

ANALISE DE AGRUPAMENTOS DA DINÂMICA FLORESTAL DE UMA DECADA NO ESTADO DO ACRE

Wesley Augusto Campanharo (139203)

Instituto Nacional de Pesquisas Espaciais – INPE
Caixa Postal 515 – 12227-010 – São José dos Campos – SP, Brasil.

RESUMO

Vários são os fatores que levam a diminuição da cobertura florestal, sejam eles antrópicos ou naturais, o seu monitoramento leva a se identificar regiões que estão sofrendo algum processo de perda de cobertura florestal e por sua vez explorar os fatores que o acarretam. O produto oficial de monitoramento da cobertura florestal brasileira é o conjunto de dados do Prodes, Deter, Degrad e Detex. Entretanto produtos como o MapBiomas vem sendo utilizado no meio acadêmico. O método de mineração de dados denominado Mapa Auto Organizável (SOM), traz ótimos resultados numa análise exploratória para usuários com pouco conhecimento em análise de dados e estatística. Desta maneira o objetivo do presente trabalho é de identificar grupos da dinâmica florestal no período de 2000 a 2015 no estado do Acre utilizando, as coberturas florestais derivadas do Prodes e do MapBiomas empregando-se o SOM. O Mapbiomas estimou 2% a mais de florestas para o estado do Acre que o produto oficial brasileiro, mostrando um caráter mais inclusivo na definição de floresta. O SOM conseguiu evidenciar falhas e *outliers* presentes em ambos os produtos de cobertura florestal utilizados.

Palavras-chave: Uso e Cobertura; KSOM; MapBiomas; Prodes.

1 INTRODUÇÃO

As mudanças de uso e cobertura da terra ocorrem de forma acelerada e geram forte impactos nos recursos naturais, na biodiversidade e nas populações, sendo as alterações em formações florestais as com maior importância uma vez que estas possuem um potencial para fixação de carbono, para proteção do solo, da água e da biodiversidade maior dentro do sistema (ROSA, 2016).

De acordo com os Serviço Florestal Brasileiro (SFB, 2017), atualmente o monitoramento dos remanescente florestais para a Amazônia é realizado pelo conjunto de quatro sistemas operacionais: PRODES, DETER, DEGRAD e DETEX, em que estes sistemas são complementares. O PRODES, é o programa de monitoramento por satélite do desmatamento por corte raso, produzindo informações desde 1988 das taxas anuais de desmatamento na região (PRODES, 2017).

Entretanto há uma emergente utilização de outros produtos de cobertura do solo para a região Amazônica, como é o caso do MapBiomas. Este produto tem como objetivo contribuir com o desenvolvimento de uma metodologia rápida e confiável para processar conjuntos de dados em grande escala, gerando séries históricas de mapas de cobertura terrestre e uso do solo a baixo custo (SOUZA JUNIOR, 2017). A coleção é gerada por anos e obtida a partir da metodologia proposta pelo Imazon (Instituto do Homem e Meio Ambiente da Amazônia) sendo os algoritmos aplicados na plataforma Google Earth Engine consumindo os produtos da série Landsat (MAPBIOMAS, 2017).

Análise de agrupamento são técnicas que consiste em dividir os dados em grupos que podem ser significativos, úteis ou ambos. Estes grupos devem capturar a estrutura natural dos dados, podendo ser considerados como ponto de partida para outros fins, como a redução do volume de dados, compreensão de padrões e o aumento de eficiência em processamentos (TAN, STEINBACH e KUMAR, 2005). Estas análises desempenham papel importante em uma ampla variedade de áreas, como: psicologia, ciências sociais, biologia, estatística, reconhecimento de padrões, recuperação de informações, *machine learning* e mineração de dados (TAN, STEINBACH e KUMAR, 2005).

O Mapa Auto Organizável da sigla em inglês “*Self-Organizing Map – SOM*” é um método automático de análise de dados por meio de machine learning, onde segundo seu idealizador, Teuvo Kohonen (2013), foi introduzido entre os anos de 1981-82 através do projeto de mapeamento não linear onde se assemelhava aos preceitos da Quantização Vetorial (*Vector Quantization -VQ*) com a diferença de que os nós se tornaram espaciais e globalmente ordenados.

Desta forma, o objetivo do presente trabalho foi identificar grupos distintos na dinâmica florestal durante o período de 2000 a 2015 no estado do Acre utilizando os dois produtos de cobertura florestal para a região.

2 MATERIAIS E MÉTODOS

A área de estudo é o estado do Acre, com cerca de 164.123,74 km², localizado na região Norte do país. O limite estadual foi obtido no IBGE em formato *shapefile* por meio da plataforma VINDE.

A Figura 1 apresenta o fluxograma resumido das etapas realizadas durante o trabalho dentro de cada programa utilizado para se atingir o objetivo.

