

Igor da Silva Freitas de Souza

Evolução temporal da área desmatada e de indicadores do desmatamento: um olhar sobre as regiões norte e centro-oeste do Brasil.

Niterói - RJ, Brasil

04 de Fevereiro de 2022

Igor da Silva Freitas de Souza

**Evolução temporal da área
desmatada e de indicadores do
desmatamento: um olhar sobre as
regiões norte e centro-oeste do
Brasil.**

Trabalho de Conclusão de Curso

Monografia apresentada para obtenção do grau de Bacharel em
Estatística pela Universidade Federal Fluminense.

Orientador: Prof. Ana Beatriz Monteiro Fonseca

Niterói - RJ, Brasil

04 de Fevereiro de 2022

Igor da Silva Freitas de Souza

Evolução temporal da área desmatada e de indicadores do desmatamento: um olhar sobre as regiões norte e centro-oeste do Brasil.

Monografia de Projeto Final de Graduação sob o título “*Evolução temporal da área desmatada e de indicadores do desmatamento: um olhar sobre as regiões norte e centro-oeste do Brasil.*”, defendida por Igor da Silva Freitas de Souza e aprovada em 04 de Fevereiro de 2022, na cidade de Niterói, no Estado do Rio de Janeiro, pela banca examinadora constituída pelos professores:

Profa. Dra. Ana Beatriz Monteiro Fonseca
Departamento de Estatística – UFF

Profa. Dra. Ludmilla da Silva Viana Jacobson
Departamento de Estatística – UFF

Profa. Dra. Márcia Marques de Carvalho
Departamento de Estatística – UFF

Ficha catalográfica automática - SDC/BIME
Gerada com informações fornecidas pelo autor

S719e Souza, Igor da Silva Freitas de
Evolução temporal da área desmatada e de indicadores do
desmatamento: um olhar sobre as regiões norte e centro-oeste
do Brasil / Igor da Silva Freitas de Souza ; Ana Beatriz
Monteiro Fonseca, orientadora. Niterói, 2022.
67 f. : il.

Trabalho de Conclusão de Curso (Graduação em
Estatística)-Universidade Federal Fluminense, Instituto de
Matemática e Estatística, Niterói, 2022.

1. Série temporal. 2. Box-Jenkins. 3. Função de
transferência. 4. Desmatamento. 5. Produção intelectual. I.
Monteiro Fonseca, Ana Beatriz, orientadora. II. Universidade
Federal Fluminense. Instituto de Matemática e Estatística.
III. Título.

CDD -

Resumo

Ao longo dos últimos anos, muito tem-se falado sobre a existência de relação entre o aumento na área desmatada, especialmente da região amazônica, com o aumento da ocupação da mesma região por atividades agropecuárias e, também, pelo aumento sucessivo da quantidade de queimadas realizadas. O presente trabalho se propõe a avaliar a série temporal do desmatamento e a correlação temporal dela com outras séries indicadoras, como a área de pastagem, área reservada à agricultura, extração de madeira, efetivo bovino e quantidade de focos de queimadas. A série usada de base para avaliar o desmatamento é representada pela área de floresta nativa, medida anualmente e avaliada no período de 1985 a 2019, possuindo um total de 35 observações. Foram construídos modelos ARIMA e funções de transferência, no software R, para gerar previsões anuais para a área desmatada para os anos de 2020 a 2030. Foram realizadas previsões para dois blocos de regiões: o primeiro, composto pela união das Regiões Norte e Centro-Oeste, e o outro, pela agregação dos dados dos estados do Pará e Mato Grosso. Na função de transferência para as Regiões Norte e Centro-Oeste, apenas a série de pastagem foi identificada como temporalmente correlacionada à área de floresta nativa, mas seu acréscimo na qualidade de ajuste e de previsão foi marginal. O melhor modelo, neste caso, estimou uma perda de 1.219.386 hectares de área da floresta nativa por ano, entre 2020 e 2030. Porém, para os estados do Pará e Mato Grosso, duas séries se mostraram temporalmente correlacionadas com a área natural preservada, pastagem e quantidade de focos de queimadas, e a função de transferência melhorou em 88% a qualidade das previsões, estimando uma média de 697.700 hectares desmatados por ano, no mesmo período.

Palavras-chave: Modelagem. Série temporal. Box-Jenkins. Função de Transferência. Previsões. Desmatamento. Agropecuária. Queimadas.

Dedicatória

Dedico este trabalho aos meus pais, que sempre me apoiaram e me deram todo o suporte necessário.

Agradecimentos

Primeiramente agradeço aos meus pais por todo o suporte para que eu chegasse aqui.

Agradeço aos meus familiares pelo suporte.

Agradeço à matemática e à ciência.

Agradeço também à minha grande orientadora Ana Beatriz.

Por fim agradeço aos meus amigos que me acompanharam nessa jornada.

Sumário

Lista de Figuras

Lista de Tabelas

1	Introdução	p. 12
1.1	Situação do desmatamento no mundo	p. 14
1.1.1	Desmatamento nas savanas e florestas tropicais	p. 16
1.2	Situação do desmatamento no Brasil	p. 17
1.2.1	O desmatamento no Norte e Centro-Oeste do Brasil	p. 18
1.3	Relação entre o desmatamento e os seus indicadores	p. 19
1.4	Objetivos	p. 21
2	Materiais e Métodos	p. 22
2.1	Bases de dados	p. 22
2.1.1	Desmatamento, agricultura e pastagem	p. 22
2.1.2	Extração de madeira	p. 23
2.1.3	Efetivo bovino	p. 24
2.1.4	Queimadas	p. 24
2.2	Modelagem de Séries Temporais	p. 24
2.2.1	Conceitos básicos	p. 25
2.2.2	Função de Autocorrelação e Autocorrelação Parcial	p. 26
2.2.3	Modelo Arima	p. 28
2.2.3.1	Modelos Box-Jenkins	p. 29

2.2.4	Critérios de Informação e Avaliação de Qualidade de Ajuste . . .	p. 32
2.2.5	Previsão	p. 33
2.2.6	Intervalo de confiança	p. 34
2.2.7	Correlação cruzada	p. 35
2.2.8	Função de Transferência	p. 36
3	Resultados	p. 38
3.1	Análises e resultados para as Regiões Norte e Centro-Oeste	p. 38
3.2	Análises e resultados para os estados de Pará e Mato Grosso	p. 48
4	Conclusões	p. 56
	Referências	p. 58
	Apêndice 1 – Comparação de modelos para as Regiões	p. 62
	Apêndice 2 – Comparações extras dos modelos	p. 64
	Apêndice 3 – Comparação de modelos para os estados	p. 65

Lista de Figuras

1	Mapa com os tipos de florestas e suas localizações no globo. Fonte: Food and Agriculture Organization (2020)	p. 12
2	Mapa das localizações das savanas no globo.	p. 13
3	Mapa dos Biomas brasileiros. Fonte: (IBGE, 2019)	p. 14
4	Estimativa anual do desmatamento e expansão florestal entre os anos de 1990 e 2020. Fonte: FAO	p. 15
5	Média anual da quantidade de hectares desmatados entre os anos de 2015 e 2020. Fonte: Ritchie (2020)	p. 15
6	Média da quantidade de hectares desmatados na América do Sul entre os anos de 1990 e 2000. Fonte: Ritchie (2020)	p. 17
7	Incremento do desmatamento na Amazônia de 2008 a 2020. Fonte: Terrabrasilis (2021)	p. 18
8	Incremento do desmatamento no Cerrado de 2002 a 2020. Fonte: Terrabrasilis (2021)	p. 18
9	Diagramas de dispersão para lags $k=1,2$. Fonte: (BOX et al., 2016)	p. 27
10	Matriz de autocorrelação e função de autocorrelação correspondente de um processo estacionário. Fonte: (BOX et al., 2016)	p. 27
11	Ciclo iterativo da metodologia Box-Jenkins para geração de previsões	p. 29
12	Séries temporais em estudo e suas tendências	p. 39
13	Séries temporais diferenciadas 1 vez	p. 40
14	Séries temporais diferenciadas 2 vezes	p. 41
15	Valores efetivos e ajustados das séries temporais	p. 43
16	Previsões das séries temporais	p. 45
17	Função de correlação cruzada para cada uma das séries	p. 46

18	Séries temporais para os estados do Pará e Mato Grosso e suas tendências	p. 48
19	Valores efetivos e ajustados das séries temporais para os estados	p. 50
20	Previsões das séries temporais para os estados	p. 51
21	Funções de correlação cruzada	p. 52
22	Comparando previsões do ARIMA puro e da FT3	p. 54
23	Previsões geradas pelo modelo FT3	p. 54

Lista de Tabelas

1	Comportamento da f_{ac} e f_{acp} nos modelos ARIMA(p,d,q). Fonte (MORETTIN; TOLOI, 2006)	p. 30
2	p-valores do teste aumentado de Dickey-Fuller para todas as séries temporais em estudo	p. 42
3	Melhores modelos baseados nos critérios de informação para combinações de p e $q \leq 2$ e métricas de qualidade de ajuste	p. 42
4	Métricas de qualidade de ajuste dos modelos	p. 43
5	Teste Ljung-Box para resíduos dos modelos ARIMA	p. 44
6	Teste Shapiro-Wilk para resíduos dos modelos ARIMA	p. 44
7	Comparação dos modelos	p. 47
8	Previsões para a área de floresta nativa para as Regiões	p. 47
9	Representatividade das séries dos estados do Pará e Mato Grosso em relação ao total das Regiões Norte e Centro-Oeste	p. 49
10	Melhores modelos, no ambiente espacial reduzido, baseados nos critérios de informação para combinações de p e $q \leq 2$ e métricas de qualidade de ajuste	p. 49
11	Teste Ljung-Box para resíduos dos modelos ARIMA	p. 50
12	Teste Shapiro-Wilk para resíduos dos modelos ARIMA	p. 51
13	Comparação dos modelos	p. 53
14	Previsões para a área de floresta nativa nos estados do Pará e Mato Grosso	p. 55
15	Comparação de modelos baseados nos critérios de informação	p. 62
16	Comparação de métricas extras para classificação de melhor modelo	p. 64
17	Comparação de modelos baseados nos critérios de informação	p. 65

1 Introdução

Neste capítulo será abordada a temática do desmatamento, com enfoque nas florestas tropicais e savanas, aprofundando a situação no mundo e no Brasil, em especial, nas regiões Norte e Centro-Oeste. Também serão apresentados os objetivos e a forma de organização do trabalho.

Segundo a Food and Agriculture Organization (FAO), o planeta possui 4,06 bilhões de hectares de florestas, o que corresponde a 31% da superfície de terra, sendo a floresta tropical a de maior presença (Food and Agriculture Organization, 2020). O clima das florestas tropicais é quente, úmido e estável, possui temperatura média entre 23 e 28 °C, os períodos de chuvas variam conforme a localidade, com média mensal de precipitação de 60mm e meses de secas são comuns, apresentando alta taxa de evaporação. A vegetação é predominantemente de árvores de grande e médio porte, e solo coberto por plantas herbáceas e samambaias, essas florestas, também, apresentam imensa diversidade de espécies de plantas e animais (GALLERY, 2014).

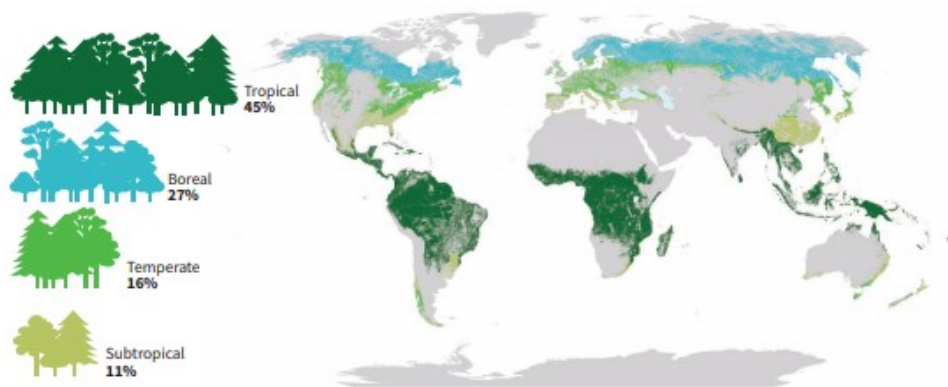


Figura 1: Mapa com os tipos de florestas e suas localizações no globo. Fonte: Food and Agriculture Organization (2020)

As savanas são o tipo mais comum de vegetação nos trópicos, o ecossistema é amplamente definido por formações contínuas de campos gramados e descontínuas de árvores e arbustos. O clima é tropical com chuvas sazonais, possuindo longos períodos de secas

que podem durar nove meses. Devido a esse fato o solo, comumente, é pobre em nutrientes e a qualidade das pastagens naturais são baixas, também possuindo baixo índice de nutrientes. Cerca de 20% da população mundial vivem em regiões que são, ou recentemente foram, cobertas por savanas, a maioria em sociedades rurais que dependem da agropecuária para a sobrevivência (SOLBRIG; MEDINA; SILVA, 1996).

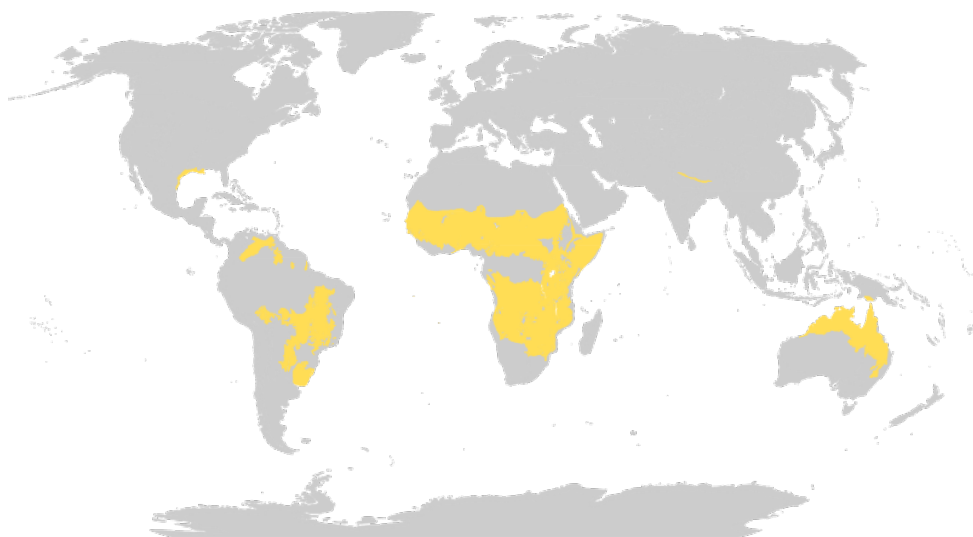


Figura 2: Mapa das localizações das savanas no globo.

A floresta Amazônica é a maior floresta tropical do mundo, e cerca de 60% dela está localizada no Brasil. O clima predominante é o equatorial úmido, com temperaturas que variam de 22 °C e 28 °C, possui alta umidade do ar e longos períodos de chuva. A floresta Amazônica possui a maior biodiversidade do mundo, a fauna conta com cerca de 30 milhões de animais e ainda hoje novas espécies são descobertas na região. Já a flora contém árvores de grande porte e matas densas, de difícil acesso, e a vegetação é dividida em 3 categorias de matas, as de igapó, de várzea e de terra firme (STOODI, 2020).

O Cerrado brasileiro é conhecido como a savana com a maior biodiversidade do mundo. Possui área aproximada de 2 milhões de quilômetros quadrados, devido a sua grande dimensão, está presente em diversos estados brasileiros, o clima predominante neste bioma é o tropical sazonal, a temperatura anual média é entre 22-23 °C, podendo chegar a mais de 40 °C. O período de outubro a março é marcado por chuvas, porém podem ocorrer tempos de seca nesse período, já de maio a setembro, as chuvas cessam e uma estação de seca se inicia, podendo durar até cinco meses. O relevo é predominantemente plano com solo seco e ácido, possuindo uma vegetação bastante diversificada, possuindo grandes campos de gramas, e, também, árvores de médio e grande porte, animais de todos os gêneros vivem na região, desde insetos e cobras a macacos e gaviões (KLEIN, 2002).

Pode-se observar na Figura 3 os Biomas brasileiros, as cores vermelho e verde delimitam o Cerrado e a Amazônia, respectivamente, eles ocupam toda a região Norte e praticamente toda a região Centro-Oeste, que são áreas muito ameaçadas e devastadas pelo desmatamento.

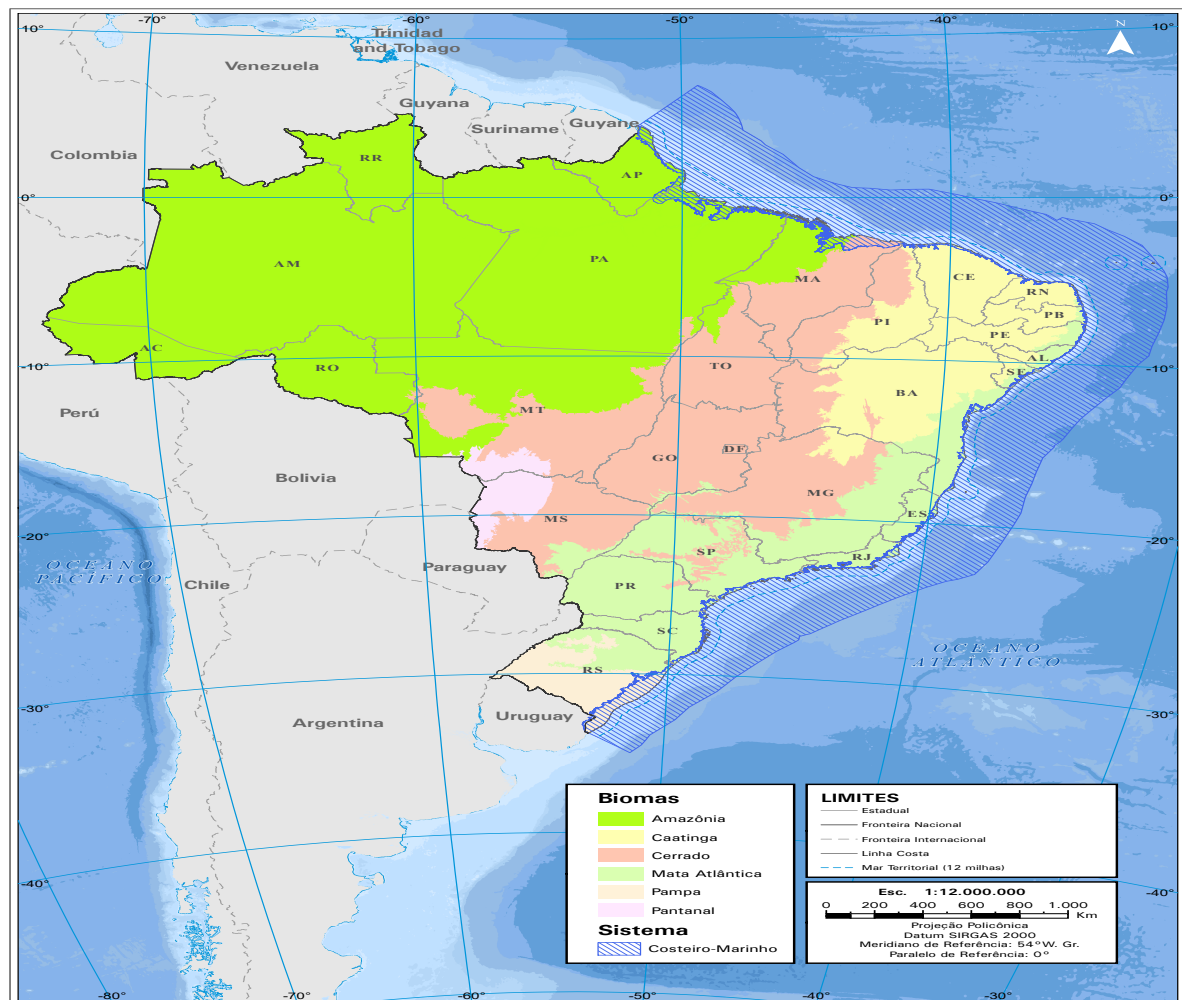


Figura 3: Mapa dos Biomas brasileiros. Fonte: (IBGE, 2019)

1.1 Situação do desmatamento no mundo

Em seu relatório, a FAO afirma que desde a década de 90 o planeta perdeu 420 milhões de hectares de florestas, definindo o desmatamento como a conversão da área florestal para o uso da terra com outros fins, entre os anos de 1990 e 2000 a estimativa foi de 16 milhões de hectares desmatados por ano (Food and Agriculture Organization, 2020). Porém a cada estimativa anual em intervalos de 5 anos a área desmatada está diminuindo como podemos observar na Figura 4.



