 4.24 mostra as imagens classificadas dos dois classificadores que utilizam apenas 4 parâmetros, onde a imagem de cima representa o classificador operacional e a imagem de baixo representa o classificador atualizado. Os 30 centroides estão distribuídos espacialmente representando os alvos classificados.

A seguir, a Figura 4.26, mostra a imagem do visível à direita e a esquerda a imagem do infravermelho, abaixo a imagem classificada referente ao novo classificador baseado no GOES-16 e que utiliza 12 parâmetros, onde os 40 centroides estão distribuídos espacialmente e representam os alvos (nuvens e superfície), no horário 15:00 UTC em ambas as figuras.

Figura 4.24 - Imagens classificadas pelo classificador operacional (cima) e atualizado (baixo) para o dia 03/03/2020 no horário 15:00 UTC.

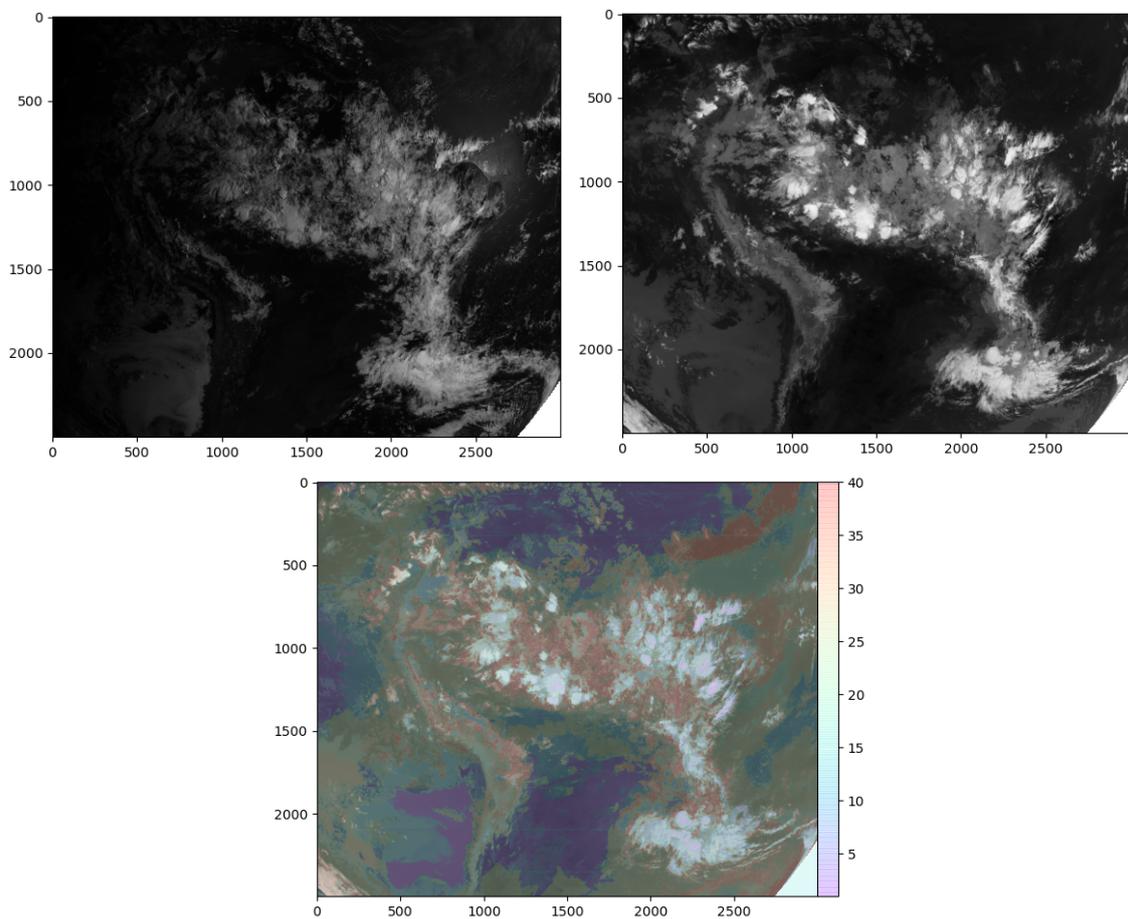


As cores indicadas nas barras de cores, representam os tipos de nuvens selecionados e as superfícies (continental e oceânica).

Fonte: Produção da autora.

Após a obtenção dos novos conjuntos de centroides para os horários das 12:00 15:00 e 18:00 UTC, iniciou-se a fase de validação do novo classificador nestes dois horários, onde os alvos de interesse foram superfície (céu claro), cumulus, cirrus, stratus e multicamadas. Nessa fase foi utilizado o dia 03/03/2020, o mesmo dia aplicado anteriormente. As [Figura 4.25](#), [Figura 4.26](#) e [Figura 4.27](#) mostram as imagens do visível a direita, a imagem do infravermelho a esquerda e abaixo a imagem classificada nos horários 12:00, 15:00 e 18:00 UTC, respectivamente, onde os 40 centroides estão distribuídos espacialmente representando os alvos (nuvem e superfície).

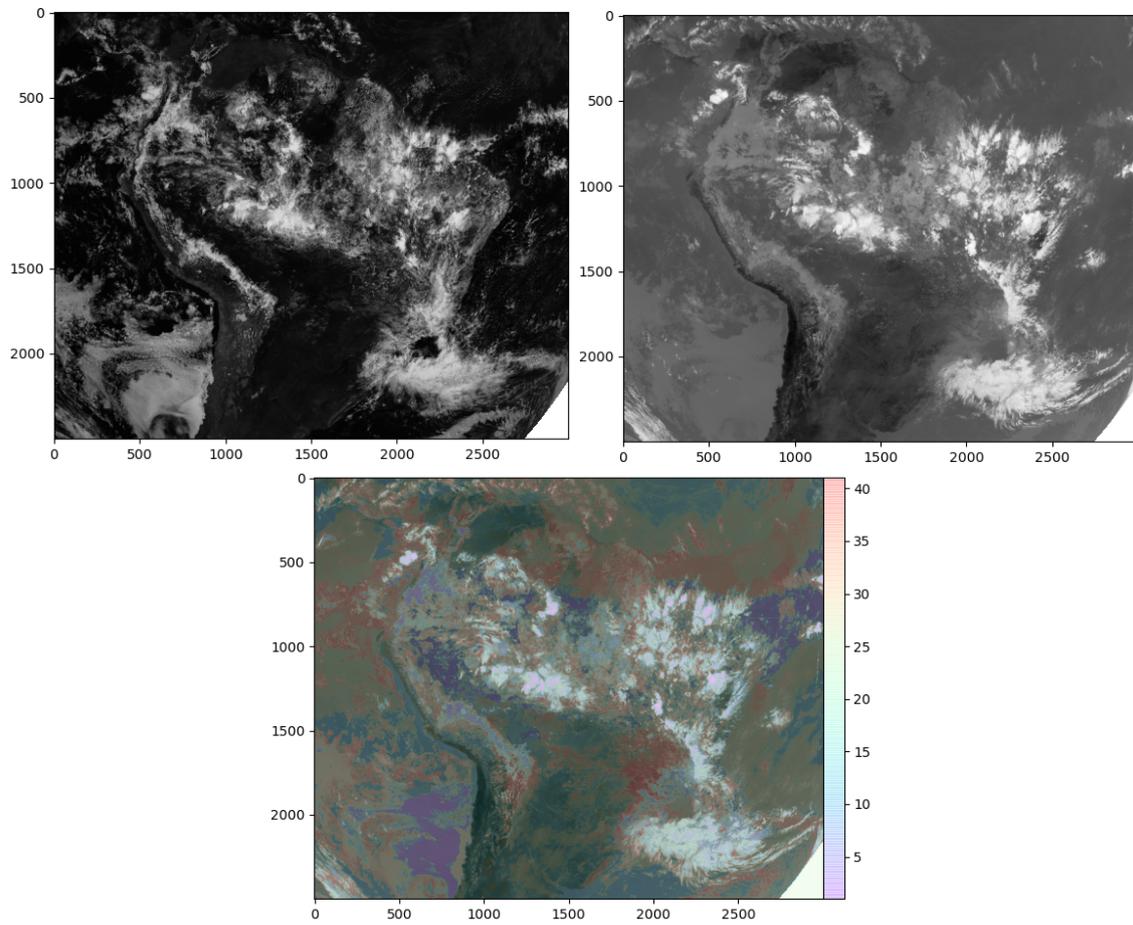
Figura 4.25 - Imagem do visível e infravermelho e imagem classificada para o dia 03/03/2020 no horário 12:00 UTC.



