



FACULDADE DE CIÊNCIAS
Departamento de Matemática e Informática

Trabalho de Licenciatura em Estatística

**Análise Espaço-Temporal da Distribuição da Gravidez
na Adolescência em Moçambique no Período
Compreendido entre 2011 à 2020**

Autora: Sónia Beatriz Carlos Nhancale

Maputo, Setembro de 2025



FACULDADE DE CIÊNCIAS
Departamento de Matemática e Informática

Trabalho de Licenciatura em Estatística

**Análise Espaço-Temporal da Distribuição da Gravidez
na Adolescência em Moçambique no Período
Compreendido entre 2011 à 2020**

Autor: Sónia Beatriz Carlos Nhancale

Supervisor: Osvaldo Francisco André Loquiha, PhD,

Maputo, Setembro de 2025

Dedicatória

Dedico este trabalho, primeiramente, a Deus, pela força e sabedoria concedidas ao longo da minha jornada. À minha família, em especial ao meu querido pai Carlos Nhancale (em memória), cuja lembrança continua a inspirar-me a ser resiliente e a lutar pelos meus sonhos, e à minha amada mãe, pelo amor, pelo apoio incondicional e pela dedicação incansável em cada etapa da minha vida. Aos meus irmãos, Edson Nhancale e Benilde Nhancale, pela confiança e suporte inestimáveis que me acompanharam ao longo do meu percurso académico. A cada um de vocês, dedico com gratidão este marco significativo na minha caminhada.

Agradecimentos

Agradeço, em primeiro lugar, ao Dr. Osvaldo Loquiha, meu supervisor, pela orientação incansável, pelas valiosas sugestões e pelo suporte académico que foram fundamentais para a realização deste trabalho. À equipa do Programa de Sistemas de Saúde, expresse a minha profunda gratidão pelo apoio técnico e intelectual, que enriqueceram significativamente este estudo. À minha família, pela paciência, amor e encorajamento incondicional durante todos os momentos da minha formação. Vocês são o meu alicerce e fonte constante de motivação.

Ao tio Jeremias que me apoiou e incentivou a continuar a estudar e ate hoje continua a incentivar e estimular o meu progresso o meu muito obrigada. Um agradecimento especial à Tia Yulla, cujo apoio financeiro foi essencial para a continuidade dos meus estudos. A sua generosidade e crença no meu potencial jamais serão por mim esquecidas. A todos aqueles que, de forma directa ou indirecta, contribuíram para a conclusão deste trabalho, o meu mais sincero obrigado.

Declaração de honra

Declaro por minha honra que o presente Trabalho de Licenciatura é resultado da minha investigação e que o processo foi concebido para ser submetido apenas para a obtenção do grau de Licenciatura em Estatística, na faculdade de Ciências da Universidade Eduardo Mondlane.

Maputo, Setembro de 2025

(Sónia Nhancale)

Resumo

A adolescência corresponde ao período que vai dos 10 aos 19 anos de idade. Esta fase é compreendida como um momento de transição entre a infância e a vida adulta, sendo marcada por diversas mudanças fisiológicas, emocionais e sociais. A gravidez nessa fase representa um grande problema, pois coloca em causa o desenvolvimento das jovens, à medida que leva à evasão escolar, marginalização social e perpetuação do ciclo de pobreza.

Este estudo teve como objectivo analisar a distribuição espaço-temporal da incidência da gravidez na adolescência nos distritos de Moçambique, no período compreendido entre 2011 à 2020, valendo-se dos dados do Inquérito Demográfico de Saúde de 2022. Para esta análise, foi utilizada a modelação inteiramente bayesiana, com aplicação do modelo de regressão de Poisson com estrutura espacial do tipo Auto-regressivo condicional (*Conditional Autoregressive-CAR*). Essa abordagem permitiu considerar a dependência espacial entre distritos vizinhos e capturar com maior precisão as variações regionais na incidência do evento ao longo do período de estudo. O software OpenBUGS versão 3.2.3 foi utilizado para ajustar o modelo bayesiano, enquanto o ArcGIS versão 10.3.1 foi empregado para o mapeamento das estimativas do risco de gravidez na adolescência. Os resultados deste estudo revelaram um padrão geográfico consistente ao longo da década analisada, mostrando um maior risco relativo de incidência de gravidez na adolescência em quase todo o país, sobretudo nas regiões norte e centro. Este padrão espacial manteve-se estável ao longo do período, sugerindo que tais áreas concentram factores de vulnerabilidade específicos que favorecem a ocorrência de gravidez na adolescência.

Palavras-chave: Gravidez na adolescência; Distribuição espaço-temporal; Modelação bayesiana; Risco relativo.

Abstract

Adolescence corresponds to the period between 10 and 19 years of age. This phase is understood as a transitional stage between childhood and adulthood, marked by various physiological, emotional, and social changes. Pregnancy during this stage poses a significant problem, as it compromises the development of young girls, often leading to school dropout, social marginalization, and the perpetuation of the cycle of poverty.

The aim of this study was to analyze the spatio-temporal distribution of the incidence of early pregnancy in the districts of Mozambique between 2011 and 2020, using data from the 2022 Demographic Health Survey. For this analysis, fully Bayesian modeling was used, applying the Poisson regression model with a spatial structure of the Conditional Autoregressive type (*Conditional Autoregressive-CAR*). This method allowed the consideration of spatial dependence among neighboring districts and enabled a more precise capture of regional variations in the incidence of early pregnancy throughout the study period. The Bayesian model was fitted using OpenBUGS version 3.2.3, while ArcGIS version 10.3.1 was used for the cartographic visualization of the results. The results of this study revealed a consistent geographical pattern throughout the decade analyzed, showing a higher relative risk of adolescent pregnancy incidence in almost the entire country, especially in the northern and central regions. This spatial pattern remained stable over time, suggesting that these areas concentrate specific vulnerability factors that contribute to the occurrence of adolescent pregnancy.

Keywords: Adolescent pregnancy; Spatio-temporal distribution; Bayesian modeling; Relative risk.

Abreviaturas e acrónimos

OMS	Organização Mundial da Saúde
UNICEF	United Nations International Children's Emergency Fund
UA	União Africana
MISAU	Ministério da Saúde
UNFPA	United Nations Population Fund
IDS	Inquérito Demográfico de Saúde
MCMC	Markov Chain Monte Carlo
BGR	Brooks, Gelman e Rubin
CAR	Conditional Autoregressive
RR	Risco Relativo

Índice

1	Introdução	1
1.1	Contextualização	1
1.2	Problema de Estudo	2
1.3	Hipóteses	3
1.4	Objectivos	3
1.4.1	Objectivo Geral	3
1.4.2	Objectivos Específicos	3
1.5	Justificação	3
1.6	Estrutura do trabalho	4
2	Revisão da Literatura	6
2.1	Adolescência	6
2.1.1	Gravidez na Adolescência	6
2.2	Estudos Relacionados	7
2.3	Epidemiologia Espacial	9
2.3.1	Análise Ecológica	9
2.3.2	Detecção de Grupos	9
2.3.3	Mapeamento de Eventos em Saúde	9
2.4	Estatística Espacial	10
2.4.1	Dados Contínuos no Espaço	10
2.4.2	Dados de Área	10
2.5	Análise Espacial de Áreas	11
2.5.1	Matriz de Proximidade Espacial	11
2.6	Autocorrelação Espacial	12
2.6.1	Índice Global de Moran	13
2.7	Autocorrelação Temporal	14
2.7.1	Teste de Durbin Watson	14
2.7.2	Teste de Breusch-Godfrey	15
2.7.3	Teste de Ljung-Box	15
2.8	Modelação Clássica dos Riscos Relativos	16
2.9	Inferência Bayesiana	17
2.9.1	Estimação Bayesiana Empírica	19
2.9.2	Estimação Inteiramente Bayesiana	20
2.10	Modelos Hierárquicos Bayesianos	21
2.10.1	Cadeias de Markov via Monte Carlo	22
2.10.2	Algoritmo de Metropolis Hastings	23
2.10.3	Redução da Dependência	24

2.10.4	Amostrador de Gibbs	24
2.11	Diagnóstico de Convergência	25
2.11.1	CrITÉrio de Gelman e Rubin	26
2.11.2	CrITÉrio de Heidelberger e Welch	26
2.11.3	CrITÉrio de Geweke	26
3	Metodologia	27
3.1	Material	27
3.1.1	Desenho de Estudo	27
3.1.2	Métodos	27
3.1.3	Hipóteses e Regra de decisão	30
4	Resultados	31
4.1	Análise Exploratória	31
4.1.1	Distribuição dos Casos de Gravidez na Adolescência	31
4.1.2	Análise Clássica do risco relativo	32
4.2	Análise Bayesiana	33
4.2.1	Autocorrelação espaço temporal	34
4.2.2	Análise da convergência do modelo	34
4.2.3	Estimativas dos parâmetros do modelo	36
4.2.4	Mapeamento das estimativas	37
4.2.5	Classificação das estimativas de Risco Relativo	38
5	Discussão	40
6	Conclusão e recomendações	44
7	Limitações	46
8	Referências	47
9	Anexos	51
9.1	Códigos	53
9.1.1	Anexo A - Código do Modelo BUGS	53
9.1.2	Anexo B - Código R para execução do modelo no OpenBUGS	54

Índice de Figuras

1	Boxplot dos casos de gravidez na adolescência por ano	32
2	Estimação clássica do risco relativo de gravidez na adolescência de 2011 a 2020	33
3	Gráficos de traços para delta e u	35
4	gráfico de correlação de delta e u	35
5	diagnostico de Gelman e Rubin para delta e u	36
6	Mapeamento das estimativas do risco relativo pela modelação bayesiana	38
7	Gráfico de traços para alfa e beta	51
8	gráfico de correlação para alfa e beta	51
9	diagnostico de Gelman e Rubin para alfa e beta	51
10	gráfico do histórico das MCMC	52

Índice de Tabelas

1	Comparação entre Estimação Empírica Bayesiana e Estimação Inteiramente Bayesiana, fonte: <i>Da Silva et al., 2009</i>	21
2	Distribuição dos casos de gravidez na adolescência por ano	31
3	Índice de Moran	34
4	Teste de LJungBox	34
5	Estimativas dos parâmetros do modelo	36
6	Top 5 de distritos com maior RR de gravidez na adolescência em 2011, 2015 e 2020	39
7	Top 5 de distritos com menor RR de gravidez na adolescência em 2011, 2015 e 2020	39

1 Introdução

1.1 Contextualização

Segundo a Organização Mundial da Saúde, a adolescência é o período de crescimento e desenvolvimento humano que ocorre após a infância e antes da idade adulta, abrangendo idades dos 10 aos 19 anos (OMS, 2009). Esta fase caracteriza-se por mudanças biológicas, cognitivas, emocionais e sociais, sendo um momento fundamental para a adopção de novas práticas e comportamentos. Durante a adolescência, os jovens tornam-se vulneráveis a comportamentos de risco que podem fragilizar a sua saúde, como alimentação inadequada, sedentarismo, tabagismo, consumo de álcool, drogas, relações sexuais desprotegidas, entre outros (Carvalho, 2013).

A vulnerabilidade dos adolescentes é agravada quando estes são expostos não apenas aos riscos inerentes aos seus comportamentos, mas também ao ambiente em que estão inseridos. Em sociedades onde normas culturais promovem uniões prematuras, os adolescentes enfrentam obstáculos que podem afectar ainda mais a sua saúde e bem-estar. Estima-se que globalmente, por ano, cerca de 14 milhões de meninas adolescentes são casadas, muitas vezes forçadas a celebrar o casamento através de arranjos feitos pelos pais, geralmente com homens muito mais velhos, o que inevitavelmente resulta em gravidez em tenra idade (União Africana, 2015).

De acordo com o Fundo das Nações Unidas para Infância, a gravidez na adolescência, também por vezes, designada de gravidez precoce, refere-se à gravidez que ocorre em jovens adolescentes, entre os 10 e os 19 anos (UNICEF, 2021). Estima-se que globalmente, cerca de 17 milhões de raparigas adolescentes dão à luz todos os anos e a maioria destes nascimentos ocorre em países de baixo e médio rendimento como é o caso de Moçambique (OMS, 2014). Esta situação, é particularmente preocupante nestes países pois a população é maioritariamente jovem e as questões de saúde reprodutiva são condicionadas por diversos factores como o acesso limitado aos serviços de saúde, a educação inadequada e normas culturais e de género e onde o rácio da mortalidade materna é ainda elevado, para o caso particular de Moçambique com 452 mortes por 100.000 nados-vivos (UNICEF, 2021; Llop-Gironés et al., 2019).

Segundo o Ministério da Saúde, em Moçambique, a gravidez na adolescência não se distribui de forma homogênea no tempo e no espaço. Há evidências de que determinadas regiões do país apresentam consistentemente taxas mais elevadas, enquanto outras mantêm níveis relativamente baixos ao longo dos anos. Essa heterogeneidade espacial reflecte as disparidades profundas no acesso a serviços de saúde sexual e reprodutiva, nos níveis de escolaridade, nas normas culturais locais e nas condições socioeconômicas (MISAU, 2024).

A literatura apresenta diversas obras que abordam a respeito das causas bem como as con-

sequências que estão por detrás desta prática, entretanto, em alguns casos pode-se ter interesse em investigar para além dos factores que contribuem para a ocorrência, a distribuição espacial e temporal de eventos em saúde, pois muitas vezes esta distribuição não é uniforme havendo diferenças que por sua vez podem ser ocasionadas por factores de natureza demográfica, socio-económica, cultural, entre outros, para além de que muitas vezes o espaço geográfico e temporal pode também ser um dos factores que contribuem para a ocorrência do evento.

Segundo Da Silva (2009), em estudos cujo interesse é analisar a distribuição espaço-temporal de eventos em saúde, diversas técnicas podem ser aplicadas. No entanto, a modelação bayesiana tem se destacado nos últimos anos pela sua capacidade de estimar efeitos espaciais e temporais com alta precisão, permitindo observar variações na incidência de forma robusta e flexível. De acordo com Bolarinwa *et al* (2022), a metodologia bayesiana é amplamente utilizada para a suavização de taxas e em contextos onde os dados são escassos ou apresentam irregularidades, pois permite incorporar conhecimentos prévios e incertezas, tornando-se ideal para modelar fenômenos de baixa frequência, como a incidência de certos eventos de saúde em áreas específicas.

1.2 Problema de Estudo

A gravidez na adolescência em Moçambique é um fenómeno crítico, especialmente em áreas rurais, onde a prática é mais prevalente. A gravidez na adolescência, está intimamente ligada ao casamento prematuro, que é visto como uma forma grave de violência contra meninas, trazendo consequências negativas para a sua saúde física e mental, além de limitar drasticamente suas oportunidades de desenvolvimento pessoal e social.

No ano de 2011, 40,2% das jovens moçambicanas tiveram filhos antes dos 18 anos, com 7,8% engravidando antes dos 15 anos. As províncias de Manica (44,9%) e Niassa (41,5%) destacaram-se pelas maiores taxas de gravidez na adolescência. Em 2022, 36% das mulheres entre 15 e 19 anos já haviam engravidado e 4% tiveram filhos antes dos 15 anos, com maiores taxas para as províncias de Cabo Delgado (55%) e Niassa (52%).

