

XVIII Congresso de Iniciação Científica da Universidade Federal de Campina Grande



Sistema computacional em nuvem para monitoramento da cobertura do solo da Caatinga

Michel Deivithy de Sousa Wanderley ¹, John Elton de Brito Leite Cunha ²

RESUMO

O Carbono Orgânico Total (COT) contido nos solos pode ser utilizado como um bom indicador de degradação. Atualmente, para identificar solos degradados é preciso ir até o local e avaliar uma amostra de solo, se tornando um processo demorado e custoso. O objetivo deste estudo foi desenvolver um sistema computacional em nuvem para o monitoramento de áreas degradadas no estado da Paraíba. Através da espectroscopia, os satélites podem identificar a assinatura espectral de áreas a partir da sua quantidade de COT e nos fornecer esses valores através de imagens, utilizando-a para identificação das áreas que estão degradadas ou conservadas a partir do treinamento de um classificador a fim de que ele consiga reconhecer padrões de COT para avaliação das áreas de interesse. Os resultados indicam que o uso de sensoriamento remoto para identificação de áreas degradadas é uma opção de baixo custo e fácil implementação. O desenvolvimento do aplicativo para compartilhar os dados de degradação dos solos do estado da Paraíba ampliou o alcance do uso deste dados para ampla comunidade de usuários.

Palavras-chave: Carbono Orgânico, Degradação dos solos, Sensoriamento Remoto.

¹ Aluno de Engenharia de Produção, Unidade Acadêmica de Engenharia de Produção, UFCG, Sumé, PB, e-mail: michel.deivithy@estudante.ufcg.edu.br

² Doutor, Professor, UAEP, CDSA, UFCG, Campina Grande, PB, e-mail: john.elton@professor.ufcg.edu.br

Cloud computing system for monitoring the Caatinga land cover

ABSTRACT

Total Organic Carbon (TOC) contained in soils can be used as a good indicator of degradation. To identify degraded soils, it is necessary to go to the site and assess a soil sample, which is a time-consuming and costly process. This study aimed to develop a cloud computing system for degraded monitoring areas in the state of Paraíba. Through spectroscopy, satellites can identify the spectral signature of regions from their amount of TOC and provide us with these values through images, using it to identify areas that are degraded or conserved from the training of a classifier so that it can recognize COT patterns for assessment of places of interest. The results indicate that the use of remote sensing to identify degraded areas is a low-cost and easy-to-implement option. The development of the application to share land degradation data from the state of Paraíba has expanded the scope of the use of this data for a wide community of users.

Keywords: Organic Carbon, Soil Degradation, Remote Sensing.

¹ Aluno de Engenharia de Produção, Unidade Acadêmica de Engenharia de Produção, UFCG, Sumé, PB, e-mail: michel.deivithy@estudante.ufcg.edu.br

² Doutor, Professor, UAEP, CDSA, UFCG, Campina Grande, PB, e-mail: john.elton@professor.ufcg.edu.br

1 INTRODUÇÃO

A Caatinga é o único bioma exclusivamente brasileiro, ele ocupa uma área equivalente a 11% do território nacional, e é caracterizada por apresentar solos, com raras exceções, pouco desenvolvidos, rico em minerais, pedregosos e pouco espessos e com fraca capacidade de retenção da água, o que limita a produção primária nessa região (ALVES; ARAÚJO; NASCIMENTO, 2009). Este ecossistema abriga uma grande biodiversidade, inclusive espécies endêmicas, que não existem em nenhuma outra parte do mundo. Independentemente de sua notoriedade, é evidente que os governos não têm dado a devida atenção para a manutenção da preservação da Caatinga, visto que atualmente apenas 53% da sua cobertura vegetal permanece preservada (MINISTÉRIO DO MEIO AMBIENTE, 2013).

A biodiversidade de lugares onde há grande presença de pessoas está ligada aos indicadores socioeconômicos da região e a resposta biofísica do uso do solo (SÁNCHEZ-AZOFEIFA *et al.*, 2005), regiões com baixos indicadores são teoricamente mais suscetíveis a degradação ambiental (Sampaio *et al.*, 2003). A degradação do solo pode ser definida como a redução da produtividade biológica do solo induzida por ações humanas ou alterações climáticas (Reynolds *et al.*, 2002). Nas regiões áridas e semiáridas, em especial, os processos de degradação são mais severos e resultam em desertificação, a maioria dessas regiões estão em risco se não forem gerenciadas corretamente (NICHOLSON; TUCKER, C. J.; BA, 1998), necessitando de monitoramento para que estas possam ter os seus mecanismos de degradação entendidos e assim, seus governos possam criar políticas de combate mais assertivas.