As cores indicadas na barra de cores, representam os tipos de nuvens selecionados e as superfícies (continental e oceânica).

Fonte: Produção da autora.

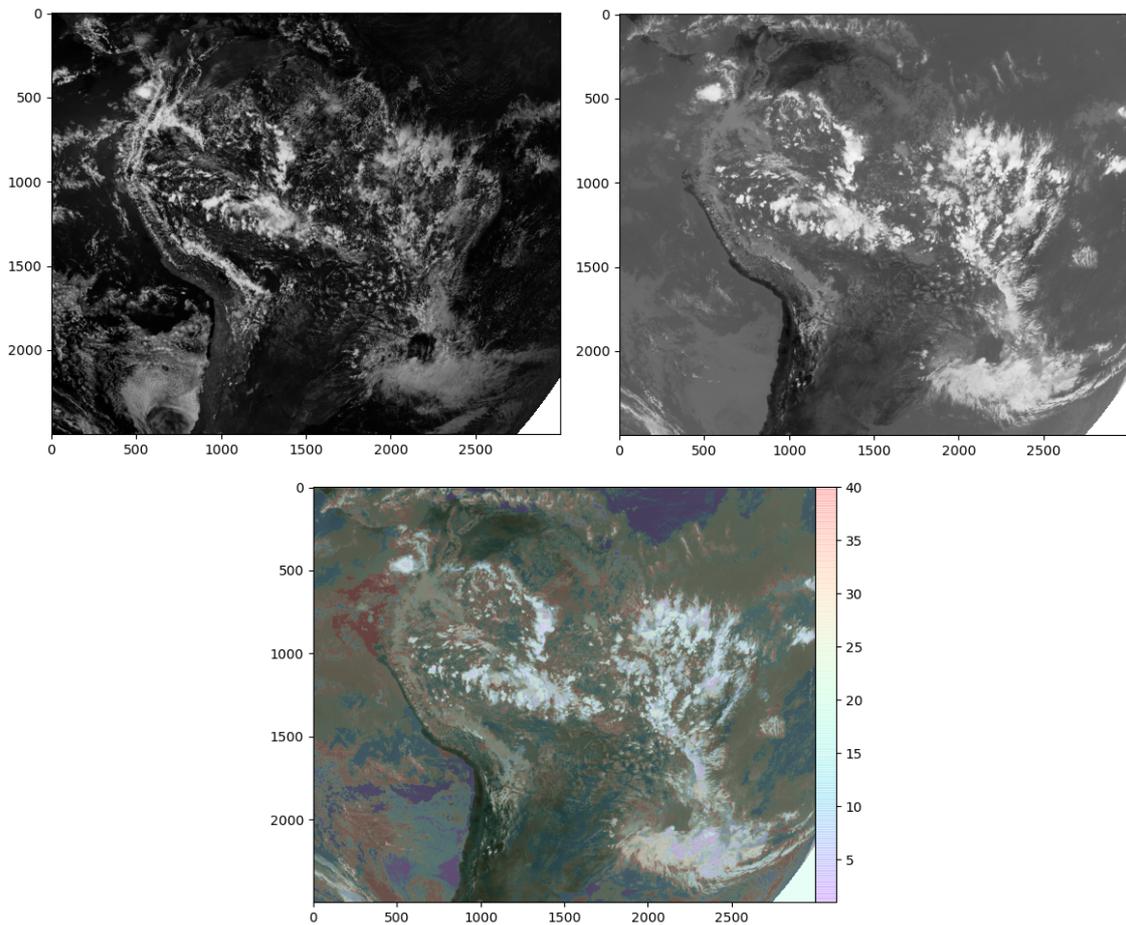
Figura 4.26 - Imagem do visível e infravermelho e imagem classificada para o dia 03/03/2020 no horário 15:00 UTC.



As cores indicadas na barra de cores, representam os tipos de nuvens selecionados e as superfícies (continental e oceânica).

Fonte: Produção da autora.

Figura 4.27 - Imagem do visível e infravermelho e imagem classificada para o dia 03/03/2020 no horário 18:00 UTC.



As cores indicadas na barra de cores, representam os tipos de nuvens selecionados e as superfícies (continental e oceânica).

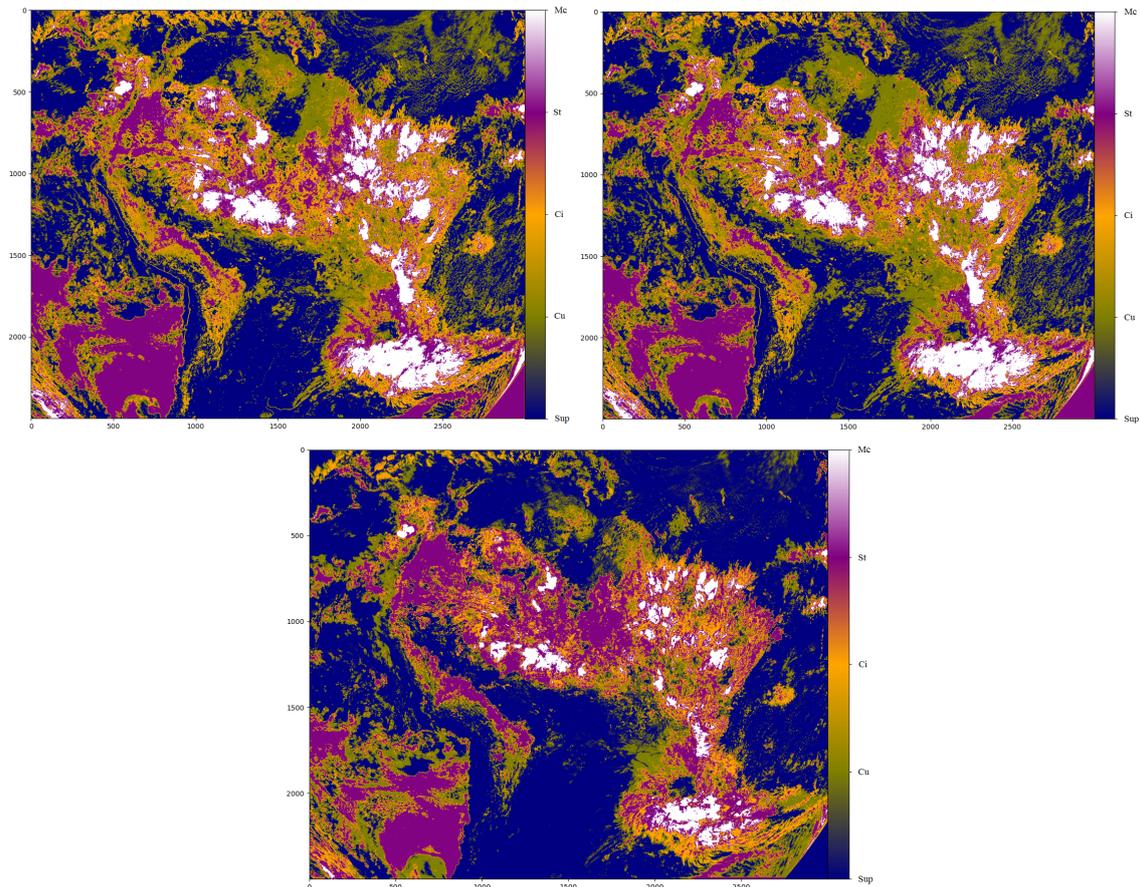
Fonte: Produção da autora.