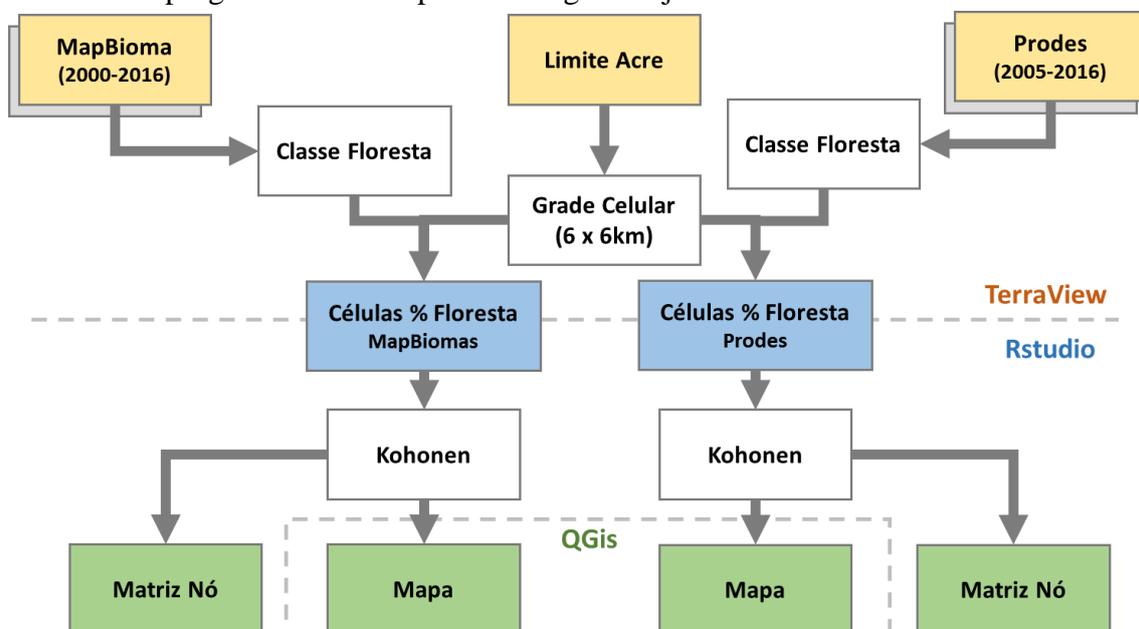


Figura 1. Fluxograma dos processos utilizados no trabalho pra se atingir o objetivo.

Os dados de cobertura do solo são oriundos do conjunto 2 do Projeto MapBioma para a Amazônia dos anos de 2000 a 2016 em formato Geotiff com pixel de 30 x 30 metros; e do

conjunto de mapas do PRODES para a Amazônia Legal em formato Geotiff com pixel de 90 x 90 metros e série temporal de 2005 a 2016.

Para se obter a porcentagem florestal, inicialmente foi necessário realizar a filtragem das classes classificadas como Floresta para ambos os produtos, em que para o MapBiomias seguiu-se os valores da legenda de nível 2 dita como “Formações Florestais Naturais”, enquanto para o Prodes seguiu a indicação para a classe Floresta.

Posteriormente gerou-se uma grade celular de aproximadamente 6 km por 6 km nas dimensões do estado do Acre, obtendo-se assim uma malha retangular de 10125 células (75 linhas e 135 colunas), porem somente 4868 estavam dentro do território do Acre. A resolução da grade celular foi baseada em números múltiplos da resolução espacial do arquivo matricial de cobertura, sendo construída utilizando o software TerraView 5.2. Utilizando este software realizou-se o preenchimento das células com as informações de porcentagem das classes de cobertura tanto para o conjunto do MapBiomias quanto para o Prodes.

O algoritmo utilizado para execução do método Mapa Auto Organizável (SOM) foram os da biblioteca *kohonen* para a plataforma R, construído por Wehrens e Kruisselbrink (2017). O SOM representa uma distribuição de itens de dados usando um conjunto finito de modelos, os quais são automaticamente associados aos nós de uma grade regular de forma ordenada, de modo que mais modelos similares se associem automaticamente a nós que são adjacentes na grade, enquanto que menos modelos similares são situados mais distantes um dos outros na grade. Esta organização, um tipo de diagrama de similaridade dos modelos, permite obter uma visão das relações topográficas dos dados, especialmente de itens com alta dimensão (KOHONEN, 2013).

3 RESULTADOS E DISCUSSÃO

De acordo com os dados do MapBiomias, a cobertura florestal no estado do Acre está em declínio chegando, atualmente, em aproximadamente 88% do estado, cerca de 144.182 km² (Figura 2 e Figura 3). Comparando com os dados do Prodes, o qual realiza indiretamente a quantificação de florestas remanescentes, para o ano de 2016 há um total de 86% de remanescente florestal, sendo valores bem próximos para as diferentes bases. A variação mais marcante entre as bases, em um contexto abrangente, é na taxa de desflorestamento, no qual a declividade da reta ajustada pelos dados do MapBiomias é superior a reta ajustada ao Prodes.