Figura 4: Estimativa anual do desmatamento e expansão florestal entre os anos de 1990 e 2020. Fonte: FAO

Apesar da área desmatada estar diminuindo, nos últimos anos, diversos países mantiveram as médias anuais de áreas desmatadas acima de 500 mil hectares, principalmente os países da América do Sul, África e Ásia, como podemos observar nas Figuras 5. Esses continentes, em especial, a América do Sul e a África, possuem grandes áreas de savanas e florestas tropicais.

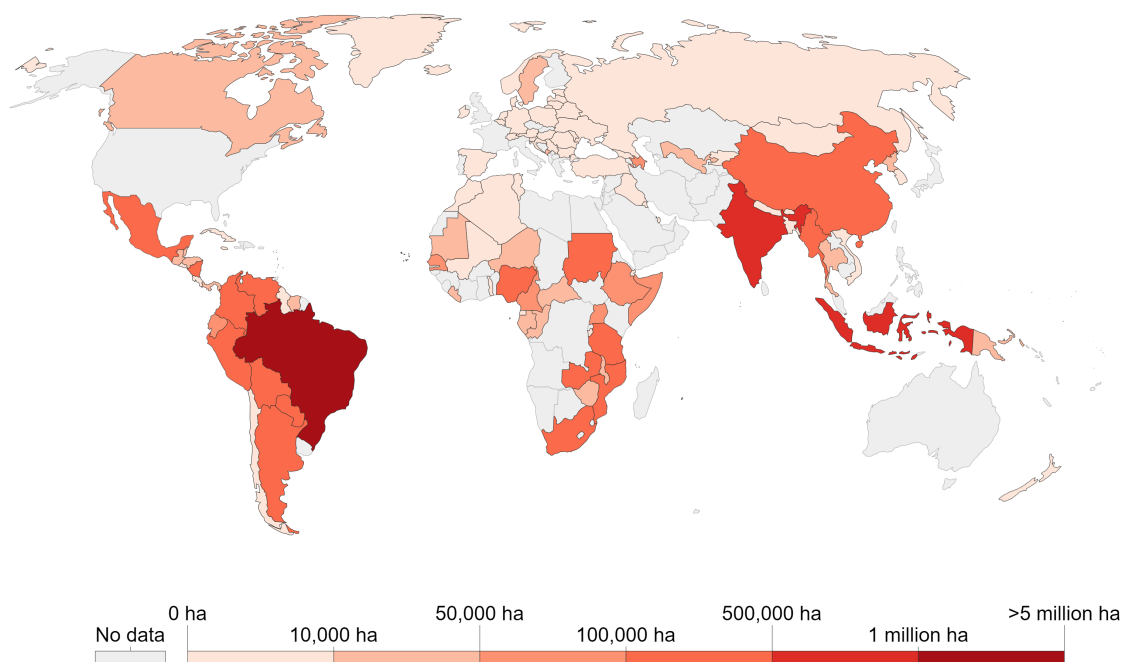


Figura 5: Média anual da quantidade de hectares desmatados entre os anos de 2015 e 2020. Fonte: Ritchie (2020)

1.1.1 Desmatamento nas savanas e florestas tropicais

O desmatamento das florestas pelo mundo está afetando tanto a flora quanto a fauna dos biomas e aumentando a quantidade de carbono na atmosfera, trazendo preocupações em relação a saúde do planeta. Nos últimos encontros de líderes mundiais o desmatamento tem sido uma pauta recorrente e trouxe a discussão sobre a necessidade da expansão de áreas de agropecuária, devido a alta demanda de uma população mundial em rápido crescimento. Porém Lamas (2017), da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), afirma que a tecnologia é o melhor aliado para o aumento da produção agrícola, não sendo necessária a expansão física da área, e, sim, a adesão de medidas sustentáveis e tecnológicas.

O presidente do Brasil, Jair Bolsonaro, afirmou na cúpula de líderes mundiais sobre o clima, que cumprirá os acordos assumidos no Acordo de Paris, que visam diminuir a emissão de gás carbônico. Para tal, Bolsonaro prometeu a eliminação do desmatamento até o ano de 2030 (Agência Brasil - Brasília, 2021). Porém, as medidas adotadas pelo governo foram no sentido oposto, obtendo, nos anos de 2019 e 2020, as maiores quantidades de área desmatada desde 2008 (INPE, 2021a).

No ano de 2020 a taxa de destruição das florestas tropicais aumentou se comparada aos últimos anos, mesmo que devido a pandemia as demandas por produtos diminuiu. A taxa de perda de florestas nativas permanece em um nível elevado, árvores são cortadas e a vegetação rasteira é queimada para abertura de terreno para pastagens. O desflorestamento cresceu durante a pandemia, devido à crescente do desemprego, muitas pessoas migraram para áreas rurais para trabalhar em fazendas (FOUNTAIN, 2021).

As savanas são comumente encontradas em regiões tropicais, onde há baixa ocorrência de chuvas, principalmente na Oceania, África e América do sul e normalmente cobrem pelo menos 25% da superfície dos países à qual pertencem. Essas áreas estão ameaçadas, seja devido ao modelo econômico desses países que desmatam essas áreas para o avanço da agropecuária ou devido ao grande número de queimadas que às atingem (FALEIRO; NETO, 2009).

As savanas provêm 85% da área global queimada, tais regiões estão experienciando uma alta taxa de mudanças do uso da terra, muito devido a esse ecossistema responder rapidamente a ação humana. O desmatamento nas savanas possui altos índices, sua vegetação é constantemente exposta a queimadas e a manuseio para criação de pastagens (LEHMANN, 2010).

1.2 Situação do desmatamento no Brasil

Com o intuito de diminuir a poluição do mundo, a maioria dos países diminuiu o desmatamento de suas florestas provocando um decaimento da média global. Porém, o Brasil é o único país no mundo que mantém a média anual de desmatamento acima dos 5 milhões de hectares. Entre os anos de 1990 e 2000 o país já tinha esta alta média de hectares desmatados e na Figura 6 observa-se que a média de desmatamento entre os anos de 2015 e 2020 permanece nos 5 milhões de hectares anuais.



Figura 6: Média da quantidade de hectares desmatados na América do Sul entre os anos de 1990 e 2000. Fonte: Ritchie (2020)

A Amazônia brasileira continua sofrendo com o desmatamento, desde 2012 o índice anual está em constante crescimento, com um alto pico nos anos de 2019 e 2020 (INPE, 2021a). Na Figura 7 pode-se observar que o desmatamento está ameaçando o interior da Amazônia, as fronteiras com o cerrado já mostram uma área desmatada bem elevada, onde a área em preto representa a supressão da vegetação nativa entre os anos de 2002 e 2020.

Já na Figura 8 observa-se que o Cerrado foi devastado pelo desmatamento, com quase sua vegetação nativa total perdida, e com isso a região Norte é o caminho "natural" do desmatamento. Observa-se que o desmatamento acumulado, área em preto, está muito

presente na fronteira entre Mato Grosso e Pará, área que requer atenção pelo potencial de devastação.

O estado do Mato Grosso é coberto parte pelo Cerrado e parte pela Amazônia, e está sendo muito afetado pelo desmatamento. A fronteira com o estado do Pará já foi ultrapassada, então esse estado também está sofrendo com as ações do desmatamento.

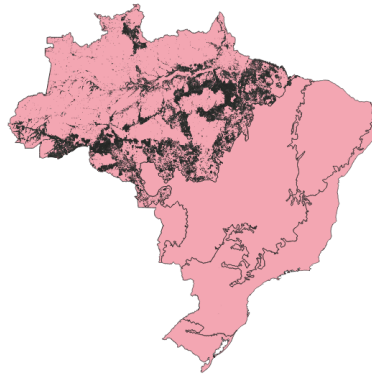


Figura 7: Incremento do desmatamento na Amazônia de 2008 a 2020. Fonte: Terrabrasilis (2021)

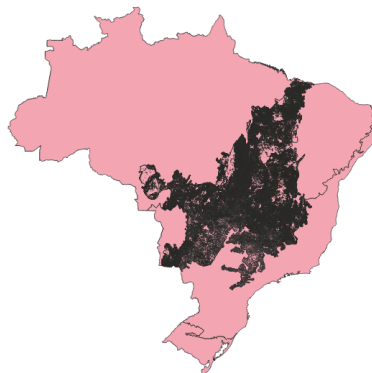


Figura 8: Incremento do desmatamento no Cerrado de 2002 a 2020. Fonte: Terrabrasilis (2021)

1.2.1 O desmatamento no Norte e Centro-Oeste do Brasil

A região Norte do Brasil é completamente coberta pela Amazônia, a maior floresta tropical do mundo, e a principal responsável pelo controle das chuvas e do gás carbônico na atmosfera, também responsável por abrigar uma grande biodiversidade de fauna e flora, porém essa região já teve 20% das suas florestas desmatadas (The Nature Conservancy).

Em 2005 o desmatamento começou a diminuir na Amazônia brasileira, após o Governo aplicar medidas com o intuito de parar o avanço das indústrias de carne e soja, que eram

responsáveis pela supressão da vegetação para abrir áreas de plantio e pastagem. Porém, em 2012, o Congresso diminuiu as proteções ambientais, o que gerou um aumento no desmatamento, tendo em 2016 o maior nível desde 2008 (TOLLEFSON, 2016).

A modelagem do incremento do desmatamento nos estados que compõem a Amazônia Legal mostra que o efetivo bovino e produtos da agricultura, em especial a soja, são significativos para sua explicação. O mesmo estudo modela a área de florestas no ano anterior, e conclui que essa variável também explica o incremento no desmatamento (OLIVEIRA et al., 2011).

A região Amazônica está constantemente sofrendo com ações humanas, sendo destruída para o avanço da urbanização, extração de madeira, mineração e abertura de terreno para a agropecuária (WWF-Brasil, 2005b).

O Cerrado, conhecido como a savana brasileira, é o 2^o maior Bioma brasileiro, e cobre uma área de 2 milhões de quilômetros quadrados, porém 50% desta área foi desmatada (AFONSO; AQUINO; CAMARGO).

O Centro-Oeste ocupa cerca de 19% do território brasileiro, tendo a maior parte de sua área coberta pelo Cerrado. Este Bioma teve 45% de sua área convertida em pastagens e lavouras, e segue ameaçada pela extração de minérios, expansão da agropecuária, poluição devido a indústrias e por construções de barragens (Instituto Chico Mendes de Biologia).

1.3 Relação entre o desmatamento e os seus indicadores

O desmatamento dificilmente ocorre de forma natural, como por exemplo queimadas geradas por raios. Sendo assim, as causas devem ser identificadas e ações promovidas para o excesso, que é prejudicial em vários aspectos, não ocorra. Alguns indicadores do desmatamento são a agropecuária, as queimadas e a extração de madeira.

Segundo o Banco Mundial, no mundo há um total de 8,9 Milhões de pessoas passando fome, porém 1/3 do total da produção de comida é perdida ou gasta de forma indevida não chegando as pessoas a qual deveriam e mesmo assim a agricultura é responsável por 4% do PIB mundial, chegando até a 25% em alguns países em desenvolvimento (Banco Mundial, 2020). Essa informação mostra a importância do uso da tecnologia nas plantações, que por si só pode aumentar em 1/3 da produção de alimento mundial.

O PIB do agronegócio correspondeu a 24,31% do PIB brasileiro no ano de 2020, porém

ao comparar o primeiro trimestre de 2020 e o primeiro trimestre de 2021, no ramo da pecuária houve um queda de 1,96%, o que equivale a 12 bilhões de Reais, enquanto no ramo da agricultura houve um aumento de 7,99%, que equivale a 136 bilhões de Reais (Centro de Estudos Avançados em Economia Aplicada, 2021).

A preocupação com a saúde do planeta é uma discussão abordada pela mídia em todo o mundo, o desmatamento é a principal causa para a piora dos problemas, então há a necessidade de investigar as suas principais causas para assim poder realizar ações preventivas e diminuir até erradicar o desmatamento .

Uma pesquisa do Sy et al. (2015), na América do Sul, aponta que a principal causa para o desmatamento é a expansão do agronegócio, em principal a pastagem, que corresponde a pelo menos 70% da área desmatada. O mesmo estudo aponta que 80% do desmatamento no Brasil é devido ao agronegócio, onde os produtores rurais devastam áreas de vegetação para abrir espaço para a agricultura e pastagens. A pecuária não é rentável na Amazônia, porém devido a subsídios do governo essa prática ainda persiste na região, já a agricultura tomou conta do cerrado, na região Centro-Oeste, e está pressionando a fronteira com a região Norte, ameaçando a Amazônia (MARGULIS, 2003).

Um estudo de Angelo e Sá (2007), mostra que existe correlação estatística entre o desflorestamento e o efetivo de rebanho bovino, preço da terra para pasto e preço da terra para lavoura, sendo o efetivo de rebanho bovino um dos fatores mais significativos para explicar o desflorestamento na Amazônia brasileira.

A queimada é uma prática utilizada para abrir áreas para plantações e pastos. Essas queimadas muitas vezes perdem controle e o fogo destrói áreas de florestas naturais (WWF-Brasil, 2005a).

O capital gerado pela extração seletiva de madeira é, por muitas vezes, investido em atividades econômicas como a criação de gado ou agricultura intensiva. Além disso, também facilitam o acesso às florestas por meios das estradas, que são realizadas para agilizar a atividade madeireira (ASNER et al., 2013).

Um estudo realizado por Asner et al. (2005) verificou a existência de correlação espacial entre a extração seletiva de madeira e o desmatamento na Amazônia brasileira. Pela pesquisa, aparentemente, a extração de madeira é uma das causas do desmatamento, porém em anos subsequentes. Desta forma, o trabalho comprova que a extração de madeira contribui para o incremento do desmatamento.

1.4 Objetivos

O objetivo geral deste trabalho é construir modelos que possam descrever a evolução anual das séries temporais referentes à área de floresta natural, aos indicadores agropecuários, sendo esses a área destinada à lavouras e pastagens e o número de cabeças de gado bovino, e à quantidade de focos de queimadas nas regiões Norte e Centro-Oeste, além de verificar a existência de relação temporal entre elas.

Os objetivos específicos, aplicados a ambas as regiões, são apresentados abaixo:

- Modelar a série temporal da área coberta por florestas naturais
- Construir um modelo para a série temporal da quantidade de área destinadas à lavouras
- Modelar a série temporal da quantidade de área do solo coberta por pastagens
- Construir um modelo para a série temporal do efetivo de gado bovino
- Modelar a série temporal da quantidade de focos de queimadas
- Construir um modelo para a série temporal da extração de madeira
- Verificar se existe correlação cruzada entre as séries temporais investigadas
- Construir modelos para o desmatamento utilizando as séries temporalmente dependentes

No capítulo 2 serão apresentadas as bases de dados referentes às séries históricas em estudo e as ferramentas de análises estatística utilizadas. As bases foram colhidas do Projeto de Mapeamento Anual do Uso e Cobertura da Terra no Brasil (MapBiomass), Instituto Brasileiro de Geografia e Estatística (IBGE) e do Instituto Nacional de Pesquisas Espaciais (INPE), já as análises foram realizadas pelo software R.

Os resultados obtidos para cumprimento dos objetivos propostos serão apresentadas no capítulo 3, serão realizadas todas as análises descritivas e inferenciais com o objetivo de descrever o comportamento das séries históricas em estudo. Também será verificada a concordância com a bibliografia e literatura já existente sobre o tema.

Por fim, no capítulo 4, serão apresentadas as conclusões finais, que discorreram sobre os principais resultados e seus efeitos, dicas e observações para futuros estudos também poderão ser comentados.

2 Materiais e Métodos

Neste Capítulo serão apresentadas as bases de dados e as metodologias estatísticas que serão utilizadas para as análises.

2.1 Bases de dados

As bases de dados deste trabalho foram recolhidos de três fontes distintas e estão representadas através de quatro séries temporais anuais que serão descritas a seguir. Os dados referem-se às Regiões Norte e Centro-Oeste do Brasil que foram tratados como 1 única Região, depois foi dado ênfase aos estados do Pará e Mato Grosso, que também foram agrupados como apenas um local.

Foram utilizados os estados do Pará e Mato Grosso devido a estes serem os locais com maior quantidade de área desmatada nas duas Regiões. A fronteira desses dois estados é o local onde o desmatamento está se concentrando, devido a este fato foi realizado o agrupamento deles como apenas um local no estudo.

O espaço temporal de todas as séries são anuais, isto é, para cada uma das séries os seus valores são correspondentes à todo o ano.

2.1.1 Desmatamento, agricultura e pastagem

Essas três séries temporais foram obtidas através do Projeto de Mapeamento Anual do Uso e Cobertura da Terra no Brasil (MapBiomias). O projeto surgiu no ano de 2015, após um seminário, onde especialistas na área de sensoriamento e mapeamento debateram sobre como criar mapas de uso e cobertura da terra no Brasil de forma barata e rápida (MAPBIOMAS, 2016).

O projeto MapBiomias utiliza imagens de satélites Landsat, que são satélites de observação da Terra, cujo objetivo é observar recursos naturais terrestres, essas imagens são

disponibilizadas na plataforma Google Earth Engine. As imagens selecionadas pelo projeto correspondem ao período de todo o ano, são analisadas pixel a pixel e possibilitam a realização de mosaicos, que são usados para produzir um mapa para cada classe de cobertura e uso do solo através do classificador baseado em aprendizagem de máquina, conhecido como random forest, assim obtendo estimativas para as áreas de cada classe (MAPBIOMAS, 2020).

Para este trabalho, o desmatamento é caracterizado como a supressão da vegetação nativa, independente da finalidade do uso futuro da terra. Essa definição é utilizada pelo projeto de estimativa do desflorestamento da Amazônia (PRODES), coordenado pelo INPE (TERRABRASILIS, 2021).