A identificação de todas as classes de nuvens, na [Figura 4.24](#) (classificador operacional e atualizado) e na [Figura 4.26](#) (novo classificador) ainda é bastante difícil. Para uma melhor representação espacial dos alvos, foi confeccionada uma imagem que mostra todas as classes agrupadas destes classificadores para uma melhor visualização na diferença entre eles no horário 15:00 UTC.

Na [Figura 4.28](#) é possível perceber as diferenças entre os classificadores discutidos no estudo. Observa-se que entre os classificadores baseados em quatro parâmetros, as diferenças são menos perceptíveis quando comparado ao novo classificador baseado

em 12 parâmetros.

Figura 4.28 - Imagens classificadas das classes agrupadas do classificador operacional (esquerda), classificador atualizado (direita) e novo classificador (baixo) para o dia 03/03/2020 no horário 15:00 UTC.



Fonte: Produção da autora.

4.4.4 Avaliação do novo classificador

O novo classificador (12 bandas) foi avaliado a partir da comparação com os produtos de topo de nuvens oficiais da NOAA, temperatura do topo da nuvem, fase da nuvem e máscara de nuvem, o mesmo conjunto de dados utilizados em Seção 4.3 referente ao ano de 2020.

A finalidade da comparação com os três produtos de nuvem foi avaliar se as classes definidas pelo novo classificador são coerentes com as características termodinâmicas, averiguando possíveis melhorias na classificação relacionado ao classificador em operação atualmente.

Previamente à comparação, foi necessário classificar todas as imagens referentes à cada um dos 20 dias selecionados, para então proceder com a comparação, a qual foi implementada dia a dia e ao final calculada a média referente aos 20 dias.

4.4.4.1 Temperatura do topo da nuvem

A comparação entre as classes de nuvens e a temperatura termodinâmica do topo das nuvens permite verificar qual é a fração das classes que provavelmente foram identificadas erroneamente, por exemplo, é esperado que uma nuvem do tipo cumulus não seja associada com temperatura abaixo de 265 K ou uma nuvem cirrus apresente temperatura acima desse valor.

A tabela relacionada à temperatura do topo da nuvem destaca oito temperaturas de brilho em Kelvin (K), para cada uma das cinco classes de alvos, onde os valores representam a porcentagem na classificação em cada uma dessas temperaturas. Além das temperaturas, há uma coluna representando porcentagens onde a comparação não foi possível, devido ao valor no produto da NOAA estar ausente, ou seja, consistir em uma situação de céu claro. Essa situação foi contabilizada na coluna indicada pelo valor "-1" (Tabela 4.11).

Ao avaliar a classificação da superfície, é possível perceber que embora uma parte tenha sido classificada como sendo nuvens, 12% como nuvens mais quentes (297 e 282 K) e 11% como sendo nuvens mais frias, ainda assim 75,6% foram classificados corretamente. Quando comparado aos classificadores operacional e atualizado, houve uma melhora de aproximadamente 4%.

A nuvem do tipo cumulus foi classificada em sua maioria entre as temperaturas 297 e 282 K, com porcentagem total de 62,5%, entretanto 20,7% foram classificadas com topos mais frios e 16,4% não foram classificados em nenhum dos grupos de nuvens. Ao comparar o novo classificador aos classificadores operacional e atualizado foi possível notar que houve um ganho de aproximadamente 8,5%.

As nuvens cirrus em sua maioria 85,8%, foram classificadas com temperatura do topo mais frias (252 a 207 K). Estes valores estão de acordo com o tipo de nuvem, já que estas nuvens se localizam em altos níveis e são compostas de gelo. Ao confrontar os classificadores operacional e atualizado ao novo classificador, nota-se que o novo classificador com centroides baseado no ABI/GOES-16 apresentou 6,7% na melhora da classificação deste tipo de nuvem.

A classificação das nuvens do tipo stratus com temperaturas do topo entre 282 e 222 K, foram de 95,4%, tendo uma melhora na classificação de aproximadamente 2% em relação aos classificadores operacional e atualizado.

As nuvens multicamadas foram as que apresentaram melhor classificação, podemos dizer que 100% dos pixels estão coerentes com a temperatura de topo estimada, com a temperatura do topo variando entre 237 e 207 K. Comparando os resultados do novo classificador aos classificadores operacional e atualizado, foi observado uma melhora em 8%.

Ao comparar a média dos 20 dias em dois horários (12:00 e 18:00 UTC) com o produto de temperatura do topo da nuvem da NOAA, os alvos que tiveram uma redução na identificação correta nestes horários em relação ao horário de 15:00 UTC, foram a superfície (céu claro), tendo uma redução de 10% e 7,5% e as nuvens cirrus, com quase 15% e 10%, respectivamente. As nuvens cumulus e stratus, tiveram um pior desempenho na classificação durante o horário de 12:00 UTC, enquanto a categorização mais acertada foi no horário de 18:00 UTC. Já as nuvens multicamadas foram constantes em todos os três horários com 100% de acerto.

Tabela 4.11 - Distribuição da temperatura do topo da nuvem em K em relação às classes de nuvens para as variantes 12:00, 15:00 e 18:00 UTC do novo classificador baseado em 12 parâmetros.