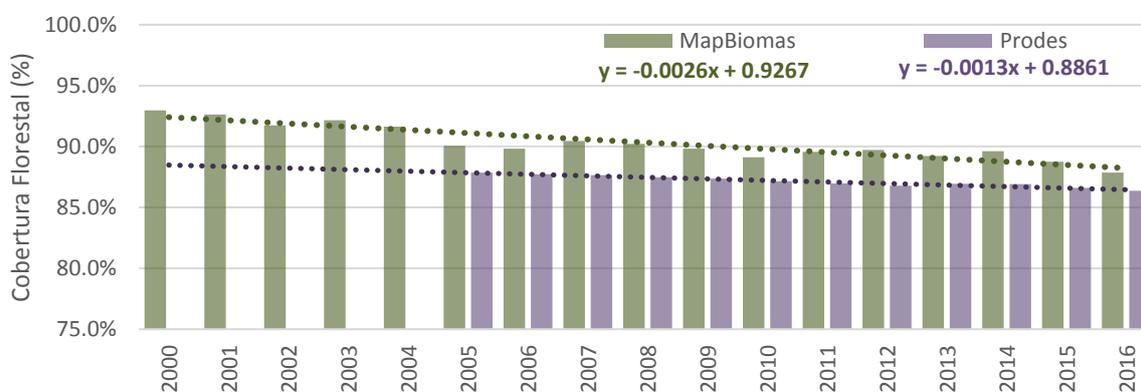


Figura 2. Histórico da cobertura florestal anual mapeada pelo MapBiomias e Prodes para o estado do Acre.

Cabe ressaltar que as metodologias dos dois produtos são diferente, o Prodes realiza a delimitação dos fragmentos de forma manual e utilizam como guia uma máscara de florestas dos anos anteriores, fazendo com que não haja regeneração vegetal que se enquadre novamente na classe florestas, enquanto o MapBiomas realiza a classificação das imagens do conjunto Landsat pixel a pixel, fazendo com que a classe floresta seja única para cada ano, possibilitando, assim, a reentrada de outras formações florestais na classe denominada Florestas em anos seguintes.

Rosa (2016) comparando o produto do MapBiomas com o produto do SOS Mata Atlântica/INPE (que apresenta metodologia semelhante ao Prodes), aponta que para o MapBiomas o conceito de floresta é mais amplo tendo como resultado um mapeamento mais inclusivo, obtendo-se valores maiores da cobertura florestal em toda a extensão do bioma. Desta forma o mesmo autor sugere que o produto seja utilizado para identificar áreas para plantio, recomposição de florestas nativas, criação de corredores locais, análise de fragmentos e análises da situação ambiental de bacias hidrográficas.

Outro fator marcante nos dados do Prodes, são as classes de não observação da cobertura devido a presença de nuvens ou resíduos de períodos anteriores, estas classes podem incluir um desvio na quantificação de Floresta, e quando agregadas em uma grade celular adicionam possíveis áreas que não estão desflorestadas ao percentual florestal.