As bases de dados sobre o desmatamento se referem a área coberta por floresta nativa nas regiões Norte e Centro-Oeste ou nos estados do Pará e Mato Grosso, composta por formações de florestas, savanas e mangues, onde consta a área de vegetação natural de 1985 até 2019 (MAPBIOMAS, 2020).

As bases de dados sobre agricultura se referem às áreas ocupadas por lavouras permanentes ou temporárias nas regiões Norte e Centro-Oeste ou nos estados do Pará e Mato Grosso. Os dados correspondem aos anos de 1985 à 2019 (MAPBIOMAS, 2020).

A área de pastagem refere-se às bases de dados sobre pastagem, a área é referente às regiões Norte e Centro-Oeste do Brasil ou aos estados do Pará e Mato Grosso, uma série histórica de 1985 à 2019 (MAPBIOMAS, 2020).

Todas as séries anuais citadas possuem o mesmo período, totalizando 35 observações para cada uma delas.

2.1.2 Extração de madeira

A pesquisa Produção da Extração Vegetal e da Silvicultura (PEVS) é uma pesquisa do Instituto Brasileiro de Geografia e Estatística (IBGE), cujo objetivo é levantar informações sobre a quantidade e o valor da produção dos recursos gerados pela exploração vegetal nativa e dos maciços florestais plantados. Os dados são coletados para cada município e teve início no ano de 1938 (IBGE, 2020c).

As bases dos dados utilizados se referem a metros cúbicos de tora de madeira, aos anos de 1986 à 2020, uma para as regiões Norte e Centro-Oeste do Brasil e a outra para os estados do Pará e Mato Grosso (IBGE, 2020d). O total de observações é de 35 para

essa série temporal.

2.1.3 Efetivo bovino

O IBGE realiza a Pesquisa da Pecuária Municipal (PPM), que investiga informações sobre efetivos de animais de diferentes espécies e seus derivados. A pesquisa tem âmbito nacional e ocorre com periodicidade anual desde o ano de 1945, os resultados são divulgados para municípios, microrregiões, mesorregiões, unidades da federação e grandes regiões (IBGE, 2020a).

As bases de dados sobre o gado consistem no número de cabeças de gado bovino presente nas regiões Norte e Centro-oeste do Brasil ou nos estados do Pará e Mato Grosso, disponibilizado pelo IBGE através da plataforma SIDRA, foram obtidos dados entre os anos de 1985 e 2019 (IBGE, 2020b). Esta série temporal possui 35 observações.

2.1.4 Queimadas

O Instituto Nacional de Pesquisas Espaciais (INPE) desenvolveu o programa Queimadas, com o intuito de realizar pesquisas e desenvolvimento tecnológico para monitorar e modelar os focos de incêndio e o risco à vegetação (INPE, 2011). O INPE recebe imagens de dez satélites com sensores ópticos que operam na faixa termal-média de 4 μ m, dentre eles há satélites polares, como as AVHRR/3, as MODIS, como as da NASA, e satélites geoestacionários, como o GOES-16. 200 imagens, de todo o Brasil, captadas pelos satélites, são processadas diariamente na divisão de geração de imagens (DGI) e na divisão de satélites e sistemas ambientais (DSA) para a detecção de focos de queimadas (INPE, 2020).

A quantidade de focos de queimadas foram os dados obtidos sobre as queimadas, consiste no resumo anual da quantidade de queimadas por estados e foram obtidos através da plataforma BDQueimadas do INPE, nos anos de 1999 até 2020 (INPE, 2021b). Há 22 observações para a série temporal de quantidade de focos de queimadas.

2.2 Modelagem de Séries Temporais

Nessa seção serão introduzidas as bases teóricas para a análise de séries temporais, ajuste do melhor modelo e a realização de estimações e previsões. O nível de significância utilizado será de 5% e as análises serão realizadas nos Programas R (R Core Team, 2020)

e Rstudio (RStudio Team, 2020) onde foram utilizados os pacotes forecast (HYNDMAN et al., 2021) e (HYNDMAN; KHANDAKAR, 2008), dplyr (WICKHAM et al., 2021), ggplot2 (WICKHAM, 2016), scales (WICKHAM; SEIDEL, 2020), gridExtra (AUGUIE, 2017) e tseries (TRAPLETTI; HORNIK, 2020).

2.2.1 Conceitos básicos

Segundo Morettin e Tolo (2006), uma série temporal é qualquer conjunto de observações que tenham uma ordenação temporal. A análise de séries temporais tem por objetivos descrever o comportamento da série, verificar a existência de periodicidades, investigar o mecanismo gerador da série temporal e realizar previsões de valores futuros.

O modelo clássico de uma série temporal, Z_t , é descrito da seguinte forma:

$$Z_t = T_t + S_t + a_t, \quad t = 1, 2, \dots, n.$$

Onde T_t representa a tendência da série, S_t a sazonalidade e a_t é um fator aleatório. A sazonalidade representa o comportamento cíclico dos eventos, relacionado ao calendário ou à horários (BROCKWELL; DAVIS, 2002). O termo t representa os anos da série temporal.

Processos estocásticos são os modelos utilizados para descrever séries temporais, esses processos são controlados por leis probabilísticas.

Definição 2.1 *Seja T um conjunto arbitrário. Um processo estocástico é uma família $Z = \{Z(t), t \in T\}$, tal que, para cada $t \in T$, $Z(t)$ é uma variável aleatória.*

Um processo Z é estacionário caso ele se desenvolva no tempo de forma que o tempo inicial escolhido não é importante, isto é, as características de $Z(t+r)$, são as mesmas de $Z(t)$, para todo r .

Modelos para descrever séries temporais, em geral, se baseiam em alguma suposição simplificadora. A mais comum entre elas é a de estacionariedade.

Definição 2.2 *Um processo estocástico $Z = \{Z(t), t \in T\}$ diz-se estritamente estacionário se todas as distribuições finito-dimensionais permanecem as mesmas sob translações no tempo, ou seja,*

$$F(z_1, \dots, z_n; t_1+r, \dots, t_n+r) = F(z_1, \dots, z_n; t_1, \dots, t_n),$$

para quaisquer t_1, \dots, t_n, r de T

Desse modo, todas as distribuições unidimensionais não variam ao longo do tempo, logo a média e a variância são constantes, isto é,

$$\mu(t) = \mu, \quad V(t) = \sigma^2,$$

$\forall t \in T$. A suposição comum é, $\mu = 0$, caso contrário, considere o processo $\{Z(t) - \mu\}$ (MORETTIN; TOLOI, 2006).

2.2.2 Função de Autocorrelação e Autocorrelação Parcial

A estacionariedade implica que a distribuição de probabilidade conjunta $p(z_{t_1}, z_{t_2})$ é a mesma para todos os tempos t_1, t_2 , que estão separados por um intervalo de mesmo tamanho k . Em particular, a covariância entre z_t e z_{t+k} , separados k defasagens, ou por lag k , deve ser a mesma para todo t sobre suposição de estacionariedade. Essa covariância é chamada de autocovariância de lag k , e é definida por

$$\gamma_k = \text{cov}[z_t, z_{t+k}] = \mathbb{E}[(z_t - \mu)(z_{t+k} - \mu)]$$

Em um processo estacionário, a variância $\sigma_z^2 = \gamma_0$ é a mesma para os tempos t e $t + k$. Então, a autocorrelação, ou fac, de lag k , que é a função correlação entre z_t e z_{t+k} , é dada por:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

Na Figura 9 pode-se observar diagramas de dispersão para lags $k = 1, 2$, à esquerda percebe-se uma correlação negativa entre z_t e z_{t+1} , já à direita, aparentemente há uma correlação positiva entre z_t e z_{t+2} .

É importante ressaltar que a função de autocorrelação (ρ_k) em um processo estacionário segue o seguinte formato $\rho_k = \rho_{-k}$ como podemos observar na Figura 10. Com isso, basta observar os valores positivos da função de autocorrelação e poder assim chegar a conclusão (BOX et al., 2016).

A função de autocorrelação parcial, ou facp (ρ_{kk}), mede a correlação entre uma observação k períodos antes e a observação atual, depois de controlar as observações intermediárias. Por exemplo, a facp mediria a correlação entre z_t e z_{t-3} depois de controlar os efeitos de z_{t-1} e z_{t-2} . Quando o lag (k) é igual a 1, fac e facp tem coeficientes iguais, já que

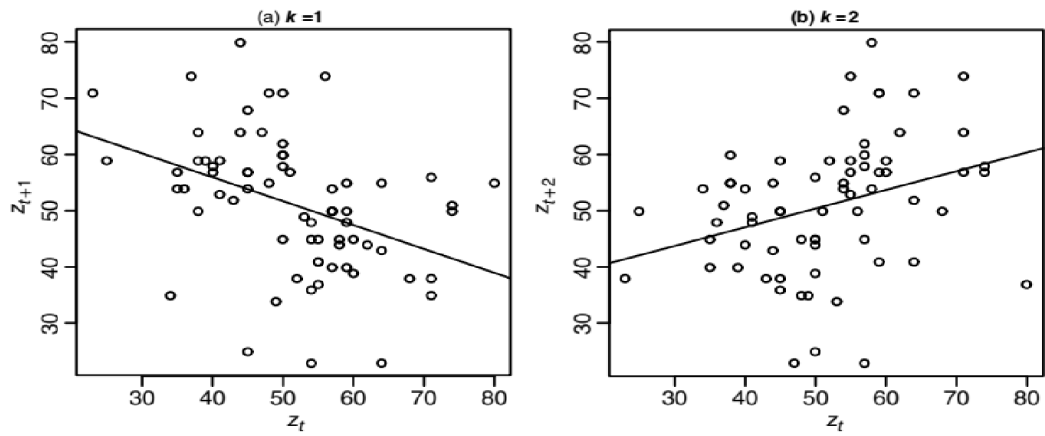


Figura 9: Diagramas de dispersão para lags $k=1,2$. Fonte: (BOX et al., 2016)

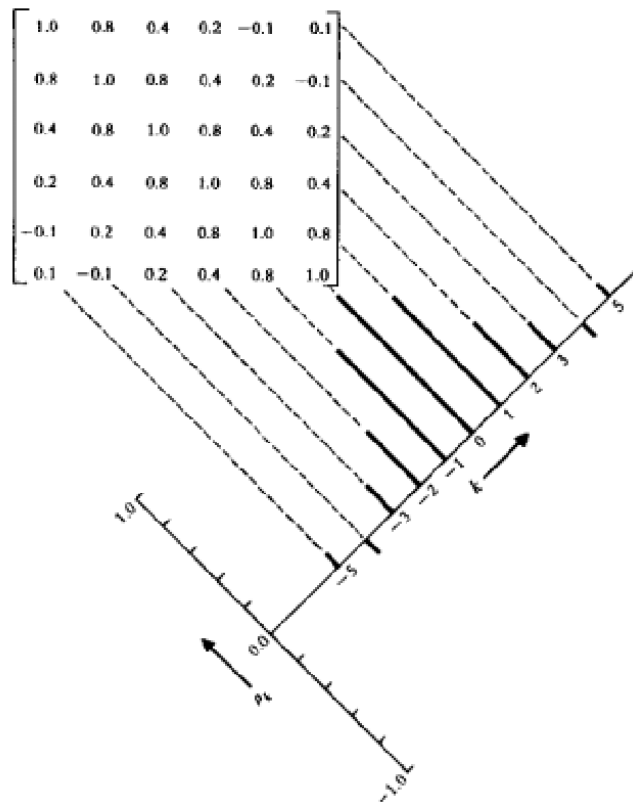


Figura 10: Matriz de autocorrelação e função de autocorrelação correspondente de um processo estacionário. Fonte: (BOX et al., 2016)

não há nenhum lag intermediário. Porém, para os outros lags, isso não necessariamente acontecerá. Por exemplo, pode-se mostrar que, para lag 2:

$$\rho_{22} = (\rho_2 - \rho_1^2) / (1 - \rho_1^2)$$

em geral, temos

$$\rho_{kk} = \text{Corr}(Z_t, Z_{t-k} | Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1})$$

2.2.3 Modelo Arima

Um modelo paramétrico, que possui número de parâmetros finito, tenta descrever o comportamento de uma série temporal no domínio do tempo. Os modelos mais comuns dessa classe são os modelos de regressão (linear ou de curva de crescimento), modelos autorregressivos e de médias móveis (ARMA), modelos autorregressivos integrados e de médias móveis (ARIMA), modelos de memória longa (ARFIMA), modelos estruturais e modelos não-lineares (MORETTIN; TOLOI, 2006).

O modelo ARMA(p,q) tem o objetivo de ser parcimonioso, fazendo a junção dos modelos das classes autorregressivos, AR(p), e médias móveis, MA(q). Este modelo é uma ferramenta muito comum em análises de séries temporais, onde tem a finalidade de entender o comportamento da série ao longo do tempo e, talvez, prever valores futuros. Seja \tilde{z}_t uma série estacionária

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

O modelo ARIMA(p,d,q) é uma generalização de um modelo ARMA(p,q), onde a série original \tilde{z}_t é não estacionária, então ela é diferenciada d vezes até se obter uma série $\Delta^d \tilde{z}_t$ estacionária. Para casos onde $d \geq 1$, $\Delta^d \tilde{z}_t = \Delta^d z_t$, então, pode-se escrever o modelo da seguinte forma (BOX; JENKINS, 1976)

$$\Delta^d z_t = \phi_1 \Delta^d z_{t-1} + \dots + \phi_p \Delta^d z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

Said e Dickey (1984) discute o teste de raiz unitária na classe de modelos ARIMA(p,d,q), o objetivo é identificar o termo d antes de estimar p e q, e assim usar a abordagem Box-Jenkins. O termo d é o número de diferenciações necessárias na série Z_t para que esta não possua raiz unitária, ou seja, a série se torne estacionária. O teste para tal, é conhecido como teste Aumentado de Dickey-Fuller (ADF).

Considera a série temporal Z_t , $t=1,2,\dots,n$ que é representada pela seguinte equação:

$$Z_t = \rho Z_{t-1} + e_t$$

e_t é um termo de ruído branco, possui média zero, variância σ^2 constante e não autocorrelacionada. Assim, as hipóteses são definidas da seguinte forma:

H_0 : $\rho = 1 \approx$ possui raiz unitária (não é estacionária)

H_1 : $\rho \neq 1 \approx$ não possui raiz unitária (é estacionária)

A estatística do teste ADF é definida pela seguinte equação:

$$\Delta^d z_t = \phi + \theta_t + \gamma z_{t-1} + \sum_{j=1}^p \rho_{j+1} + \Delta^d z_{t-j} + a_t$$

Assim, com a realização do teste ADF, primeiro encontra-se o valor de d , no qual a série seja estacionária, e então serão estimados os valores de p e q .

2.2.3.1 Modelos Box-Jenkins

Em uma série temporal sem componente da sazonalidade, a metodologia de Box-Jenkins aborda a previsão com base em modelos da classe auto-regressivo integrado e de médias móveis, ou, ARIMA(p,d,q). O método passa pelas etapas de identificação, estimação e validação do modelo. A partir da escolha do melhor modelo, pode-se avaliar a capacidade preditiva do modelo adotado (BOX; JENKINS, 1976).

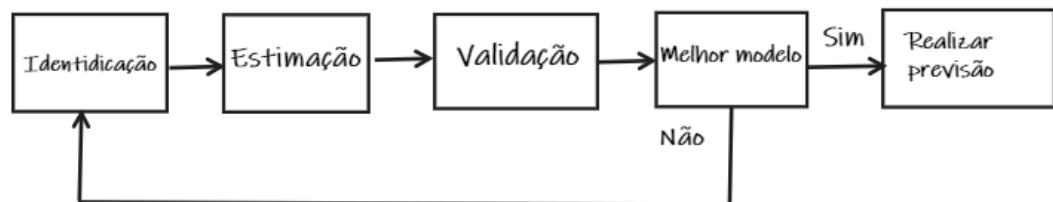


Figura 11: Ciclo iterativo da metodologia Box-Jenkins para geração de previsões

A etapa de identificação é a fase crítica, ela visa determinar os valores de p , d e q do modelo ARIMA(p,d,q), esse processo é constituído por três partes (MORETTIN; TOLOI, 2006):

- Verificar a necessidade de transformação da série original como o intuito de estabilizar a variância. Tal identificação pode ser feita de forma descritiva, através dos gráficos dependentes do tempo

- diferenciar a série obtida no item anterior até obter uma série estacionária, a fim de reduzir o processo $\Delta^d Z_t$ a uma ARMA(p,q). O teste de raiz unitária será de grande utilidade para esta verificação
- Identificar valores p e q da ARMA(p,q) através das funções de autocorrelação e autocorrelação parcial

Em um modelo auto-regressivo de médias móveis (ARMA), o fac e facp são de grande utilidade. Observando o comportamento de decaimento ao longo dos lags pode-se verificar possíveis valores para p e q da ARMA(p,q), onde p e q são o número de defasagens da série estacionária Z_t e do fator aleatório a_t , respectivamente (BROOKS, 2008). A relação da fac e facp com o modelo auto-regressivo integrado e de médias móveis (ARIMA) é representada na Tabela 1.

Tabela 1: Comportamento da fac e facp nos modelos ARIMA(p,d,q). Fonte (MORETTIN; TOLOI, 2006)

Ordem	(1,d,0)	(0,d,1)
Comportamento da fac	Decaimento exponencial	Somente $\rho_1 \neq 0$
Comportamento da facp	Somente $\rho_{11} \neq 0$	Decaimento exponencial
Ordem	(2,d,0)	(0,d,2)
Comportamento da fac	Mistura de exponenciais ou ondas senóides amortecidas	Somente $\rho_1 \neq 0$ e $\rho_2 \neq 0$
Comportamento da facp	Somente $\rho_{11} \neq 0$ e $\rho_{22} \neq 0$	Dominada por mistura de exponenciais ou senóides amortecidas
Ordem	(1,d,1)	
Comportamento da fac	Decaimento exponencial após lag 1	
Comportamento da facp	Dominada por decaimento exponencial após o lag 1	

A FAC e FACP podem não levar a boas estimativas para os parâmetros p e q da ARIMA(p,d,q). A razão é que, quando dados reais "confusos" são usados, os padrões apresentados na Tabela 1 não são respeitados. Isso torna a FAC e a FACP muito difíceis de interpretar e, portanto, é difícil especificar um modelo para os dados. Outra técnica, que remove a subjetividade envolvida na interpretação da FAC e FACP, é usar os critérios de informação. O objetivo é escolher o número de parâmetros que minimizem o valor dos critérios de informação. Portanto, adicionar um termo extra reduzirá o valor dos critérios apenas se a queda na soma dos quadrados dos resíduos é mais do que suficiente para compensar o termo de penalidade. Os dois critérios de informação mais populares são o critério de informação de akaike, critério de informação Bayesiano (BROOKS, 2008).

A etapa de estimação é aquela em que serão obtidas estimativa para os vários valores dos coeficientes $\boldsymbol{\phi} = (\phi_1, \dots, \phi_p)$ e $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)$. Os estimadores para $\boldsymbol{\phi}$ e $\boldsymbol{\theta}$ podem ser encontrados através de algumas formas, como pelo método dos momentos, método da máxima verossimilhança ou estimação não linear.