	312	297	282	267	252	237	222	207	-1
12:00 UTC									
su	0,0	11,9	5,4	0,5	0,2	0,9	4,3	11,2	65,6
cu	0,0	14,7	40,7	8,3	2,2	3,8	11,2	10,8	8,3
ci	0,0	0,5	10,4	12,8	4,8	14,7	33,9	17,5	5,4
st	0,0	0,3	54,8	8,3	3,7	8,0	10,2	1,6	13,1
mc	0,0	0,0	0,0	0,0	0,0	6,7	58,6	34,7	0,0
15:00 UTC									
su	1,4	10,0	2,0	0,2	0,2	0,7	3,6	6,3	75,6
cu	0,4	20,3	42,2	2,8	0,6	2,0	6,8	8,5	16,4
ci	0,0	1,1	5,6	7,2	3,0	13,9	47,9	21,0	0,3
st	0,0	0,5	39,5	19,6	7,3	13,8	15,2	3,9	0,2
mc	0,0	0,0	0,0	0,0	0,0	15,0	59,5	25,5	0,0
18:00 UTC									
su	2,4	12,8	3,7	0,2	0,2	0,9	4,0	7,7	68,1
cu	0,1	15,9	54,1	4,6	0,8	2,3	7,4	7,7	7,1
ci	0,0	10,8	4,6	3,6	2,2	12,0	41,8	22,4	2,6
st	0,0	0,4	18,2	22,5	12,3	25,8	18,4	2,3	0,1
mc	0,0	0,0	0,0	0,0	0,2	8,8	58,9	32,1	0,0

Fonte: Produção da autora.

Ao comparar o novo classificador com o produto oficial de temperatura do topo da nuvem, foi observado que os alvos com maiores ocorrência de identificações errôneas foram a nuvem cumulus e a superfície com 62,5 e 75,6 %, respectivamente. Já as nuvens cirrus, stratus e multicamada foram melhores classificadas. Ao comparar o novo classificador baseados no GOES-16 e os dois classificadores operacional e atualizado, houve uma melhora significativa na classificação dos alvos analisados.

4.4.4.2 Fase da nuvem

A fim de investigar situações onde possam ocorrer classificações discrepantes quanto à fase da nuvem, procedemos com a comparação entre esse parâmetro estimado para o topo das nuvens e as classes definidas pelo novo classificador baseado no GOES-16. Foram classificadas as fases das nuvens para os 20 dias do ano de 2020 e posteriormente realizado a média.

As fases da nuvem consideradas foram: água líquida, água líquida super-resfriada, fase mista e gelo, além destas fases também foi considerado céu claro, a fim de avaliar os casos definidos como superfície.

Observando a [Tabela 4.12](#) é possível verificar que a classe de nuvens mc (multi-camadas) coincide com praticamente 100% da classe "gelo", o que é perfeitamente coerente.

A classe superfície coincide em pouco mais de 50% dos casos com a classe "céu claro" 44,4%, como o pior caso, do classificador no horário de 18:00 UTC.

Nuvens cumulus tiveram uma boa concordância com este produto apresentando os máximos de concordância com a classe "água líquida", 58,1%, 69% e 73,8% para os horários 12:00, 15:00 e 18:00 UTC, respectivamente.

Uma fração considerável, em torno de 20% a 30%, coincidiu com a classe "gelo" indicando que pode haver um erro de classificação provavelmente em regiões de transição entre topos mais baixos e topos mais altos.

Nuvens cirrus deveriam coincidir predominantemente com a classe "gelo" e apesar da boa concordância para o horário de 15:00 UTC (87,6%), em 12:00 UTC apresentou péssimos resultados concordando em cerca de 68% dos casos, mas coincidiu em 14,5% dos casos com a classe "água líquida".

Esse erro pode ter sido induzido pelo efeito de semi-transparência de algumas nuvens cirrus. Também, nessa porção do dia a convecção ainda não atingiu o seu máximo e a quantidade de nuvens cirrus é menor em relação às demais classes.

Quanto às nuvens do tipo stratus, as mesmas podem ser consideradas como nuvens que contém desde água líquida até gelo. Geralmente as nuvens stratus são nuvens baixas e por estarem abaixo da isoterma de 0 °C são compostas exclusivamente por água. Entretanto, como ser visualizado na [Figura 2.1](#), há algumas classes de nuvens estratiformes, as quais são consideradas neste estudo, que vão além da nuvem stratus composta somente por água líquida. Há também as nuvens nimbostratus, por exemplo, que cobrem uma grande extensão horizontal, e que apesar de ter sua base em baixos níveis, possuem uma extensão vertical considerável, podendo ter gelo em seu topo (([HOUBE JUNIOR, 1993](#); [WOLDE; VALI, 2002](#); [YU et al., 2004](#); [WALLACE](#);

HOBBS, 2006).

Portanto, é esperado que essa classe de nuvens coincida predominantemente com todas as classes exceto a classe "céu claro", o que foi observado no horário das 18:00 UTC do novo classificador. Para o horário das 15:00 UTC há um separação quase igual entre as classes "água líquida" e "gelo", compreendendo mais de 80% das amostras, o que é perfeitamente coerente com a estrutura microfísica dessas nuvens.

O pior caso foi no horário das 12:00 UTC em que quase 12% das amostras coincidiram com a classe "céu claro" e a taxa de concordância com a classe "gelo" foi a menor entre os horários. Essa tendência de aumento gradativo da taxa de concordância entre as nuvens classificadas como stratus e a classe "gelo", pode estar associada as alterações na distribuição espacial dos tipos de nuvens durante o ciclo diurno da convecção, em que ao final do dia, com o máximo da atividade convectiva a proporção entre nuvens formadas por cristais de gelo é maior em relação às nuvens com topos mais baixos e formadas por água líquida ou super-resfriada.

A comparação do produto fase de nuvem da NOAA com o novo classificador, dois dos cinco alvos analisados obtiveram uma melhora em relação aos classificadores original e atualizado baseados em 4 parâmetros, foram as nuvens cirrus e as nuvens multicamadas, tendo uma melhora de aproximadamente 15% e 5,6%, respectivamente (Tabela 4.5).

A classe de nuvens stratus apresentou resultado muito similar. Já as nuvens cumulus e a superfície apresentaram uma piora em relação aos resultados obtidos pelos classificadores operacional e atualizado baseado nas medidas do sensor GOES-8/*Imager*.

Tabela 4.12 - Distribuição da fase termodinâmica do topo da nuvem em relação às classes de nuvens para os horários 12:00, 15:00 e 18:00 UTC do novo classificador baseado em 12 parâmetros.

	Céu Claro	Água Líquida	Água Líquida Super-resfriada	Fase Mista	Gelo
12:00 UTC					
su	52,6	28,4	0,0	0,0	19,0
cu	3,1	58,1	2,3	2,5	34,0
ci	4,9	14,5	6,0	6,0	68,6
st	11,9	55,9	5,2	4,0	23,0
mc	0,0	0,0	0,0	0,0	100,0
15:00 UTC					
su	55,3	31,5	0,0	0,0	13,2
cu	9,1	69,0	0,5	0,4	21,0
ci	0,1	9,0	0,8	2,5	87,6
st	0,2	41,1	8,1	7,9	42,7
mc	0,0	0,0	0,0	0,5	99,5
18:00 UTC					
su	44,4	40,7	0,0	0,0	14,9
cu	4,5	73,8	1,0	0,8	19,9
ci	0,8	17,8	0,3	1,0	80,1
st	0,2	22,9	12,6	13,4	50,9
mc	0,0	0,0	0,0	0,4	99,6

Fonte: Produção da autora.

4.4.4.3 Máscara de nuvem

Esta comparação permitiu verificar se o novo classificador pode ser utilizado como uma máscara de nuvens. O produto máscara de nuvem foi utilizado para analisar a classificação dos alvos entre céu claro e céu nublado, em outras palavras, presença ou não de nuvens.