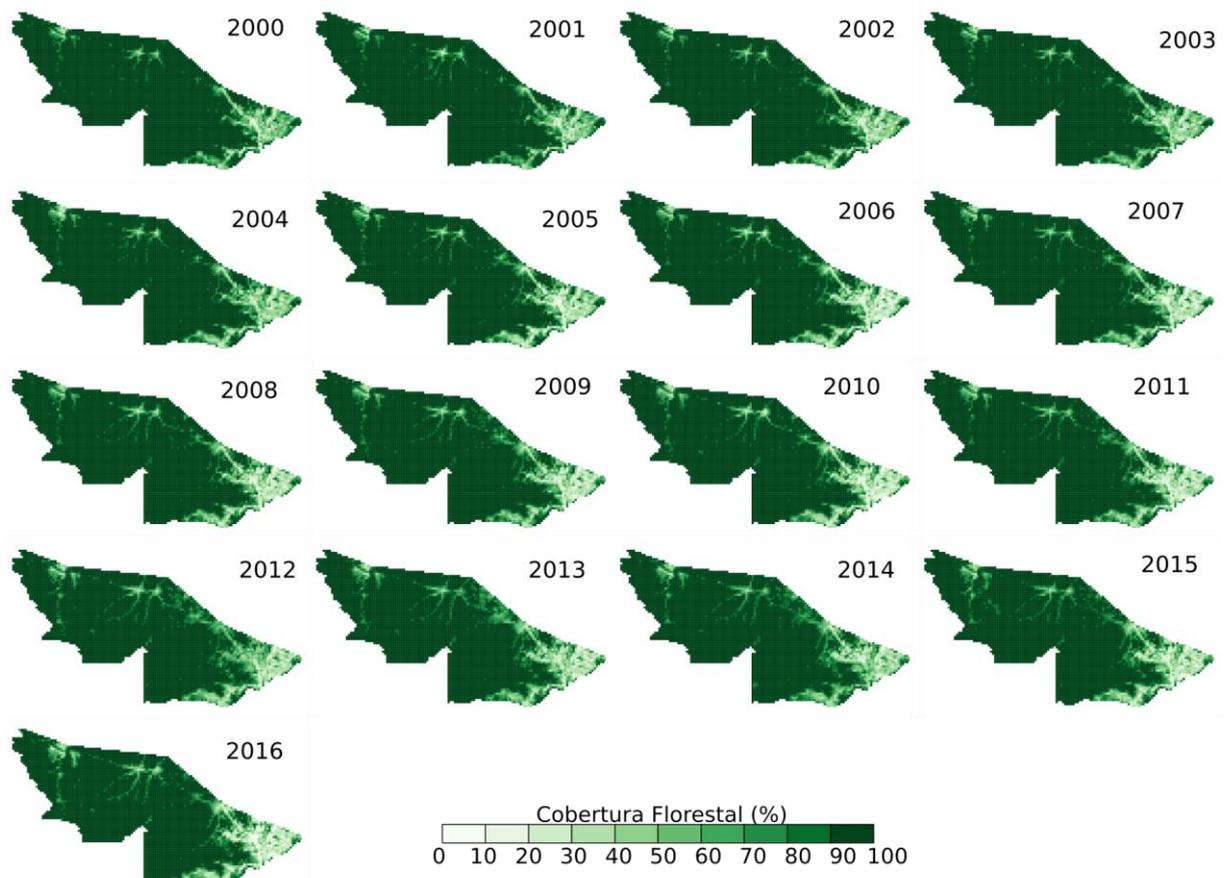


Figura 3. Mapas de porcentagem de cobertura florestal com base no MapBiomas para o período analisado para o estado do Acre.

Como verificado na Figura 4, a área Oeste do estado é a que está mais prejudicada com a presença de nuvens e resíduos provenientes do produto Prodes. Estes artefatos em uma

abordagem geral são facilmente contornados, entretanto quando se necessita de uma avaliação espacial eles influenciam tanto na própria visualização do dado, como pode ser observado no conjunto de mapas da evolução da cobertura florestal no Acre (Figura 5) quanto na análise de agrupamentos que será abordado nos próximos tópicos.

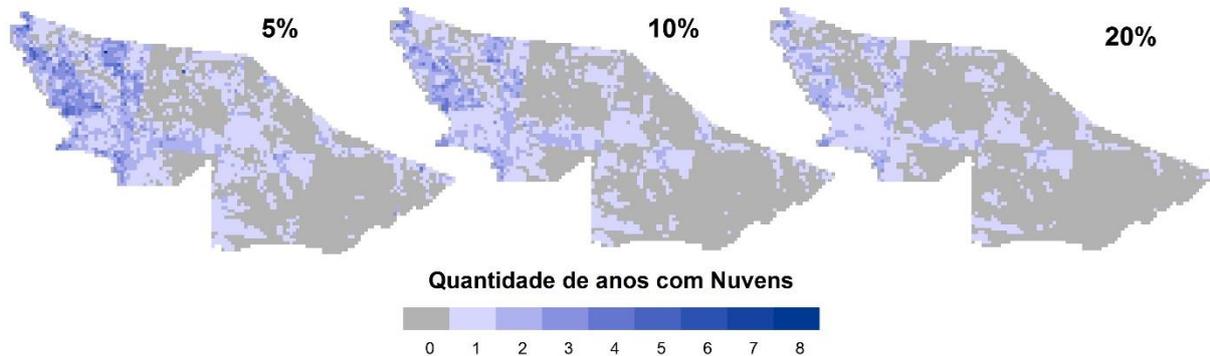


Figura 4. Distribuição das células com 5,10 e 20% de cobertura de nuvem pelo PRODES e sua frequência durante o período analisado.

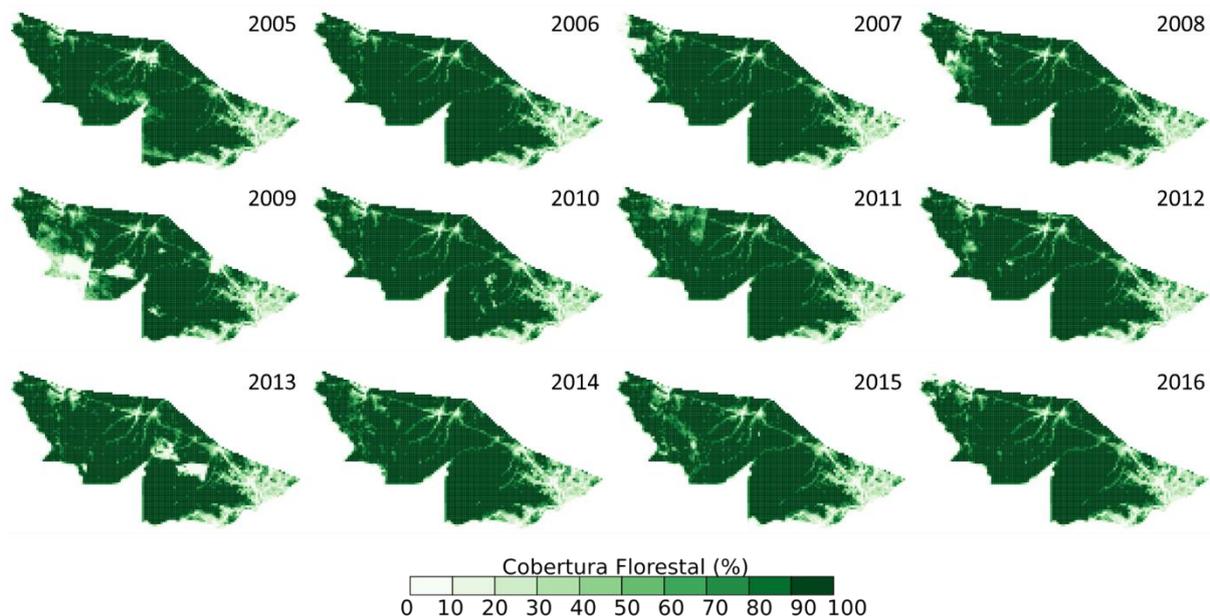


Figura 5. Porcentagem de cobertura florestal oriunda do PRODES para o período de 2005 a 2016.