Apesar de um decréscimo de 4,2% na taxa global de gravidez antes dos 18 anos e de 3,3% para a gravidez antes dos 15 anos entre 2011 e 2022, verifica-se disparidades na distribuição das taxas de gravidez na adolescência nas diferentes províncias de Moçambique, pois, para além destas não serem uniformes, enquanto que em algumas províncias estas taxas diminuem gradualmente, o cenário inverso se verifica em outras províncias, como é o caso da província de Niassa.

Embora a gravidez na adolescência seja um fenómeno de conhecida relevância para a saúde

pública, a compreensão de sua distribuição temporal e espacial no país, permanece limitada. Compreender se essa distribuição apresenta variações ao longo do tempo, entre regiões ou em ambas as dimensões é fundamental para o delineamento de estratégias de prevenção mais eficazes. Nesse contexto, este estudo propõe-se a responder a seguinte pergunta de estudo: a distribuição da gravidez na adolescência é simultaneamente determinada pelo tempo e pelo espaço, ou está associada predominantemente a uma dessas dimensões?

1.3 Hipóteses

H_0 : A distribuição da gravidez na adolescência em Moçambique não varia significativamente ao longo do tempo nem entre diferentes regiões do país

H_1 : A distribuição da gravidez na adolescência em Moçambique varia significativamente ao longo do tempo e entre diferentes regiões do país.

1.4 Objectivos

1.4.1 Objectivo Geral

Analisar a distribuição espaço-temporal da gravidez na adolescência em Moçambique no período compreendido entre 2011 à 2020

1.4.2 Objectivos Específicos

- Descrever a evolução temporal da incidência de gravidez na adolescência em Moçambique no período compreendido entre 2011 à 2020;
- Modelar o risco relativo de gravidez na adolescência entre os distritos de Moçambique;
- Comparar os resultados obtidos pela estimação clássica e o método de estimação inteiramente bayesiana.

1.5 Justificação

A gravidez na adolescência ocorre de forma diferenciada entre regiões, entre países, dentro de um país e mesmo entre idades e grupos económicos (Fundo das Nações Unidas para a População [UNFPA], 2013). Moçambique apresenta grandes diferenças no que diz respeito tanto à distribuição da economia quanto ao acesso à educação, aos serviços de saúde e tem uma diversidade cultural entre regiões, factores que contribuem para que haja diferenças nos diversos aspectos entre as comunidades (Banco Mundial, 2016).

Embora seja notória uma tendência temporal de diminuição acentuada das taxas de gravidez na adolescência em todo o mundo, ainda que em ritmo lento nalgumas regiões, esta diminuição é desigual entre países, (Lima et al., 2024). Em Moçambique não é diferente, pois a taxa de gravidez na adolescência tem uma tendência temporal decrescente lenta a nível de todo o país; entretanto, algumas províncias têm registado um aumento na taxa de gravidez na adolescência, o que sublinha a importância do presente estudo, cujo intuito é entender padrões de distribuição destas taxas.

A relevância deste estudo decorre do facto de que ao empregar a modelação bayesiana para analisar a distribuição da gravidez na adolescência com base nos dados do IDS, poderá enriquecer a literatura existente e fornecer subsídios importantes para estudos futuros. Ao analisar a distribuição das taxas de gravidez na adolescência ao longo do tempo e do espaço, tornar-se-á possível a detecção de padrões espaço-temporais ao longo do período de estudo, a identificação de zonas críticas para a ocorrência deste evento e, deste modo, ajudará a fornecer informações relevantes para desenvolver intervenções específicas e eficazes, que considerem as particularidades de cada província, contribuindo para o planeamento estratégico de recursos e políticas públicas, visando a redução da incidência de gravidez na adolescência e melhorar a concepção de programas baseados em evidências.

1.6 Estrutura do trabalho

Este trabalho está organizado em nove (09) capítulos, que se complementam e permitem uma compreensão clara deste estudo, nomeadamente: Introdução, Revisão de Literatura, Metodologia, Resultados, Discussão, Conclusão e Recomendações, limitações, Referências Bibliográficas e Anexos.

- **Capítulo I – Introdução:** Este capítulo apresenta o enquadramento geral do estudo, o problema de estudo, os objectivos gerais e específicos, a justificativa e a estrutura do trabalho.
- **Capítulo II – Revisão da Literatura:** Neste capítulo são apresentados os principais conceitos, teorias e estudos prévios relacionados a temática em análise, fornecendo a base teórica necessária para sustentar o estudo. Também aborda o que é apresentado na literatura sobre a gravidez na adolescência e a utilização de modelos estatísticos em análises espaço-temporais.
- **Capítulo III – Metodologia:** Descreve os procedimentos adoptados para o desenvolvimento deste estudo, incluindo o delineamento metodológico e a fonte dos dados.
- **Capítulo IV – Resultados:** Apresenta os resultados obtidos a partir da aplicação da metodologia adoptada, seguidos de sua interpretação.

- **Capítulo V - Discussão:** Neste capítulo é feita uma breve discussão dos resultados deste estudo e de seguida uma comparação com a literatura e suas implicações.
- **Capítulo VI – Conclusões e Recomendações:** Este capítulo sintetiza os principais resultados do estudo, responde aos objetivos formulados e sugere recomendações para estudos futuros.
- **Capítulo VII – Limitações:** Neste capítulo são apresentadas as principais limitações deste estudo.
- **Capítulo VIII – Referências bibliográficas:** apresenta as obras consultadas no âmbito da realização do trabalho.
- **Capítulo IX – Anexos:** Este capítulo apresenta gráficos, tabelas e códigos complementares do trabalho.

2 Revisão da Literatura

2.1 Adolescência

Segundo Neto (2007), a adolescência é uma fase de crescimento marcada por mudanças fisiomorfológicas e adaptações psicossociais que são decisivas para saúde e bem-estar na vida adulta. Para Silva *et al.*, (2021) a adolescência é uma fase do desenvolvimento humano que apresenta como delimitação tanto critérios cronológicos e físicos, como também sociais e culturais, caracterizada pela integração social, enfrentamento de limitações e desenvolvimento saudável. Segundo a OMS (2009), a adolescência, que corresponde ao período entre 10 e 19 anos, divide-se em duas fases: a pré-adolescência (10-14 anos), marcada por mudanças físicas, hormonais, cognitivas e sexuais iniciais, e a adolescência propriamente dita (15-19 anos), caracterizada pela integração social, enfrentamento de limitações e desenvolvimento saudável.

Este, constitui um período importante de desenvolvimento que depende da maneira como factores biológicos, psicológicos e sociais se combinam para moldar os padrões de maturidade do indivíduo requeridos no seio da sociedade. Na visão de Neto (2007), adolescência é uma fase da vida que pode ser definida em sua dimensão psico-biológica e em sua dimensão histórica, política, económica, social e cultural. No campo psicossocial são comuns alterações nas relações escolares, a experimentação intensa de diversos sentimentos, a busca pela autonomia e independência familiar, bem como a predileção em experimentar novos comportamentos e vivências acompanhadas com frequência de contestação de ideias e conceitos preestabelecidos (Silva *et al.*, 2021).

Deste modo pode-se afirmar que a adolescência caracteriza-se por diversas transformações não apenas no âmbito social e físico mas também no âmbito psicológico ao longo das fases que a compõem. Embora o organismo das adolescentes ainda não esteja totalmente preparado para a gravidez, estas podem engravidar, entretanto a gravidez nesta fase da vida expõe as jovens a diversos riscos o que amplia o cenário de vulnerabilidade enfrentado pelas adolescentes.

2.1.1 Gravidez na Adolescência

De acordo com Martins *et al.*, (2011) gravidez na adolescência constitui uma ocorrência de repercussão mundial, cuja dimensão diverge nas diferentes culturas e contextos devido a vários factores que por sua vez também diferem de região para região. A gravidez na adolescência constitui um problema de saúde pública em todo o mundo, especialmente entre os países africanos de baixa e média renda, uma vez que estes ocupam as primeiras posições do ranking com as maiores taxas de gravidez na adolescência (Salhan, 2007).

A gravidez na adolescência pode gerar consequências imediatas no emocional dos jovens en-

volvidos. Alguns sentimentos experimentados por estes jovens são: medos, insegurança, desespero, sentimento de solidão, principalmente no momento da descoberta da gravidez.

As consequências derivadas da gravidez na adolescência são grandes especialmente para as raparigas, uma vez que as raparigas que engravidam na adolescência tem maior probabilidade de enfrentar complicações na gravidez, incluindo aborto inseguro, e de se tornarem jovens mães pela segunda vez, pois de acordo com a UNICEF (2021), entre crianças com o primeiro nascimento aos 14 anos ou menos, quase três quartos também têm um segundo parto na adolescência, e 40% das que têm dois filhos progridem para um terceiro antes de sair da adolescência. Os bebês nascidos de mães adolescentes enfrentam um risco substancialmente maior de morte do que aqueles nascidos de mulheres com idade entre 20 e 24 anos, dentre estes riscos estão a desnutrição, baixo desenvolvimento mental e físico, conexão social inadequada com os pais e educação deficiente (Ganchimeg *et al.*, 2014).

Adicionalmente, estas adolescentes podem desenvolver problemas psicológicos devido ao estigma social, sofrer violência física e doméstica na tentativa de atender às demandas da gravidez e da maternidade e têm maior probabilidade de abandonar a escola e muitas vezes sem chance de retornar o que sem duvidas perpetua um ciclo intergeracional de pobreza e vulnerabilidade sócio-económica (Aransiola *et al.*, 2013).

Diversos estudos sugerem uma série de factores que influenciam para a ocorrência da gravidez nesta faixa etária. Alguns destes estudos associam a questões sociais como a taxa de participação de adolescentes, a desigualdade, os problemas de saúde, o baixo nível de despesas com a saúde pública, o baixo número de mulheres com emprego remunerado e o difícil acesso a educação de qualidade (Bolarinwa *et al.*, 2022; Chodota *et al.*, 2021; Martins *et al.*, 2011).

2.2 Estudos Relacionados

Partindo do pressuposto que a gravidez na adolescência não acontece de forma aleatória no seio da sociedade, torna se obvio que existem aspectos que estejam a influenciar para que tais disparidades ocorram.

De acordo com Bolarinwa (*et al.*, 2022), o nível educacional exerce uma influência directa na gravidez na adolescência, pois adolescentes com mais anos de escolaridade tendem a ter melhor acesso a informações sobre saúde sexual e reprodutiva, métodos contraceptivos e riscos associados à gravidez na adolescência . Além disso, a educação promove a consciencialização sobre a importância do planeamento familiar e pode fortalecer o senso de autonomia e empoderamento, permitindo que as adolescentes tenham maior controle sobre suas escolhas e retardem a maternidade. Em contraste, baixos níveis de escolaridade limitam o acesso a esse conheci-

mento essencial e contribuem para comportamentos de risco, aumentando a vulnerabilidade à gravidez na adolescência.

Um estudo intitulado casamento prematuro e gravidez na adolescência em Moçambique desenvolvido por Nhampoca e Maritz (2022) concluiu que o casamento prematuro e a gravidez na adolescência são potenciais factores de abandono escolar, baixa escolaridade e resultados adversos em termos de saúde na gravidez e no parto e que a educação das raparigas pode actuar como um factor de protecção para retardar o casamento precoce e a gravidez na adolescência e portanto, a retenção escolar das raparigas deve ser monitorada, contribuindo assim para a paridade de género na educação.

Em outro estudo desenvolvido por Chodota *et al.* (2021), na Tanzânia no distrito de Njombe concluiu-se que os factores que Influenciam a gravidez de alto risco são a educação, a profissão, o local de residência e a cultura. Um outro estudo feito pela UNICEF em coordenação com o FNUAP em 2015, sobre casamento prematuro e gravidez na adolescência em Moçambique, destacou que a zona de residencia é um factor determinante para a ocorrência da gravidez na adolescência onde constatou-se que raparigas de áreas urbanas eram significativamente mais propensas a engravidar antes dos 15 anos do que as residentes nas áreas rurais, apesar de as raparigas na área urbana terem maior acesso à informação, quer através da rádio ou de outros meios (UNICEF e UNFPA, 2015).

Todavia, a metodologia adoptada nestes estudos baseou-se nos métodos da estatística clássica com o objectivo final de identificar os factores associados a gravidez na adolescência. Embora a análise dos factores associados a gravidez na adolescência seja de extrema importância, o enfoque deste estudo está no estudo da distribuição espaço-temporal deste evento. Modelos espaço-temporais são amplamente utilizados para capturar padrões de risco relativos entre diferentes regiões e ao longo do tempo, o que permite identificar áreas prioritárias para intervenções, mapeando os locais com maior risco e analisando a evolução das taxas ao longo de vários anos.

Em estudos cujo intuito era de fazer uma análise espaço-temporal, vários outros métodos foram utilizados, dos quais pode-se destacar o estudo desenvolvido por Fernandes *et al.*, (2024), no Brasil cujo objectivo era de avaliar a distribuição da proporção de mães adolescentes no tempo e no espaço e sua relação com indicadores socioeconómicos e de vulnerabilidade social, este estudo revelou que embora tenha havido uma redução nas taxas de gravidez na adolescência, essa redução não foi uniforme em todas as regiões, revelou ainda uma relação consistente entre a gravidez na adolescência e indicadores de vulnerabilidade social e económica. O estudo destaca que áreas com maior vulnerabilidade materna e piores condições socioeconômicas apresentam maior proporção de mães adolescentes, e que essa relação se manteve estável ao longo do tempo.

Outro estudo de Bolarinwa *et al.* (2022) feito na Nigéria revelou igualmente disparidades no que diz respeito a distribuição das taxas de gravidez na adolescência. Uma alta proporção de gravidez na adolescência foi localizada nas partes do norte da Nigéria, o estudo mostrou ainda que a idade de início da atividade sexual, o nível educacional, o estado civil, a etnia e a situação de trabalho foram os fatores associados à gravidez na adolescência.

Nestes estudos foi possível detectar áreas de risco para este fenómeno, levando em consideração a modelação bayesiana. Entretanto, em relação à gravidez na adolescência, observa-se que poucos estudos utilizaram a modelação bayesiana para analisar a distribuição da gravidez na adolescência ao longo do tempo e espaço.

2.3 Epidemiologia Espacial

O estudo da distribuição geográfica da incidência de eventos em saúde, tem ganhado destaque através da aplicação e desenvolvimento de métodos e modelos estatísticos. De acordo com Nogueira (2004), a análise espacial com a informação da localização dos dados é uma técnica que busca descrever os padrões existentes nos dados espaciais e estabelecer a relação entre variáveis geo referenciadas. De acordo com a distribuição espacial dos dados observados e de acordo com os objectivos da análise a epidemiologia espacial está dividida em três grandes áreas, nomeadamente análise ecológica, mapeamento de doenças e detecção de grupos (Lawson, 2003).

2.3.1 Análise Ecológica

De acordo com Da Silva (2009), a análise ecológica refere-se ao estudo da associação entre fatores ambientais, sociais ou demográficos e a ocorrência de doenças ou condições de saúde, levando em consideração áreas geográficas agregadas ou seja, avaliam as relações entre a distribuição espacial da variável em estudo e factores explicativos.

2.3.2 Detecção de Grupos

De acordo com Caumo (2006), a detecção de grupos refere-se à identificação de agrupamentos de casos de doenças ou condições de saúde que ocorrem em uma área geográfica ou período de tempo específico, mais frequentemente do que seria esperado ao acaso. Esta abordagem avalia se o evento mapeado forma um agrupamento de casos em determinados locais ou se estes são aglomerados como um todo que pode ser monitorado ao longo do tempo.