Ao averiguar a média dos 20 dias do ano 2020, pode ser notado que as nuvens cirrus, cumulus e stratus tiveram uma melhora em sua classificação, ao serem categorizados como céu nublado, em aproximadamente 8,7%, 8,7%, 0,3%, respectivamente. Enquanto a multicamada se manteve similar aos classificadores operacional e atualizado baseados no GOES-8, enquanto isso, a superfície apresentou uma queda em relação aos classificadores anteriores de aproximadamente 4%.

Analisando o novo classificador nos horários 12: e 18:00 UTC (Tabela 4.13), foi possível perceber que ao comparar a classificação em 15:00 UTC, houve um declínio em categorizar corretamente a superfície (céu claro) em aproximadamente 8%. O inverso foi observado para as nuvens do tipo cumulus, onde houve uma melhora na categorização em quase 9%. Já os outros tipos de nuvens (cirrus, stratus e multicamadas) a identificação foi semelhante nos três horários.

Tabela 4.13 - Distribuição da cobertura de nuvens em relação às classes de nuvens nos horários 12:00, 15:00 e 18:00 UTC do novo classificador baseado em 12 parâmetros

	Céu Limpo	Céu com Nuvens
12:00 UTC		
su	65,9	34,1
cu	8,4	91,6
ci	5,2	94,8
st	13,7	86,3
mc	0,0	100,0
15:00 UTC		
su	73,9	26,1
cu	16,1	83,9
ci	0,2	99,8
st	0,2	99,8
mc	0,0	100,0
18:00 UTC		
su	65,6	34,4
cu	7,3	92,7
ci	2,5	97,5
st	0,3	99,7
mc	0,0	100,0

Fonte: Produção da autora.

Após a comparação do novo classificador aos três produtos oficiais de nuvem da NOAA, constatou-se que houve uma melhora na classificação das nuvens, principalmente da nuvem cirrus semi-transparentes em relação aos classificadores operacional e atualizado (baseados em 4 parâmetros). Por serem nuvens finas, são bastante difíceis de detectar e classificar devido à semi-transparência mas ao adicionar o canal 4 ($1,37 \mu m$), bem como a dTB5 ($11,2-12,3 \mu m$), verificou-se que houve uma melhora

na classificação deste tipo de nuvem, corroborando com (GAO et al., 1993; FREY et al., 2008).

Os produtos de topos de nuvens da NOAA utilizados nessa análise dos classificadores baseados em 4 ou 12 parâmetros, foram empregados para verificar se os tipos de nuvens identificados estariam ou não coerentes com as variáveis temperatura do topo, fase termodinâmica e, em último caso, se havia nuvens ou não no mesmo ponto em cada um dos produtos. Esses produtos de topos de nuvens possuem erros, pois também são estimativas feitas a partir de observações remotas, portanto é conveniente considerar os possíveis erros durante a interpretação das estatísticas apresentadas nas Tabelas 4.13, 4.12 e 4.11.

Verificar objetivamente a qualidade das classificações obtidas nesse estudo foi um grande desafio. O uso dos produtos de nuvens da NOAA permitiu realizar uma comparação quantitativa entre as classes de nuvens e os parâmetros microfísicos e termodinâmicos dos topos das nuvens estimados a partir do mesmo sensor, ABI/GOES-16, associados à alguns dados auxiliares como previsões numéricas de tempo. Essa metodologia de análise permitiu identificar as situações em que as classes de nuvens podem ser aprimoradas, como as nuvens do tipo cirrus ou aquelas em que o classificador identifica com alta precisão, como as nuvens multicamadas representando Cbs e pode ser utilizada rotineiramente na avaliação desse tipo de produto em um cenário operacional.

5 CONCLUSÕES

Esse trabalho buscou aprimorar o modelo de classificação de nuvens operacional do CPTEC/INPE baseado em imagens do satélite geoestacionário GOES-16. O classificador foi originalmente desenvolvido para ser aplicado em imagens obtidas pelo satélite GOES-8. Portanto, foi necessário atualizá-lo, seguindo a mesma metodologia aplicada na sua criação, ao conjunto de imagens similares que são obtidas pelo satélite atual e verificar as deficiências e as potencialidades de ambas as versões.

O processo de validação desse produto foi desafiador pelo fato da comparação não poder ser feita diretamente, cada classe de nuvem atribuída a cada pixel da imagem teve de ser comparada com outras variáveis como temperatura estimada do topo das nuvens presentes no pixel, a fase termodinâmica ou mesmo uma máscara de nuvens.

Durante a validação dos dois classificadores baseados em 4 parâmetros, encontrou-se diferenças entre as classes de nuvens para os três horários, onde o horário das 12:00 UTC foi o que atingiu valores entre 14% e 16% em vários casos (Figura 4.11). Esse valor compreendia as discrepâncias entre todas as classes possíveis (mc, cu, st, sup, ci). Ao analisar as diferenças entre cada classe separadamente foi possível observar que as discrepâncias foram maiores para algumas classes, como a classe cu ou ci.

Durante essa análise foi identificado que a versão atualizada do classificador de quatro parâmetros identifica nuvens que, a partir de nefanálise, indicam estar equivocadas, como é o exemplo de pixels associados às nuvens cumulus sendo classificados como stratus (Figura 4.15), mesmo que em quantidade muito pequena. Essa análise não permitiu concluir que o classificador atualizado obteve maior ou menor êxito em relação ao classificador operacional, mas permitiu verificar a validade da metodologia de classificação e a escolha dos parâmetros originais aplicados às imagens disponíveis atualmente.

Buscando verificar a qualidade das classificações em relação à alguma inferência independente, foi realizada uma análise dos dois classificadores, e posteriormente replicada para a análise do novo classificador baseado em 12 parâmetros, com alguns dos produtos oficiais de topo de nuvens fornecidos pela NOAA. Essa comparação permitiu avaliar a coerência entre o tipo de nuvem e alguma característica termodinâmica esperada para essa classe, como presença de cristais de gelo, temperatura dentro de uma certa faixa ou a presença ou não de nuvens. Os produtos utilizados

para tal foram a temperatura do topo da nuvem, a fase do topo da nuvem e a máscara de nuvem. Outros produtos estavam disponíveis mais essa seleção atendeu às expectativas.

Com relação à temperatura do topo da nuvem, os tipos de alvos avaliados foram categorizados em diferentes faixas de temperaturas, indo desde temperaturas mais quentes (312 K), até temperaturas mais frias (207 K). Como as classes dos alvos existem nos classificadores, os valores indefinidos do produto de temperatura do topo das nuvens foi considerado e indicado pela coluna de valor *-1*. Foi possível verificar que a diferença entre os dois classificadores (operacional e atualizado) em relação à temperatura do topo da nuvem houve uma melhora de 3% no classificador atualizado na categorização da superfície.

A fase da nuvem é um produto que fornece informações sobre quatro fases: água líquida, água super-resfriada, fase mista e gelo. Foi também adicionado na análise a condição de céu claro, a fim de verificar a classificação da superfície, mas também onde ocorreram classificações errôneas de nuvens, como por exemplo, nuvens cirrus ou bordas identificadas como superfície e vice-versa. Ao analisar os dois classificadores (operacional e atualizado), foi possível verificar, que assim como o produto anterior, neste a diferença entre os dois foi bastante similar, e de modo geral também houve uma melhora em relação ao classificador atualizado, o qual categorizou a nuvem cirrus em aproximadamente 4% a mais em relação ao classificador operacional.