O processo de desertificação sempre se inicia com o desmatamento e substituição da sua vegetação nativa por outra cultivável, de porte ou ciclo diferente, assim a falta da reposição de nutrientes no solo causa a perda de fertilidade (PEREZ-MARIN; MENDONÇA; CAVALCANTE, 2012). Apesar da degradação nessas áreas ser evidente, segundo Perez-Marin (2012), ainda não existem indicadores capazes de medir o nível de degradação dessas áreas de maneira consistente. Além disso, os métodos existentes não são aplicáveis para grandes áreas.

O uso de dados sobre a cobertura vegetal (CV) é necessário para determinar as taxas de alteração dos recursos biológicos (TUCKER, Compton J.; TOWNSHEND, J. R. G.; GOFF, 1985). A saúde do ecossistema tem sido relacionada a quantidade de vegetação presente e por isso, mudanças em diversos tipos de cobertura vegetal vêm sendo estudadas utilizando técnicas de sensoriamento remoto (SONG *et al.*, 2014), uma vez que as imagens de satélite facilitam a investigação dos componentes da vegetação por meio das suas refletâncias espectrais (TUCKER, Compton J., 1979).

A análise de dados de satélite a partir de séries temporais permitem a observação de tendências anuais e sazonais da cobertura florestal (SCHMIDT; KARNIELI, 2000), em regiões áridas e semiáridas é aconselhável o seu uso uma vez que essas áreas sofrem uma grande variabilidade nos seus índices de precipitação (TUCKER, Compton J.; TOWNSHEND, J. R. G.; GOFF, 1985). No entanto, devido a dinâmica da cobertura vegetal da Caatinga, a avaliação da degradação a partir de técnicas tradicionais de análise da CV se torna inviável, uma vez que a vegetação perde suas folhas na estiagem para se preservar.

Uma alternativa para a detecção de áreas degradadas é a mensuração da quantidade de Carbono Orgânico Total (COT) contido no solo. Medições acuradas do COT permitem a obtenção de informações importantes para a qualidade do solo (MCBRIDE *et al.*, 2011). A utilização de reflectância espectral é uma ótima estratégia para as medições de COT, uma vez que podem ser utilizados dados de satélite para obtenção dessas informações, oferecendo uma análise rápida e não destrutiva (SORENSEN; QUIDEAU; RIVARD, 2018).

Pensando na escalabilidade dessa aplicação, o uso do Machine Learning (ML) poderia fornecer a capacidade de processamento adequada para a identificação de áreas degradadas a partir do COT. Tendo em vista a complexidade da classificação de uma imagem, a utilização de artifícios que tornem automática e imparcial a determinação do grau de degradação de uma área de interesse é essencial para atingir os resultados desejados de classificação. O ML é um ramo da inteligência artificial focado na construção de aplicações que aprendem a partir de dados e melhoram sozinhas com a precisão com o tempo (IBM, 2020). Ele pode ser

definido como conjunto de métodos computacionais que utilizam a experiência para melhorar sua performance ou fazer previsões precisas (MOHRI, 2018), podendo otimizar sua performance utilizando dados de exemplo (ALPAYDIN, 2020).

Apesar de ser muito útil o monitoramento da cobertura do solo utilizando satélites, são necessários uma grande capacidade computacional para serem processados em escala regional. Sendo assim, é de grande ajuda a utilização de plataformas computacionais para processamento de grandes volumes de dados, como o Google Earth Engine (GEE), a união da capacidade de processamento do GEE, aliado ao ML supervisionado, ou seja, o treinamento do algoritmo a partir dos valores padrões de áreas com pouco COT para identificação, pode transformar as ideias descritas acima em uma poderosa ferramenta de classificação de áreas degradadas..