De acordo com Van Hulle (2012) a aplicação mais utilizada do SOM é em relação à sua matriz de nós (também denominada matriz de *codebooks*, ou mapa topográfico), o qual é utilizada na análise de agrupamento, uma vez que o método realiza o particionamento do conjunto completo dos dados em subconjuntos semelhantes sem utilizar um conhecimento prévio do número de grupos existentes ou da sua natureza, possibilitando a interpretação mais rápida dos dados por usuários sem exigir conhecimento a fundo sobre análise de dados e estatística.

Para os dados provenientes do MapBiomas, a Figura 6, mostra dois gráficos com informações do mapa topográfico, um no qual evidencia a distância média entre um nó e seus vizinhos, e o outro a quantidade de células presente em cada nó. Já a Figura 7, é uma variação do mapa topográfico o qual são incluídos as séries temporais pertencentes a cada nó e a quantidade de elementos presentes. Com isso, é possível notar que existe poucos grupos de

células distintas, devido à baixa variação da distância entre os vizinhos. O padrão com a maior predominância é o de alta cobertura e constância florestal durante todo o período. Outros grupos evidenciados, são dos extremos da matriz de nós, onde a presença de *outliers* é nítida, devido ao alto contraste entre a distância com os vizinhos, a baixa quantidade de células presente no nó e o seu padrão florestal desconexo.

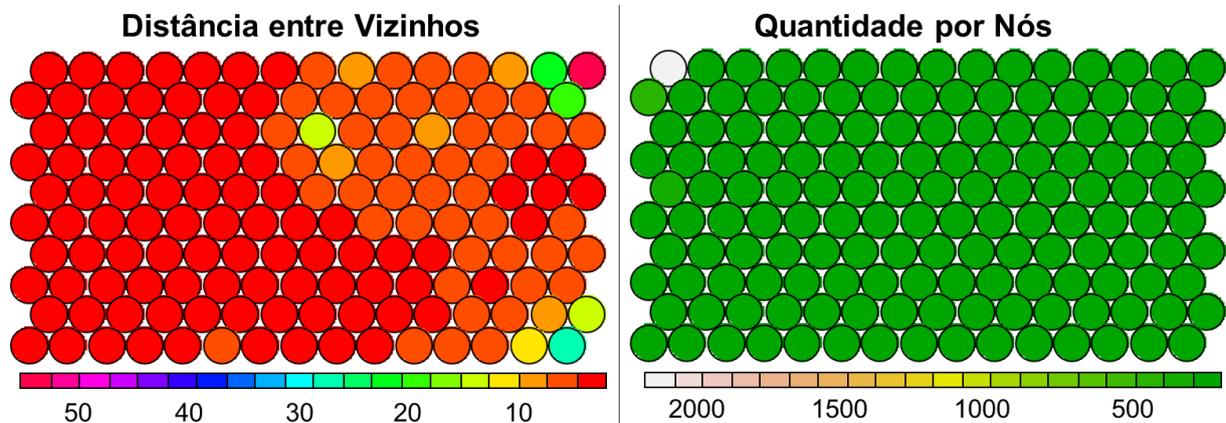


Figura 6. Distância média entre os nós vizinhos e quantidade de eventos por nó para os dados do MapBiomias.