Outro produto utilizado foi máscara de nuvem, este produto nos dá informações com relação à nebulosidade. Foram analisados as condições de céu claro, ou seja, sem presença de nuvens e condições de nebulosidade (com presença de nuvens). Os dois classificadores (operacional e atualizado) tiveram resultados bastante similares entre si, nesta avaliação notou-se que as nuvens que foram mais bem classificadas foram stratus, multicamadas e cirrus. Dentre as três a nuvem cirrus teve em média sua classificação como céu claro em aproximadamente 6%, o que pode indicar presença de cirrus semi-transparentes, as quais foram erroneamente classificadas. Esse resultado é importante e mostra que o classificador atualizado (baseado em quatro parâmetros) pode ser utilizado como uma máscara de nuvens tal qual o produto fornecido pela NOAA.

Ao eliminar os parâmetros redundantes através da técnica PCA, chegou-se ao resul-

tado de 12 parâmetros multiespectrais, oito a mais que o conjunto de parâmetros utilizado no classificador operacional [canal 2 (0,64 μm), canal 4 (1,37 μm), canal 8 (6,2 μm), canal 11 (8,4 μm), canal 13 (10,3 μm), textura do canal 2, textura do canal 8, textura do canal 13, dTB1, dTB5, dTB6 e dTB7].

Estes parâmetros nos dão uma gama de informações que variam desde reflectância, temperatura de brilho, bem como outras características como a diferença entre as temperaturas de brilho dTB6, por exemplo, que é utilizada para identificar nuvens altas, principalmente nuvens cirrus (FREY et al., 2008). Utilizando um conjunto de 6×10^6 (seis milhões) de amostras para o treinamento, correspondente à 10 dias de cada mês durante o ano de 2019 foi possível realizar a aplicação do método *k-means*, o qual gerou novos centroides baseados no GOES-16.

A validação desse novo classificador seguiu a metodologia de comparação com os três produtos de topo de nuvens fornecidos pela NOAA, que foram aplicados na validação dos classificadores baseados em quatro bandas.

Em relação à temperatura do topo da nuvem, concluiu-se que, apesar de terem ocorrido classificações errôneas como por exemplo, nuvens do tipo cumulus, as quais são nuvens quentes, e cerca de 8,5 % foram categorizadas com temperatura do topo igual a 207 K. De forma geral, todos os alvos analisados tiveram uma classificação coerente acima dos 60%.

Houveram algumas classificações errôneas entre o produto da fase da nuvem e o novo classificador, como foi o caso da nuvem cumulus, classificada como gelo em 21%. Contudo, as nuvens cirrus tiveram uma melhora na classificação de 15%. Na comparação com o produto máscara de nuvem, dos cinco alvos analisados, somente a classificação da superfície como céu claro foi similar aos classificadores operacional e atualizado.

Os horários 12:00 e 18:00 UTC do novo classificador baseado em 12 parâmetros, apresentaram erros geralmente superiores em relação ao horário das 15:00 UTC na classificação de alguns tipos de alvos, como por exemplo, superfície (céu claro). Para a aplicação do método *k-means* utilizou-se sementes iniciais extraídas do horário 15:00 UTC, isso pode ter ocasionado um efeito negativo na classificação correta dos alvos nos outros dois horários. Recomendamos a seleção de sementes iniciais diferentes nestes dois intervalos de horários.

De modo geral, é possível afirmar que o classificador atualizado de nuvens baseado no sensor ABI/GOES-16, cujo centroides originais, baseados no sensor *Imager*/GOES-8, foram utilizados como sementes iniciais, houve uma pequena melhora na classificação quando analisada visualmente a partir de estudos de caso. Esta melhora foi reforçada ao compará-lo aos produtos de topos de nuvens da NOAA.

Com a adição de novos parâmetros ao conjunto original, e ao utilizar um novo grupo de centroides baseados no ABI/GOES-16, houve uma melhora ainda maior na classificação de nuvens, como pôde ser visto ao comparar o novo classificador aos produtos de nuvens da NOAA.

5.1 Sugestões de trabalhos futuros

- Avaliar o impacto de matrizes maiores que 3 x 3 para no cálculo do parâmetro de textura;
- Verificar e testar outras possíveis combinações entre as diferenças da temperaturas de brilhos que ajudem a melhorar a classificação das nuvens;
- Escolher novas sementes iniciais para a geração do classificador baseado em 12 parâmetros nos horários diurnos das 12:00 e 18:00 UTC;
- Gerar classificadores para o período noturno, resultando em uma classificação de nuvens contínua.
- Realizar uma classificação a partir de um sistema de votação a partir de um *ensamble* de classificadores.

REFERÊNCIAS BIBLIOGRÁFICAS

- ACKERMAN, S. A.; STRABALA, K. I.; MENZEL, W. P.; FREY, R. A.; MOELLER, C. C.; GUMLEY, L. E. Discriminating clear sky from clouds with MODIS. **Journal of Geophysical Research: Atmospheres**, v. 103, n. D24, p. 32141–32157, dec 1998. ISSN 01480227. Disponível em: <<http://doi.wiley.com/10.1029/1998JD200032>>. 4, 10, 33
- BAGIROV, A. M.; UGON, J.; WEBB, D. Fast modified global k-means algorithm for incremental cluster construction. **Pattern Recognition**, v. 44, n. 4, p. 866–876, apr 2011. ISSN 00313203. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0031320310005029>>. 27
- BOTTINO, M. J.; CEBALLOS, J. C. Daytime cloud classification over South American region using multispectral GOES-8 imagery. **International Journal of Remote Sensing**, v. 36, n. 1, p. 1–19, jan 2015. ISSN 0143-1161. Disponível em: <<http://dx.doi.org/10.1080/01431161.2014.978953><https://www.tandfonline.com/doi/full/10.1080/01431161.2014.978953>>. 1, 2, 3, 4, 7, 10, 11, 15, 28, 29, 30, 33, 39, 74, 75
- CHU, H.-J.; LIAU, C.-J.; LIN, C.-H.; SU, B.-S. Integration of fuzzy cluster analysis and kernel density estimation for tracking typhoon trajectories in the Taiwan region. **Expert Systems with Applications**, Elsevier Ltd, v. 39, n. 10, p. 9451–9457, aug 2012. ISSN 09574174. Disponível em: <<http://dx.doi.org/10.1016/j.eswa.2012.02.114><https://linkinghub.elsevier.com/retrieve/pii/S0957417412003776>>. 10, 11, 12
- COTTON, W. R.; BRYAN, G.; HEEVER, S. C. van den. Storm and cloud dynamics: the dynamics of clouds and precipitating mesoscale systems. **International Geophysics.**, v. 10, p. S0074–6142, 2011. 66, 69, 74
- DESSLER, A.; YANG, P. The distribution of tropical thin cirrus clouds inferred from terra modis data. **Journal of Climate**, v. 16, n. 8, p. 1241–1247, 2003. 32, 67
- DOWNS, G.; BARNARD, J. **Hierarchical and non-hierarchical clustering**. 1995. Disponível em: <<https://www.daylight.com/meetings/mug96/barnard/E-MUG95.html>>. Acesso em: 28 jan. 2020. 11