Esse estudo tem como objetivo o desenvolvimento de um sistema computacional em nuvem para o monitoramento da Paraíba, que seja capaz de detectar diferentes níveis de degradação do solo para oferecer uma solução tecnológica, acessível e de fácil utilização para os membros da academia, bem como os cidadãos comuns, para que tenham acesso à informações sobre degradação em qualquer área de interesse na Paraíba.

2 MATERIAIS E MÉTODOS

Esta pesquisa utiliza como fonte de dados, imagens de satélite e amostras de campo de solos degradados. Estas informações foram analisadas a partir da análise comparativa da resposta das bandas do espectro eletromagnético à quantidade de carbono orgânico contido nos solos degradados (amostras). Esta relação foi utilizada para treinar um algoritmo de maneira que fosse capaz de identificar solos degradados em qualquer lugar do estado da Paraíba.

2.1 COLETA DE IMAGENS E AMOSTRAS DE CAMPO

As imagens de satélite utilizadas para o desenvolvimento desse estudo foram obtidas do satélite Sentinel-2, pertencente ao programa Global Monitoring for Environment and Security (GMES) e administrado pela União Europeia. Os sensores instalados nesse satélite são capazes de capturar imagens de alta resolução

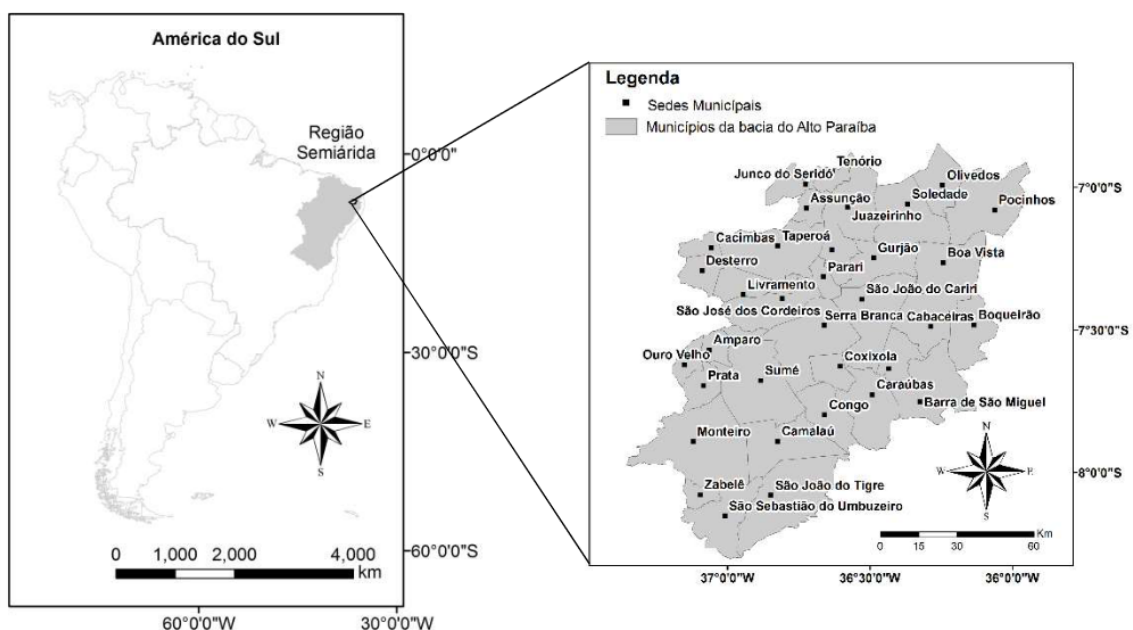
espacial (10m), além de conseguir dividir em 13 as faixas (bandas) do espectro eletromagnético.

As amostras de campo foram frutos do estudo desenvolvido por pesquisadores do Instituto Nacional do Semiárido (INSA), consistem em uma tabela de coordenadas de diversos pontos da Paraíba, em cada uma delas é indicada o COT e seu estado de conservação. As mesmas foram gentilmente cedidas para o desenvolvimento desse projeto. São dados sobre as amostras de solo dos núcleos de desertificação da Paraíba, contendo informações pertinentes para o modelo de classificação.