As estimativas obtidas pelo método dos momentos(MM), consistem em relacionar as autocorrelações(ou autocovariâncias) e os parâmetros do modelo. No modelo ARMA(p,q) o processo é realizado em duas etapas:

(i) Obter o estimador de $\boldsymbol{\phi}$

$$r_j = \hat{\phi}_1 MM r_{j-1} + \dots + \hat{\phi}_p MM r_{j-p}, \quad j = q + 1, \dots, q + p$$

onde

$$\hat{\phi} MM = R_p^{-1} \mathbf{r}_p$$

$$\mathbf{r}_p = (r_1, r_2, \dots, r_p)' \text{ e } R_p \text{ é a matriz formada pelos elementos } (r_{i-j})_{i,j=1}^p$$

vale ressaltar que r_i é a estimativa da fac ρ_i , para todo $i = 1, 2, \dots, p$

(ii) Obter o estimador de $\boldsymbol{\theta}$

$$\gamma_j = \phi_1 \gamma_{j-1} + \dots + \phi_p \gamma_{j-p} + \gamma_{za}(j) - \theta_1 \gamma_{za}(j-1) - \dots - \theta_q \gamma_{za}(j-q)$$

onde $\gamma_{za}(j)$ é a covariância cruzada entre Z_t e a_t , definida por:

$$\gamma_{za}(j) = \mathbb{E}(a_t Z_{t-j})$$

A etapa de validação consiste na verificação de independência e normalidade dos resíduos. Uma forma de verificar a não autocorrelação dos resíduos, avaliando se estes são ruídos brancos, é através do teste de Ljung-Box, uma extensão do teste de Box-Pierce, cuja estatística é:

$$Q(K) = n(n+2) \sum_{j=1}^K \frac{\hat{r}_j^2}{(n-j)},$$

onde \hat{r}_j^2 são as autocorrelações do modelo estimado (MORETTIN; TOLOI, 2006).

Caso o modelo for apropriado, a estatística Q(K) terá aproximadamente uma distribuição qui-quadrado com $K - p - q$ graus de liberdade.

O teste mais comum para verificar a normalidade dos resíduos é o Shapiro-Wilk, a estatística W do teste é definida por (SHAPIRO; WILK, 1965):

$$W = \frac{(\sum_{i=1}^n a_i z_i^2)}{\sum_{i=1}^n (z_i - \bar{z})^2}$$

onde a_i são os melhores coeficientes lineares imparciais normalizados, tabulados por Sarhan e Greenberg (1959)

Pequenos valores para a estatística W indicam não normalidade.

2.2.4 Critérios de Informação e Avaliação de Qualidade de Ajuste

Vários procedimentos de identificação de modelos ARMA foram propostos, alguns consistem em funções penalizadoras. O objetivo é encontrar as ordens de k e l que minimizem a quantidade

$$P(k, l) = \ln \hat{\sigma}_{k,l}^2 + (k + l) \frac{C(N)}{N}$$

em que $\hat{\sigma}_{k,l}^2$ é a estimativa para a variância residual do modelo ARMA(k, l) ajustado às N observações da série temporal e $C(N)$ é uma função do tamanho da série.

O termo penalizador $(k + l) \frac{C(N)}{N}$, aumenta quando novos parâmetros são adicionados, enquanto a variância residual diminui. Assim, o intuito é identificar as ordens k e l que equilibrem esse comportamento. Procedimentos de identificação que minimizam funções penalizadoras particulares são usualmente utilizadas para a escolha do melhor modelo.

O critério de informação de akaike(AIC) escolhe o melhor modelo cujas ordens k e l minimizem o seguinte critério

$$AIC(k, l) = \ln \hat{\sigma}_{k,l}^2 + \frac{2(k+l)}{N}$$

O critério de informação Bayesiano(BIC) sugere que o melhor modelo é aquele que minimiza o seguinte critério

$$BIC(k, l) = \ln \hat{\sigma}_{k,l}^2 + (k + l) \frac{\ln N}{N}$$

onde $\hat{\sigma}_{k,l}^2$ é a estimativa de máxima verossimilhança da variância do modelo ARMA(k, l) (MORETTIN; TOLOI, 2006).

O erro é a diferença entre o valor observado na série e o valor ajustado pelo modelo. Então, erro é a parte não previsível de uma observação, pode ser escrito como

$$e_t = z_t - \hat{z}_t$$

Através de algumas métricas, pode-se verificar se o modelo representa bem o comportamento da série, ou seja, a diferença do valor efetivo e do valor ajustado pelo modelo é bem pequena.

A média dos erros absolutos é uma métrica muito utilizada, representada pelo Erro Absoluto Médio (MAE)

$$MAE = \frac{\sum_{t=1}^n |e_t|}{n}$$

Há também a Raiz do Erro Quadrático Médio (RMSE), que é escrita como

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (e_t)^2}{n}}$$

Erros em porcentagem tem a vantagem de não estarem atrelados a alguma unidade de medida, a métrica mais comum para realizar esta comparação é o Erro Médio Absoluto Percentual (MAPE)

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{z_t} \right|$$

Com essas métricas já é possível verificar a bondade do ajuste do modelo escolhido, e assim comparar a série original e a série ajustada (HYNDMAN; ATHANASOPOULOS, 2018).

2.2.5 Previsão

Uma das principais utilizações de modelos de séries temporais é gerar previsões com o intuito de preparar-se para o futuro. A realização dessas previsões, em um modelo ARIMA, é feito baseado em toda a série histórica.

Considerando um modelo da classe ARIMA, sem sazonalidade, há 3 formas explícitas para realizar previsões para z_{t+v} , $v \geq 1$, sendo elas (BOX et al., 2016):

1. Em termos direto da diferença

$$\hat{z}_{t+v} = \phi_1 z_{t-1+v} + \dots + \phi_{p+d} z_{t-p-d+v} - \theta_1 a_{t-1+v} - \dots - \theta_q a_{t-q+v} + a_{t+v}$$

2. Uma soma infinita dos fatores aleatórios, considerando pesos

$$\hat{z}_{t+v} = \sum_{j=0}^{\infty} \Psi_j a_{t-j+v},$$

onde $\Psi_0 = 1$, e os demais pesos são calculados da seguinte maneira

$$\phi(\mathbf{B})(1 + \Psi_1 \mathbf{B} + \Psi_2 \mathbf{B}^2 + \dots) = \theta(\mathbf{B})$$

logo Ψ_j pode ser calculado da seguinte forma

$$\Psi_j = \phi_1 \Psi_{j-1} + \dots + \phi_{p+d} \Psi_{j-p-d} - \theta_j$$

3. Uma soma infinita de observações passadas mais o fator aleatório, também considerando pesos

$$\hat{z}_{t+v} = \sum_{j=1}^{\infty} \pi_j z_{t-j+v} + a_{t+v}$$

os pesos π podem ser calculados por

$$\phi(\mathbf{B}) = (1 - \pi_1 \mathbf{B} - \pi_2 \mathbf{B}^2 + \dots) \theta(\mathbf{B})$$

sendo que

$$\sum_{j=1}^{\infty} \pi_j = 1$$

Uma forma de analisar o modelo que produz as melhores previsões é através da validação cruzada, comparando o MAPE de cada modelo. As etapas ocorrem da seguinte maneira:

- Primeiro passo é separar em amostras treino e teste, onde a amostra treino corresponde a maior parte da série temporal e a teste apenas aos últimos anos
- Os modelos serão construídos utilizando apenas usando a amostra treino
- Para cada modelo será gerado previsões para os anos presentes na amostra teste
- Por fim, as previsões de cada modelo serão comparadas aos valores reais, presentes na amostra teste, gerando assim um MAPE para cada modelo, o modelo com menor MAPE será considerado o melhor para gerar previsões

O operador de defasagem \mathbf{B} é definido da seguinte maneira (BOX et al., 2016):

$$\mathbf{B} = \frac{z_{t-1}}{z_t}$$

2.2.6 Intervalo de confiança

Com o intuito de gerar um intervalo de confiança para as previsões, é necessário fazer as seguintes suposições sobre a_t para qualquer t :

$$\begin{aligned}\mathbb{E}(a_t) \\ \text{Var}(a_t) &= \sigma_a^2 \\ \mathbb{E}(a_t a_s), t \neq s \\ a_t &\sim N(0, \sigma_a^2)\end{aligned}$$

O intervalo de confiança para uma previsão qualquer, z_{t+v} , é calculado da seguinte forma (MORETTIN; TOLOI, 2006):

$$\hat{z}_t(v) - \mu_{(1-\alpha/2)} \hat{\sigma}_a \left[1 + \sum_{j=1}^{v-1} \Psi_j^2 \right]^{1/2} \leq z_{t+v} \leq \hat{z}_t(v) + \mu_{(1-\alpha/2)} \hat{\sigma}_a \left[1 + \sum_{j=1}^{v-1} \Psi_j^2 \right]^{1/2}$$

2.2.7 Correlação cruzada

Supondo k séries temporais, denotadas por $x_{1,t}, x_{2,t}, \dots, x_{k,t}$, então tem-se o seguinte vetor de séries temporais no tempo t , $\mathbf{X}_t = (x_{1,t}, x_{2,t}, \dots, x_{k,t})$. Múltiplas séries temporais, em determinados contextos, podem ser correlacionadas.

O objetivo de estudar múltiplas séries temporais é modelá-las conjuntamente, identificando possíveis relações cruzadas entre elas. Há dois motivos-chaves para analisar e modelar essas séries, são eles:

1. Entender a dinâmica das relações das séries ao decorrer do tempo
2. Melhorar a acurácia das previsões das séries individualmente, utilizando informações das séries relacionadas

O vetor $\{\mathbf{X}_t\}$ é estacionário se as distribuições de probabilidade dos vetores aleatórios, $(\mathbf{X}_{t_1}, \mathbf{X}_{t_2}, \dots, \mathbf{X}_{t_m})$ e $(\mathbf{X}_{t_1+l}, \mathbf{X}_{t_2+l}, \dots, \mathbf{X}_{t_m+l})$, são os mesmos para os tempos arbitrários t_1, t_2, \dots, t_m , para todo m e todo lag.

Para um processo estacionário $\{\mathbf{X}_t\}$ a covariância entre $x_{i,t}$ e $x_{j,t+l}$ deve depender apenas do lag l . A covariância cruzada entre as séries $x_{i,t}$ e $x_{j,t+l}$ no lag l , é dado por:

$$\gamma_{ij}(l) = \text{cov}[x_{i,t}, x_{j,t+l}] = \mathbb{E}[(x_{i,t} - \mu_i)(x_{j,t+l} - \mu_j)]$$

A função de correlação cruzada (CCF) correspondente para o lag l , é (BOX et al., 2016):

$$\rho_{ij}(l) = \text{corr}[x_{i,t}, x_{j,t+l}] = \frac{\gamma_{ij}(l)}{\{\gamma_{ii}(0)\gamma_{jj}(0)\}^{1/2}}$$

onde $\gamma_{ii}(0) = \text{var}[x_{i,t}]$

2.2.8 Função de Transferência

A metodologia de Box-Jenkins permite integrar variáveis auxiliares ao modelo ARIMA. A forma que essas variáveis influenciam na variável de interesse é dada por uma função de transferência de ordem (r,s) (BOX et al., 2016):

$$(1 - \delta_1 \mathbf{B} - \dots - \delta_r \mathbf{B}^r) Y_t = (\omega_0 - \omega_1 \mathbf{B} - \dots - \omega_s \mathbf{B}^s) X_{i,t-b}$$

A série de saída, Y_t , será influenciada pelas séries temporalmente correlacionadas $X_{i,t}$.

Para construir a função de transferência seguindo a metodologia Box-Jenkins há alguns passos a serem seguidos:

1. Ajustar um modelo ARIMA para $X_{i,t}$;
2. Filtrar Y_t pelo modelo encontrado para $X_{i,t}$;
3. Calcular a função de correlação cruzada entre os resíduos dos modelos ARIMA estimados;
4. Identificar os valores para r, s e b;
5. Estimar o modelo Box-Jenkins com função de transferência;
6. Encontrar o melhor modelo.

A primeira etapa é apenas ajustar um modelo ARIMA da forma usual para cada uma das séries $X_{i,t}, i \geq 1$.

Já a segunda etapa trata de calcular um modelo ARIMA para a série temporal de interesse Y_t , porém acrescentando o modelo já estimado para $X_{i,t}$, onde este modelo é ajustado a Y_t sem reestimar qualquer parâmetro. Logo serão construídos i modelos ARIMA nesta etapa.

Para realizar a terceira etapa basta calcular a função de correlação cruzada entre os resíduos dos 2 modelos estimados nas etapas anteriores. Nesta etapa será verificada a existência de relação entre $Y_t, X_{i,t}$ e suas defasagens.

Na quarta etapa é necessário avaliar o gráfico da CCF e identificar os valores para r, s e b. Essa identificação pode ser feita da seguinte forma:

1. b representa o primeiro lag significativo

2. s corresponde ao número de defasagens crescentes após b
3. r é baseado no comportamento da CCF, por exemplo: $r=0$, caso a correlação seja constante para todos os lags após b , $r=1$, caso a correlação seja crescente após b , $r=2$, caso a correlação possua comportamento de onda após b .

Uma vez identificado as ordens r , s e b de todas as séries auxiliares é possível realizar a quarta etapa. O primeiro passo dela é defasar cada uma das variáveis $X_{i,t}$ b vezes. Feito isso, é só calcular o modelo ARIMA se utilizando da função de transferência, onde deve ser apresentada às ordens de r e s para cada série auxiliar (FERREIRA; SOUZA; MATTOS, 2015) e (BOX et al., 2016).

Supondo um modelo hipotético cujos parâmetros encontrados fossem $b=1$, $r=2$ e $s=3$, o modelo seria representado pela seguinte equação:

$$Y_t = \frac{(\omega_0 - \omega_1 \mathbf{B} - \omega_2 \mathbf{B}^2 - \omega_3 \mathbf{B}^3)}{(1 - \delta_1 \mathbf{B} - \delta_2 \mathbf{B}^2)} X_{t-1}$$

3 Resultados

Vale ressaltar que para as análises as regiões de estudo foram agrupadas. Isto é, as Regiões Norte e Centro-Oeste estão sendo representadas como apenas uma Região, e os estados de Pará e Mato Grosso também foram agrupados como uma só região.

3.1 Análises e resultados para as Regiões Norte e Centro-Oeste

Os comportamentos das séries históricas são apresentados na Figura 12. Aparentemente as séries não são estacionárias, e, elas possuem tendência, seja crescente ou decrescente, representada pela linha azul, o que indica que o cálculo da diferença pode ser eficaz para tornar as séries estacionárias. A série de floresta nativa é a única que apresenta tendência exclusivamente decrescente durante todo o período.

As séries referentes às áreas de agricultura, pastagem e efetivo bovino apresentam tendência exclusivamente crescente, o que pode ser um indício que esses fatores afetam no decréscimo da área de floresta nativa. A série sobre as quantidade de focos de queimadas possui um intervalo de tempo diferente das demais séries, e o comportamento é diferente ao decorrer dos anos, porém percebe-se que a partir de 2012 a série apresenta uma tendência crescente.

A série dos focos de queimadas possui variabilidade alta em alguns anos iniciais, porém para realizar uma transformação a interpretação seria muito alterada, então, devido a maior parte da série aparentemente ter uma variabilidade equilibrada nenhuma transformação foi feita. Já as outras séries aparentam ter um comportamento linear por quase todo o período, o que sugere que a variabilidade deve ser constante, então não necessita de nenhuma transformação.

Comportamentos cíclicos não foram identificados, o que sugere que as séries não possuem sazonalidade. Devido a grandeza das unidades de medidas das variáveis, foi realizado

uma transformação a fim de melhorar o entendimento dos comportamentos das séries. As áreas de floresta nativa, pastagem e agricultura estão apresentadas em 10.000 hectares, o efetivo bovino está em 10.000 cabeças de bois e a quantidade de focos de queimadas está em 1.000 focos.

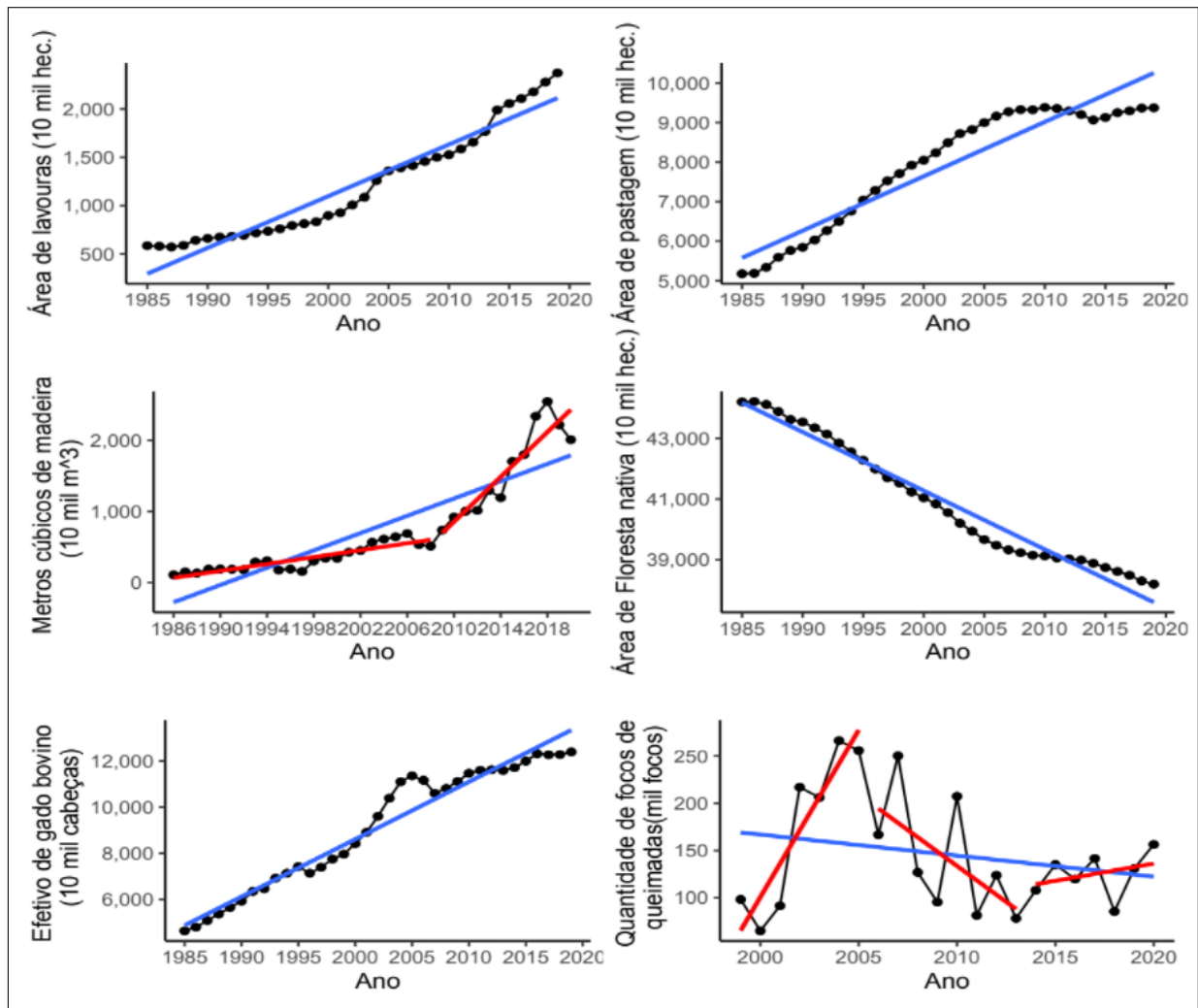


Figura 12: Séries temporais em estudo e suas tendências

Para verificar se a diferenciação pode tornar a média das séries constantes, podemos verificar os gráficos. Na Figura 13 tem-se as séries diferenciadas 1 vez e percebe-se que apenas a quantidade de focos de queimadas aparenta ter a média constante, já as outras podem necessitar uma nova diferenciação.