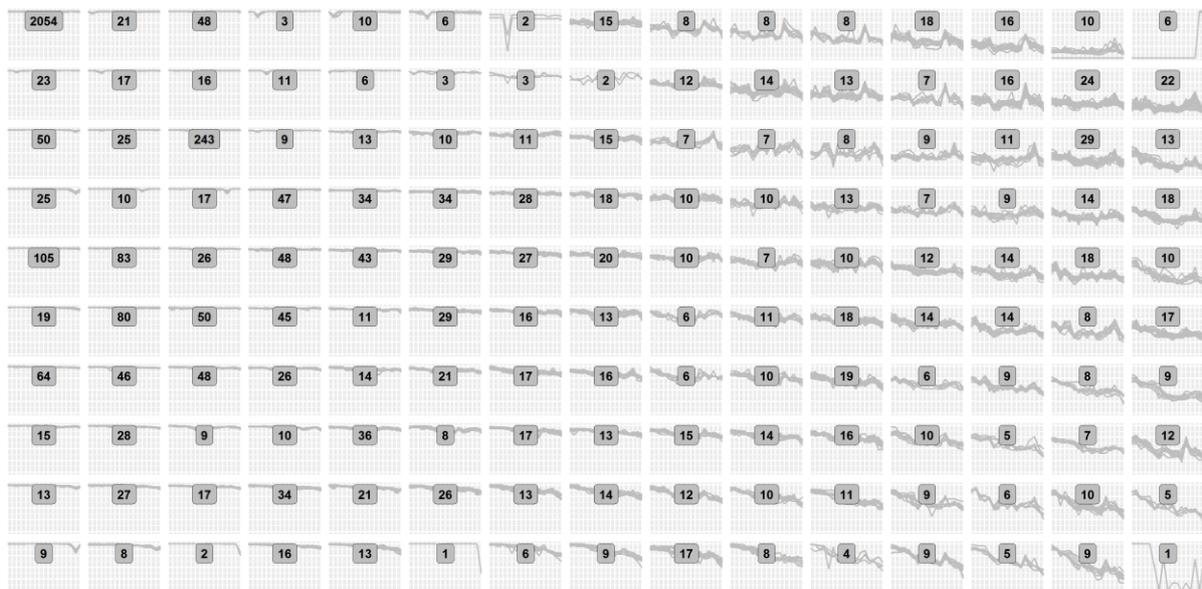


Figura 7. Mapa topográfico para os dados do MapBiomias obtido pelo SOM, com detalhe para as séries temporais e a quantidade de células que cada nó possui.

Analisando o comportamento da matriz de nós (Figura 8) para os dados do Prodes, observa-se uma maior variação na distância entre vizinhos, mostrando um possível agrupamento de classes mesmo que tenham quantidades de indivíduos relativamente próximos. Diferentemente da matriz de nós encontrados para o conjunto de dados do MapBiomias o do Prodes apresentou nós mortos, ou seja, nós que não tiveram nenhuma célula com grau de similaridade significativa, além disso a distribuição de células por nós foi muito mais homogênea que em comparação ao MapBiomias, sendo constatado pela escala do gráfico de quantidade. Ao se incluir as séries temporais pertencentes a cada nó (Figura 9) se torna evidente

que a presença dos artefatos existentes nos dados do Prodes criando vazios dentro da série temporal provocou dois grupos de nós com distâncias muito discrepantes nos dados.

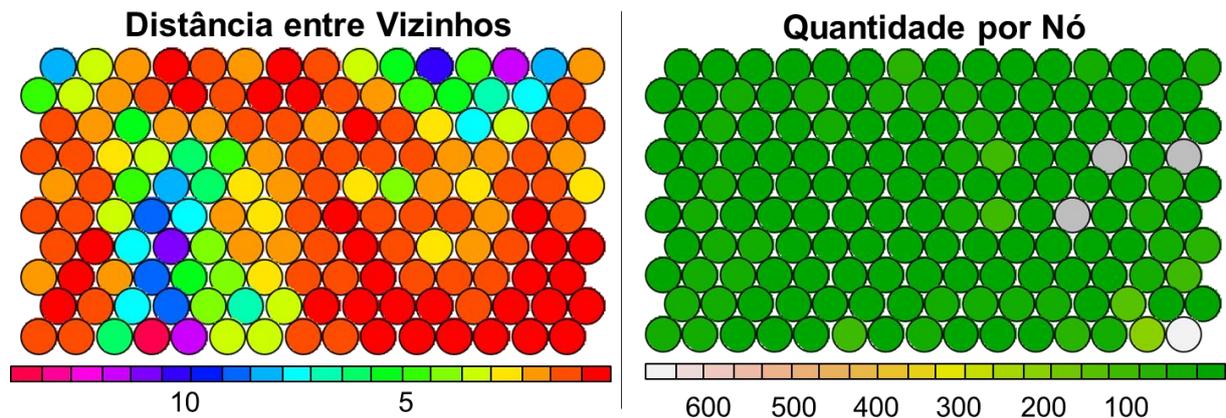


Figura 8. Distância média entre os nós vizinhos e quantidade de eventos por nó para os dados do Prodes.