2.2 ÁREA DE ESTUDO

A bacia do Alto Paraíba (Figura 1) está localizada no estado da Paraíba e está situada em um dos núcleos de desertificação da Paraíba. A região drena suas águas para o reservatório Epitácio Pessoa, conhecido por Boqueirão, e é o responsável por abastecer um grande número de cidades, a mais importante delas, Campina Grande. A degradação dos solos nessa área pode causar um grande desequilíbrio socioambiental, tendo em vista que resultará na diminuição da capacidade de armazenamento de água pelo reservatório, o que posteriormente fará com que haja uma redução da disponibilidade de água nos períodos de seca.

Figura 1 – Bacia do Alto Paraíba



3 DESENVOLVIMENTO

3.1 CLASSIFICAÇÃO E PROCESSAMENTO DAS IMAGENS

A Figura 2 descreve o fluxo de atividades necessárias para fazer a classificação das imagens de satélite. A plataforma do Google Earth Engine irá fornecer a capacidade computacional necessária, bem como as coleções de imagens do satélite Sentinel-2. Os valores das bandas do espectro eletromagnético que melhor respondem ao COT contido nas amostras de solo serão utilizados como parâmetro para o treinamento do classificador por meio de Machine Learning supervisionado, aplicando a técnica random forest e utilizando 10 árvores de decisão.

3.2 DIVISÃO DAS AMOSTRAS

As amostras de solo fornecidas foram divididas em conservado e degradado seguindo a classificação já feita pelos pesquisadores do INSA, assim como foi descrito anteriormente. A tabela de coordenadas foi responsável por indicar o local para que o satélite possa capturar as respostas das bandas das imagens daquele local e posteriormente, sejam utilizadas para o treinamento do classificador.

3.3 PREPARAÇÃO DAS IMAGENS DE SATÉLITE

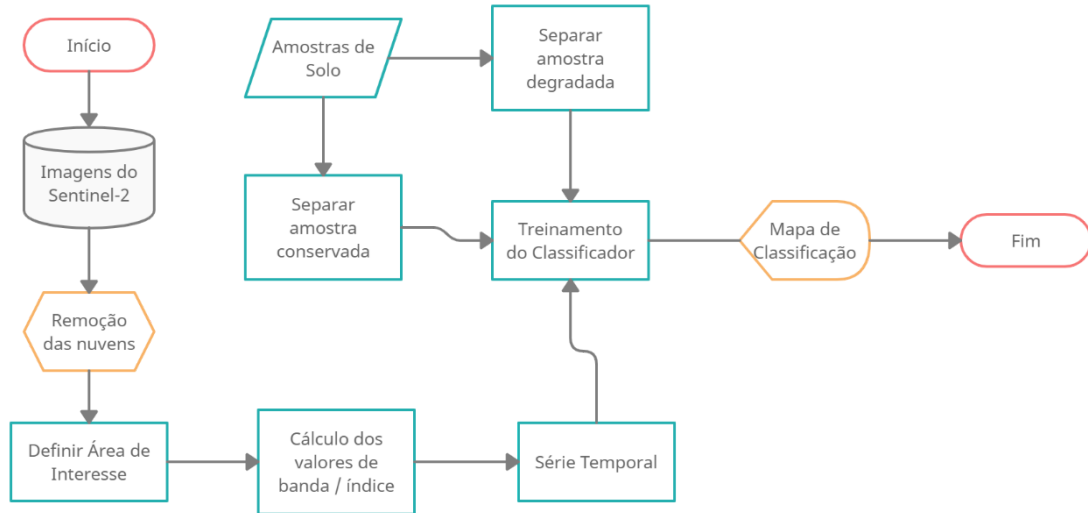
Foi definido uma janela temporal (período de interesse), logo após as imagens passaram por uma filtragem para retirada das nuvens e sombras de nuvens. Finalmente, utilizando as coordenadas das amostras, foram obtidas as respostas das bandas a partir dessas imagens.

3.4 CLASSIFICAÇÃO

A etapa de classificação consiste no cruzamento das informações obtidas das imagens indicadas pelas coordenadas, e as que estão sendo classificadas, o classificador é responsável por reconhecer os padrões de cada estado de conservação e atribuir o valor de degradado ou conservado para as imagens que estão sendo analisadas. O classificador terá como base os valores das bandas para reconhecer a assinatura espectral e determinar o estado de conservação do solo da

paraíba em conservado ou degradado, assim como mostra a Figura 2.