ESCRIG, H.; BATLLES, F.; ALONSO, J.; BAENA, F.; BOSCH, J.; SALBIDEGOITIA, I.; BURGALETA, J. Cloud detection, classification and motion estimation using geostationary satellite imagery for cloud cover forecast. **Energy**, v. 55, p. 853–859, jun 2013. ISSN 03605442. Disponível em: <<http://dx.doi.org/10.1016/j.energy.2013.01.054><https://linkinghub.elsevier.com/retrieve/pii/S0360544213000856>>. 2, 7, 9

FIGUEIREDO FILHO, D. B.; SILVA JÚNIOR, J. A. Desvendando os mistérios do coeficiente de correlação de Pearson (r). **Revista Política Hoje**, v. 18, n. 1, p. 115–146, 2009. ISSN 0104-7094. 82

FREY, R. A.; ACKERMAN, S. A.; LIU, Y.; STRABALA, K. I.; ZHANG, H.; KEY, J. R.; WANG, X. Cloud detection with MODIS. part I: improvements in the MODIS cloud mask for collection 5. **Journal of Atmospheric and Oceanic Technology**, v. 25, n. 7, p. 1057–1072, jul 2008. ISSN 1520-0426. Disponível em: <<http://journals.ametsoc.org/doi/10.1175/2008JTECHA1052.1>>. 2, 4, 9, 10, 33, 103, 107

GAO, B.-C.; GOETZ, A. F. H.; WISCOMBE, W. J. Cirrus cloud detection from Airborne Imaging Spectrometer data using the 1.38 μm water vapor band. **Geophysical Research Letters**, v. 20, n. 4, p. 301–304, feb 1993. ISSN 00948276. Disponível em: <<http://doi.wiley.com/10.1029/93GL00106>>. 9, 33, 67, 103

GONG, X.; RICHMAN, M. B. On the application of cluster analysis to growing season precipitation data in North America east of the rockies. **Journal of Climate**, v. 8, n. 4, p. 897–931, apr 1995. ISSN 0894-8755. 10, 11

GOODMAN, A.; HENDERSON-SELLERS, A. Cloud detection and analysis: a review of recent progress. **Atmospheric Research**, v. 21, n. 3-4, p. 203–228, may 1988. ISSN 01698095. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/0169809588900270>>. 5

GOVENDER, P.; SIVAKUMAR, V. Application of k-means and hierarchical clustering techniques for analysis of air pollution: a review (1980–2019). **Atmospheric Pollution Research**, v. 11, n. 1, p. 40–56, jan 2020. ISSN 13091042. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1309104219304556>>. 11, 28

- HAN, S.; LEE, H.; KIM, J.; LEE, W. On the connectivity in opportunistic D2D networks with hierarchical and non-hierarchical clustering. In: **Globecom Workshops, 2016**. Proceedings... IEEE, 2016. p. 1–6. ISBN 978-1-5090-2482-7. Disponível em: <<http://ieeexplore.ieee.org/document/7848977/>>. 11
- HEIDINGER, A. **Algorithm theoretical basis document ABI cloud height**. 2012. 1–79 p. Disponível em: <https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_Cloud%20Height_v3.0_July%202012.pdf>. 21, 23, 24
- HEIDINGER, A.; STRAKA III, W. C. **Algorithm theoretical basis document ABI cloud mask**. 2012. 1–106 p. Disponível em: <https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_Cloud_Mask_v3.0_July%202012.pdf>. 26, 27
- HONGYU, K.; SANDANIELO, V. L.; OLIVEIRA JUNIOR, G. J. Análise de componentes principais: resumo teórico, aplicação e interpretação principal component analysis: theory, interpretations and applications. **ES - Engineering and Science**, v. 5, n. 1, p. 83–90, 2015. ISSN 2358-5390. 31
- HOUZE JUNIOR, R. A. **Cloud dynamics**. [S.l.]: Academic Press, 1993. 1, 5, 67, 90, 99, 100
- HOWARD, L. I. On the modifications of clouds, and on the principles of their production, suspension, and destruction; being the substance of an essay read before the Askesian Society in the session 1802–3. **The Philosophical Magazine**, v. 17, n. 65, p. 5–11, oct 1803. ISSN 1941-5796. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/14786440308676365>>. 5
- INOUE, T. On the temperature and effective emissivity determination of semi-transparent cirrus clouds by bi-spectral measurements in the 10 μ m window region. **Journal of the Meteorological Society of Japan Serie II**, v. 63, n. 1, p. 88–99, 1985. 32
- JAIN, A. K. Data clustering: 50 years beyond K-means. **Pattern Recognition Letters**, v. 31, n. 8, p. 651–666, jun 2010. ISSN 01678655. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0167865509002323>>. 10, 11, 12, 13, 27
- JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering. **ACM Computing Surveys (CSUR)**, v. 31, n. 3, p. 264–323, sep 1999. ISSN 0360-0300. Disponível em: <<http://dl.acm.org/doi/10.1145/331499.331504>>. 10, 11, 27, 28

JIN, W.; WANG, L.; ZENG, X.; LIU, Z.; FU, R. Classification of clouds in satellite imagery using over-complete dictionary via sparse representation. **Pattern Recognition Letters**, v. 49, p. 193–200, nov 2014. ISSN 01678655. Disponível em: <<http://dx.doi.org/10.1016/j.patrec.2014.07.015><https://linkinghub.elsevier.com/retrieve/pii/S0167865514002323>>. 1, 2, 3, 5, 7, 9

LAZRI, M.; AMEUR, S.; BRUCKER, J. M.; OUALLOUCHE, F. Convective rainfall estimation from MSG/SEVIRI data based on different development phase duration of convective systems (growth phase and decay phase). **Atmospheric Research**, v. 147-148, p. 38–50, 2014. ISSN 01698095. Disponível em: <<http://dx.doi.org/10.1016/j.atmosres.2014.04.019>>. 32

LAZRI, M.; AMEUR, S.; MOHIA, Y. Instantaneous rainfall estimation using neural network from multispectral observations of SEVIRI radiometer and its application in estimation of daily and monthly rainfall. **Advances in Space Research**, v. 53, n. 1, p. 138–155, 2014. ISSN 02731177. Disponível em: <<http://dx.doi.org/10.1016/j.asr.2013.10.005>>. 32

LI, C.; MA, J.; YANG, P.; LI, Z. Detection of cloud cover using dynamic thresholds and radiative transfer models from the polarization satellite image. **Journal of Quantitative Spectroscopy and Radiative Transfer**, v. 222-223, p. 196–214, jan 2019. ISSN 00224073. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0022407318304254>>. 3

LIKAS, A.; VLASSIS, N.; VERBEEK, J. J. The global k-means clustering algorithm. **Pattern Recognition**, v. 36, n. 2, p. 451–461, feb 2003. ISSN 00313203. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0031320302000602>>. 27

LIU, Y.; KEY, J. R.; FREY, R. A.; ACKERMAN, S. A.; MENZEL, W. Nighttime polar cloud detection with MODIS. **Remote Sensing of Environment**, v. 92, n. 2, p. 181–194, aug 2004. ISSN 00344257. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0034425704001683>>. 32

MAĆKIEWICZ, A.; RATAJCZAK, W. Principal components analysis (PCA). **Computers Geosciences**, v. 19, n. 3, p. 303–342, mar 1993. ISSN 00983004. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/009830049390090R>>. 31