2.3.3 Mapeamento de Eventos em Saúde

Ainda de acordo com Da Silva (2009), mapeamento de eventos envolve a representação gráfica da sua distribuição em um determinado espaço geográfico cujo objectivo principal é descrever

a distribuição geográfica do evento, gerar hipóteses, controlar áreas de risco elevado e estimar o risco de ocorrência do evento controlando suas oscilações. Para este efeito geralmente utilizam-se modelos estatísticos que geram mapas sem perturbações externas com vista a descobrir a estrutura subjacente aos dados isto é a verdadeira distribuição da população em estudo.

Esta técnica apresenta grande interesse e utilidade dentro da sociedade, visto que pode ser usada para avaliar a alocação de recursos para a saúde de acordo com a variação geográfica, bem como em estudos de investigação da distribuição da incidência e prevalência da doença (Lawson *et al.*, 2003).

2.4 Estatística Espacial

Segundo Caumo (2006), a estatística espacial é um campo que se dedica a estudar a distribuição espacial de dados associados a fenômenos geográficos. Ela investiga métodos científicos para colectar, descrever, visualizar e analisar informações que estão ligadas a coordenadas geográficas. Quando o objectivo é examinar a evolução de um fenómeno não apenas no espaço, mas também ao longo do tempo, pode-se recorrer a análise espaço-temporal, combinando as dimensões geográficas e temporais na investigação dos dados.

Na análise espacial, dois tipos de dados georreferenciados são principalmente levados em consideração dependendo do tipo e da natureza do estudo, nomeadamente: dados de área e dados contínuos no espaço, entretanto o tipo de dados que serão usados como objecto deste estudo são os dados de área.

2.4.1 Dados Contínuos no Espaço

O objectivo da análise deste tipo de dados é de modelar uma superfície espacial que represente o comportamento da variável em estudo no espaço geográfico ou seja, examinar se o conjunto de eventos apresenta algum tipo de padrão sistemático ou aleatório, buscando entender em que escala este padrão ocorre (Getis, 1992).

De acordo com Câmara *et al.* (2002), um padrão pontual refere-se a um conjunto de dados que consiste em uma série de localizações pontuais (P1, P2, etc.), representando a ocorrência de eventos relevantes dentro de uma área específica de estudo. Nesse tipo de análise, apenas a ocorrência dos eventos é levada em consideração, embora as observações também possam estar associadas a outros atributos além da localização geográfica.

2.4.2 Dados de Área

Dados de área referem-se a dados agregados ou resumidos sobre uma determinada região geográfica. Em vez de representar informações pontuais ou individuais os dados de área se referem

a unidades espaciais maiores, como distritos, bairros, municípios, ou outras divisões geográficas. segundo Assunção (2001), dados de área envolvem a conexão de um mapa geográfico a uma base de dados, onde o mapa é subdividido em áreas, e cada área é associada a uma ou mais variáveis aleatórias $(y_1, \dots, y_n)$ que representam um valor agregado para toda a área, em vez de um valor para um ponto específico. Nesse tipo de dado, a localização exacta dos eventos não é conhecida, mas sim um valor que representa a área como um todo (Assunção, 2001).

Ao se analisar este tipo de dados, geralmente tem-se também o interesse de verificar a existência de um padrão espacial nos valores observados. Quando este padrão é observado, torna-se necessário verificar se é aleatório ou apresenta uma agregação definida ou ainda se esta distribuição pode ser associada a causas mensuráveis (Caumo, 2006). Quando se aplicam técnicas de estatística espacial para analisar dados de área, os resultados costumam ser apresentados por meio de índices que medem a associação espacial, como o Índice de Moran, gráficos de dispersão, como o Box Map, e mapas temáticos que ilustram a distribuição espacial dos dados. Esses índices são fundamentais para identificar agrupamentos de áreas com valores altos ou baixos de uma variável, bem como para detectar zonas de transição e valores atípicos.

2.5 Análise Espacial de Áreas

De acordo com Caumo (2006) nos dados de área, considera-se uma região dividida em áreas contíguas, disjuntas e bem definidas, para cada área é associada uma quantia, um indicador de acordo com a característica em estudo, isto é lida-se com eventos agregados em espaços delimitados por polígonos fechados. Os valores associados as áreas não estão relacionados a localização pontual específica de um evento no espaço, mas correspondem a uma quantidade que representa um padrão global para a área em relação a variável em estudo.

O objectivo da análise de dados de área não é prever valores para áreas não observadas, mas sim identificar padrões ou configurações espaciais relacionadas à variável aleatória de interesse, além de explorar possíveis relações com covariáveis (Caumo, 2006).

2.5.1 Matriz de Proximidade Espacial

Em Estatística, a correlação é uma medida que indica o grau de variação conjunta entre duas variáveis. No entanto, em análise espacial, o foco geralmente não é apenas a correlação numérica entre duas variáveis, mas sim investigar como o valor observado de uma única variável em uma determinada área geográfica depende dos valores dessa mesma variável nas áreas vizinhas (Luzardo *et al.*, 2017). O ponto central é incorporar a noção de proximidade espacial em uma medida de autocorrelação. Para isso, é fundamental capturar a relação espacial entre todos os pares de áreas.

Para estimar esta variabilidade espacial de dados de área, uma ferramenta básica é a *Matriz de Proximidade Espacial*, também chamada *Matriz de Vizinhança*. Dado um conjunto de n áreas $\{A_1, \dots, A_n\}$, constrói-se a matriz $W^{(1)}$ ($n \times n$), onde cada um dos elementos w_{ij} representa uma medida de proximidade entre A_i e A_j . (Getis et al. , 1992)

de acordo com (Luzardo *et al.*, 2017) para se atribuir valores a cada elemento da matriz podem ser usados diferentes critérios como:

- **Critério da distância:** $w_{ij} = 1$, se o centro geométrico do polígono de A_i está dentro de uma determinada distância (estabelecida pelo investigador) do centro geométrico do polígono de A_j ; caso contrário, $w_{ij} = 0$.
- **Critério da contiguidade ou adjacência:** $w_{ij} = 1$, se A_i compartilha um lado comum com A_j ; caso contrário, $w_{ij} = 0$.
- **Critério da vizinhança:** $w_{ij} = 1$, se A_j é um dos p vizinhos mais próximos de A_i (o valor p é estabelecido pelo investigador); caso contrário, $w_{ij} = 0$.

Baumont (2004) propõe que a matriz de vizinhança seja construída pelo método dos k vizinhos mais próximos, definida da seguinte forma:

$$w_{ij}(k) = \begin{cases} 1, & \text{se } d_{ij} < d_{ij}(k) \\ 0, & \text{caso contrário} \end{cases} \quad (1)$$

onde w_{ij} é um elemento da matriz normalizada, d_{ij} é a menor distância entre a região i e todas as outras unidades, de modo que cada unidade tenha exactamente k vizinhos. Uma vez determinada a estrutura espacial de análise, expressa pela matriz W , qualquer medida particular de autocorrelação pode ser concebida mediante a definição de um modo de mensurar a diferença entre os valores do atributo associado às áreas.

2.6 Autocorrelação Espacial

A autocorrelação espacial é semelhante à autocorrelação temporal, onde a dependência ocorre apenas entre vizinhos próximos, como em séries temporais com dependência unidirecional. No entanto, de acordo com Vieira (2009), em dados espaciais, esse cenário é raro, já que as áreas geralmente estão dispostas em uma superfície bidimensional, resultando em uma dependência bidirecional entre as regiões.

A autocorrelação espacial ocorre quando as observações distribuídas no espaço têm uma influência mútua entre si. Conforme Câmara *et al.* (2002), a autocorrelação espacial avalia o grau de dependência entre o valor de uma variável em uma determinada área e os valores dessa

mesma variável em regiões vizinhas, e pode ser positiva ou negativa, sendo que, quando há um nível elevado de autocorrelação espacial positiva, os valores observados em uma região tendem a ser parecidos com os valores das áreas adjacentes, o que significa que a ocorrência de um fenômeno em uma área específica pode influenciar a presença desse mesmo fenômeno nas regiões próximas. De acordo com Vieira (2009), em situações extremas, o valor em uma localização pode ser utilizado para prever precisamente os valores nas áreas vizinhas. Entre os principais indicadores de autocorrelação espacial global está o índice global de Moran.

2.6.1 Índice Global de Moran

O Índice Global de Moran oferece uma medida abrangente da associação espacial presente em um conjunto de dados. Dados com baixa associação espacial tendem a gerar um índice reduzido. Valores positivos indicam uma autocorrelação espacial positiva, enquanto valores negativos reflectem uma autocorrelação espacial negativa.

Segundo Camara et al. (2002), esse indicador é uma medida de correlação espacial utilizada para identificar desvios em relação a uma distribuição espacial aleatória, permitindo a observação da existência de padrões espaciais e a detecção de semelhanças entre áreas de primeira ordem. O Índice de Moran avalia se as áreas vizinhas têm uma maior semelhança em relação ao indicador em análise do que seria esperado em um padrão aleatório.

Uma vez calculado o índice de Moran é necessário testar a sua validade estatística, para tal, duas abordagens são possíveis:

- **Teste de Pseudo-significância:** Diferentes permutações dos valores dos atributos associados às regiões são geradas, resultando em novos arranjos espaciais, uma vez que os valores são redistribuídos entre as áreas. Dentre esses arranjos, apenas um reflecte a situação observada, permitindo a construção de uma situação empírica para o índice de Moran (I). Se o valor de I se alinhar com um extremo da distribuição simulada, isso indica a ocorrência de um evento significativo. Essa abordagem é fundamental para avaliar a relevância estatística dos padrões espaciais identificados, possibilitando a comparação entre a estrutura observada e a variabilidade esperada sob um modelo aleatório. Assim, o teste de significância oferece suporte para a interpretação dos resultados, ajudando a determinar se a autocorrelação espacial observada é devida a fatores aleatórios ou se reflecte um verdadeiro padrão subjacente nos dados (Câmara *et al.*, 2002).
- **Distribuição aproximada:** Para um número adequado de sub-regiões, e considerando que as variáveis aleatórias associadas a cada localização do atributo são independentes e seguem uma distribuição normal, pode-se assumir que o índice I tem uma distribuição

aproximadamente normal. Essa suposição é essencial para a aplicação de testes estatísticos, permitindo a utilização de métodos paramétricos que facilitam a interpretação dos resultados. Além disso, essa normalidade aproxima as propriedades do índice a uma distribuição de referência, possibilitando a comparação dos valores observados com aqueles esperados sob condições de aleatoriedade. Assim, a análise estatística torna-se mais robusta, oferecendo maior confiabilidade nas inferências sobre a autocorrelação espacial e nos padrões subjacentes nos dados (Câmara *et al.*, 2002).

2.7 Autocorrelação Temporal

A autocorrelação temporal refere-se ao grau de correlação entre os valores de uma mesma variável ao longo do tempo. Podem-se destacar vários testes de autocorrelação temporal, nomeadamente o teste de Durbin-Watson, o teste de Breusch-Godfrey, o teste de Ljung-Box. Os testes de autocorrelação serial revelam o grau de dependência entre os termos de uma série.

2.7.1 Teste de Durbin Watson

O teste de Durbin-Watson é usado para detectar a presença de autocorrelação nos resíduos. De acordo com King (1981), o teste de Durbin-Watson é uma medida estatística que verifica a presença de autocorrelação nos resíduos de um modelo de regressão. A estatística de teste é calculada como a razão da soma das diferenças, elevadas ao quadrado, entre os resíduos sucessivos e a soma dos quadrados dos resíduos (SQE).

A estatística de teste é dada por (King, 1981):

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (2)$$

Onde:

- e_i são os resíduos do modelo.
- n é o número de observações.

Uma grande vantagem da estatística d é que ela utiliza os resíduos estimados, que são geralmente calculados durante a análise de regressão. Devido a essa vantagem, tornou-se prática comum interpretar a estatística de teste de Durbin-Watson juntamente com outras medidas, como o coeficiente de determinação (R^2), o R^2 ajustado, estatísticas t e estatísticas F , cujas hipóteses do teste são:

- H_0 : Não há autocorrelação nos resíduos.
- H_1 : Há autocorrelação nos resíduos.

2.7.2 Teste de Breusch-Godfrey

O Teste de Breusch-Godfrey, também chamado de teste de Multiplicador de Lagrange, é utilizado para identificar a presença de autocorrelação nos resíduos de um modelo de regressão. A autocorrelação, comum em análises de séries temporais, ocorre quando os erros de previsão estão correlacionados, comprometendo a validade dos resultados e das inferências estatísticas. Este teste é especialmente relevante para avaliar a dependência temporal nos resíduos.

O procedimento envolve a estimação do modelo original e, em seguida, a regressão dos resíduos quadráticos em relação às variáveis independentes e suas defasagens. Sob a hipótese nula de ausência de autocorrelação, a soma dos quadrados dos resíduos deve estar dentro do esperado. O valor do teste é comparado a uma distribuição qui-quadrado para determinar sua significância estatística, indicando se há evidência de autocorrelação nos dados.

A estatística do teste é dada por:

$$LM = T \left(\frac{e'X_0(X_0'X_0)^{-1}X_0'e}{e'e} \right) = TR_0^2 \quad (3)$$

onde X_0 é a matriz X original aumentada por P colunas adicionais contendo os resíduos de MQO atrasados, $e_{t-1}, \dots, e_{t-p}$. O teste pode ser realizado simplesmente regredindo os resíduos de mínimos quadrados ordinários e x_{t0} (preenchimento de valores ausentes para resíduos atrasados com zeros) e referente a TR_0^2 ao valor crítico apresentado para a distribuição qui-quadrado com P graus de liberdade.

Dado que $X'e = 0$, o teste é equivalente a regressão e_t parte dos resíduos defasados que não é explicada por X . Existe, portanto, uma lógica para isso, se algum ajuste for encontrado, é devido à correlação entre a corrente e a defasagem dos resíduos (Greene, W. H., 2002).

Esta estatística tem uma distribuição qui-quadrado com dois graus de liberdade (χ^2) (Student-mund, A. H., 2001).

As hipóteses do teste são:

$$\begin{cases} H_0 : \rho_1 = \rho_2 = \dots = \rho_p = 0 \\ H_1 : \text{pelo menos um } \rho_j \neq 0, \quad j = 1, \dots, p \end{cases} \quad (4)$$

2.7.3 Teste de Ljung-Box

O teste de Ljung-Box (LB) analisa nos resíduos a presença de autocorrelação, a hipótese nula (H_0) é de não existir evidências de que os resíduos sejam autocorrelacionados, enquanto a hipótese alternativa (H_1) é de ter evidências de que os resíduos sejam autocorrelacionados e a estatística deste teste, apresentada na Eq. 5, converge para uma distribuição χ^2 com $k - p - q$

graus de liberdade (Burns, 2002).

$$LB = T(T+2) \left(\frac{\hat{\rho}_1^2}{T-1} + \frac{\hat{\rho}_2^2}{T-2} + \dots + \frac{\hat{\rho}_k^2}{T-k} \right) \quad (5)$$

em que: T é o tamanho da série, $\hat{\rho}_j$ é uma estimativa de autocorrelação com j defasagens ($j = 1, 2, \dots, k$).

2.8 Modelação Clássica dos Riscos Relativos

De maneira a fazer comparações entre diferentes populações no espaço e no tempo, as contagens dos eventos devem ser padronizadas com vista a gerar taxas livres do efeito do tamanho da população (Da Silva, 2009). Neste contexto, é geralmente utilizado o risco relativo, onde, pelo método de máxima verossimilhança, são atribuídas as áreas do mapa estimativas para o risco relativo que podem ser estimadas usando a razão de incidência padronizada (*Standardized incidence ratio-SIR*).