Figura 2 – Fluxograma do processo de classificação.



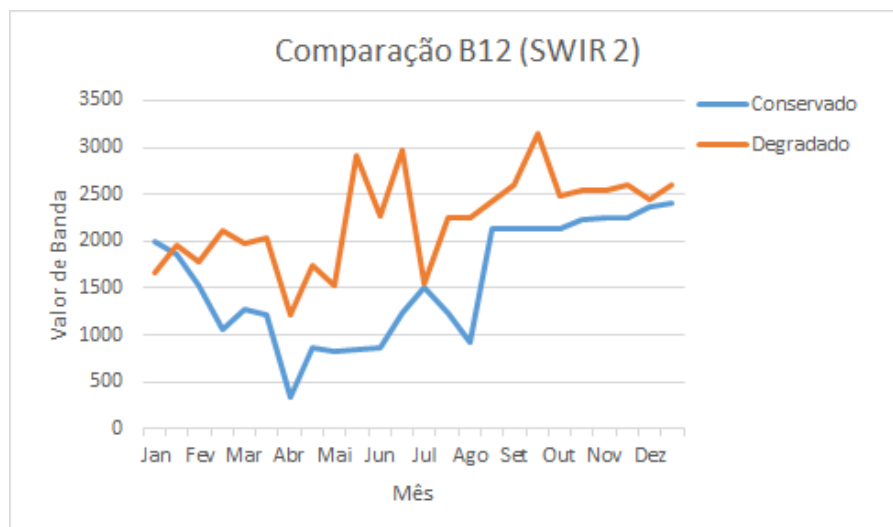
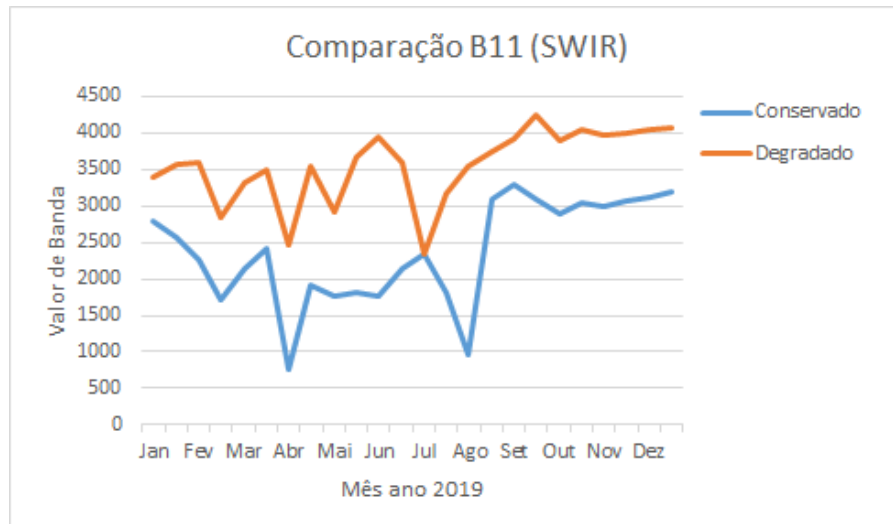
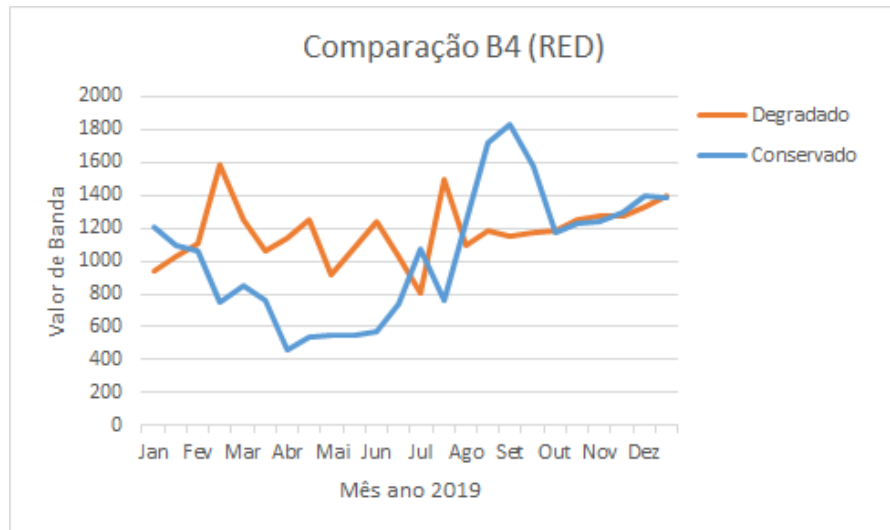
Fonte: Autor (2021)