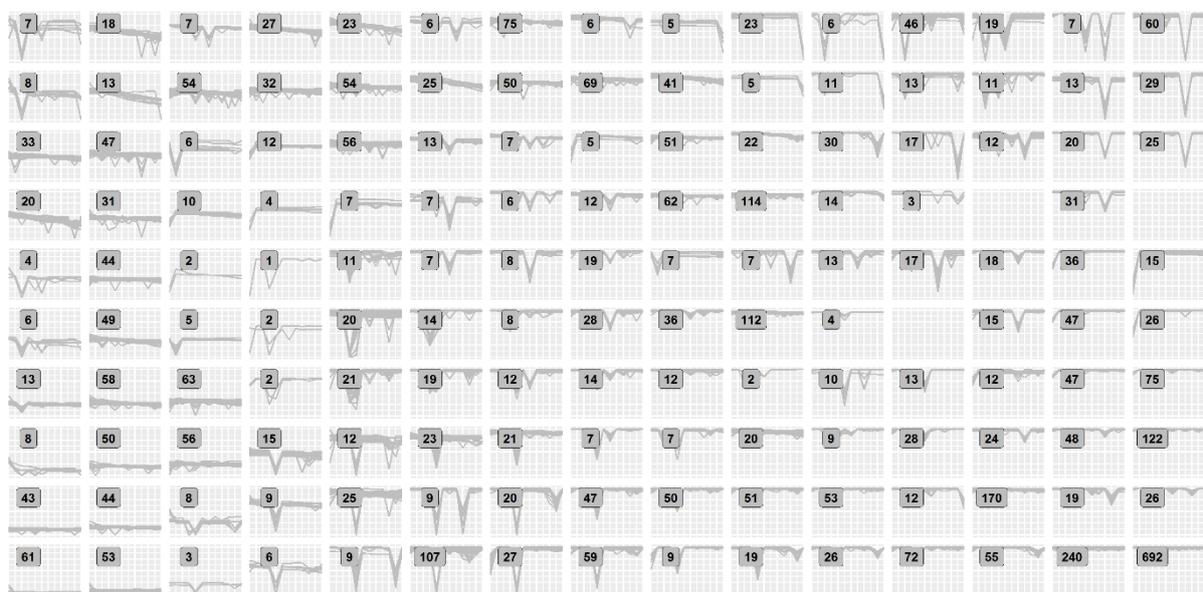


Figura 9. Mapa topográfico para os dados do Prodes obtido pelo SOM, com detalhe para as séries temporais e a quantidade de células que cada nó possui.

O SOM projeta os pontos de dados de alta dimensão para um espaço de dimensões mais baixas (normalmente uma grade bidimensional) de maneira que objetos vizinhos em espaço de alta dimensão permaneçam vizinhos em espaço de baixa dimensão (MOOSAVI, 2017). A fim de se verificar a continuidade destes padrões ao se retornar as classes para o espaço original, realizou-se a espacialização das células com os valores de cada nó, podendo ser observados na Figura 10. Onde nota-se que existe padrões semelhantes que acompanham a malha hídrica e de padrões com grande variações de classes dentro dos grandes centros urbanos.

No agrupamento com o produto do Prodes, mais uma vez verificasse os padrões de influencias de nuvens principalmente na região central e extremo oeste do Estado. Assim como para o MapBiomias encontrou-se grupos ao longo de estradas e rios divergindo do seu entorno.