O uso da SIR é um método bem estabelecido para mapeamento de doenças, permitindo a identificação do risco relativo em diferentes áreas geográficas. A SIR é calculada dividindo o número observado de incidentes pelo número de casos "esperados" naquela região, e é dada por:

$$SIR_i = R_i = \frac{y_i}{e_i} \quad (6)$$

onde:

- y_i é a contagem de eventos na área i ;
- N_i é a população em risco na área i ;
- $e_i = r \cdot N_i$ é o número esperado de eventos na área i .

E a taxa global de risco em toda a área estudada é dada por:

$$r = \frac{\sum y_i}{\sum N_i} \quad (7)$$

Se a SIR for igual a 1 a área i apresentou tantos casos observados quanto seria esperado caso o seu risco fosse idêntico ao de toda área avaliada, a SIR varia de zero a infinito com variância da SIR_i (y_i/e_i^2) que é inversamente proporcional ao número esperado de casos na área i (e_i) e diretamente proporcional ao tamanho da população na área i , isto é, quanto menor o tamanho da população em risco, maior a incerteza na estimativa do risco (Olsen *et al.*, 1996).

Essa incerteza causa instabilidade ao modelo, tornando difícil a interpretação dos mapas, e pode se tornar ainda maior em estudos com baixa incidência. Todavia, existem na literatura

três métodos alternativos para contornar esse problema, nomeadamente agregar áreas, usar mapas de probabilidade e usar modelação Bayesiana. Actualmente as duas primeiras não estão sendo mais tão utilizadas em função das vantagens demonstradas pela modelação Bayesiana (Assunção, 2001).

2.9 Inferência Bayesiana

A inferência estatística desempenha um papel fundamental na análise de dados, permitindo tirar conclusões sobre uma população com base em uma amostra. Existem dois paradigmas principais de inferência estatística: a inferência clássica e a inferência bayesiana.

De acordo com Caumo (2006), embora em situações práticas, a estatística Bayesiana apareça com destaque somente nos últimos anos, sua história é muito antiga e tem origem associada ao trabalho original do reverendo Thomas Bayes de 1763, sendo que no passado, a análise estatística baseada no teorema de Bayes era muitas vezes intimidadora devido às integrações numéricas necessárias, porém, a complexidade do processo analítico de muitos problemas cujo enfoque Bayesiano lhes é atribuído, fez com que a aplicação prática dos mesmos estagnasse. Contudo, métodos computacionais desenvolvidos recentemente revolucionaram a aplicação dos métodos bayesianos, e tais métodos agora oferecem uma abordagem abrangente para a estimação de modelos complexos. Eles fornecem uma maneira de melhorar a estimação em conjuntos de dados esparsos e permitem inferências em amostras finitas sem apelo a argumentos de grandes amostras, como no método de máxima verossimilhança e em outros métodos clássicos.

Para Silva (2009), enquanto a abordagem clássica fundamenta-se na base de que a informação a respeito do fenómeno aleatório que se deseja compreender é oriunda dos dados a serem obtidos pelo investigador, ou seja, da amostra, sendo assim, toda e qualquer inferência é embasada apenas naquilo que se observou, seja através da obtenção de uma amostra ou da realização de um experimento. Isto é, enquanto a abordagem clássica considera os parâmetros populacionais como valores fixos, tratando a incerteza como resultado da variabilidade das amostras, a inferência bayesiana adopta uma perspectiva diferente, nesta, os parâmetros são tratados como variáveis aleatórias, e a incerteza é quantificada por meio de distribuições de probabilidade.

A teoria bayesiana, que fundamenta a inferência bayesiana, baseia-se no Teorema de Bayes. Segundo O'Hagan A. (1994) este teorema descreve como actualizar a probabilidade de uma hipótese com base em novas evidências, integrando conhecimento prévio (distribuição *a priori*) com informações obtidas a partir dos dados observados (verossimilhança) para gerar a distribuição *a posteriori* e descreve como a probabilidade de uma hipótese deve ser ajustada à luz de novas evidências, permitindo um processo dinâmico de actualização. A fórmula que representa essa actualização é dada pela equação 7:

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{\int p(\theta)p(y|\theta)d\theta} \quad (8)$$

Onde:

$p(\theta)$ é a distribuição *a priori*, que reflecte o conhecimento inicial ou suposições sobre o parâmetro θ .

$p(y|\theta)$ é a verossimilhança, que mede como os dados observados y são consistentes com diferentes valores do parâmetro.

$p(\theta|y)$ é a distribuição *a posteriori*, que combina a priori e a verossimilhança para reflectir o estado atualizado do conhecimento sobre θ , dado os dados.

De acordo com Gelman *et al.*(1995), a inferência bayesiana é o processo de ajustar um modelo de probabilidade para um grupo de dados e resumir os dados por uma distribuição de probabilidade dos parâmetros do modelo e em quantidades não observáveis como previsões para novas observações. A Inferência Estatística Bayesiana é o processo de diminuição de incerteza que se baseia em dados estatísticos e evidências pessoais.

Dependendo da escolha da distribuição a priori, a inferência Bayesiana pode ser subjectiva, quando a priori é baseada em opiniões ou suposições do investigador ou objectiva, quando usa uma distribuição a priori não informativa.

Segundo Bernadinelli e Montonolli (1992), no contexto da inferência bayesiana, a especificação das distribuições a priori é uma etapa fundamental que pode influenciar significativamente os resultados. No entanto, em situações em que não se dispõe de conhecimento prévio substancial sobre os parâmetros do modelo, é comum recorrer a prioris não informativas ou objectivas, com o objectivo de permitir que os dados observados exerçam maior influência na distribuição posterior. As distribuições priors não informativas são construídas para representar um estado de conhecimento mínimo sobre os parâmetros, buscando minimizar a introdução de viés na inferência.

Desta forma, autores como Gelman *et al.* (1995), recomendam o uso de prioris fracamente informativas, que mantêm a neutralidade mas ajudam a evitar problemas numéricos e inferências extremas em modelos complexos.

Entretanto, a inferência Bayesiana pode gerar controvérsias, especialmente quando se trata da escolha da distribuição a priori, que pode influenciar os resultados. Todavia, a inferência clássica é crítica em relação à metodologia Bayesiana uma vez que considera que ao adoptar uma abordagem Bayesiana para inferência do parâmetro introduz-se um elemento de subjectividade e arbitrariedade na inferência estatística do parâmetro (Caumo, 2006).

Uma das principais vantagens da inferência Bayesiana é a flexibilidade em incorporar informações prévias, o que é particularmente útil quando se tem poucos dados ou conhecimento prévio relevante. Sendo que a abordagem também oferece uma maneira natural de lidar com a incerteza, já que permite modelar e quantificar essa incerteza explicitamente por meio de probabilidades (Natário, 2004).

O'Hagan (1994) afirma que, em termos gerais, alguns dos argumentos a favor do enfoque Bayesiano são os que a análise Bayesiana é fundamentalmente completa, muito flexível, produz conclusões claras e directas e faz uso de toda informação disponível. Em contraste, o enfoque clássico ignora a informação *a priori*. Segundo Natário (2004), uma das vantagens reconhecidas na utilização da abordagem Bayesiana é a flexibilidade de modelação de informação e a possibilidade de incorporar nesta modelação parâmetros adicionais como, por exemplo, a variabilidade espacial, elemento fundamental nos estudos de estatística espacial cujo objecto de estudo é compreender a distribuição dos dados provenientes de um determinado acontecimento, ocorrido num dado espaço.

Para além das vantagens em termos de flexibilidade e incorporação de informação prévia, métodos de estimação bayesiana baseados em amostragem fornecem um perfil completo da densidade de um parâmetro, de modo que qualquer não-normalidade clara se torne evidente, e permitem que uma gama de hipóteses sobre os parâmetros seja simplesmente avaliada usando a coleção de amostras da distribuição a posteriori (Caumo, 2006).

Os métodos bayesianos também podem superar os estimadores clássicos em termos de precisão das estimativas. Isso ocorre porque especificar a distribuição a priori traz informações extras ou dados com base no conhecimento acumulado, e a estimativa posterior é baseada nas fontes combinadas de informação (a priori e verossimilhança), o que proporciona maior precisão.

A teoria bayesiana apresenta para estimação de modelos bayesianos duas diferentes abordagens: a modelação bayesiana empírica e a modelação inteiramente bayesiana.

2.9.1 Estimação Bayesiana Empírica

A estimação Bayesiana empírica é uma abordagem híbrida entre o método clássico e o método Bayesiano. Segundo Da Silva (2009), a modelação Bayesiana Empírica assume hiperparâmetros desconhecidos (parâmetros indexadores da distribuição *a priori*), porém utiliza estimativas pontuais de máxima verossimilhança para tais hiperparâmetros e, como uma forma de aproximação, os parâmetros da distribuição a priori são estimados a partir dos dados.

A ideia básica consiste na utilização dos próprios dados da amostra para estimar os valores dos hiperparâmetros e, então, utilizar estas estimativas como valores fixos para os mesmos. A estimação bayesiana empírica ignora a variabilidade introduzida pelos hiperparâmetros, justamente por não considerar a incerteza na estimação dos mesmos, ou seja, por considerá-los fixos (Caumo, 2006).

Esta modelação, no entanto, pode subestimar a variabilidade dos riscos estimados (além de não lidar bem com a variabilidade das estimativas produzidas), essencialmente quando *a priori* condicionada aos hiperparâmetros apresentar alta dispersão. Além disso, não permite generalizações para situações mais complexas, como casos de análise espaço-temporal. Ou seja, a estimativa Bayesiana Empírica ignora a variabilidade introduzida pelos hiperparâmetros justamente por não considerar a incerteza na estimativa dos mesmos.

2.9.2 Estimação Inteiramente Bayesiana

Segundo Bernadinelli e Montonolli (1992), o método de estimação inteiramente Bayesiano, por outro lado, toda a análise é conduzida de forma puramente Bayesiana, ou seja, deve-se especificar uma distribuição *a priori* para todos os parâmetros antes de observar os dados. Esta *a priori* é combinada com a função de verossimilhança, obtida a partir dos dados observados, para gerar a distribuição a posteriori.

De acordo com Assunção (2001), a modelação inteiramente bayesiana considera que, por assumir que os hiperparâmetros são quantidades aleatórias, *hyperpriors* devem ser especificados aos mesmos, pois esta considera toda a variabilidade através da imposição de distribuições de probabilidade para os hiperparâmetros. A especificação de *hyperpriors* torna-se possível através da inserção de uma estrutura hierárquica na modelação.

Esta metodologia é por vezes apresentada na literatura como modelação hierárquica bayesiana, uma vez que a mesma compreende pelo menos três estágios de hierarquia:

- Primeiro nível: a verossimilhança $p(y | \theta)$, que relaciona os dados observados com os parâmetros do modelo;
- Segundo nível: a priori $\pi(\theta | \phi)$, onde ϕ representa os hiperparâmetros que caracterizam a distribuição a priori;
- Terceiro nível: a hiperpriori $\pi(\phi)$, que atribui distribuições de probabilidade aos hiperparâmetros.

Então, ao considerarmos que os hiperparâmetros são quantidades aleatórias, caracterizamos:

- ϕ é o vetor de hiperparâmetros que parametriza a distribuição a priori;
- $\pi(\theta | \phi)$ é a distribuição a priori do vetor de parâmetros θ , condicionada aos hiperparâmetros.

Deste modo, a modelação inteiramente bayesiana para os riscos das áreas em análise não só se apresenta superior à metodologia de estimação bayesiana empírica como supera as desvantagens de agregar áreas e dos mapas de probabilidade no combate ao problema das pequenas áreas, reduzindo consideravelmente o problema da variabilidade dos riscos estimados.

Tabela 1: Comparação entre Estimação Empírica Bayesiana e Estimação Inteiramente Bayesiana, fonte: *Da Silva et al., 2009*

Aspecto	Estimação Empírica Bayesiana	Estimação Inteiramente Bayesiana
Natureza	Semi-bayesiana (mistura frequentista e bayesiana)	Totalmente bayesiana
Tratamento da <i>a priori</i>	Derivada empiricamente dos dados	Definida explicitamente; hiperparâmetros podem ser aleatórios
Complexidade	Relativamente simples e computacionalmente eficiente	Mais complexa, requer maior poder computacional
Incerteza	Parcialmente incorporada	Completamente incorporada, incluindo a incerteza da <i>priori</i>
Flexibilidade	Limitada a modelos mais simples	Muito flexível, podendo lidar com modelos hierárquicos e não lineares

fonte: *Da Silva et al., 2009*

2.10 Modelos Hierárquicos Bayesianos

A modelação bayesiana, como referido anteriormente, é aquela que atribui informações a priori para os riscos relativos das áreas. Consequentemente, as inferências são derivadas da distribuição *a posteriori* resultante da combinação entre a informação *a priori* com a oriunda dos dados. Ao se considerar uma distribuição *a priori* para os riscos relativos, já se entra em um caso de modelação hierárquica, tal que o primeiro estágio da hierarquia está associado à verossimilhança $l(y|\theta)$, o segundo à *a priori* $\pi(\theta|\phi)$ condicionada aos hiperparâmetros.

Com fins de formalização, temos que:

- $y = (y_1, \dots, y_n)$: é o vector das contagens dos eventos;
- $\theta = (\theta_1, \dots, \theta_n)$: é o vector de parâmetros (riscos relativos) a serem estimados;
- n : é o número de áreas da região.

Considera-se padrão o procedimento de atribuir uma distribuição de Poisson para cada y_i , supondo independência entre as contagens das áreas. A hipótese de independência é realista quando se trata de um evento não contagioso. Sendo assim, a função de verossimilhança associada ao vector y corresponde a:

$$l(y|\theta) = \prod_{i=1}^n \frac{e^{-\lambda_i(\theta_i)} y_i!}{y_i!}. \quad (9)$$

Quanto à determinação das *prioris*, percebe-se que, na prática, *prioris* especificadas com todos os hiperparâmetros conhecidos são raramente utilizadas. A abordagem usual considera que as distribuições *a priori* para os riscos têm estrutura composta por hiperparâmetros desconhecidos. Então, posteriormente à definição de *prioris* para os riscos relativos, é preciso que os hiperparâmetros sejam estabelecidos. A especificação a respeito dos hiperparâmetros deve ser realizada subjectivamente de modo a retratar o conhecimento do investigador a respeito do evento em questão.

Segundo Bernardinelli *et al.* (1995), o número de casos observados na i -ésima área, e t -ésima unidade de tempo, definida por y_{it} é modelado através da distribuição de Poisson com média $E_{it}\theta_{it}$ onde θ_{it} é o risco relativo e E_{it} é o número de casos esperado, considerando para ambos a i -ésima área e a t -ésima unidade de tempo, caso o risco seja constante em toda a região e ao longo do tempo. As estimativas de θ_{it} são obtidas através da sua decomposição em efeitos espaciais, temporais e espaço temporais utilizando a função de ligação logarítmica entre os números de casos observados e os componentes do modelo.

Por forma a combater o problema geral da inferência Bayesiana, que consiste na impossibilidade de derivação da distribuição *a posteriori*, principalmente para modelos mais complexos, a literatura sugere o uso de métodos computacionais que foram desenvolvidos para facilitar os cálculos neste processo.