4 RESULTADOS E DISCUSSÕES

4.1 DEFININDO BANDAS UTILIZADAS

Para definir qual banda deveria ser utilizada na classificação do COT, foram feitos gráficos de séries temporais comparando as respostas de cada banda para a região degradada e região conservada, assim como mostra a Figura 3. O objetivo era analisar qual delas tinham maior diferença entre respostas para conservado ou degradado.

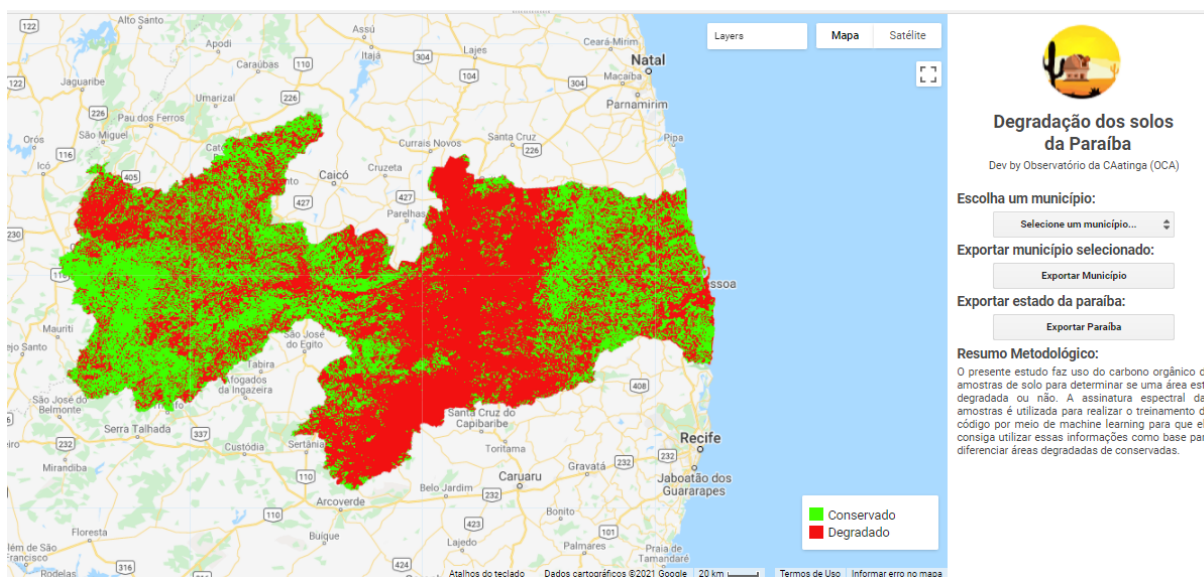
Figura 3 – Gráficos de serie temporal das bandas B4, B11, B12



Fonte: Autor (2021)

O desenvolvimento da UI surgiu como uma necessidade de facilitar as análises feitas a partir dessa classificação. Utilizando as ferramentas de desenvolvimento presentes no Google Earth Engine, foi desenvolvido uma simples interface que permitirá aos usuários realizar algumas operações, assim como mostra a Figura 5.

Figura 5 – Representação da UI desenvolvida durante o projeto.



Fonte: Autor (2021)

A partir da UI, o operador consegue exportar uma imagem da paraíba para realizar análises futuras em outros softwares, bem como filtrar a área de interesse por limites municipais, exportar a imagem gerada do município e ler um breve resumo metodológico.

5 CONCLUSÃO

O modelo computacional desenvolvido foi capaz de atingir o objetivo de indicar áreas degradadas no estado da Paraíba, como proposto pelo estudo. A possibilidade de fazer uma análise para qualquer área da Paraíba faz com que o modelo seja uma ótima ferramenta de trabalho para os profissionais do ramo da preservação ambiental, principalmente atividades relacionadas à conservação dos solos.