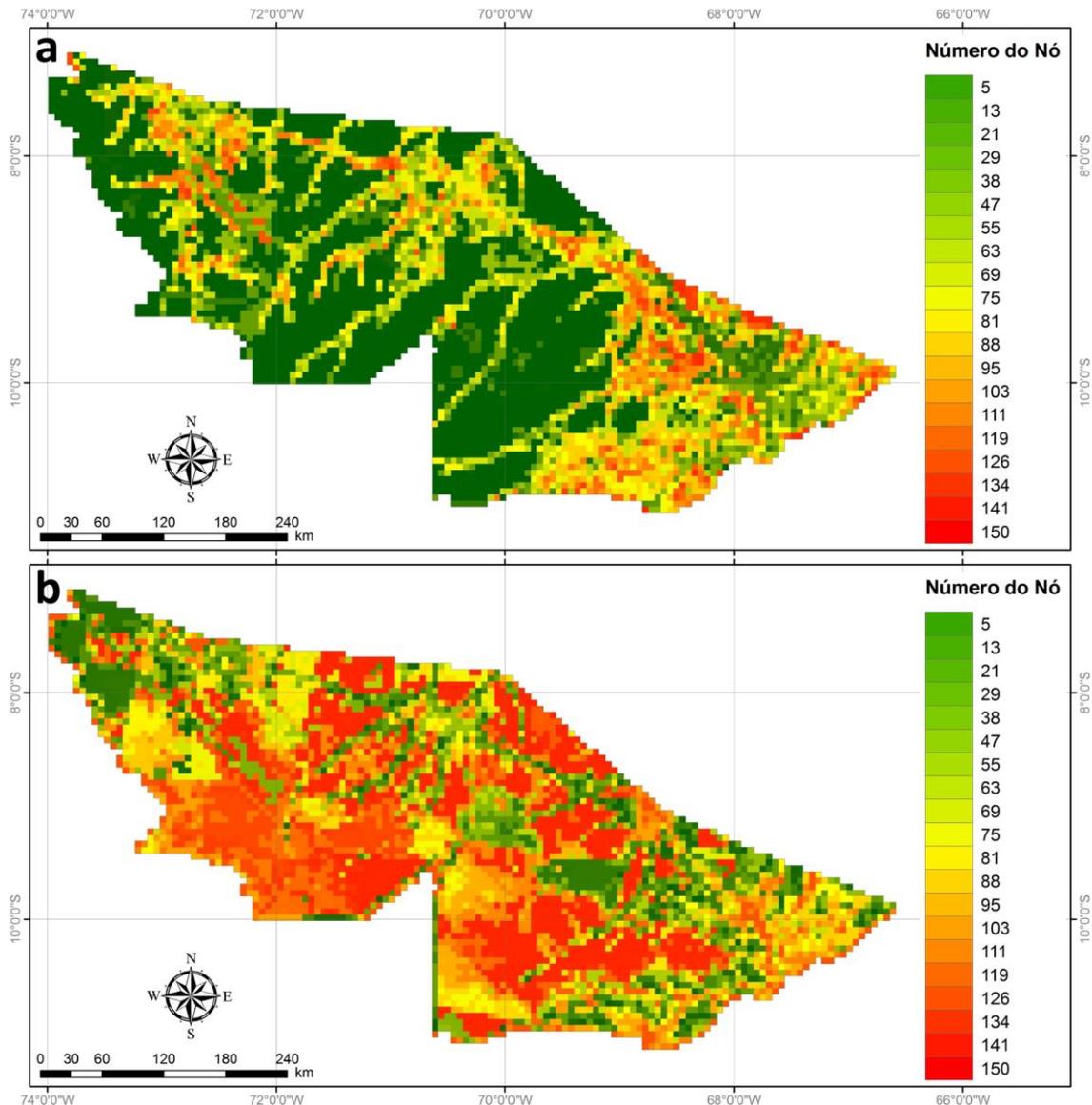


Figura 10. Espacialização das células do produto do MapBiomias (a) e do Prodes (b) de acordo com sua classe de nó obtido pelo SOM.

4 CONCLUSÃO

Existe uma diferença de 2% na cobertura florestal entre o MapBiomias e o Prodes para o ano de 2016 para o estado do Acre, evidenciando o caráter mais inclusivo deste produto que está sendo mais comumente utilizado pela comunidade científica .

Nuvens e artefatos podem se configurar como um empecilho durante a análise espacial da série temporal, uma vez que os métodos de agrupamento utilizados conseguiu identifica-los em grupos claros tanto na matriz de nós quanto na espacialização dos dados.

O padrão da dinâmica florestal com maior concentração no Acre é o de alta cobertura florestal e constância durante todo o período analisado, seja pelo produto do MapBiomias quanto pelo do Prodes. Entretanto, não se encontrou grupos com dinâmica de regeneração, obtendo-se apenas decréscimo da cobertura florestal em diferentes níveis, principalmente em torno dos grandes centros e rodovias.

O uso do SOM como método de agrupamento ,neste caso específico, mostrou-se de grande utilidade e dinamismo, principalmente numa análise exploratória e visual de possíveis grupos de dinâmica florestal, evidenciando falhas e *outliers* presentes nos produtos utilizados.

5 REFERÊNCIAS

KOHONEN, T. Essential of the self-organizing map. **Neural Networks**, n. 37, p. 52-65, 2013.

MAPBIOMAS. **ATBD_R - Algorithm Theoretical Basis Document & Results**: Amazon Biome. [S.l.]: [s.n.], 2017. Disponível em: <<https://s3.amazonaws.com/mapbiomas-ecostage/Base+de+dados/Metodologia/ATBD-R+++Bioma+Amaz%C3%B4nia+++2017-03-16.pdf>>. Acesso em: julho 2017.

MOOSAVI, V. Contextial mapping: Visualization of high-dimensional spatial patterns in a single geo-map. **Computers, Environment and Urban Systems**, v. 61, p. 1-12, 2017.

PRODES. Monitoramento da Floresta Amazônica Brasileira por Satélite. **Observação da Terra** - **INPE**, 2017. Disponível em: <<http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>>. Acesso em: dez. 2017.

ROSA, M. R. Comparação e análise de diferentes metodologias de mapeamento da cobertura florestal da Mata Atlântica. **Boletim Paulista de Geografia**, v. 95, p. 25-34, 2016.

SFB. Perda da Cobertura Florestal. **Sistema Nacional de Informações Florestais**, 2017. Disponível em: <http://www.florestal.gov.br/snif/recursos-florestais/index.php?option=com_k2&view=item&layout=item&catid=14&id=166>. Acesso em: dez. 2017.

SOUZA JUNIOR, C. **MapBiomias General "Handbook"**. 1. ed. [S.l.]: [s.n.], 2017. Disponível em: <<https://s3.amazonaws.com/mapbiomas-ecostage/Base+de+dados/Metodologia/ATBD+MapBiomias+Geral+2017-03-22.pdf>>. Acesso em: julho 2017.

TAN, P.-N.; STEINBACH, M.; KUMAR, V. Cluster Analysis: Basic Concepts and Algorithms. In: TAN, P.-N.; STEINBACH, M.; KUMAR, V. **Introduction to Data Mining**. 1ª. ed. [S.l.]: Pearson Addison-Wesley, 2005. Cap. 8, p. 487-568.

VAN HULLE, M. M. Self-organizing Maps. In: ROZENBERG, G.; BÄCK, T.; KOK, J. N. **Handbook of Natural Computing**. Berlin, Heidelberg: Springer, 2012. Cap. 19, p. 585-622.

WEHRENS, R.; KRUISSELBRINK, J. Supervised and Unsupervised Self-Organising Maps. **CRAN R-project**, 2017. Disponível em: <<https://cran.r-project.org/web/packages/kohonen/kohonen.pdf>>.