2.10.1 Cadeias de Markov via Monte Carlo

Segundo Caumo (2006), trata-se essencialmente de um procedimento de simulação estocástica que permite uma boa aproximação para a distribuição *posterior* desconhecida sendo de extrema utilidade sobre a óptica Bayesiana para problemas práticos, principalmente por não restringir o número de parâmetros no modelo e possuir um algoritmo de simulação relativamente simples. Os métodos *Markov Chain Monte Carlo* (MCMC) possibilitam a geração de amostras de distribuições *a posteriori* com função densidade de probabilidade desconhecida ou demasiadamente complicada.

Segundo Gamerman (1997), as cadeias de Markov são processos estocásticos de tempo discreto. Processos estocásticos por sua vez referem-se a processos aleatórios que ocorrem segundo uma função aleatória. São uma classe de geradores de amostras aleatórios cujas execuções convergem para as chamadas distribuições estacionárias.

Nos casos em que através do processo analítico ou até mesmo por aproximação numérica, não

se consegue obter a distribuição a posteriori de interesse a metodologia de integração MCMC surge como solução.

No presente trabalho, serão apresentados os principais métodos MCMC nomeadamente Metropolis-Hastings e Amostragem de Gibbs que por sua vez representa um caso particular do algoritmo de Metropolis-Hastings.

2.10.2 Algoritmo de Metropolis Hastings

Este procedimento é usado em situações em que a distribuição a posteriori condicional completa não apresenta uma forma conhecida, fechada sendo assim não identificável. Este método utiliza uma função densidade auxiliar (Vicente, 2011).

De acordo com Lawson *et al* (2003), a ideia geral consiste na geração de valores a partir de uma distribuição $p(\theta|y)$, chamada de distribuição proposta, que pode ser arbitrariamente escolhida. Os valores gerados θ^* são candidatos a formar a amostra da distribuição *posterior* de interesse $p(\theta)$, onde θ pode ser escalar (um parâmetro de interesse) ou corresponder a um vetor (mais de um parâmetro de interesse). Os candidatos θ^* farão parte da amostra de $p(\theta)$ com probabilidade específica que garante a equivalência entre a amostra resultante dos valores candidatos selecionados e uma amostra derivada da distribuição de interesse $p(\theta)$. A esta probabilidade atribui-se o nome de probabilidade de aceitação. Resumidamente, trata-se de um procedimento que garante a possibilidade de se gerar amostras de distribuições não conhecidas.

Mais especificamente, constrói-se uma cadeia de Markov onde, a cada tempo $t \geq 0$, gera-se um candidato para o próximo estado, θ^* , a partir de uma distribuição proposta $q(\theta|\theta_t)$. O valor gerado dependerá do actual valor da cadeia, e será “aceite” e adicionado à cadeia de forma que $\theta_{t+1} = \theta^*$ com probabilidade de aceitação dada por:

$$\alpha(\theta_t, \theta^*) = \min \left(1, \frac{\pi(\theta^*)q(\theta_t|\theta^*)}{\pi(\theta_t)q(\theta^*|\theta_t)} \right) \quad (10)$$

O mecanismo funciona com o intuito de rejeitar com maior probabilidade os pontos simulados que forem muito discrepantes em relação aos esperados pela distribuição $\pi(\theta)$. Percebe-se também, de forma intuitiva, que, quando a distribuição proposta e a distribuição de interesse forem similares, tal probabilidade tende a ser próxima de 1, aumentando a possibilidade de aceitação do valor gerado (Gelman, 1997).

O algoritmo de Metropolis-Hastings é formalizado pelos seguintes passos:

1. Define-se $t = 0$ e escolhe-se um valor arbitrário para θ_0 (valor inicial);
2. Gera-se um valor θ^* a partir de $q(\theta|\theta_t)$;
3. Gera-se um valor U de uma distribuição Uniforme(0, 1);
4. Se $U \leq \alpha(\theta_t, \theta^*)$, faz-se $\theta_{t+1} = \theta^*$. Caso contrário, $\theta_{t+1} = \theta_t$;

5. Incrementa-se t ;
6. Retorna-se ao passo 2, continuando as iterações até que o tamanho n seja suficientemente grande.

Sob certas condições de regularidade, a cadeia de Markov esquece os estados iniciais e converge para a distribuição estacionária $\pi(\theta)$. Sendo assim, para o processo inferencial, não são considerados os m primeiros valores gerados de θ (desconsidera-se o período de burn-in, ou seja, aquecimento). Obtém-se então, após a verificação da convergência da cadeia, em $\theta = (\theta_{m+1}, \dots, \theta_n)$, uma amostra não independente da distribuição $\pi(\theta)$.

2.10.3 Redução da Dependência

Para reduzir a dependência, um procedimento usual é considerar um intervalo que, em inglês, é conhecido como “thin interval”. Este procedimento consiste em aproveitar valores espaçados da cadeia de acordo com um intervalo pré-fixado. A ideia que fundamenta este conceito acredita que valores mais distantes no tempo estão menos autocorrelacionados.

Há possibilidade de se utilizar uma probabilidade de aceitação que depende da distribuição desconhecida de interesse $\pi(\theta)$. Porém a probabilidade de aceitação depende de $\pi(\theta)$ somente através do quociente $\pi(\theta^*)/\pi(\theta_t)$. Por definição, e considerando θ escalar, a posteriori é dada por:

$$\pi(\theta) = \frac{\pi(\theta|x)}{p(x)} \quad (11)$$

onde $p(x)$ é uma constante de normalização que não depende de θ . Desta forma, pode-se perceber que, tanto para $\theta = \theta^*$ quanto para $\theta = \theta_t$, o denominador não se altera. Logo, o processo depende apenas da prior $\pi(\theta)$ e da verossimilhança $l(x|\theta)$.

Tem-se então, por fim, uma amostra da distribuição de interesse de onde, através da integração do método de Monte Carlo, podemos obter estimativas de momentos da posteriori desconhecida de interesse. Além disso, pode-se fazer uso dos valores amostrados para obtenção de uma estimativa da própria posteriori $\pi(\theta|x)$. Uma técnica muito utilizada para esta finalidade corresponde ao método não paramétrico denominado “Kernel smoothing”.

2.10.4 Amostrador de Gibbs

O amostrador de Gibbs (*Gibbs Sampling*) é um caso particular do algoritmo de Metropolis-Hastings, adequado para ocasiões em que θ é um vetor de dimensão p , tal que $p \geq 2$. Logo, compartilha dos mesmos princípios básicos. Para a aplicação desta metodologia, faz-se necessário que a posteriori conjunta, mesmo que desconhecida, possa, pelo menos, ser traduzida em *posteriors* condicionais completas (Gamerman, 1997).

Por exemplo, quando se tiver interesse em uma distribuição que compreende três parâmetros, sendo $\pi(\theta) = \pi(\theta^1, \theta^2, \theta^3)$, necessitaríamos conhecer $\pi(\theta^1|\theta^2, \theta^3)$, $\pi(\theta^2|\theta^1, \theta^3)$ e $\pi(\theta^3|\theta^1, \theta^2)$,

ou quaisquer outras possibilidades de condicionais completas, tal qual $\pi(\theta^{(i)}|\theta^{(-i)})$, para que se faça possível a implementação do amostrador de Gibbs.

Visto que o amostrador de Gibbs corresponde a um caso particular do algoritmo de Metropolis-Hastings com distribuições propostas iguais às condicionais completas, pode-se demonstrar que a probabilidade de aceitação de cada valor gerado é igual a 1, fazendo com que a cadeia sempre aceite os novos candidatos. Isso faz com que, na prática, o amostrador de Gibbs seja mais eficiente e rápido.

Segundo Gamerman (1997), o amostrador de Gibbs é formalizado pelos seguintes passos:

1. Define-se $t = 0$ e escolhem-se valores iniciais $\theta_0 = (\theta_0^1, \theta_0^2, \dots, \theta_0^p)$;
2. Obtém-se o próximo estado da cadeia, $\theta_{t+1} = (\theta_{t+1}^1, \theta_{t+1}^2, \dots, \theta_{t+1}^p)$, com base nas gerações sucessivas oriundas das condicionais completas, tal que:

$$\theta_{t+1}^1 = \pi(\theta^1 | \theta_t^2, \theta_t^3, \dots, \theta_t^p) \quad (12)$$

$$\theta_{t+1}^2 = \pi(\theta^2 | \theta_{t+1}^1, \theta_t^3, \dots, \theta_t^p) \quad (13)$$

$$\theta_{t+1}^3 = \pi(\theta^3 | \theta_{t+1}^1, \theta_{t+1}^2, \dots, \theta_t^p) \quad (14)$$

$$\vdots$$

$$\theta_{t+1}^p = \pi(\theta^p | \theta_{t+1}^1, \theta_{t+1}^2, \dots, \theta_{t+1}^{p-1}) \quad (15)$$

3. Incrementa-se t ;

4. Retorna-se ao passo 2, procedendo às iterações até que o tamanho n seja suficientemente grande e satisfatório.

2.11 Diagnóstico de Convergência

Segundo Vicente (2011), um dos principais problemas na aplicação dos métodos MCMC é determinar se uma cadeia já convergiu para a distribuição alvo, um aspecto importante para essa avaliação de convergência é o período de aquecimento, conhecido como *burn-in*, esse período inicial é caracterizado por valores da cadeia que ainda não alcançaram sua distribuição estacionária. Nessa fase, as observações podem variar amplamente e apresentar uma forte dependência dos valores iniciais da cadeia.

Segundo Turkman (2000), a quantidade ideal de iterações depende da velocidade com que a cadeia se aproxima do estado de equilíbrio e da estrutura de correlação entre as componentes. Para garantir a confiabilidade da amostra, é necessário selecionar apenas os estados obtidos após o período de aquecimento, quando a cadeia já atingiu o equilíbrio. Além disso, se houver uma correlação significativa entre os estados gerados, recomenda-se escolher amostras não consecutivas, mas espaçadas por um intervalo adequado, que assegure a independência entre os estados selecionados.

2.11.1 Critério de Gelman e Rubin

Segundo Vicente (2011), este método se baseia na comparação da variabilidade entre as diferentes cadeias de simulação, realizadas em paralelo. Para aplicá-lo, escolhem-se valores iniciais amplamente dispersos pelo espaço paramétrico, permitindo que cada cadeia comece em um ponto distinto. Para cada cadeia, são feitas $2n$ iterações, sendo que as primeiras n são descartadas para evitar que os valores iniciais, ou período de aquecimento, influenciem o resultado final. Quando as inferências feitas a partir das diferentes cadeias mostram-se suficientemente próximas, isso sugere que a convergência foi alcançada, ou está prestes a ser. Assim, caso haja convergência, os valores finais das cadeias (50% das iterações) compõem uma amostra representativa da distribuição marginal desejada. De acordo com Gelman *et al.* (1995) deve-se aceitar que houve convergência quando o factor de redução de escala estiver compreendido entre 1 e 1.1.

2.11.2 Critério de Heidelberger e Welch

De acordo com Roy (2020), este método propõe testar a hipótese de estacionariedade da sequência gerada, por meio de testes estatísticos. Uma vez que a cadeia tenha passado no teste de estacionariedade, este teste calcula um intervalo de confiança com uma "largura da metade" para a média. Se o intervalo de confiança for suficientemente pequeno, a cadeia é considerada convergida. Caso contrário, é necessário continuar gerando mais amostras até que se alcance a convergência pretendida. Este teste permite inferir um valor para o tamanho do período de aquecimento.

2.11.3 Critério de Geweke

De acordo com Nogueira (2004), este critério, por sua vez, propõe um diagnóstico de convergência para cadeias de Markov baseadas no teste de igualdade de médias da primeira e última parte da cadeia de Markov, geralmente os primeiros 10% e dos últimos 50%. Calcula-se a média e a variância de ambas partes e se as médias das duas partes forem suficientemente próximas, conclui-se que a cadeia alcançou a convergência.

3 Metodologia

3.1 Material

Para a realização deste estudo, foi utilizada uma base de dados secundária proveniente do Inquérito Demográfico de Saúde (IDS) 2022-2023, com foco em mulheres, disponível no site do The DHS Program (<https://dhsprogram.com/data/available-datasets.cfm>). Essa base contém informações abrangentes sobre diversos aspectos demográficos, socioeconômicos e de saúde. Adicionalmente, foi utilizada uma base com informações geoespaciais, também disponibilizada mediante solicitação no mesmo portal, a qual foi essencial para a análise espacial desenvolvida no presente estudo. A elaboração do relatório foi realizada no software $\text{\LaTeX}$, o processamento e análise dos dados foram conduzidos no RStudio, versão 3.5.1 e o OpenBUGS versão 3.2.3 e Para a produção dos mapas temáticos, foi utilizado o software ArcGIS, versão 10.2.1.

3.1.1 Desenho de Estudo

Trata-se de um estudo transversal retrospectivo com abordagem quantitativa que visou analisar a distribuição espaço-temporal da gravidez na adolescência em Moçambique no período de 2011 a 2022. A população de estudo é composta por adolescentes do sexo feminino com idade igual ou inferior a 19 anos. A área de estudo corresponde ao território nacional coberto pela colecta de dados do IDS-2022.

O processo de amostragem do IDS foi desenhado para ser representativo a nível provincial pelo que alguns distritos não foram incluídos na amostra. Além disso, devido ao contexto de conflito armado na região norte do país, certas áreas foram excluídas por razões de segurança dos inquiridores. No total, o inquérito abrangeu 133 distritos de Moçambique. Como o levantamento foi realizado em um único momento (transversal), foi necessário estimar o ano da primeira gravidez das mulheres entrevistadas. Para isso, adoptou-se uma abordagem retrospectiva, que permitiu identificar, com base nas informações fornecidas pelas respondentes, o período aproximado em que ocorreram as primeiras gestações, pela seguinte fórmula:

$$\text{Ano da 1ª gravidez} = \text{Ano do inquérito} - (\text{Idade à data do inquérito} - \text{Idade à 1ª gravidez})$$

3.1.2 Métodos

Por forma a alcançar os objectivos propostos para o presente estudo foi feita em primeiro momento uma análise exploratória dos dados. A análise exploratória de dados é uma etapa fundamental na estatística e na investigação científica, pois permite uma compreensão inicial dos padrões, tendências e variabilidade presente nos dados. Através de técnicas descritivas e gráficas, é possível identificar algumas inconsistências, valores atípicos e estruturas subjacentes que

podem influenciar a modelação estatística e a interpretação dos resultados. Além disso, esta análise facilita a escolha de métodos adequados para investigações subsequentes, garantindo que as inferências estatísticas sejam baseadas em suposições devidamente avaliadas (Martinez *et al.*, 2017).

De seguida, foi feita a estimação do risco relativo de gravidez na adolescência para cada distrito e posteriormente o respectivo mapeamento destas estimativas. O risco relativo foi estimado usando duas metodologias, a estimação clássica e a estimação bayesiana. A estimação clássica empregou fórmulas tradicionais de padronização de taxas, enquanto que a estimação Bayesiana, por sua vez, consistiu na estimação de um modelo hierárquico para incorporar a variabilidade entre distritos e os anos, suavizando as estimativas.

A convergência das cadeias de Markov foi avaliada utilizando o método de Gelman e Rubin, simulando três cadeias em paralelo com valores iniciais dispersos. Conforme Lawson *et al.* (2003), a convergência foi avaliada por meio de diagnósticos combinados com inspeção visual de gráfico de traços e diagramas de autocorrelação. Correlogramas foram usados para controlar a autocorrelação entre os estados das cadeias.