O desenvolvimento do sistema pode ser ampliado para outros estados do nordeste, a partir da ampliação das amostras de treinamento para outros estados do

nordeste, fazendo com que a classificação tenha maior confiabilidade. O uso de ferramentas para fazer medição de acurácia do classificador pode fazer com que o modelo se torne mais robusto para aplicações que exigem um maior grau de confiabilidade.

AGRADECIMIENTOS

O presente trabalho foi realizado com apoio do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil, ao PIBITI/CNQP-UFCG, Programa Institucional de Bolsas de Iniciação em Desenvolvimento Tecnológico e Inovação.

REFERÊNCIAS

- ALVES, J. J. A.; ARAÚJO, M. A. De; NASCIMENTO, S. S. Do; Queima Com Pastagem. *Revista Caatinga*, 2009. v. 22, n. 3, p. 126–135.
- NICHOLSON, S. E.; TUCKER, C. J.; BA, M. B. Desertification, Drought, and Surface Vegetation: An Example from the West African Sahel. *Bulletin of the American Meteorological Society*, 1998. v. 79, n. 5, p. 815–829.
- MCBRIDE, A. C. et al. Indicators to support environmental sustainability of bioenergy systems. *Ecological Indicators*, 2011. v. 11, n. 5, p. 1277–1289. Disponível em: <<http://dx.doi.org/10.1016/j.ecolind.2011.01.010>>.
- SORENSEN, P. T.; QUIDEAU, S. A.; RIVARD, B. High resolution measurement of soil organic carbon and total nitrogen with laboratory imaging spectroscopy. *Geoderma*, 2018. v. 315, n. August 2017, p. 170–177. Disponível em: <<https://doi.org/10.1016/j.geoderma.2017.11.032>>.
- TUCKER, Compton J; TOWNSHEND, J. R. G.; GOFF, T. E. African land-cover classification using satellite data. *Science*, 1985. v. 227, n. 4685, p. 369–375.
- TUCKER, Compton J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 1979. v. 8, n. 2, p. 127–150.
- SCHMIDT, H.; KARNIELI, A. Remote sensing of the seasonal variability of vegetation in a semi-arid environment. *Journal of Arid Environments*, 2000. v. 45, n. 1, p. 43–59.
- SÁNCHEZ-AZOFEIFA, G. A. et al. Research priorities for neotropical dry forests. *Biotropica*, 2005. v. 37, n. 4, p. 477–485.
- SAMPAIO, E. V. S. B.; SAMPAIO, Y.; VITAL, T.; ARAÚJO, M. S. B.; SAMPAIO, G. R. Desertificação no Brasil: conceitos, núcleos e tecnologias de recuperação e convivência. Recife: Ed. Universitária da UFPE, 2003.
- Reynolds J F, Stafford Smith M. 2002. Global desertification: do humans create deserts? In: Stanford-Smith M, Reynolds J F. *Do Humans Create Deserts?* Berlin: Dahlem University Press, 1–22
- PEREZ-MARIN, A. M.; MENDONÇA, A. De; CAVALCANTE, B. Núcleos de

desertificação no semiárido brasileiro: ocorrência natural ou antrópica? [s.d.].

SONG, X. P. et al. Annual detection of forest cover loss using time series satellite measurements of percent tree cover. *Remote Sensing*, 2014. v. 6, n. 9, p. 8878–8903.

MINISTÉRIO DO MEIO AMBIENTE (Brasil). **Em defesa da Caatinga**. [S. l.], 2013. Disponível em: <https://www.gov.br/mma/pt-br/noticias/em-defesa-da-caatinga>. Acesso em: 23 mar. 2021.

MOHRI, Mehryar; ROSTAMIZADEH, Afshin; TALWALKAR, Ameet. **Foundations of machine learning**. MIT press, 2018.

ALPAYDIN, Ethem. **Introduction to machine learning**. MIT press, 2020.

IBM (US). IBM Cloud Education. **What is Machine Learning?**. [S. l.], 2020. Disponível em: <https://www.ibm.com/cloud/learn/machine-learning>. Acesso em: 30 mar. 2021.