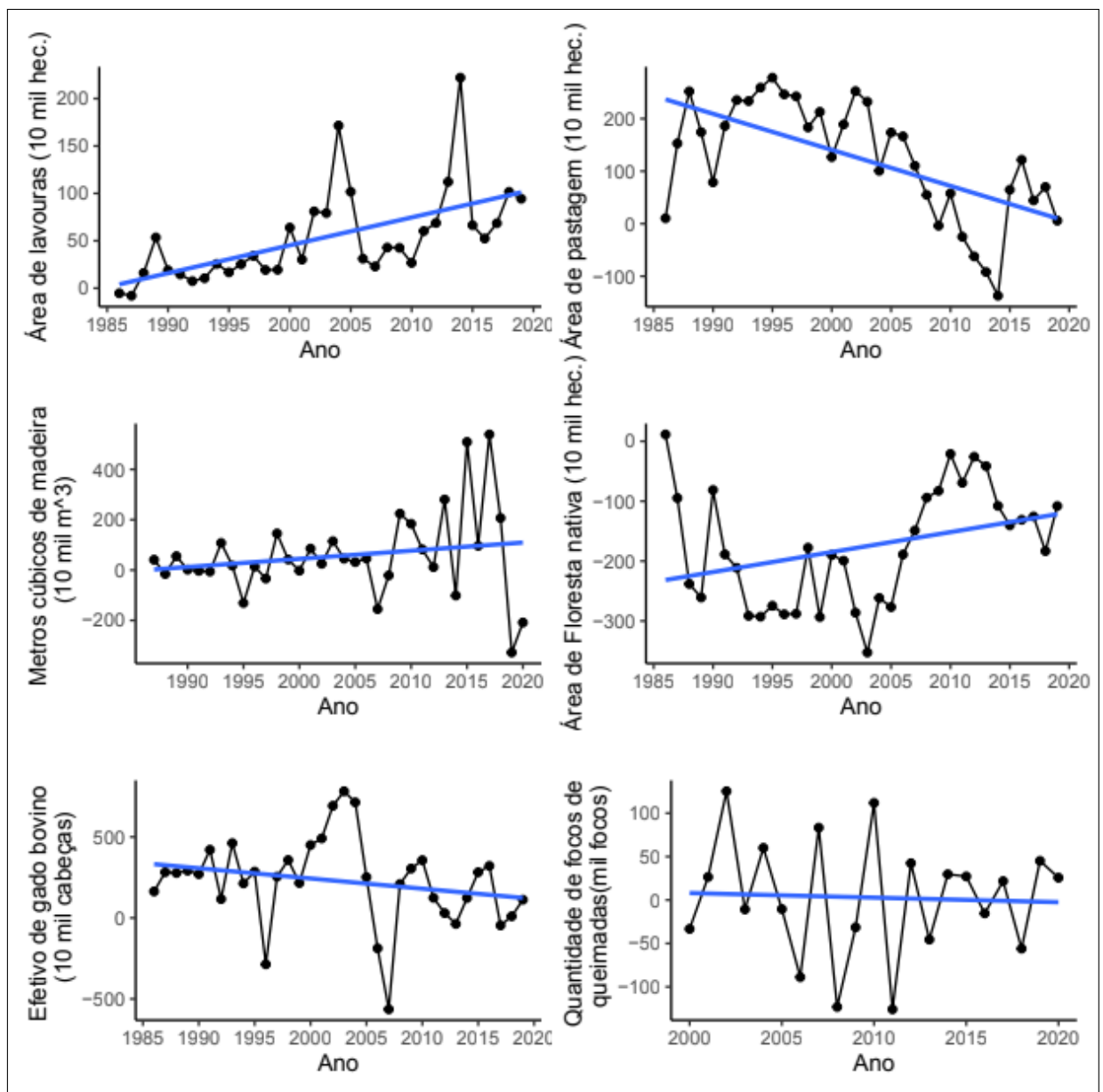


Figura 13: Séries temporais diferenciadas 1 vez

A Figura 14 apresenta as séries diferenciadas 2 vezes, não há grandes mudanças na quantidade de focos de queimadas, o que sugere que 1 diferenciação é suficiente, mas as outras séries aparentam ter médias mais próximas de uma constante na segunda diferenciação, o que sugere que elas necessitam de tal para serem estacionárias. É necessário realizar o teste aumentado de Dickey-Fuller para verificar a veracidade da aparência visual.

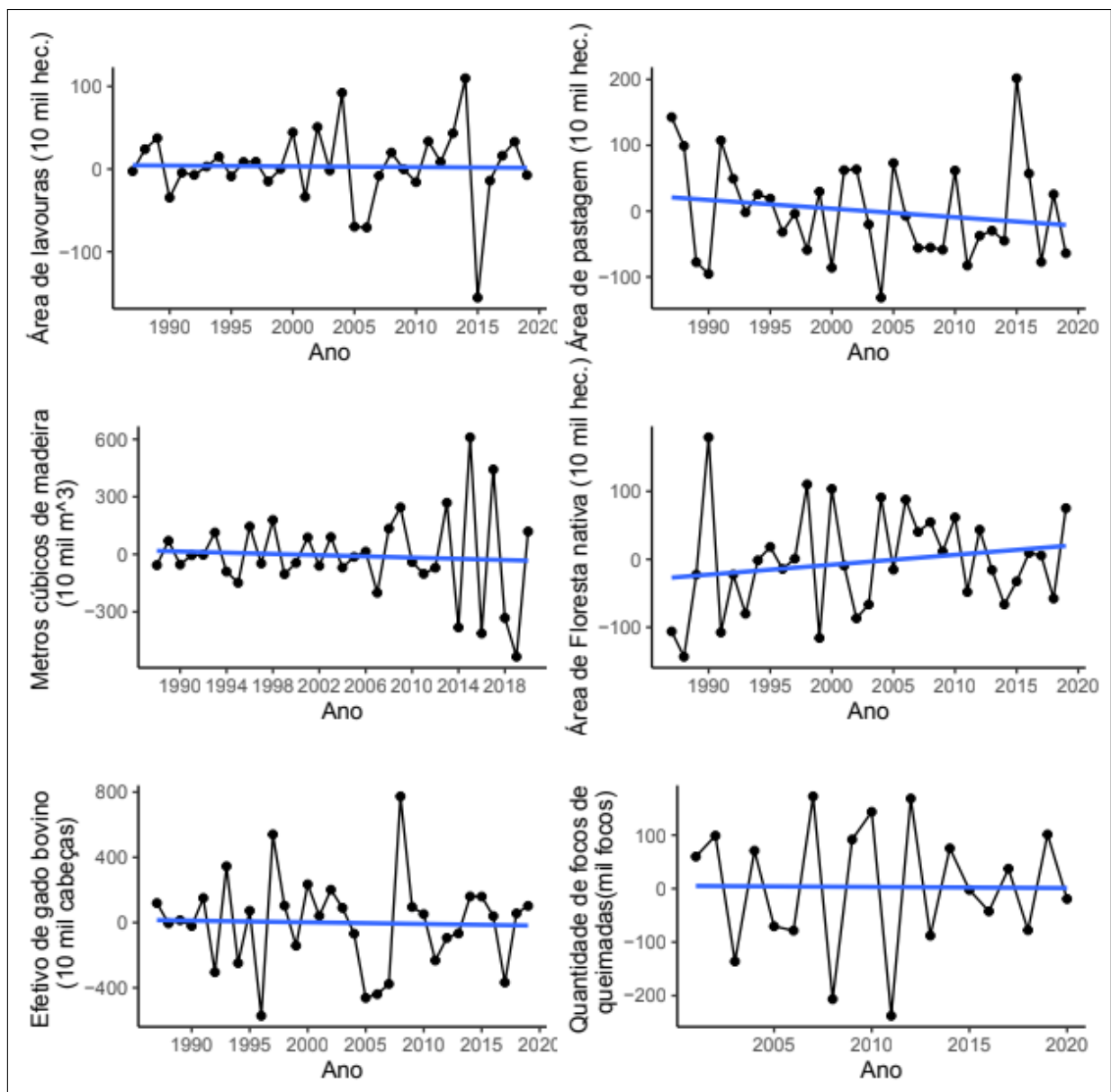


Figura 14: Séries temporais diferenciadas 2 vezes

Com o intuito de verificar o termo integrado(d) do modelo ARIMA, foi usado o teste aumentado de Dickey-Fuller para verificar a estacionariedade das séries. As séries foram diferenciadas até o teste não rejeitar hipótese de estacionariedade, sendo assim, as séries temporais relacionadas ao Efetivo Bovino, a Extração de Madeira e às áreas de Pastagem e Floresta Nativa necessitaram de duas diferenciações para se tornarem estacionárias, então o termo d do modelo ARIMA(p,d,q) será igual a 2 nesses casos. Já as séries temporais da área de Agricultura e quantidade de Focos de Queimadas foram diferenciadas uma vez até que o teste não rejeita-se a hipótese de estacionariedade, logo o valor do termo d será igual a 1 nesses casos. A série da área de Agricultura não convergiu com o resultado visual, onde, segundo o teste, a 1ª diferenciação já é suficiente para a estacionariedade.

Tabela 2: p-valores do teste aumentado de Dickey-Fuller para todas as séries temporais em estudo

Séries	Diferenciações		
	d=0	d=1	d=2
Área de Floresta nativa	0,93	0,16	<0,01
Área de Pastagem	0,96	0,16	<0,01
Extração de Madeira	0,65	0,33	<0,01
Efetivo bovino	0,56	0,08	0,02
Área de Agricultura	0,87	0,03	Não necessária
Quantidade de focos de queimadas	0,31	0,02	Não necessária

Os valores para p e q da $ARIMA(p,d,q)$ foram escolhidos através dos critérios de informação, onde combinações com p e $q \leq 2$ estão apresentadas no Apêndice 1. Para algumas séries temporais, devido ao fato dos critérios de informação serem bem próximos em diferentes modelos, foram utilizadas algumas métricas de qualidade de ajuste para realizar a escolha, que são apresentadas no Apêndice 2. Os melhores modelos podem ser observados na Tabela 3

Tabela 3: Melhores modelos baseados nos critérios de informação para combinações de p e $q \leq 2$ e métricas de qualidade de ajuste

Série	Melhor Modelo
Floresta Nativa	ARIMA(0,2,1)
Pastagem	ARIMA(2,2,0)
Extração de Madeira	ARIMA(0,2,1)
Efetivo Bovino	ARIMA(1,2,1)
Agricultura	ARIMA(1,1,0)
Focos de Queimadas	ARIMA(2,1,1)

Para avaliar a qualidade do ajuste a Tabela 4 apresenta os valores das métricas mais comuns, em que a de melhor compreensão é a MAPE, apresenta em termos percentuais. Aparentemente, os modelos se ajustam bem aos dados, pois apresentam pontual médio de erro menor que 3%, com exceção da série de focos de queimadas, que apresenta, em média,

mais de 30% de diferença absoluta entre os valores observados e os valores ajustados pelo modelo.

Tabela 4: Métricas de qualidade de ajuste dos modelos

Série	RSME	MAE	MAPE
Floresta Nativa	70,323	56,652	0,137%
Pastagem	67,947	54,417	0,704%
Extração de Madeira	166,521	104,234	16,258%
Efetivo Bovino	230,104	173,837	1,910%
Agricultura	42,806	28,188	2,329%
Focos de Queimadas	50,332	39,912	31,706%

Na Figura 15 há a comparação gráfica dos valores observados das séries e os valores ajustados pelos correspondentes modelos ARIMA. Percebe-se que a série temporal da quantidade de focos de queimadas apresenta um comportamento descompassado em relação ao modelo, onde os ajustes parecem estar deslocados no tempo, com um certo atraso.

A série de extração de madeira apresenta comportamento dos valores observados bem ajustados pelo modelo, porém, a partir de 2008 esses ajustes apresentam erros maiores, não aparentando ter uma boa qualidade de ajuste. Já as outras séries têm o comportamento muito bem ajustado pelos seus respectivos modelos, quase não havendo valores ajustados discrepantes dos valores efetivos.

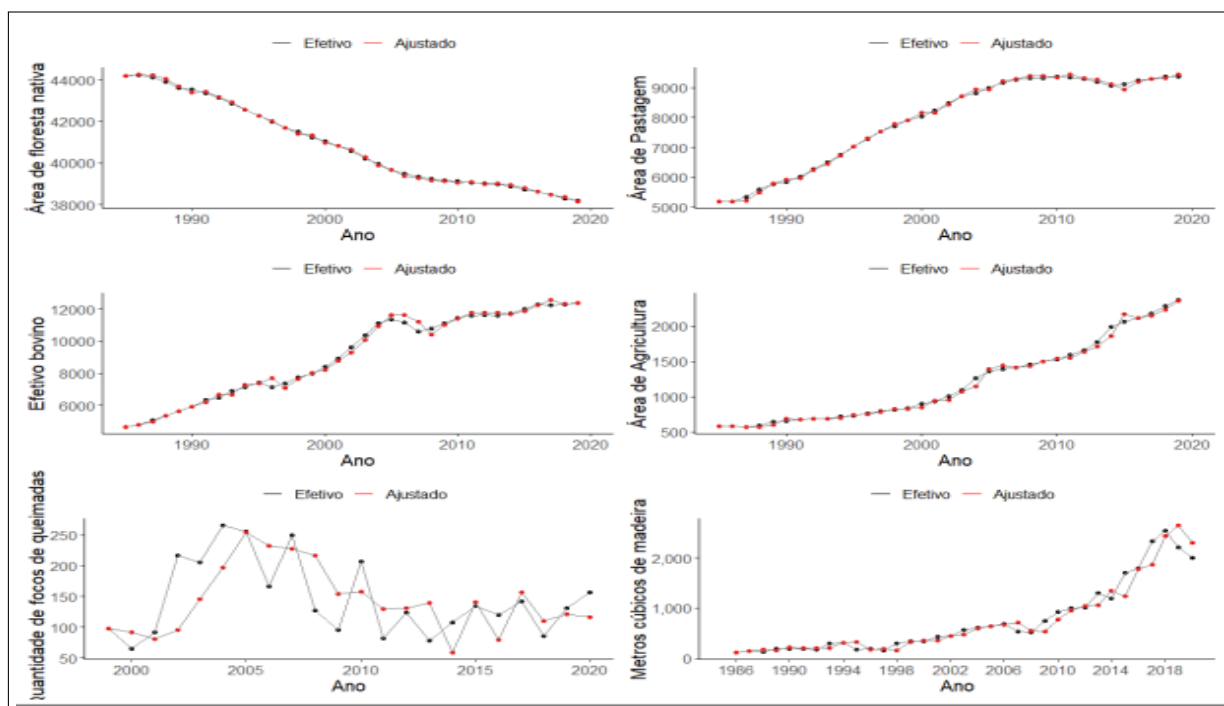


Figura 15: Valores efetivos e ajustados das séries temporais

Através do teste de Ljung-Box, apresentado na Tabela 5, percebe-se que todos os modelos possuem resíduos independentes. Porém pela Tabela 6, as séries temporais referente à área ocupada com Agricultura e Extração de Madeira não apresentam normalidade dos resíduos, o que indica que os modelos não seriam adequados para a realização de previsões. Já os modelos para as demais séries foram considerados adequados.

Tabela 5: Teste Ljung-Box para resíduos dos modelos ARIMA

Série	p-valor
Floresta Nativa	0,41
Pastagem	0,78
Extração de Madeira	0,80
Efetivo Bovino	0,77
Agricultura	0,30
Focos de Queimadas	0,61

Tabela 6: Teste Shapiro-Wilk para resíduos dos modelos ARIMA

Série	p-valor
Floresta Nativa	0,930
Pastagem	0,346
Extração de Madeira	0,002
Efetivo Bovino	0,085
Agricultura	0,008
Focos de Queimadas	0,952

Fazendo a análise da não normalidade das séries, foram encontradas poucas observações influentes, que geravam a não normalidade dos resíduos. Com o descarte delas, a normalidade passou a ser aceita. Não há nenhuma razão aparente para a exclusão dessas observações, então elas serão mantidas, mas a aproximação pela normal será considerada adequada e, assim, os modelos poderão ser utilizados na realização de previsões.

As previsões para os modelos ARIMA estão apresentados na Figura 16. Aparentemente todas as séries seguem o mesmo comportamento indicado na literatura, isto é, a série da Floresta Nativa decai ao longo dos anos, enquanto as outras tendem a aumentar ou permanecer aproximadamente constantes. Percebe-se que os intervalos de confiança crescem a cada ano, mostrando um erro maior na previsão.

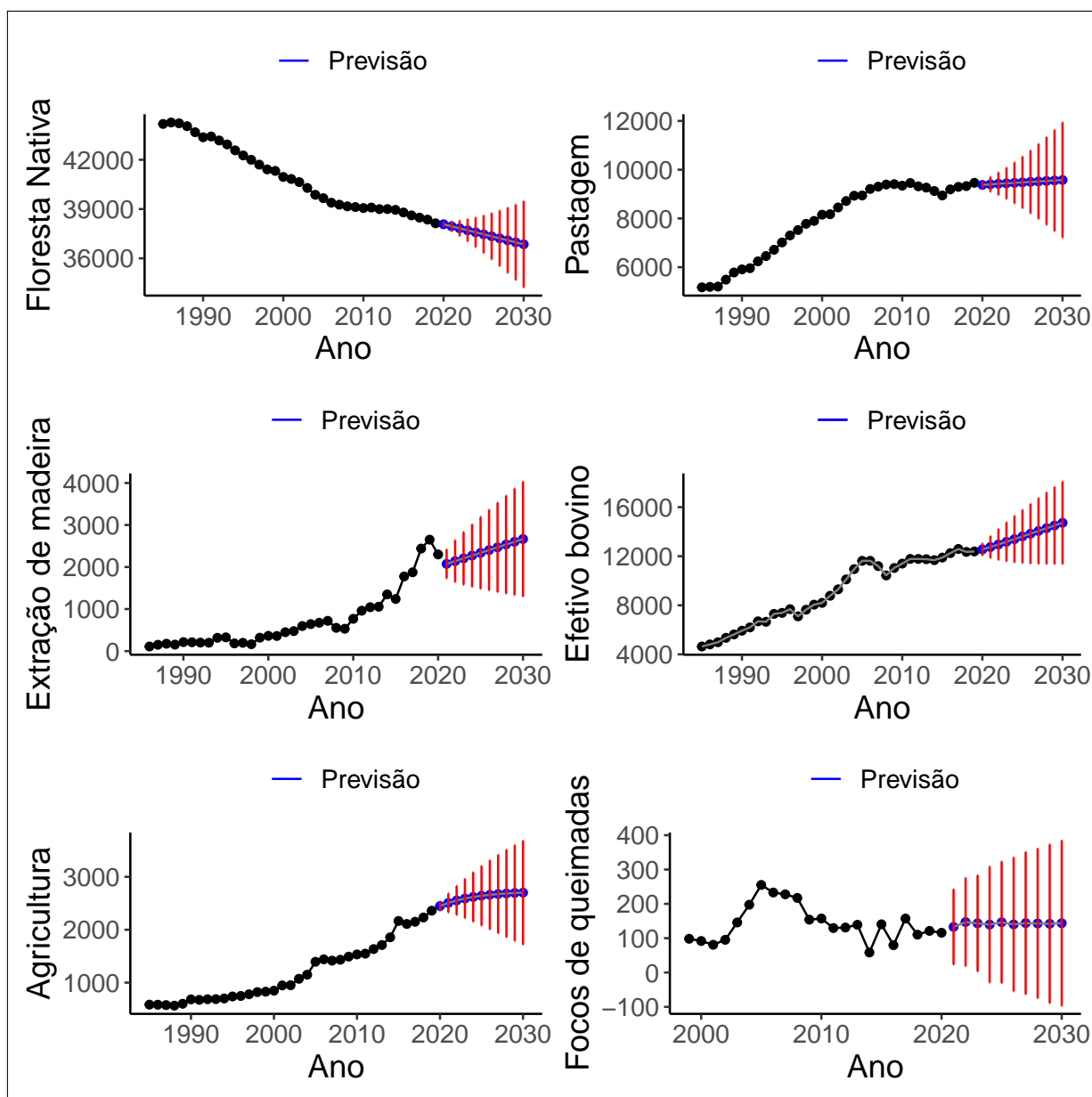


Figura 16: Previsões das séries temporais

A fim de comparar o desempenho do modelo de função de transferência para estimar a área desmatada a partir de seus indicadores, os últimos 5 anos, de 2015 a 2019, da série da Floresta Nativa, foram separados do processo de estimação do modelo. O objetivo é prever através de cada modelo os valores para esses anos, e medir através do MAPE qual seria o melhor modelo para realizar previsões.