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In: **BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY**, 5., 1967. [S.l.]: Proceedings..., 1967. p. 281–297. 11

MCCORKEL, J.; EFREMOVA, B.; HAIR, J.; ANDRADE, M.; HOLBEN, B. GOES-16 ABI solar reflective channel validation for earth science application. **Remote Sensing of Environment**, v. 237, p. 111438, feb 2020. ISSN 00344257. Disponível em: <<https://doi.org/10.1016/j.rse.2019.111438><https://linkinghub.elsevier.com/retrieve/pii/S0034425719304572>>. 18

METOFFICE. **Cloud names and classifications**. n.d. Disponível em: <<https://www.metoffice.gov.uk/weather/learn-about/weather/types-of-weather/clouds/cloud-names-classifications>>. Acesso em: 24 nov. 2020. 5

MINNIS, P.; HECK, P. W. **GOES-R Advanced Baseline Imager (ABI) algorithm theoretical basis document for nighttime cloud optical depth, cloud particle size, cloud ice water path, and cloud liquid water path**. [s.n.], 2012. 84 p. Disponível em: <https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_Cloud_NCOMP_v3.0_Jul2012.pdf>. 22

OUKALI, S.; LAZRI, M.; LABADI, K.; BRUCKER, J. M.; AMEUR, S. Development of a hybrid classification technique based on deep learning applied to MSG / SEVIRI multispectral data. **Journal of Atmospheric and Solar-Terrestrial Physics**, v. 193, p. 105062, oct 2019. ISSN 13646826. Disponível em: <<https://doi.org/10.1016/j.jastp.2019.105062><https://linkinghub.elsevier.com/retrieve/pii/S1364682619300331>>. 3, 7, 10, 33

PAVOLONIS, M. **GOES-R Advanced Baseline Imager (ABI) algorithm theoretical basis document for cloud type and cloud phase**. 2010. 1–60 p. Disponível em: <https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_Cloud_Phase_Type_v2.0_Sep2010.pdf>. 2, 17, 21, 25, 26

PETER, D. **Introductory statistics with R**. [S.l.]: Springer, 2008. 40

RASCHKE, E. The International Satellite Cloud Climatology Project, ISCCP, and its European Regional Experiment ICE (International Cirrus Project). **Atmospheric Research**, v. 21, n. 3-4, p. 191–201, may 1988. ISSN 01698095.

Disponível em:

<<https://linkinghub.elsevier.com/retrieve/pii/0169809588900269>>. 5

SCHMIT, T. J.; GRIFFITH, P.; GUNSHOR, M. M.; DANIELS, J. M.; GOODMAN, S. J.; LEBAIR, W. J. A closer look at the ABI on the goes-r series. **Bulletin of the American Meteorological Society**, v. 98, n. 4, p. 681–698, 2017. ISSN 00030007. 17, 18, 20

SCHMIT, T. J.; GUNSHOR, M. M. ABI imagery from the GOES-R series. In: GOODMAN, S. J.; SCHMIT, T. J.; DANIELS, J.; REDMON, R. J. (Ed.). **The GOES-R series**. [S.I.]: Elsevier, 2020. p. 23–34. ISBN 978-0-12-814327-8.

Disponível em: <<http://www.sciencedirect.com/science/article/pii/B9780128143278000044>>. 18

SCHMIT, T. J.; GUNSHOR, M. M.; FU, G.; RINK, T.; BAH, K.; WOLF, W. **GOES-R Advanced Baseline Imager (ABI) algorithm theoretical basis document for cloud and moisture imagery product**. 2010. 62 p. Disponível em: <https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_ABI_CMI_KPP_v3.0_July2012.pdf>. 19

STRABALA, K. I.; ACKERMAN, S. A.; MENZEL, W. P. Cloud properties inferred from 8-12- μ m data. **Journal of Applied Meteorology**, v. 33, n. 2, p. 212–229, feb 1994. ISSN 0894-8763. 1, 3, 5, 7, 8, 10, 32

TAPAKIS, R.; CHARALAMBIDES, A. Equipment and methodologies for cloud detection and classification: a review. **Solar Energy**, v. 95, p. 392–430, sep 2013. ISSN 0038092X. Disponível em:

<<https://linkinghub.elsevier.com/retrieve/pii/S0038092X12004069>>. 67

THIES, B.; NAUSS, T.; BENDIX, J. Precipitation process and rainfall intensity differentiation using Meteosat Second Generation Spinning Enhanced Visible and Infrared Imager data. **Journal of Geophysical Research**, v. 113, n. D23, p. D23206, dec 2008. ISSN 0148-0227. Disponível em:

<<http://doi.wiley.com/10.1029/2008JD010464>>. 4, 7, 8, 10, 33

WALLACE, J. M.; HOBBS, P. V. **Atmospheric science: an introductory survey**. [S.I.]: Elsevier, 2006. 66, 67, 69, 74, 75, 90, 99, 100

WALTHER, A.; CIMSS, U. W.; STRAKA, W.; HEIDINGER, A. K.; STAR, N. N. **Center for satellite applications and ABI algorithm theoretical basis document for Daytime Cloud Optical and Microphysical Properties (**

- DCOMP). [s.n.], 2013. 66 p. Disponível em:
 <https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_Cloud_DCOMP_v3.0_Jun2013.pdf>. 22
- WEBB, A. R. **Statistical pattern recognition**. [S.l.]: John Wiley & Sons, 2003. 10, 11, 31
- WILKS, D. S. **Statistical methods in the atmospheric sciences**. [S.l.]: Academic Press, 2011. 31, 40, 82
- WMO, W. M. O. **International cloud atlas, volume I: manual on the observation of clouds and other meteors**. Geneva: [s.n.], 1975. 180 p. Disponível em: <https://cloudatlas.wmo.int/docs/wmo_407_en-v1.pdf>. 5
- _____. **International cloud atlas**. 2017. Disponível em: <<https://cloudatlas.wmo.int/en/clouds.html>>. Acesso em: 24 nov. 2020. 5
- WOLDE, M.; VALI, G. Cloud structure and crystal growth in nimbostratus. **Atmospheric Research**, v. 61, n. 1, p. 49–74, 2002. 99, 100
- XIA, M.; LU, W.; YANG, J.; MA, Y.; YAO, W.; ZHENG, Z. A hybrid method based on extreme learning machine and k-nearest neighbor for cloud classification of ground-based visible cloud image. **Neurocomputing**, Elsevier, v. 160, p. 238–249, jul 2015. ISSN 09252312. Disponível em: <<http://dx.doi.org/10.1016/j.neucom.2015.02.022https://linkinghub.elsevier.com/retrieve/pii/S092523121500171X>>. 1, 5
- XIE, J.; JIANG, S. A simple and fast algorithm for global K-means clustering. In: **INTERNATIONAL WORKSHOP ON EDUCATION TECHNOLOGY AND COMPUTER SCIENCE, 2010**. Proceedings... IEEE, 2010, 2010. p. 36–40. ISBN 978-1-4244-6388-6. Disponível em: <<http://ieeexplore.ieee.org/document/5460041/>>. 10, 11
- YU, R.; WANG, B.; ZHOU, T. Climate effects of the deep continental stratus clouds generated by the tibetan plateau. **Journal of Climate**, v. 17, n. 13, p. 2702–2713, 2004. 99, 100