O modelo bayesiano adoptado para este estudo baseia-se num modelo de interação espaço-tempo concebido para a análise da variação geográfica do risco de eventos de saúde. Este modelo pressupõe que o logaritmo do risco relativo evolui linearmente ao longo do tempo, com coeficientes específicos para cada área geográfica (Kelsall *et al.*, 1999). Além disso, é imposta uma distribuição *a priori* que assume que áreas geograficamente próximas apresentam perfis de risco semelhantes, capturando assim a dependência espacial inerente aos dados. Esta modelagem foi conduzida assumindo uma estrutura de contagem que segue uma distribuição de Poisson:

$$Y_{it} \sim \text{Poisson}(\mu_{it}) \quad (16)$$

onde:

$$\mu_{it} = E_{it}\theta_{it} \quad (17)$$

Onde Y_{it} representa o número observado de casos de gravidez na adolescência na área i e no tempo t , E_{it} é o número esperado de casos, e θ_{it} é o risco relativo. O risco relativo é modelado no logaritmo da seguinte forma:

$$\log(\mu_{it}) = \log(E_{it}) + \log(\theta_{it}) \quad (18)$$

$$\log(\theta_{it}) = \alpha + u_i + (\beta + \delta_i)t \quad (19)$$

Ou seja:

$$\log(\theta_{it}) = \alpha + u_i + \beta t + \delta_{it} \quad (20)$$

onde:

- α é o termo de intercepto;
- u_i é o efeito espacial específico da área i ;
- β_t representa o efeito temporal;
- δ_{it} é o termo de interação espaço-tempo.

A distribuição *a posteriori* dos parâmetros foi obtida por amostragem via métodos MCMC, especificamente o Amostrador de Gibbs, no qual a distribuição proposta é gerada a partir da distribuição condicional de cada parâmetro, dado os demais, o que exigiu a especificação dos valores iniciais dos parâmetros (α e β) e dos hiperparâmetros desconhecidos (τ_δ e τ_u) para cada cadeia em simulação. Foram pré-definidos valores não informativos como sugere a literatura de modo que os dados observados predominem na inferência posterior.

$$\alpha \sim \text{dnorm}(0.01, 0.001) \quad (21)$$

$$\beta \sim \text{dnorm}(0.01, 0.001) \quad (22)$$

$$\tau_u \sim \text{dgama}(0.05, 0.0005) \quad (23)$$

$$\tau_\delta \sim \text{dgama}(0.05, 0.0005) \quad (24)$$

Para a realização da amostragem, foram utilizadas três cadeias de Markov independentes. Cada cadeia passou inicialmente por um período de aquecimento (burn-in) de 2000 iterações, durante o qual as amostras, potencialmente influenciadas pelas condições iniciais, foram descartadas, permitindo que a cadeia alcançasse sua distribuição estacionária. Após esse período, foi aplicado um thin-interval (subamostragem) com intervalo de 10, ou seja, apenas uma a cada 10 iterações foi retida, a fim de reduzir a autocorrelação entre amostras consecutivas. Ao final do processo, foram obtidas 68400 amostras no total, correspondendo a 22800 amostras por cadeia.

Por fim, compararam-se as estimativas de risco relativo obtidas pelos métodos clássico e Bayesiano, buscando identificar qual dos métodos forneceria melhores estimativas, ou seja, estimativas que sejam de melhor interpretação.

3.1.3 Hipóteses e Regra de decisão

Hipóteses do teste de Moran:

- H_0 : Não há autocorrelação espacial significativa
- H_1 : Existe autocorrelação espacial significativa.

Hipóteses do teste de Ljung-Box:

- H_0 : Não há autocorrelação nos resíduos da série temporal.
- H_1 : Há autocorrelação nos resíduos da série temporal.

$$\text{Regra de Decisão: } \begin{cases} \text{Se } p < \alpha = 0.05 \text{ rejeita-se } H_0. \\ \text{Se } p \geq \alpha = 0.05 \text{ não se rejeita } H_0. \end{cases} \quad (25)$$

4 Resultados

4.1 Análise Exploratória

Nesta secção, é apresentada a distribuição dos casos de gravidez na adolescência ao longo do período em estudo e nos diferentes distritos que compõem a área de análise, destacando padrões espaciais e temporais relevantes. Serão utilizadas medidas descritivas, representações gráficas e mapas temáticos para ilustrar a variação da incidência da gravidez na adolescência em Moçambique. Posteriormente, proceder-se-á ao mapeamento da estimativa clássica do risco relativo, permitindo identificar as áreas com maior vulnerabilidade e possíveis concentrações de risco. Esta análise contribuirá para uma melhor compreensão da dinâmica da gravidez na adolescência no país.

4.1.1 Distribuição dos Casos de Gravidez na Adolescência

A Tabela 2, que apresenta a distribuição dos casos de gravidez na adolescência por ano, revela que, ao longo de todo o período em estudo, mais de 50% das mulheres entrevistadas que tiveram a primeira gravidez durante este intervalo engravidaram na adolescência. Este resultado sugere um índice elevado de gravidez na adolescência no país, situando-se em torno dos 60%, com excepção do ano de 2018, em que a estimativa foi de 58%.

Adicionalmente, a análise do boxplot, na figura 1 que ilustra a média dos casos de gravidez na adolescência para cada ano, indica que, embora a média mais elevada por distrito não tenha atingido cinco casos, verifica-se um número considerável por ano de distritos com valores extremos para a incidência. Este padrão de distribuição, caracterizado pela presença de distritos com uma frequência consideravelmente elevada de casos, mantém-se consistente ao longo de todos os anos em estudo, evidenciando possíveis desigualdades regionais na ocorrência da gravidez na adolescência.

Tabela 2: Distribuição dos casos de gravidez na adolescência por ano

Ano	Casos de 1ª Gravidez	Gravidez na adolescência	%
2011	404	263	65%
2012	400	255	64%
2013	434	293	68%
2014	498	322	65%
2015	535	335	63%
2016	492	335	68%
2017	659	446	68%
2018	732	423	58%
2019	667	418	63%
2020	542	323	60%

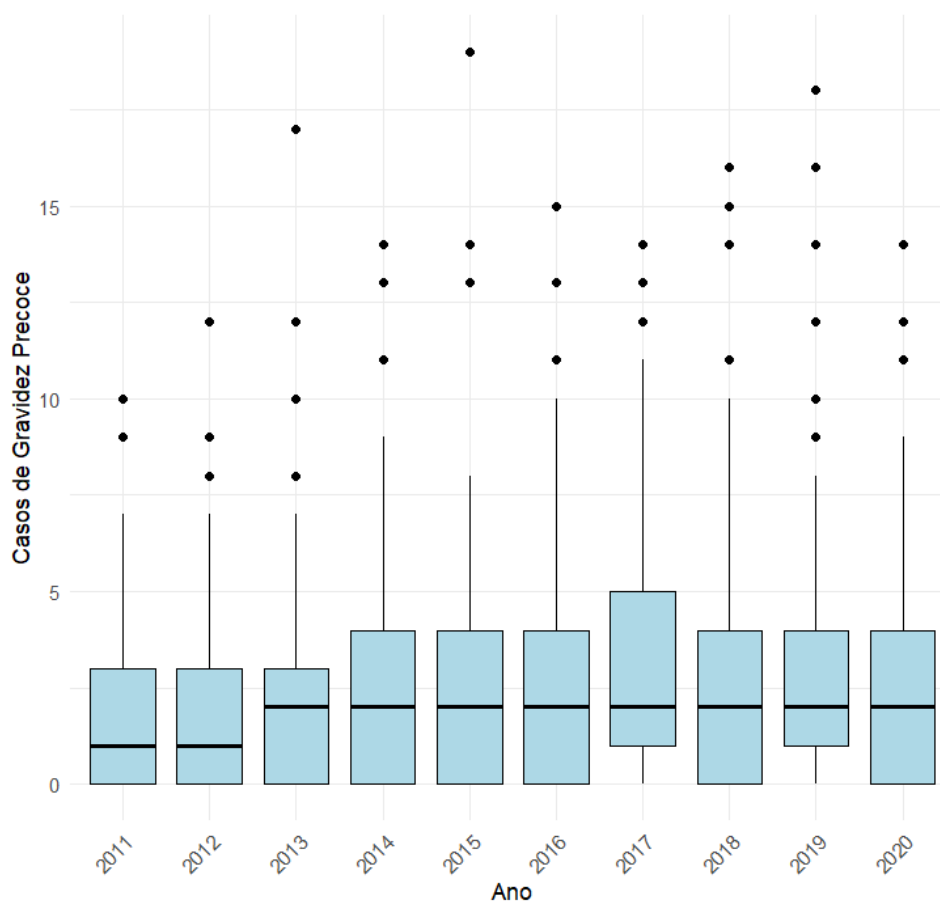


Figura 1: Boxplot dos casos de gravidez na adolescência por ano

4.1.2 Análise Clássica do risco relativo

A análise clássica do risco relativo consistiu no mapeamento de taxas padronizadas de incidência da gravidez na adolescência para explorar a distribuição espacial e temporal das mesmas. Considerando os mapas apresentados na Figura 2, pode-se verificar que alguns distritos das zonas sul, centro e norte apresentam um risco relativo elevado no período de estudo. Observa-se que os cinco distritos com as maiores taxas de gravidez na adolescência variam ao longo dos anos analisados. Em 2011, destacam-se os distritos de Machanga, Macossa, Zumbu, Ribáuè e Cheringoma. No ano de 2015, os maiores índices foram registados nos distritos de Chinde, Govuro, Cidade de Inhambane, Cheringoma e Funhalouro. Já em 2020, os distritos com as maiores taxas foram Boane, Zumbu, Lalaua, Marracuene e Zavala. Entretanto, nota-se uma grande dispersão do padrão de cores, o que torna a interpretação dos mapas tanto quanto difícil. A razão de se estar a observar este cenário pode-se atribuir ao problema das pequenas áreas, como referido anteriormente. Deste modo, recorreu-se à modelação bayesiana.

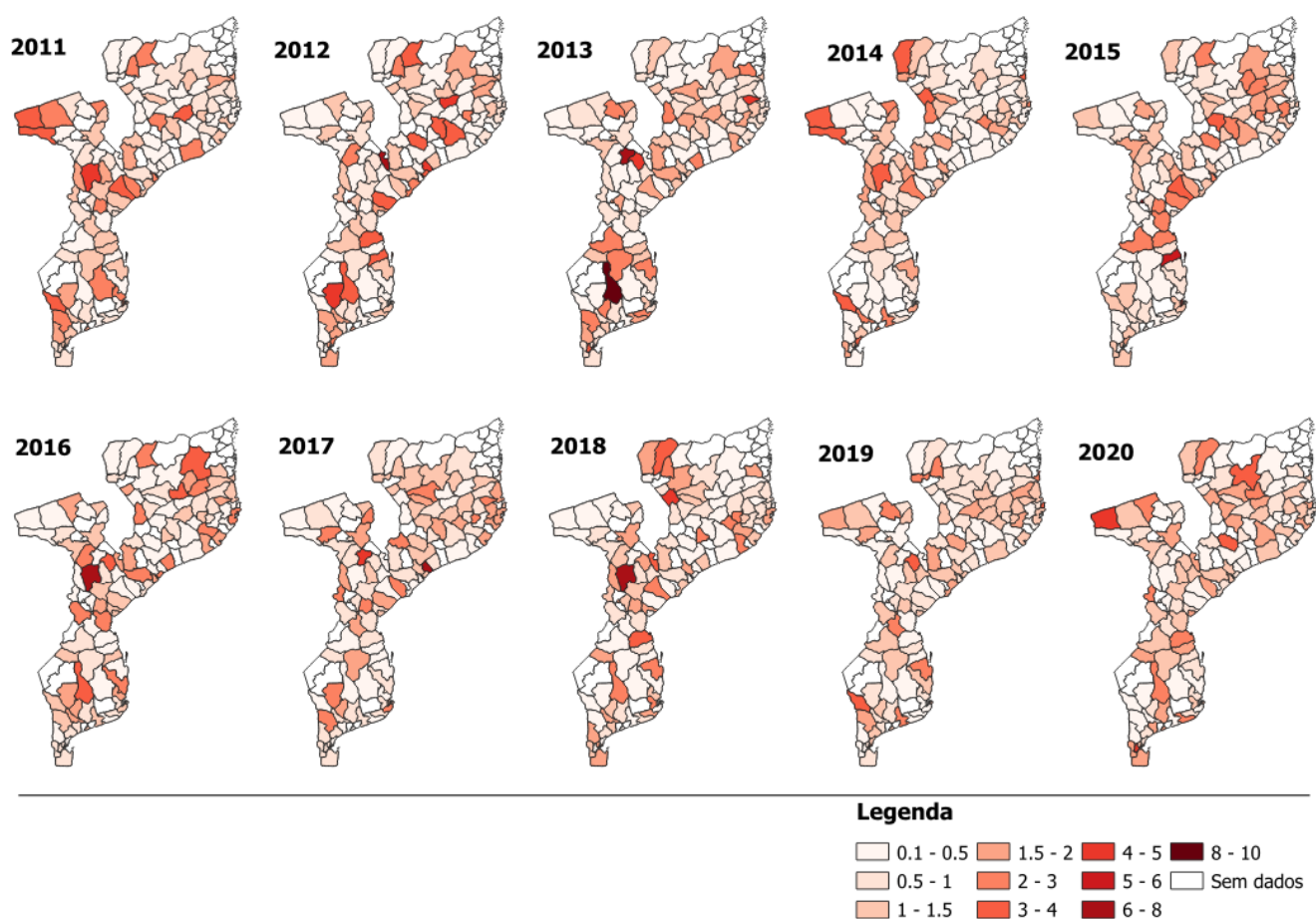


Figura 2: Estimação clássica do risco relativo de gravidez na adolescência de 2011 a 2020

4.2 Análise Bayesiana

Visto que os mapas feitos por meio da estimaco clssica apresentada em 4.1.2 no fornecem uma interpretao clara da distribuico da gravidez na adolescncia para os distritos, sobretudo ao longo do perodo em estudo, nesta seco, foi feita a estimaco de um modelo espaco-temporal hierrquico bayesiano, para a estimaco do risco relativo de gravidez na adolescncia, por forma a que se tivessem taxas mais suavizadas e que considerassem tanto a variabilidade espacial, bem como a variabilidade temporal do evento.

Para que esta tcnica seja considerada uma opo vlida para contornar a limitao observada na estimaco clssica,  necessrio que exista nos dados uma dependncia espacial entre as observaes e a presena de autocorrelao no perodo de estudo, deste modo, foi usado o ndice de Moran para avaliar a dependncia espacial das observaes e o teste de L-Jung Box para verificar a presena de autocorrelao nos dados.

4.2.1 Autocorrelação espaço temporal

De acordo com as tabelas 3 e 4, pode-se observar que existe uma dependência espacial entre os casos de gravidez na adolescência entre os diferentes distritos de Moçambique, pois para cada um dos anos, com um p-valor menor que o nível de significância de 5%, rejeita-se a hipótese nula de que não existe dependência espacial entre as observações. O mesmo se verifica para a autocorrelação temporal, pois, de acordo com o teste de Ljung-Box, que apresenta um p-valor menor que o nível de significância estabelecido para este estudo, conclui-se assim que os casos de gravidez na adolescência estão correlacionados ao longo do período em estudo.