Através da função de autocorrelação cruzada, apresentada na Figura 17, foi percebido que a única série auxiliar que possui lag significativo com a série alvo, é a série de Pastagem, sendo este o lag 0. Isto indica que a influência da pastagem no desmatamento ocorre no mesmo ano. Com intuito de verificar influências descritas na revisão bibliográfica, as

demais séries auxiliares serão utilizadas em modelos secundários, para analisar se podem melhorar as previsões, e o lag considerado como significativo também será o 0.

O passo seguinte foi identificar os valores de r e s para a função de transferência. Os valores encontrados para cada série auxiliar foram $r=2$ e $s=0$ para todas as séries.

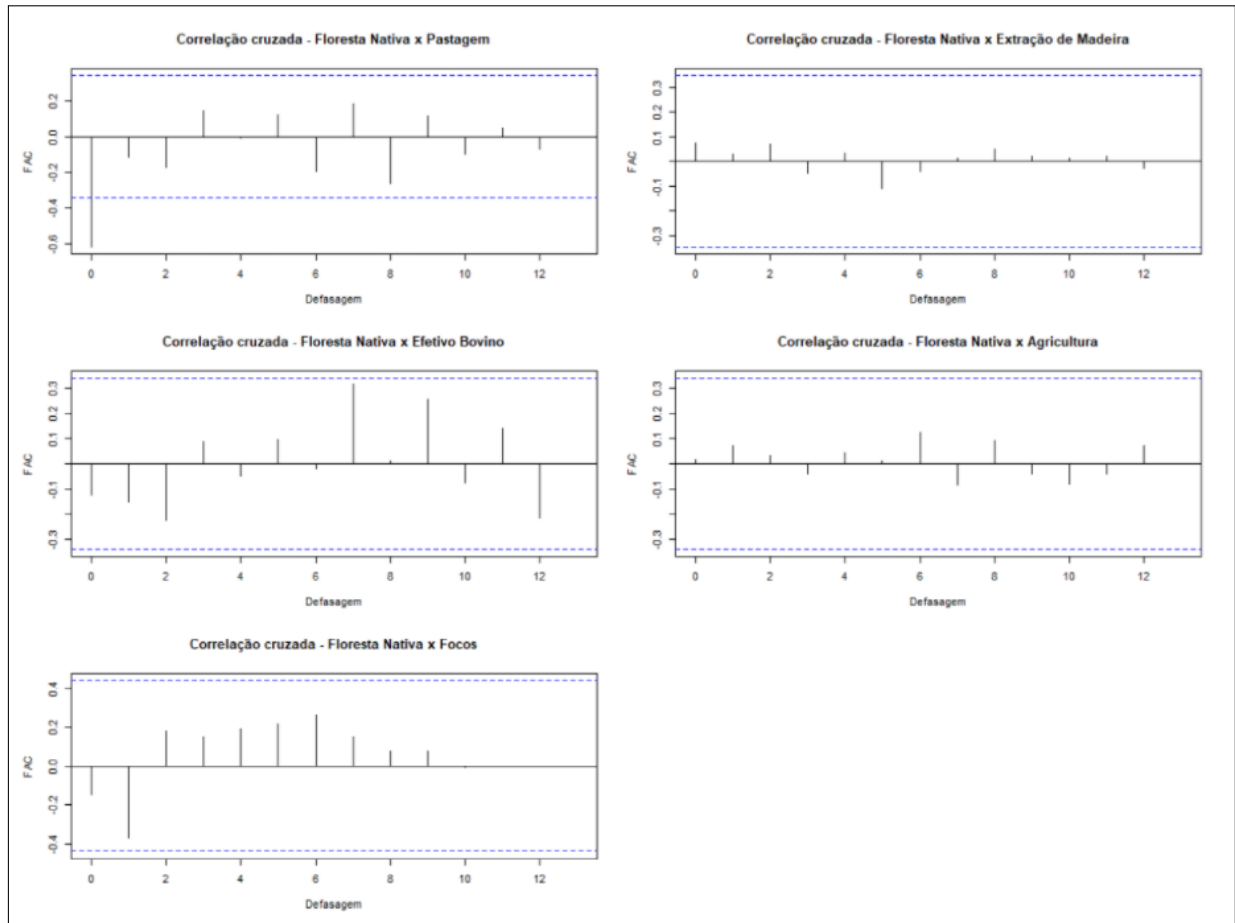


Figura 17: Função de correlação cruzada para cada uma das séries

A seguir, na Tabela 7, podemos verificar as medidas de AIC e MAPE para comparar a adequabilidade dos modelos e das previsões, respectivamente. A denominação dos modelos considerados é a seguinte:

- ARIMA = Série ARIMA sem Função de Transferência;
- FT1 = Série ARIMA com Função de Transferência utilizando a série de Pastagem;
- FT2 = Série ARIMA com Função de Transferência utilizando a série de Pastagem, Extração de Madeira, Efetivo Bovino e Agricultura;
- FT3 = Série ARIMA com Função de Transferência utilizando todas as séries.

Primeiramente foi incluído o modelo temporalmente correlacionado. Então, foram incluídos todos os modelos cuja qualidade de ajuste foi considerada parcialmente boa e por fim todas as séries foram incluídas.

Utilizando a função de transferência nenhuma das séries mostrou uma melhor acurácia na precisão do que o modelo ARIMA inicial. Também observa-se que todos os modelos com função de transferência, apesar de apresentarem menor AIC que o ARIMA puro, apresentaram previsões menos precisas, com exceção do modelo FT3.

Apesar da melhora na previsão, esta é muito pequena, e devido à complexidade do modelo que conta com 6 séries temporais ao todo. Desta forma, o ARIMA puro será utilizado como base para as previsões.

Tabela 7: Comparação dos modelos

Modelo	AIC	MAPE
ARIMA	325,69	0,353%
FT1	288,39	1,295%
FT2	262,98	0,443%
FT3	37,79	0,318%

Logo, as previsões de valores futuros da Floresta Nativa, correspondentes aos anos de 2020 a 2030, são os mesmos representados na Figura 16, e estes valores estão na Tabela 8.

Tabela 8: Previsões para a área de floresta nativa para as Regiões

Ano	Previsões	Intervalo de 95%
2020	380.732.257	(379.290.789; 382.173.726)
2021	379.512.867	(376.570.618; 382.455.115)
2022	378.293.476	(373.575.273; 383.011.679)
2023	377.074.085	(370.334.938; 383.813.232)
2024	375.854.695	(366.874.423; 384.834.966)
2025	374.635.304	(363.212.808; 386.057.800)
2026	373.415.913	(359.365.115; 387.466.712)
2027	372.196.523	(355.343.502; 389.049.544)
2028	370.977.132	(351.158.051; 390.796.213)
2029	369.757.741	(346.817.288; 392.698.195)
2030	368.538.351	(342.328.543; 394.748.158)

Com o intuito de gerar uma análise mais pontual, foram realizadas modelagens para os estados dessas Regiões que mais foram afetados pelo desmatamento, sendo estes os estados do Pará e Mato Grosso.

3.2 Análises e resultados para os estados de Pará e Mato Grosso

Ainda seguindo resultados da bibliografia, será realizada uma análise para os estados do Pará e Mato Grosso, que são relatados como os estados em que o desmatamento está mais presentes nas Regiões Norte e Centro-Oeste. Os estados foram agrupados pois a fronteira destes é o atual caminho do desmatamento.

Os dados para os estados são das mesmas fontes e possuem o mesmo espaço temporal dos dados para as Regiões. O intuito é analisar se utilizar modelos com função de transferência podem trazer previsões com maior acurácia ao se olhar para uma região reduzida.

Já pela Figura 18 percebe-se que os comportamentos de quase todas as séries seguem os padrões observados para as Regiões Norte e Centro-Oeste. A única que difere é a série temporal da Extração de Madeira, sendo assim tiramos as mesmas conclusões que para as Regiões. Ou seja, a área de floresta nativa diminuindo, enquanto os demais indicadores aumentam.

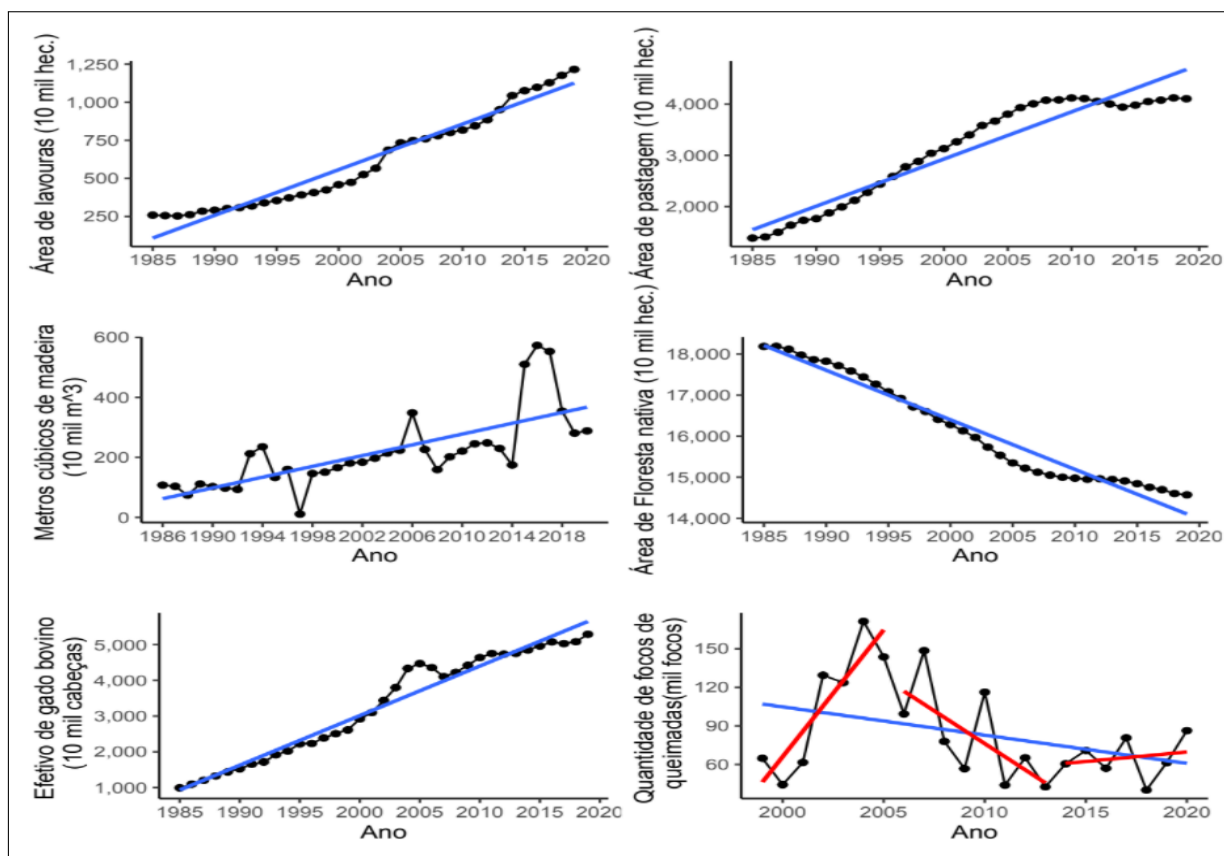


Figura 18: Séries temporais para os estados do Pará e Mato Grosso e suas tendências

Aparentemente, pode-se concluir que os estados do Pará e Mato Grosso são os prováveis responsáveis pelos comportamentos das séries para as Regiões Norte e Centro-Oeste. Pela Tabela 9 observa-se que, na maioria dos casos, pelo menos 40% dos valores observados para o ano inicial ou o ano final das Regiões são provenientes desses estados.

Tabela 9: Representatividade das séries dos estados do Pará e Mato Grosso em relação ao total das Regiões Norte e Centro-Oeste

Ano	Série	Representatividade
1985	Floresta Nativa	41%
2019	Floresta Nativa	38%
1985	Pastagem	27%
2019	Pastagem	44%
1985	Agricultura	44%
2019	Agricultura	51%
1986	Extração de Madeira	99%
2020	Extração de Madeira	14%
1985	Efetivo Bovino	21%
2020	Efetivo Bovino	44%
1999	Focos de Queimadas	66%
2020	Focos de Queimadas	55%

Realizando os mesmos procedimentos para encontrar o melhor modelo, os modelos ARIMA que obtiveram as melhores métricas são os que estão na Tabela 10. A comparação dos modelos foi apresentada no Apêndice 3.

Tabela 10: Melhores modelos, no ambiente espacial reduzido, baseados nos critérios de informação para combinações de p e $q \leq 2$ e métricas de qualidade de ajuste

Série	Melhor Modelo
Floresta Nativa	ARIMA(1,2,0)
Pastagem	ARIMA(1,2,0)
Extração de Madeira	ARIMA(1,0,0)
Efetivo Bovino	ARIMA(0,1,2)
Agricultura	ARIMA(1,2,0)
Focos de Queimadas	ARIMA(2,1,1)

Pela Figura 19 percebe-se que os modelos se comportam de forma parecida aos obtidos para as Regiões Norte e Centro-Oeste, apresentando o mesmo problema para a série de Focos de Queimada, onde parece apresentar atraso do ajuste se comparado ao valor observado.

A série de Extração de Madeira apresenta qualidade de ajuste um pouco superior, se comparado ao modelo para as Regiões, e só perde a qualidade de ajuste a partir de 2014. Já as demais séries apresentam bons ajustes.

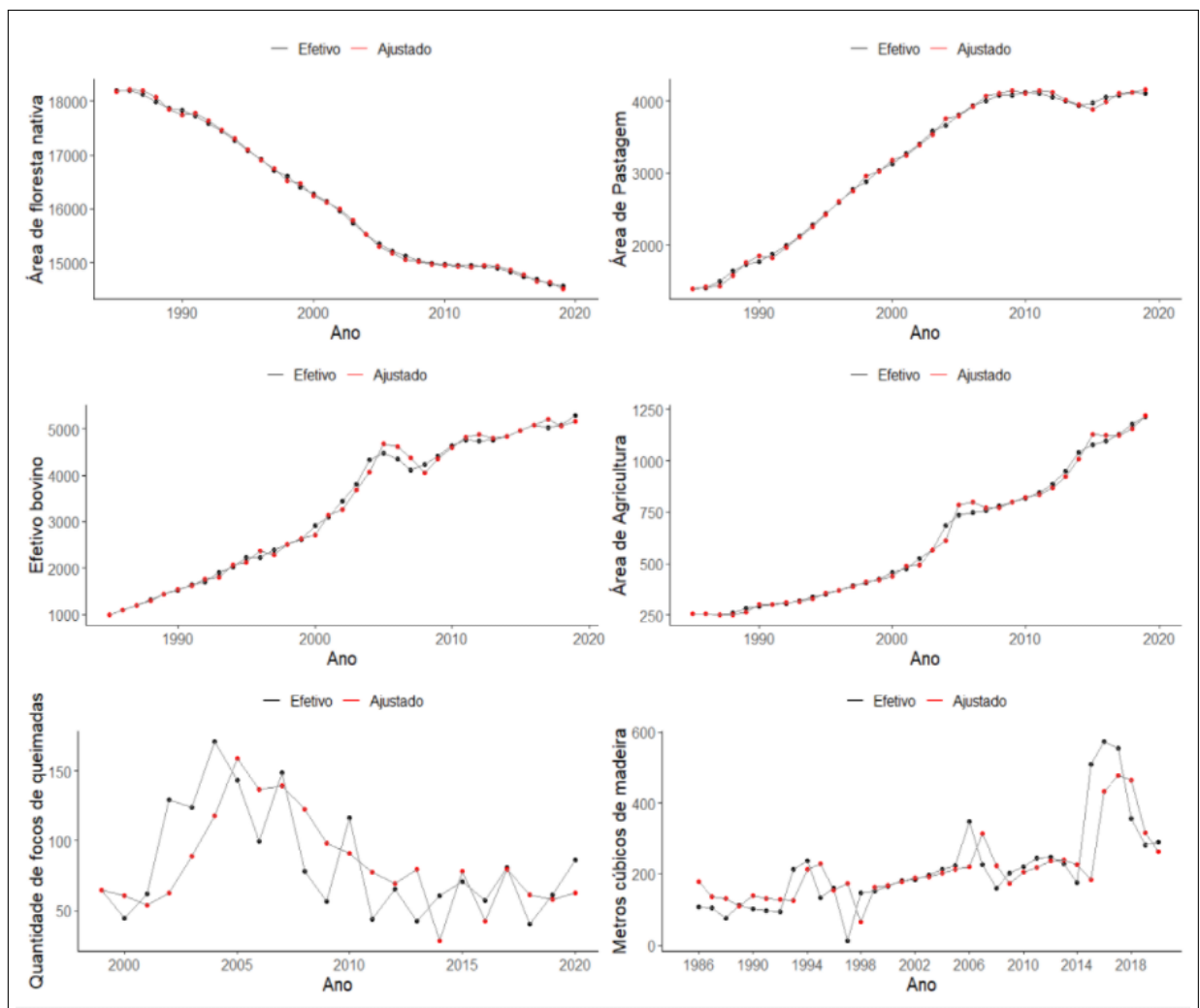


Figura 19: Valores efetivos e ajustados das séries temporais para os estados

Pelos testes de Ljung-Box e Shapiro-Wilk, apresentados nas Tabelas 11 e 12, apenas duas séries não apresentam resíduos com comportamentos normais.

Tabela 11: Teste Ljung-Box para resíduos dos modelos ARIMA

Série	p-valor
Floresta Nativa	0,878
Pastagem	0,945
Extração de Madeira	0,638
Efetivo Bovino	0,829
Agricultura	0,591
Focos de Queimadas	0,672

Novamente duas observações influentes foram encontradas, e com suas exclusões a normalidade não foi rejeitada. Sendo assim, o mesmo procedimento adotado para as Regiões foi adotado aqui, ou seja, os valores não serão, de fato, excluídos, porém os modelos serão considerados válidos para realização de previsões.

Tabela 12: Teste Shapiro-Wilk para resíduos dos modelos ARIMA

Série	p-valor
Floresta Nativa	0,7298
Pastagem	0,5704
Extração de Madeira	0,3242
Efetivo Bovino	0,0004
Agricultura	0,6453
Focos de Queimadas	0,0011

Com a validação dos modelos as previsões foram realizadas e podem ser observadas na Figura 20. Observa-se que o comportamento das previsões aparentam seguir o comportamento histórico, porém a série de Pastagem parece estabilizar em um patamar bem alto. Novamente o intervalo de confiança cresce muito ao decorrer do tempo, devido ao alto erro para previsões muitos distantes.

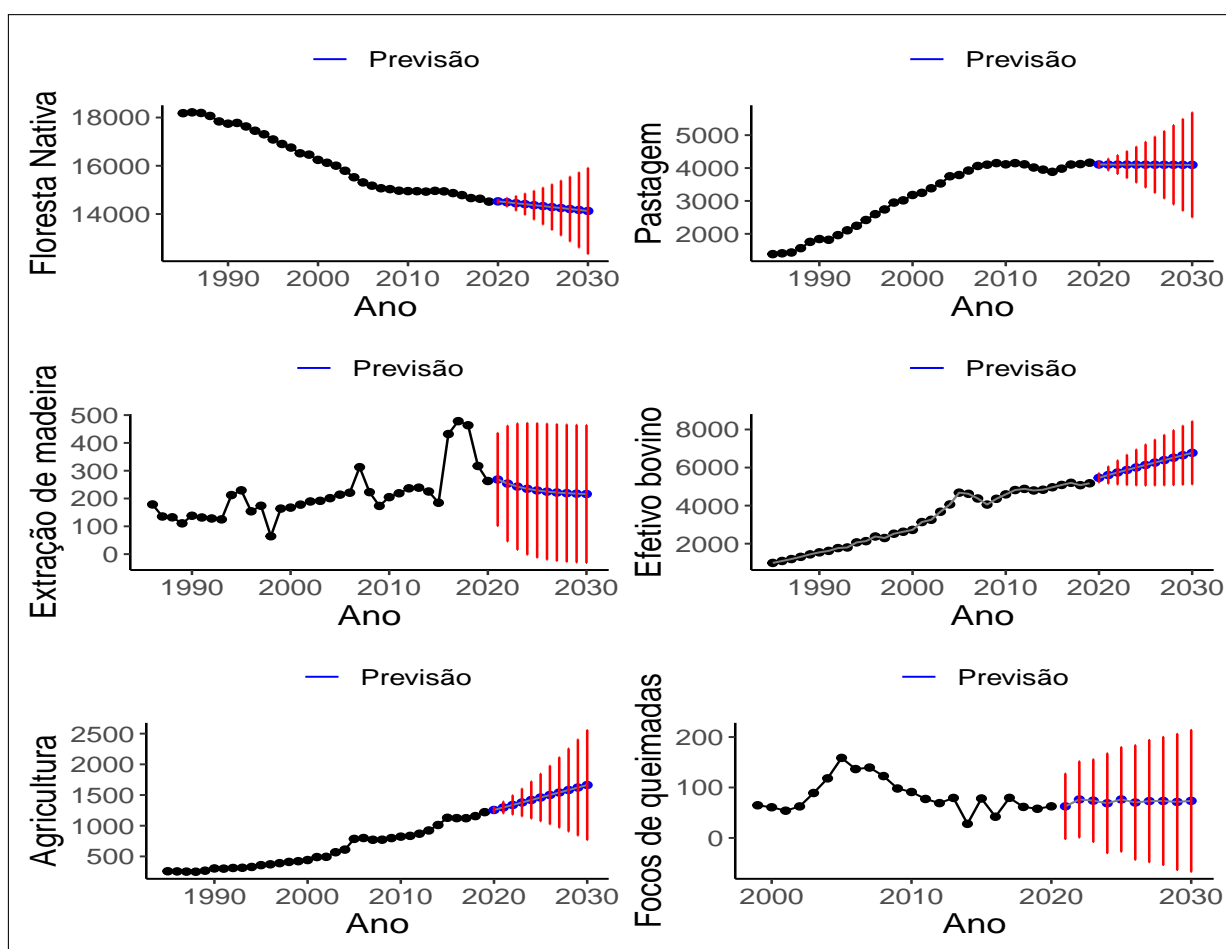


Figura 20: Previsões das séries temporais para os estados

A fim de verificar se há um modelo que possa gerar melhores estimativas para as previsões do que o ARIMA puro, serão utilizados, novamente, modelos ARIMA com

função de transferência.

Através da função de autocorrelação cruzada, apresentada na Figura 21, foi percebido que as únicas séries auxiliares que possuem lag significativo com a série alvo, são as de Pastagem e a de Focos de Queimadas, com lags 0 e 1, respectivamente. Com intuito de verificar influências descritas na bibliografia, as demais séries auxiliares serão utilizadas em modelos secundários para analisar se podem melhorar as previsões, e o lag considerado como significativo será 0.

O passo seguinte foi identificar os valores de r e s para a função de transferência, e os valores encontrados para cada série auxiliar foram $r=2$ e $s=0$ para todas as séries.

Vale ressaltar que devido ao fato da série de focos de queimadas estar correlacionada com a série de floresta nativa com lag 1, o primeiro significativo, foi necessário trabalhar com os valores desta série atrasados em 1 ano.

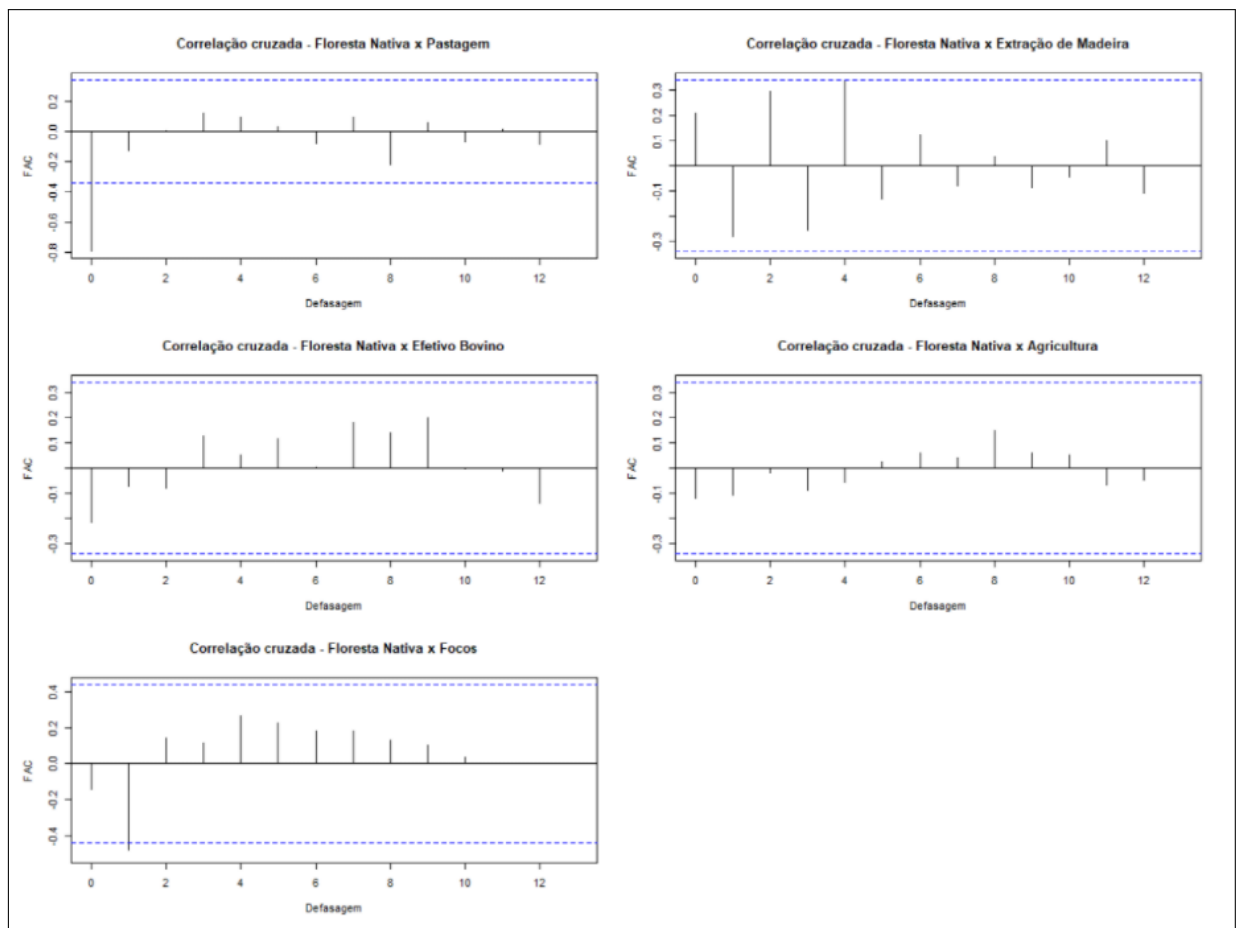


Figura 21: Funções de correlação cruzada

A seguir, na Tabela 13, podemos verificar as medidas de AIC e MAPE para comparar a adequabilidade dos modelos e das previsões, respectivamente. Os modelos são os seguintes:

- ARIMA = Série ARIMA sem Função de Transferência;
- FT1 = Série ARIMA com Função de Transferência utilizando a série de Pastagem;
- FT2 = Série ARIMA com Função de Transferência utilizando a série de Focos de Queimadas;
- FT3 = Série ARIMA com Função de Transferência utilizando a série de Pastagem e Focos de Queimadas;
- FT4= Série ARIMA com Função de Transferência utilizando todas as séries.

Primeiro foram adicionadas as séries temporalmente correlacionadas, separadamente. Então, as duas séries correlacionadas foram conjuntamente adicionadas. Por fim, todas as séries foram utilizadas.

Utilizando a função de transferência, o modelo FT3 apresentou melhores métricas que o modelo ARIMA puro, melhorando consideravelmente as previsões.

Tabela 13: Comparação dos modelos

Modelo	AIC	MAPE
ARIMA	297,91	0,81%
FT1	252,32	1,19%
FT2	121,75	0,65%
FT3	123,43	0,18%
FT4	72,83	0,76%

Pela Figura 22 é visualmente perceptível essa melhora nas previsões, onde as estimativas oriundas do Modelo FT3 se aproximam muito mais dos valores reais da série histórica.

As previsões para os anos futuros, de 2020 a 2030, estão apresentados visualmente na Figura 23, e esses valores são apresentados na Tabela 14.

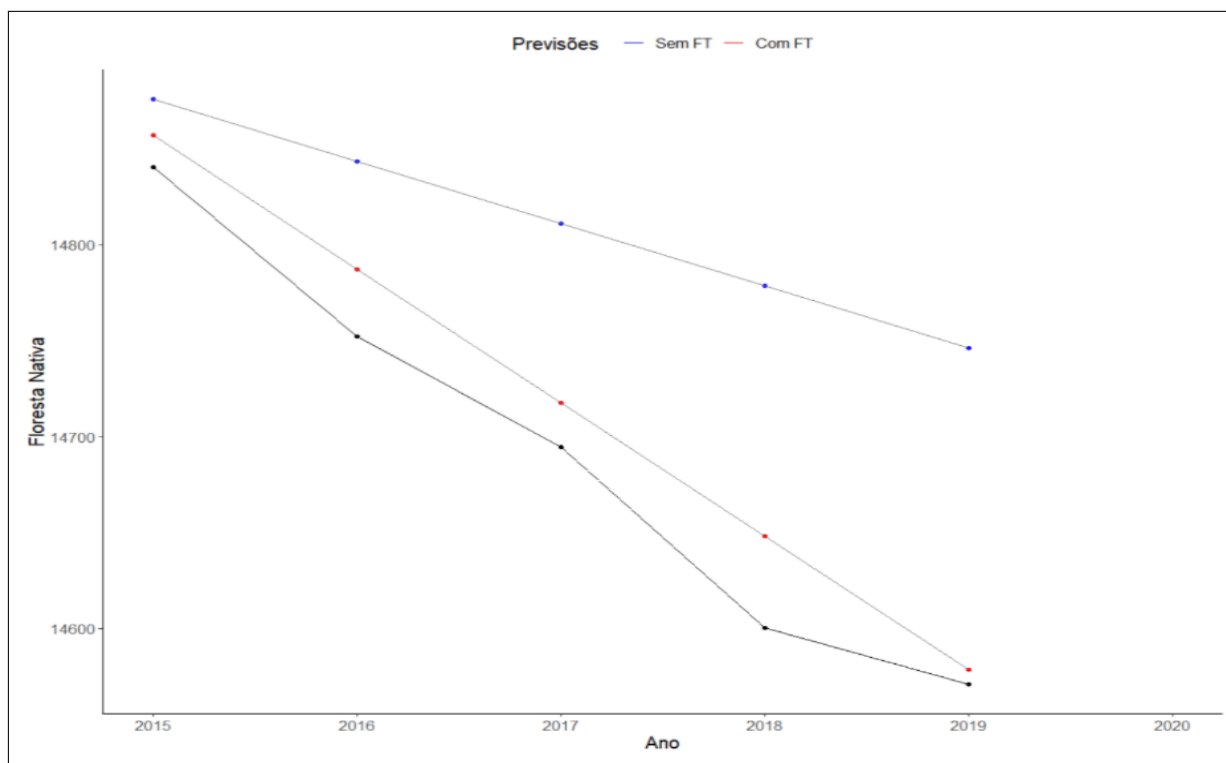


Figura 22: Comparando previsões do ARIMA puro e da FT3

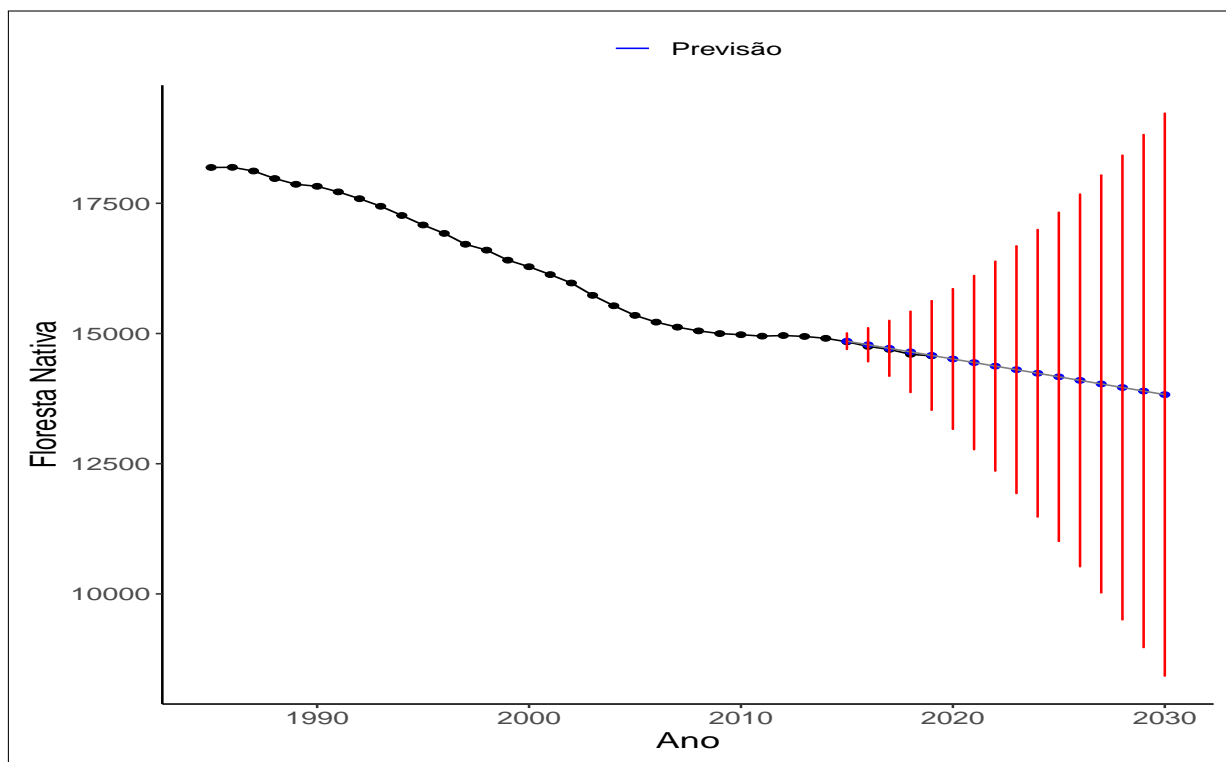


Figura 23: Previsões geradas pelo modelo FT3

Tabela 14: Previsões para a área de floresta nativa nos estados do Pará e Mato Grosso

Ano	Previsões	intervalo de 95%
2020	145.104.777	(131.634.305;158.575.250)
2021	144.420.970	(127.737.343;161.104.596)
2022	143.737.162	(123.624.120;163.850.203)
2023	143.053.354	(119.308.927;166.797.781)
2024	142.369.546	(114.803.674;169.935.418)
2025	141.685.738	(110.118.485;173.252.992)
2026	141.001.931	(105.262.093;176.741.768)
2027	140.318.123	(100.242.125;180.394.120)
2028	139.634.315	(95.065.301;184.203.329)
2029	138.950.507	(89.737.587;188.163.427)
2030	138.266.699	(84.264.314;192.269.085)

O desmatamento, aparentemente, seguirá uma tendência linear nos próximos anos, não apresentando nenhuma diminuição desse ritmo ao longo do tempo.

4 Conclusões

Para as regiões Norte e Centro-Oeste foi possível identificar uma correlação temporal entre as séries do Desmatamento e a de Pastagem. Porém modelos mais complexos, utilizando de funções de transferência, não foram capazes de, significativamente, gerar previsões melhores que um ARIMA simples.

Logo, para essas regiões, entre os anos de 2020 e 2030, é esperado uma média de 1.219.386 hectares desmatados a cada ano. O patamar elevado de desmatamento deve, de fato, ser algo a se preocupar. Porém, o nível ficará estável, onde a média será a área desmatada a cada ano, o que facilitaria que ações governamentais contivessem o excesso da área desmatada.

Já para os estados do Pará e do Mato Grosso houve duas séries temporalmente correlacionadas com a série do Desmatamento, sendo elas a de Pastagem e a de Focos de Queimadas. O modelo com função de transferência gerou previsões 88% melhores que o ARIMA simples.

O valor esperado para os anos de 2020 à 2030 é de 696.700 hectares por ano. Comprovadamente, esses estados serão responsáveis por mais de 50% do desmatamento das regiões Norte e Centro-Oeste.

Apesar da bibliografia sugerir correlação do desmatamento com diversas outras características, como gado, agricultura e extração de madeira, apenas a área de pastagem e a quantidade de focos de queimadas foram apontadas como temporalmente correlacionadas com a área de floresta nativa.