Tabela 3: Índice de Moran

Ano	Moran I Statistic	p-value
2011	0.117818177	0.01354
2012	0.113912188	0.01558
2013	0.027643049	0.2615
2014	0.065708853	0.0966
2015	0.196391620	0.00011
2016	0.198359954	0.00013
2017	0.203488372	0.00010
2018	0.192224270	0.00020
2019	0.216904379	3.02e-05
2020	0.121471940	0.01077

Tabela 4: Teste de LjungBox

Estatística Qui-Quadrado	p-value
22293516	2.2e-16

4.2.2 Análise da convergência do modelo

Para que os parâmetros do modelo e por fim, as estimativas do risco relativo ajustadas pelo método bayesiano sejam considerados válidos, é necessário que a convergência no modelo tenha sido alcançada. Para garantir estimativas confiáveis, esta etapa não deve ser negligenciada. Para efeitos de avaliação da convergência, recomenda-se que sejam geradas duas ou mais cadeias paralelas pois, de acordo com Gelman e Rubin (1992), a geração de múltiplas cadeias paralelas é uma prática recomendada para avaliar a convergência em simulações MCMC, permitindo uma comparação mais consistente e precisa entre as cadeias e dentro das cadeias.

De acordo com o teste de Gelman e Rubin (1992), que se baseia na comparação da variância entre cadeias e a variância dentro das cadeias, o valor do PSRF (Potential Scale Reduction Factor) deve situar-se entre 1 e 1.1. Valores dentro desse intervalo indicam que as cadeias estão

a convergir para a mesma distribuição. Por outro lado, desvios acentuados entre as diferentes cadeias sugerem falta de convergência.

Os gráficos de traços na Figura 3 mostram a evolução dos valores gerados pela MCMC ao longo das iterações. Para o parâmetro delta, os valores simulados apresentam oscilações em torno das suas médias, um comportamento típico de cadeias de Markov que atingiram a convergência. Embora o gráfico de traços para o parâmetro u mostre uma média relativamente mais baixa ao longo do tempo, ele também tende a se estabilizar e se aproximar da distribuição estacionária, indicando que a convergência foi alcançada, embora de forma mais gradual. Em ambos os casos, os padrões observados são sinais de que os valores simulados atingiram a estabilidade, confirmando a convergência dos parâmetros.

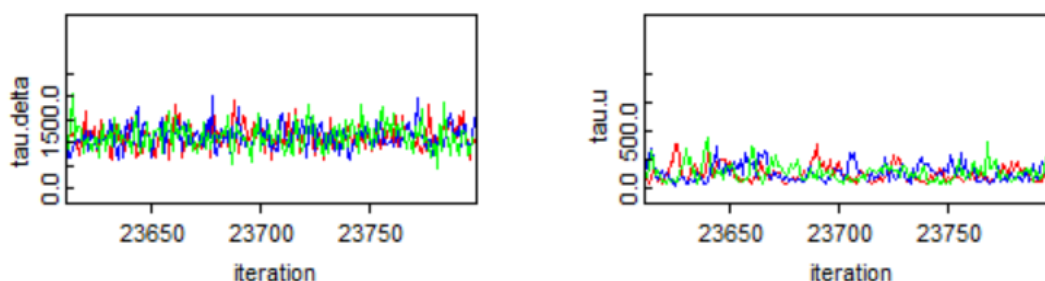


Figura 3: Gráficos de traços para delta e u

Os gráficos de correlação apresentados na figura 4, indicam a ausência de autocorrelação nos estados das cadeias simuladas, pois os coeficientes de autocorrelação apresentam quedas acentuadas e seus valores tendem a zero à medida que o número de defasagens aumenta. Esse comportamento sugere que as amostras geradas pelas cadeias MCMC são independentes e que a convergência foi alcançada.

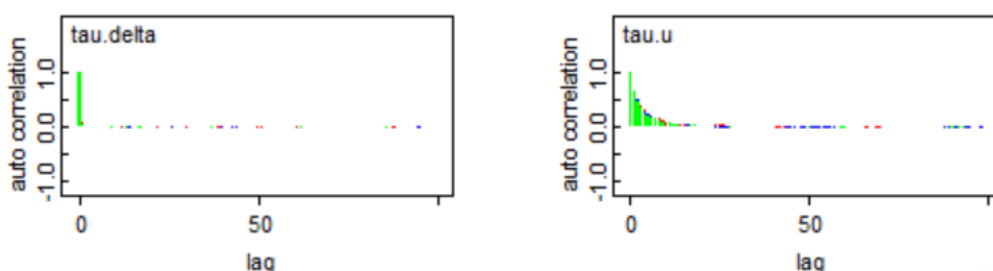


Figura 4: gráfico de correlação de delta e u

Nos gráficos de diagnóstico de Gelman e Rubin para os parâmetros delta e u que são apresentados na Figura 5, ambos parâmetros tau.delta e tau.u, mostram valores de BGR próximos de 1, sem grandes flutuações ao longo das iterações. Isso sugere que as cadeias alcançaram convergência de forma satisfatória, conferindo confiabilidade aos resultados inferidos pelo modelo.

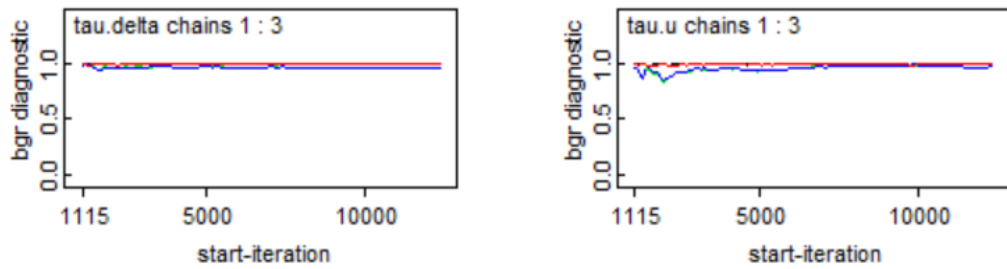


Figura 5: diagnostico de Gelman e Rubin para delta e u

4.2.3 Estimativas dos parâmetros do modelo

Tabela 5: Estimativas dos parâmetros do modelo

Parâmetro	Estimativa	Desvio Padrão	2.5%	97.5%	Erro MC
α	-3.433	0.04149	-3.514	-3.352	0.0001859
β	0.1915	0.006483	0.1789	0.2043	0.00002866
τ_δ	1122	266.5	675.7	1720	1.096
τ_u	153.4	93.9	43.51	398.5	0.9626

Através da Tabela 5, pode-se verificar pelo intervalo de credibilidade que, com 95% de credibilidade, todos os parâmetros do modelo são significativos, pois as médias estimadas para cada um dos parâmetros estão dentro do intervalo estimado e as estimativas são mais de duas vezes maiores que o respectivo desvio padrão. Em relação às constantes alfa e beta, observa-se que o modelo bayesiano modela risco relativo como uma função exponencial dos efeitos estruturais e aleatórios, tal que:

$$\theta_{it} = \exp(\alpha + u_i + \beta_t + \delta_{it}) \quad (26)$$

Portanto, $\exp(\alpha) = 0.03$ representa o nível base estimado do risco relativo global de gravidez na adolescência. Este valor indica que, na ausência de efeitos espaciais e temporais adicionais, o risco relativo é significativamente inferior a 1, sugerindo um risco bastante reduzido em relação à média geral.

Por outro lado, $\exp(\beta) = 1.21$ representa o efeito temporal global. Este valor indica que, a cada ano, há um aumento de aproximadamente 21% no risco relativo de gravidez precoce em comparação com o ano anterior, assumindo que os efeitos espaciais permaneçam constantes.

Os parâmetros τ_u e τ_δ representam a precisão do efeito espacial aleatório e a precisão do efeito espaço-temporal, respectivamente. Uma vez que a precisão é a inversa da variância, tal que:

$$\tau = \frac{1}{\sigma^2} \quad (27)$$

ou equivalentemente:

$$\sigma^2 = \frac{1}{\tau} \quad (28)$$

valores altos de precisão indicam menor variabilidade, enquanto valores baixos sugerem maior variação, portanto, observa-se variabilidade tanto espacial quanto espaço-temporal, contudo, no modelo, sendo $\tau_u = 153.4$ menor que $\tau_\delta = 1122$, pode-se dizer que a variabilidade no risco relativo é mais acentuada entre os distritos do que ao longo do tempo dentro de cada distrito.

Tendo sido verificada a confiabilidade das estimativas do modelo inteiramente bayesiano, foram mapeadas as estimativas suavizadas do risco relativo, permitindo uma interpretação mais clara e precisa dos padrões espaciais e espaço-temporais da incidência da gravidez na adolescência em Moçambique.

4.2.4 Mapeamento das estimativas

De acordo com o mapeamento apresentado na Figura 6, referente à estimação bayesiana do risco relativo de gravidez na adolescência em Moçambique durante o período de 2011 a 2020, observa-se um padrão geográfico consistente na incidência do evento. De forma geral, observa-se um aumento progressivo em todo o país ao longo dos anos, evidenciado pela transição de tonalidades mais claras para mais escuras nos mapas. Observa-se ainda que este aumento incide mais rapidamente sobre os distritos das regiões norte e centro do país, sendo que para alguns distritos da região sul ocorre de forma relativamente mais lenta e apresentando um risco relativamente mais baixo durante o período analisado quando comparado com as regiões norte e centro.

Observou-se que, em 2011, os cinco distritos com maior risco estimado foram Montepuez, Cahora Bassa, Cidade de Quelimane, Pemba e Ancuabe; já em 2015, destacaram-se os distritos de Montepuez, Cidade de Quelimane, Pemba, Ancuabe e Cahora Bassa; e em 2020 mantiveram-se nos primeiros lugares Cidade de Pemba, Montepuez, Cidade de Quelimane, Dondo e Pemba. Esse padrão evidencia a persistência dos distritos de Montepuez, Cidade de Quelimane e Pemba entre os maiores valores de risco de gravidez na adolescência ao longo do período em estudo. Observou-se também que o menor risco de gravidez na adolescência concentra-se nos distritos de Ka-Mpfumo, Ka-Mubukuane, Ka-Maxaquene, Ka-Chamanculo e Ka-Tembe, padrão que também verifica-se de forma persistente por praticamente todo período de estudo.

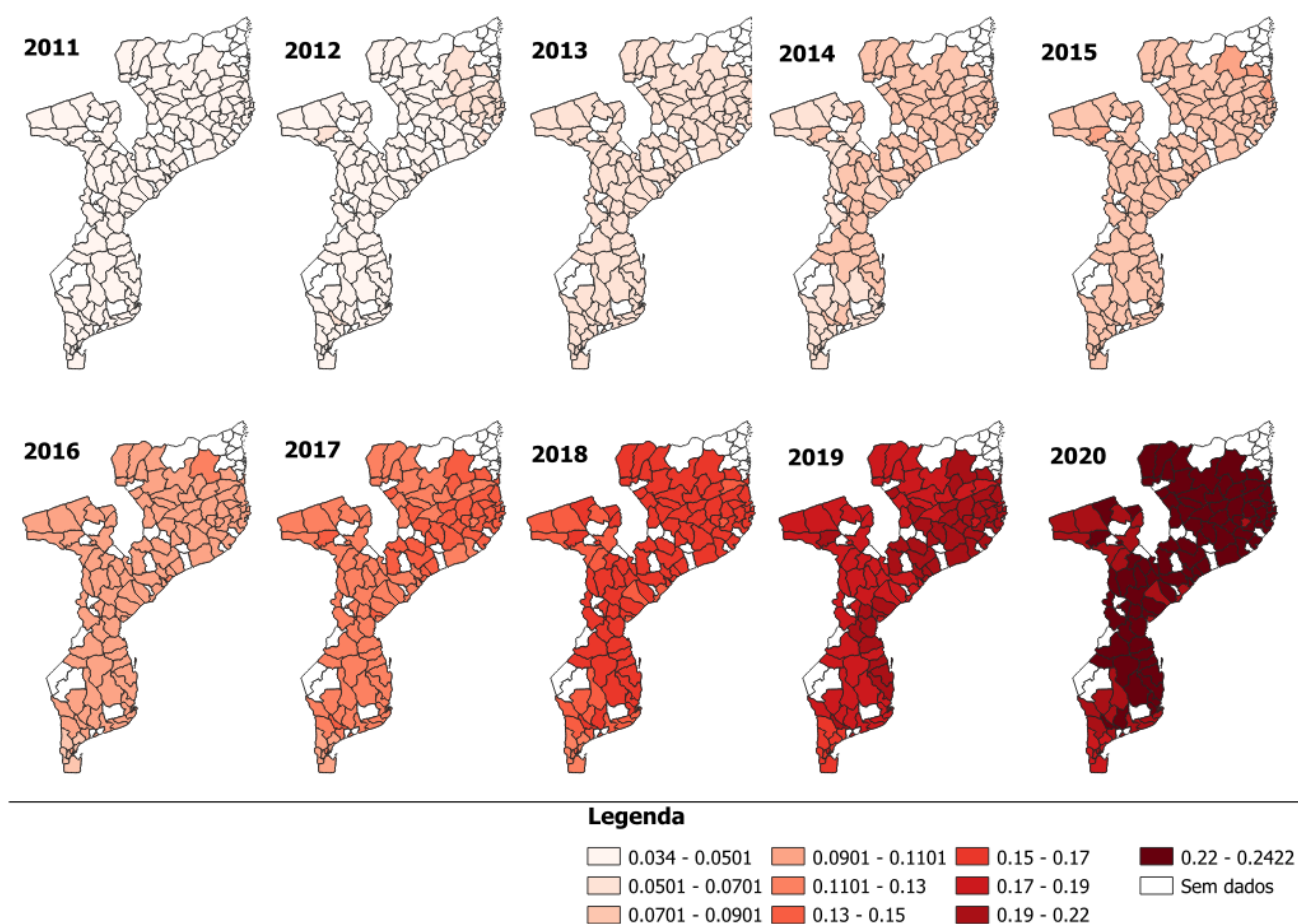


Figura 6: Mapeamento das estimativas do risco relativo pela modelação bayesiana

4.2.5 Classificação das estimativas de Risco Relativo

As tabelas 6 e 7 apresentam os cinco distritos com maior e menor risco relativo de gravidez na adolescência, juntamente com seus respectivos intervalos de credibilidade a um nível de confiança de 95%. As tabelas mostram diferenças significativas no risco relativo de gravidez na adolescência entre os distritos ao longo dos anos analisados. Em 2011, os cinco distritos com maior θ_{it} apresentaram valores entre 0,04209 e 0,04226. Já os distritos com menor risco apresentaram θ_{it} de 0,03493 e 0,03504, com intervalos de credibilidade estreitos e que não se sobrepõem aos dos distritos de maior risco, indicando elevada certeza na estimativa.

Em 2015, os distritos com maior risco apresentaram estimativas entre 0,09077 e 0,09147, enquanto os de menor risco variaram entre 0,07172 e 0,07207. Os intervalos de credibilidade mantiveram-se relativamente estreitos, o que indica confiança nas estimativas e reforça as disparidades entre os distritos. Em 2020, observou-se um aumento geral nos valores de θ_{it} , com os distritos de maior risco atingindo valores em torno de 0,2383 a 0,2422 e os de menor risco variando entre 0,1772 e 0,1786. Contudo, a separação dos intervalos de credibilidade continua evidente, sugerindo persistência das disparidades territoriais no risco de gravidez na adolescência.

cia ao longo do tempo.