A parcial discrepância em relação às conclusões da revisão bibliográfica pode ser devida a alguns fatores, como o fato de se ter trabalhado com séries serem anuais. Valores mensais para as séries temporais seriam mais adequados, porém esses dados não estavam disponíveis e a busca ativa por esses dados é uma sugestão para trabalhos futuros.

A função de transferência se mostrou, em alguns casos, mais eficiente para previsões

que o ARIMA simples, isso mostra a importância de estudos mais profundos para identificação das causas do desmatamento, para assim conseguir abordar a solução de maneira mais ampla e objetiva.

Para trabalhos futuros, sugestões incluem a utilização de Modelos Dinâmicos Bayesianos e Modelos Hierárquicos, com o objetivo de tentar melhorar o poder de explicação e também melhorar a identificação da relação entre as séries temporais estudadas.

Por fim, para análises futuras, avaliar as Regiões e os estados de forma separada é uma última sugestão.

Referências

- AFONSO, A. M.; AQUINO, F. de G.; CAMARGO, A. J. A. de. *Cerrado*. Acesso em: 17 jul 2021. Disponível em: <https://www.embrapa.br/contando-ciencia/bioma-cerrado>.
- Agência Brasil - Brasília. *Confira discurso do presidente Bolsonaro na Cúpula do Clima*. 2021. Acesso em: 01 jul 2021. Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2021-04/confira-discurso-do-presidente-bolsonaro-na-cupula-do-clima>.
- ANGELO, H.; Sá, S. P. P. de. O desflorestamento na amazônia brasileira. *Ciência Florestal*, 2007. Disponível em: <https://www.redalyc.org/articulo.oa?id=53417304>.
- ASNER, G. et al. Selective logging and its relation to deforestation. In: _____. [S.l.: s.n.], 2013. v. 186, p. 25–42.
- ASNER, G. et al. Selective logging in the brazilian amazon. *Science (New York, N.Y.)*, v. 310, p. 480–2, 11 2005.
- AUGUIE, B. *gridExtra: Miscellaneous Functions for "Grid" Graphics*. [S.l.], 2017. R package version 2.3. Disponível em: <https://CRAN.R-project.org/package=gridExtra>.
- Banco Mundial. Agriculture and food. Acesso em: 01 jul 2021. 2020. Disponível em: <https://www.worldbank.org/en/topic/agriculture/overview>.
- BOX, G. E. P.; JENKINS, G. M. *Time Series Analysis: Forecasting and Control*. Revisada. [S.l.]: Holden-Day, 1976.
- BOX, G. E. P. et al. *Time Series Analysis: Forecasting and Control*. 5. ed. [S.l.]: John Wiley and Sons, Inc., 2016.
- BROCKWELL, P. J.; DAVIS, R. A. *Introduction to Time Series and Forecasting*. 2. ed. [S.l.]: Springer texts in statistics, 2002.
- BROOKS, C. *Introductory Econometrics for Finance*. 2. ed. [S.l.]: Cambridge University Press, 2008.
- Centro de Estudos Avançados em Economia Aplicada. *PIB DO AGRONEGÓCIO BRASILEIRO*. 2021. Acesso em: 22 ago 2021. Disponível em: <https://www.cepea.esalq.usp.br/br/pib-do-agronegocio-brasileiro.aspx>.
- FALEIRO, F. G.; NETO, A. L. de F. (Ed.). *Savanas: Demandas para Pesquisa*. 1. ed. [S.l.]: Embrapa, 2009.
- FERREIRA, P. G. C.; SOUZA, R. C.; MATTOS, D. M. Box & jenkins com função de transferência usando o r: Um estudo para a série de ena do subsistema sul. *XLVII SBPO*, 2015.

Food and Agriculture Organization. Global forest resources assessment 2020 – key findings. Acesso em: 22 jul 2021. 2020. Disponível em: <http://www.fao.org/documents/card/en/c/ca8753en>.

FOUNTAIN, H. Tropical forest destruction accelerated in 2020. *The New York Times*, 2021.

GALLERY, R. E. Ecology of tropical rain forests. In: _____. [S.l.]: Ecology and the Environment, 2014. p. 247–272.

HYNDMAN, R. et al. *forecast: Forecasting functions for time series and linear models*. [S.l.], 2021. R package version 8.15. Disponível em: <https://pkg.robjhyndman.com/forecast/>.

HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: principles and practice*. 2. ed. [S.l.]: OTexts: Melbourne, Australia, 2018.

HYNDMAN, R. J.; KHANDAKAR, Y. Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, v. 26, n. 3, p. 1–22, 2008. Disponível em: <https://www.jstatsoft.org/article/view/v027i03>.

IBGE. *Biomass*. 2019. Acesso em: 22 ago 2021. Disponível em: <https://www.ibge.gov.br/geociencias/cartas-e-mapas/informacoes-ambientais/15842-biomass.html?=&t=acesso-ao-produto>.

IBGE. *Pesquisa da Pecuária Municipal*. 2020. Acesso em 22 ago 2021. Disponível em: <https://sidra.ibge.gov.br/pesquisa/ppm/tabelas/brasil/2019>.

IBGE. *Pesquisa da Pecuária Municipal*. 2020. Acesso em: 22 ago 2021. Disponível em: <https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9107-producao-da-pecuaria-municipal.html?=&t=o-que-e>.

IBGE. *Produção da Extração Vegetal e da Silvicultura*. 2020. Acesso em: 25 nov 2021. Disponível em: <https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9105-producao-da-extracao-vegetal-e-da-silvicultura.html?=&t=o-que-e>.

IBGE. *Produção da Extração Vegetal e da Silvicultura*. 2020. Acesso em: 25 nov 2021. Disponível em: <https://sidra.ibge.gov.br/tabela/291>.

INPE. *Apresentação*. 2011. Acesso em: 22 ago 2021. Disponível em: <https://queimadas.dgi.inpe.br/queimadas/portal/informacoes/apresentacao>.

INPE. *Quais satélites são usados e onde são recebidas e processadas as imagens?* 2020. Acesso em: 22 ago 2021. Disponível em: <https://queimadas.dgi.inpe.br/queimadas/portal/informacoes/perguntas-frequentes#p6>.

INPE. *Monitoramento do Desmatamento da Floresta Amazônica Brasileira por Satélite*. 2021. Acesso em: 22 ago 2021. Disponível em: <http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>.

- INPE. *Monitoramento dos Focos Ativos por Região*. 2021. Acesso em: 22 ago 2021. Disponível em: https://queimadas.dgi.inpe.br/queimadas/portal-static/estatisticas_estados/.
- Instituto Chico Mendes de Biologia. *Ameaças*. Acesso em: 22 ago 2021. Disponível em: <https://www.icmbio.gov.br/cbc/conservacao-da-biodiversidade/ameacas.html>.
- KLEIN, A. L. (Ed.). *Eugen Warming e o cerrado brasileiro: um século depois*. [S.l.]: Editora UNESP, 2002.
- LAMAS, F. M. A tecnologia na agricultura. 2017. Acesso em: 01 jul 2021. Disponível em: <https://www.embrapa.br/busca-de-noticias/-/noticia/30015917/artigo-a-tecnologia-na-agricultura>.
- LEHMANN, C. E. R. Savannas need protection. *Science*, v. 327, 2010.
- MAPBIOMAS. *O PROJETO*. 2016. Acesso em: 22 ago 2021. Disponível em: <https://mapbiomas.org/o-projeto>.
- MAPBIOMAS. *CONHEÇA OS PASSOS DA METODOLOGIA MAPBIOMAS*. 2020. Acesso em: 22 ago 2021. Disponível em: <https://mapbiomas.org/atbd---entenda-cada-etapa>.
- MAPBIOMAS. *ESTATÍSTICAS*. 2020. Acesso em 22 ago 2021. Disponível em: <https://mapbiomas.org/estatisticas>.
- MARGULIS, S. *Causas do Desmatamento da Amazônia Brasileira*. 1. ed. [S.l.]: Banco Mundial, 2003.
- MORETTIN, P. A.; TOLOI, C. M. de C. *Análise de séries temporais*. 2. ed. [S.l.]: Edgard Blucher, 2006.
- OLIVEIRA, R. C. et al. Desmatamento e crescimento econômico no Brasil: uma análise da curva de Kuznets ambiental para a Amazônia legal. *Sociedade Brasileira de Economia e Sociologia Rural*, v. 49, 2011.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2020. Disponível em: <http://www.R-project.org/>.
- RITCHIE, H. *Deforestation and Forest Loss*. 2020. Acesso em: 01 jul 2021. Disponível em: <https://ourworldindata.org/deforestation>.
- RStudio Team. *RStudio: Integrated Development Environment for R*. Boston, MA, 2020. Disponível em: <http://www.rstudio.com/>.
- SAID, S. E.; DICKEY, D. A. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, [Oxford University Press, Biometrika Trust], v. 71, n. 3, p. 599–607, 1984. Disponível em: <http://www.jstor.org/stable/2336570>.
- SARHAN, A. E.; GREENBERG, B. G. Estimation of location and scale parameters for the rectangular population from censored samples. *Journal of the Royal Statistical Society. Series B (Methodological)*, [Royal Statistical Society, Wiley], v. 21, n. 2, p. 356–363, 1959.

- SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). *Biometrika*, [Oxford University Press, Biometrika Trust], v. 52, n. 3/4, p. 591–611, 1965.
- SOLBRIG, O.; MEDINA, E.; SILVA, J. *Biodiversity and Savanna Ecosystem Processes: A Global Perspective*. [S.l.]: John Wiley and Sons Ltd, 1996.
- STOODI. *Amazônia: características, impactos ambientais e mais!* 2020. Acesso em: 17 jul 2021. Disponível em: <https://www.stoodi.com.br/blog/geografia/amazonia-entenda-mais/>.
- SY, V. D. et al. Land use patterns and related carbon losses following deforestation in south america. 2015. Disponível em: <https://iopscience.iop.org/article/10.1088/1748-9326/10/12/124004/meta>.
- TERRABRASILIS. *Downloads*. 2021. Acesso em: 22 ago 2021. Disponível em: <http://terrabrasilis.dpi.inpe.br/downloads/>.
- TERRABRASILIS. *Incremento no desmatamento do Bioma Amazônia à partir de 2008*. 2021. Acesso em: 22 ago 2021. Disponível em: <http://terrabrasilis.dpi.inpe.br/geonetwork/srv/eng/catalog.search#/metadata/b75b83db-8026-43f9-9537-ee1dfa308158>.
- The Nature Conservancy. *Amazônia Um gigantesco patrimônio brasileiro*. Acesso em: 22 ago 2021. Disponível em: https://www.tnc.org.br/sobre-a-tnc/onde-trabalhamos/amazonia/?gclid=Cj0KCQjwiqWHBhD2ARIsAPCDzakXIANLeHFextAw3R7E5DMzLUI3oyQvEMhxq6ie2Y0w_6DajK8VF3kaApnOEALw_wcB.
- TOLLEFSON, J. Deforestation spikes in brazilian amazon. *Nature*, 2016. Disponível em: <https://doi-org.ez24.periodicos.capes.gov.br/10.1038/nature.2016.21083>.
- TRAPLETTI, A.; HORNIK, K. *tseries: Time Series Analysis and Computational Finance*. [S.l.], 2020. R package version 0.10-48. Disponível em: <https://CRAN.R-project.org/package=tseries>.
- WICKHAM, H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN 978-3-319-24277-4. Disponível em: <https://ggplot2.tidyverse.org>.
- WICKHAM, H. et al. *dplyr: A Grammar of Data Manipulation*. [S.l.], 2021. R package version 1.0.7. Disponível em: <https://CRAN.R-project.org/package=dplyr>.
- WICKHAM, H.; SEIDEL, D. *scales: Scale Functions for Visualization*. [S.l.], 2020. R package version 1.1.1. Disponível em: <https://CRAN.R-project.org/package=scales>.
- WWF-Brasil. *Agricultura e pecuária*. 2005. Acesso em: 6 jul 2021. Disponível em: [https://www.wwf.org.br/natureza_brasileira/areas_prioritarias/amazonia1/ameacas_riscos_amazonia/](https://www.wwf.org.br/natureza_brasileira/areas_prioritarias/amazonia1/ameacas_riscos_amazonia/desmatamento_na_amazonia/agropecuaria_na_amazonia/).
- WWF-Brasil. *Ameaças*. 2005. Acesso em: 6 jul 2021. Disponível em: https://www.wwf.org.br/natureza_brasileira/areas_prioritarias/amazonia1/ameacas_riscos_amazonia/.

APÊNDICE 1 – Comparação de modelos para as Regiões

Tabela 15: Comparação de modelos baseados nos critérios de informação

Modelos	Critérios de informação	
	AIC	BIC
Floresta Nativa		
ARIMA(0,2,0)	378,78	380,28
ARIMA(0,2,1)	379,59	382,59
ARIMA(1,2,0)	379,80	382,79
ARIMA(1,2,1)	381,39	385,88
ARIMA(0,2,2)	381,41	385,90
ARIMA(2,2,0)	381,52	386,01
ARIMA(1,2,2)	382,57	388,56
ARIMA(2,2,1)	383,39	389,37
ARIMA(2,2,2)	383,21	390,69
Pastagem		
ARIMA(0,2,0)	380,20	381,70
ARIMA(0,2,1)	381,99	384,98
ARIMA(1,2,0)	382,14	385,13
ARIMA(1,2,1)	382,48	386,97
ARIMA(0,2,2)	380,38	384,87
ARIMA(2,2,0)	380,29	384,78
ARIMA(1,2,2)	382,37	388,35
ARIMA(2,2,1)	382,06	388,04
ARIMA(2,2,2)	384,02	391,50
Extração de Madeira		
ARIMA(0,2,0)	452,67	454,16
ARIMA(0,2,1)	439,41	442,40
ARIMA(1,2,0)	443,40	446,39
ARIMA(1,2,1)	441,04	445,53
ARIMA(0,2,2)	441,13	445,62
ARIMA(2,2,0)	445,27	449,76
ARIMA(1,2,2)	443,39	449,37
ARIMA(2,2,1)	442,43	448,42
ARIMA(2,2,2)	444,09	451,57

Modelos	Critérios de informação	
	AIC	BIC
Efetivo Bovino		
ARIMA(0,2,0)	466,47	467,96
ARIMA(0,2,1)	467,52	470,52
ARIMA(1,2,0)	467,73	470,72
ARIMA(1,2,1)	463,01	467,50
ARIMA(0,2,2)	464,92	469,41
ARIMA(2,2,0)	469,68	474,17
ARIMA(1,2,2)	464,77	470,75
ARIMA(2,2,1)	464,48	470,46
ARIMA(2,2,2)	466,58	474,07
Agricultura		
ARIMA(0,1,0)	388,52	390,04
ARIMA(0,1,1)	371,90	374,95
ARIMA(1,1,0)	357,91	360,96
ARIMA(1,1,1)	358,99	363,57
ARIMA(0,1,2)	366,54	371,12
ARIMA(2,1,0)	366,54	371,12
ARIMA(1,1,2)	358,16	364,26
ARIMA(2,1,1)	358,20	364,30
ARIMA(2,1,2)	359,93	367,56
Focos de Queimadas		
ARIMA(0,1,0)	237,85	238,89
ARIMA(0,1,1)	235,41	237,50
ARIMA(1,1,0)	235,20	237,29
ARIMA(1,1,1)	236,93	240,07
ARIMA(0,1,2)	236,19	239,33
ARIMA(2,1,0)	236,57	239,71
ARIMA(1,1,2)	238,17	242,35
ARIMA(2,1,1)	234,26	238,43
ARIMA(2,1,2)	234,72	239,94

APÊNDICE 2 – Comparações extras dos modelos

Tabela 16: Comparação de métricas extras para classificação de melhor modelo

Modelos	Métricas de qualidade de ajuste		
	RSME	MAE	MAPE
Floresta Nativa			
ARIMA(0,2,0)	71,62243	56,67334	0,1375159
ARIMA(0,2,1)	70,32309	56,65228	0,1372716
ARIMA(1,2,0)	70,55569	56,54707	0,1370017
Pastagem			
ARIMA(0,2,0)	72,39075	58,49444	0,7703727
ARIMA(0,2,2)	68,05578	55,78643	0,7255450
ARIMA(2,2,0)	67,94725	54,41728	0,7043519

APÊNDICE 3 – Comparação de modelos para os estados

Tabela 17: Comparação de modelos baseados nos critérios de informação

Modelos	Critérios de informação	
	AIC	BIC
Floresta Nativa		
ARIMA(0,2,0)	424,59	426,08
ARIMA(0,2,1)	426,04	429,03
ARIMA(1,2,0)	426,19	429,19
ARIMA(1,2,1)	420,47	424,96
ARIMA(0,2,2)	421,48	425,97
ARIMA(2,2,0)	428,09	432,58
ARIMA(1,2,2)	421,94	427,93
ARIMA(2,2,1)	421,12	427,11
ARIMA(2,2,2)	424,33	431,81
Pastagem		
ARIMA(0,2,0)	353,74	355,24
ARIMA(0,2,1)	351,69	354,68
ARIMA(1,2,0)	351,67	354,66
ARIMA(1,2,1)	353,63	358,12
ARIMA(0,2,2)	353,67	358,15
ARIMA(2,2,0)	353,66	358,14
ARIMA(1,2,2)	355,28	361,26
ARIMA(2,2,1)	354,95	360,93
ARIMA(2,2,2)	357,22	364,70
Extração de Madeira		
ARIMA(0,0,0)	441,18	444,29
ARIMA(0,0,1)	425,23	429,89
ARIMA(1,0,0)	415,54	420,20
ARIMA(1,0,1)	417,30	423,52
ARIMA(0,0,2)	418,76	424,98
ARIMA(2,0,0)	417,27	423,49
ARIMA(1,0,2)	418,70	426,48
ARIMA(2,0,1)	419,27	427,05
ARIMA(2,0,2)	420,62	429,95

Modelos	Critérios de informação	
	AIC	BIC
Efetivo Bovino		
ARIMA(0,1,0)	345,14	346,67
ARIMA(0,1,1)	330,96	334,01
ARIMA(1,1,0)	316,45	319,50
ARIMA(1,1,1)	317,42	322,00
ARIMA(0,1,2)	321,56	326,14
ARIMA(2,1,0)	317,65	322,22
ARIMA(1,1,2)	317,56	323,67
ARIMA(2,1,1)	319,31	325,41
ARIMA(2,1,2)	321,17	328,80
Agricultura		
ARIMA(0,2,0)	348,79	350,29
ARIMA(0,2,1)	348,86	351,85
ARIMA(1,2,0)	348,45	351,44
ARIMA(1,2,1)	349,23	353,72
ARIMA(0,2,2)	350,56	355,05
ARIMA(2,2,0)	350,06	354,55
ARIMA(1,2,2)	351,15	357,13
ARIMA(2,2,1)	351,16	357,14
ARIMA(2,2,2)	352,34	359,82
Focos de Queimadas		
ARIMA(0,1,0)	215,37	216,42
ARIMA(0,1,1)	213,19	215,28
ARIMA(1,1,0)	213,15	215,24
ARIMA(1,1,1)	214,81	217,94
ARIMA(0,1,2)	213,90	217,03
ARIMA(2,1,0)	214,30	217,43
ARIMA(1,1,2)	215,89	220,07
ARIMA(2,1,1)	212,26	216,44
ARIMA(2,1,2)	216,77	221,99