Tabela 6: Top 5 de distritos com maior RR de gravidez na adolescência em 2011, 2015 e 2020

Ano	Distrito	RR	IC 95%
2011	Montepuez	0.042226	0.03472 – 0.05019
	Cahora Bassa	0.04216	0.03415 – 0.05092
	Cidade de Quelimane	0.04215	0.03400 – 0.05134
	Pemba	0.04214	0.03431 – 0.05076
	Ancuabe	0.04209	0.03445 – 0.05015
2015	Montepuez	0.09147	0.08160 – 0.10280
	Cidade de Quelimane	0.09115	0.07711 – 0.10730
	Pemba	0.09086	0.07806 – 0.10520
	Ancuabe	0.09081	0.07972 – 0.10300
	Cahora Bassa	0.09077	0.07813 – 0.10480
2020	Cidade de Pemba	0.2422	0.1932 – 0.3025
	Montepuez	0.2407	0.2064 – 0.2796
	Cidade de Quelimane	0.2402	0.1847 – 0.3060
	Dondo	0.2386	0.1967 – 0.2896
	Pemba	0.2383	0.1893 – 0.2952

Tabela 7: Top 5 de distritos com menor RR de gravidez na adolescência em 2011, 2015 e 2020

Ano	Distrito	RR	IC 95%
2011	KaMavota	0.03493	0.02599 – 0.04324
	KaMubukwana	0.03494	0.02620 – 0.04316
	Nlhamankulu	0.03498	0.02615 – 0.04329
	KaMpfumu	0.03502	0.02633 – 0.04324
	KaMaxakeni	0.03504	0.02633 – 0.04331
2015	KaMubukwana	0.07172	0.05987 – 0.08311
	KaMavota	0.07185	0.05922 – 0.08416
	Nlhamankulu	0.07187	0.05953 – 0.08414
	KaMpfumu	0.07202	0.05982 – 0.08386
	KaMaxakeni	0.07207	0.06008 – 0.08370
2020	KaMubukwana	0.1772	0.1390 – 0.2199
	Nlhamankulu	0.1779	0.1363 – 0.2246
	KaMavota	0.1781	0.1357 – 0.2264
	KaMpfumu	0.1784	0.1376 – 0.2238
	KaMaxakeni	0.1786	0.1384 – 0.2239

5 Discussão

O modelo empregado neste estudo é um modelo espaço-temporal com estrutura paramétrica que captura a evolução do risco ao longo do tempo e entre diferentes unidades espaciais. Esta abordagem baseia-se na premissa de que tanto o tempo quanto o espaço contribuem conjuntamente para a variabilidade observada nos dados, sendo, portanto, inadequado tratá-los separadamente. Segundo Pinheiro (2009), modelos espaço-temporais permitem que a inferência seja melhorada por meio do compartilhamento de informação entre regiões vizinhas e entre períodos próximos, oferecendo estimativas mais estáveis e realistas.

Na literatura, diversas classes de modelos têm sido propostas para lidar com dados espaço-temporais. Modelos puramente espaciais, como o modelo de Besag-York-Mollié (BYM) abordado em Morris *et al.* (2019), consideram apenas a estrutura de vizinhança entre regiões, assumindo independência temporal. Já os modelos unicamente temporais, como os modelos autorregressivos (AR) e outros modelos de séries temporais, ignoram a dependência espacial, tratando cada unidade geográfica de forma isolada. Essas abordagens, embora simples, perdem eficiência em contextos onde existe dependência clara entre o espaço e o tempo.

Modelos espaço-temporais, com estrutura paramétrica como o modelo utilizado neste estudo, oferecem uma estrutura computacionalmente eficiente e são úteis para capturar padrões regulares, assumindo que o risco evolui segundo funções lineares ou quadráticas ao longo do tempo, com coeficientes estruturados espacialmente (Griffith e Paelinck, 2011). Contudo, essas suposições podem ser inadequadas para períodos mais longos ou padrões não lineares, limitando a capacidade de adaptação do modelo. Como alternativa, modelos mais flexíveis, como o proposto por Knorr-Held (2000), introduzem componentes estruturais que permitem a decomposição do risco em efeitos espaciais e temporais independentes, bem como em termos de interação espaço-tempo. Esses modelos podem incluir componentes suavizados para representar a evolução temporal e espacial sem pressupor formas funcionais específicas. Sua vantagem reside na adaptabilidade a diferentes padrões de risco, entretanto, requerem maior esforço computacional e cuidados na especificação das distribuições priors e da estrutura de dependência.

Modelos bayesianos dinâmicos, baseados em cadeias de Markov com estados latentes, oferecem também uma alternativa robusta para lidar com séries espaço-temporais em contextos altamente incertos. Eles permitem a inclusão de choques estruturais, sazonalidades e outros componentes latentes que variam ao longo do tempo. Embora poderosos, tais modelos requerem técnicas de amostragem mais avançadas, como filtros de partículas ou MCMC mais complexos (Pacca, 2015).

Nesse contexto, embora o modelo adotado neste estudo não seja um modelo dinâmico com

estados latentes propriamente ditos, ele compartilha a filosofia bayesiana ao incorporar estrutura hierárquica espaço-temporal que captura variações ao longo das regiões e do tempo. A modelação da dependência espacial por meio de componentes Auto-regressivo condicional (*Conditional Autoregressive-CAR*) e a inclusão de efeitos temporais e interativos permitem representar de forma flexível a heterogeneidade da distribuição do risco. Assim, mesmo sem recorrer a filtros de partículas ou técnicas de amostragem sequencial, a abordagem empregada oferece uma estrutura suficientemente robusta para lidar com a complexidade dos dados, desde que acompanhada de uma avaliação criteriosa da convergência das cadeias de Markov utilizadas na inferência.

Para que as estimativas do risco relativo encontradas a partir do modelo bayesiano utilizado neste estudo fossem consideradas confiáveis e as inferências não fossem consideradas sem validade estatística, foi necessário testar a convergência das cadeias de Markov à distribuição estacionária e a confiabilidade dos parâmetros do modelo. Esta verificação foi feita usando uma combinação de critérios de verificação. Pois, segundo Roy (2020), existem vários métodos para aferir se foi alcançada ou não a convergência das cadeias de Markov. Estes métodos de diagnóstico podem ser vistos sob duas perspectivas, isto é, algumas ferramentas de diagnóstico são concebidas para avaliar a convergência da cadeia de Markov à distribuição estacionária, enquanto outras verificam a convergência das estatísticas sumárias, como as médias e os quantis das amostras, para as quantidades populacionais correspondentes. Estes dois critérios são utilizados de forma combinada pois fornecem uma avaliação mais eficiente da convergência.

Estes critérios apontaram para a convergência do modelo, pois as MCMC para todos os parâmetros do modelo convergiram à distribuição estacionária. Os gráficos de correlação indicaram a não existência de autocorrelação dos estados das cadeias simuladas, o que é desejável pois a correlação alta entre os estados de uma cadeia MCMC pode dificultar a convergência, tornando a exploração do espaço amostral menos eficiente e exigindo mais amostras para obter estimativas confiáveis, o gráfico de diagnóstico de Gelman e Rubin apresentou valores do factor de potencial de escala reduzida próximos de 1, reforçando a ideia de que a convergência foi alcançada.

A utilização de um modelo Bayesiano com estrutura espaço-temporal permitiu analisar a incidência da gravidez na adolescência de forma abrangente, levando em consideração as variações geográficas e temporais. O modelo possibilitou identificar o risco médio global de ter gravidez na adolescência sem levar em consideração factores externos $\exp(\alpha)$, a tendência temporal global dos casos de gravidez na adolescência sem levar em consideração factores externos $\exp(\beta)$, o efeito espacial τ_u e as interações espaço-temporais τ_δ . Isso representa um avanço relevante em relação às análises tradicionais que vêm sendo feitas para a gravidez na adolescência em Moçambique, pois permitiu superar em termos metodológicos estas abordagens, trazendo re-

sultados que consideram não apenas o tempo ou apenas o espaço.

A análise da distribuição espaço-temporal do risco relativo de gravidez na adolescência em Moçambique entre 2011 e 2020 revelou padrões geográficos consistentes ao longo do tempo. Observou-se que os distritos com maior risco estão concentrados predominantemente nas regiões norte e centro do país, com destaque para os distritos de Montepuez, Pemba e Cidade de Quelimane. Esses distritos apresentaram um risco acima da média nacional em comparação com os demais distritos em todo o período de estudo.

A alta incidência de gravidez na adolescência sobretudo no norte e centro de Moçambique pode estar intimamente ligada a factores socioculturais prevalentes nas regiões. Os ritos de iniciação, comuns em grupos étnicos como macuas, macondes e muanis, são práticas que conferem às jovens o estatuto de adultas, independentemente da idade. Esses ritos desempenham um papel central na construção da identidade de género e na preparação para os papéis tradicionais ligados à vida conjugal, frequentemente legitimando o casamento prematuro como transição social esperada para a fase adulta (UNICEF, 2021). Apesar do valor cultural atribuído, estas práticas constituem um problema social por violarem os direitos humanos de mulheres e crianças. Aliado a essas práticas, podem-se destacar a Pobreza, a baixa escolaridade feminina e o acesso restrito aos serviços de saúde sexual, que são factores que impulsionam a ocorrência de mais casos de gravidez na adolescência (UNFPA, 2013). Em contraste, diversas áreas da região sul mantiveram-se com risco abaixo da média por quase todo o período de estudo, possivelmente associadas a factores como maior urbanização, acesso a serviços de saúde ou estratégias locais de prevenção.

De forma específica, observou-se que o menor risco de gravidez na adolescência concentra-se nos cinco distritos urbanos do município de Maputo – Ka-Mpfumo, Ka-Mubukuane, Ka-Maxaquene, Ka-Chamanculo e Ka-Tembe por praticamente todo o período em estudo. Estas áreas são caracterizadas por maior densidade de serviços de saúde reprodutiva, melhor oferta e acesso a educação, infraestruturas públicas, transporte e comunicação. Nesses distritos, a proximidade a centros de saúde e acesso a campanhas de educação sexual nas escolas favorece o acesso atempado a métodos contraceptivos e a informação sobre saúde sexual e reprodutiva, reduzindo assim a ocorrência de gestações na adolescência. Além disso, o perfil socioeconómico relativamente mais elevado e a maior inserção das jovens nas escolas, contribuem para retardar tanto o início da vida reprodutiva quanto o casamento prematuro que por sua vez está intimamente ligado a gravidez na adolescência.

Em consonância ao que se observa em Moçambique, estudos realizados no Brasil, Etiópia e Nigéria destacam que as áreas marcadas por vulnerabilidade económica e social, aliadas a valores culturais profundamente enraizados, são as mais propensas a apresentar maiores taxas de

gravidez na adolescência. Um estudo feito no Brasil, por Fernandes *et al.* (2024), destacou que áreas com maior vulnerabilidade materna e piores condições socioeconômicas apresentam maior proporção de mães adolescentes, e que essa relação se manteve ao longo do tempo. Essa dinâmica também se reflecte no contexto moçambicano, especialmente na região norte do país, onde se observam características socioeconômicas semelhantes.

Um estudo realizado por Bolanriwa *et al.* (2022) sobre a gravidez na adolescência na Nigéria revelou que as maiores taxas ocorrem nas regiões do Norte, especialmente nas cidades de Sokoto, Kebbi, Zamfara, Katsina, Kano, Jigawa, Bauchi e Niger. Esses locais apresentam alta prevalência devido a factores como início precoce da vida sexual (entre 15 e 19 anos), casamentos prematuros e influências étnicas, como entre os Igbo, que demonstram maior risco. O nível educacional e a situação de emprego também influenciam pois adolescentes que trabalham têm menor probabilidade de engravidar.

Outro estudo desenvolvido por Tigabu *et al.* (2021), na Etiópia, identificou as regiões de Somali, Afar, Oromia e Hareri como zonas críticas de gravidez na adolescência. Essas áreas compartilham características como baixos níveis de escolaridade, pobreza, uso limitado de contraceptivos modernos e forte influência de práticas culturais e religiosas. São regiões predominantemente rurais com difícil acesso aos serviços de saúde e a educação.

Nos últimos anos, em Moçambique, diversas intervenções foram implementadas com o objectivo de reduzir a gravidez na adolescência, especialmente nas províncias das regiões norte e centro, onde os níveis de risco são mais elevados. Programas como o Geração Biz, actuando desde 2000 com cobertura nacional (MISAU, 2015), e o Rapariga Biz, focado em Nampula, Zambézia, Niassa e Cabo Delgado (UNFPA, 2013), representam esforços significativos nesse sentido. Além disso, a integração da educação sexual nos currículos escolares (MINEDH, 2020) e a criação de Serviços de Atendimento a Adolescentes e Jovens como os Serviços Amigos dos Adolescentes e Jovens (SAAJ) contribuíram para ampliar o acesso à informação e aos serviços de saúde sexual e reprodutiva.

Todavia, apesar das intervenções implementadas, o risco de gravidez na adolescência apresenta uma tendência de aumento ao longo do tempo. Essa situação pode indicar um alcance limitado ou insuficiente das ações de combate à gravidez na adolescência, pois mesmo com políticas públicas vigentes, ainda há barreiras à sua efectividade, sugerindo que as estratégias adoptadas até o momento não têm sido suficientemente eficazes para produzir mudanças significativas nos níveis de incidência, especialmente em contextos socioculturais de vulnerabilidade.

6 Conclusão e recomendações

O presente estudo teve como principal objectivo analisar a distribuição da incidência da gravidez na adolescência a nível dos distritos de Moçambique no período de 2011 a 2020, considerando sua distribuição no tempo e no espaço. Para tal, foi utilizado um modelo Bayesiano de regressão de Poisson com estrutura espacial CAR, buscando identificar padrões temporais e geográficos na incidência desse evento.

Em relação à aplicação deste modelo para analisar a distribuição do evento em estudo, pode-se dizer que, embora uma das principais características ou vantagens desta abordagem seja a possibilidade de inclusão de informação a priori, esta também pode ser usada para a suavização de taxas, sobretudo quando se lida com pequenas áreas de estudo em que algumas apresentam valores muito baixos ou até mesmo não tem registo de ocorrência do evento. Este método forneceu estimativas ajustadas para a componente de variabilidade presente nos dados que deveu-se tanto pelo factor espacial bem como pelo factor temporal e permitiu fazer uma interpretação mais consistente da incidência da gravidez na adolescência a nível dos distritos de Moçambique para o período de estudo.

Com base nos resultados obtidos, rejeita-se a hipótese de que a distribuição da gravidez na adolescência em Moçambique não varia significativamente no tempo e nem no espaço. As evidências trazidas neste estudo, apontaram para um padrão espacial consistente, com maior risco observado sobretudo nas regiões norte e centro e alguns distritos da região sul, ao longo de todo o período analisado. Além disso, foi identificada uma tendência de aumento ao longo de todo período de estudo, o que indica variação temporal relevante. Esses resultados também confirmam a existência de uma interação espaço-temporal significativa, apoiando a hipótese de que o risco de gravidez na adolescência não apenas varia entre distritos e ao longo do tempo, mas que essas variações estão interligadas.

Esta metodologia permitiu ainda apresentar estimativas mais suaves e de fácil interpretação diferentemente da estimação clássica, permitiu que os mapas tivessem uma interpretação mais clara da distribuição da gravidez na adolescência. A modelação bayesiana mostrou supremacia e ser mais poderoso para análise da distribuição de eventos em saúde.

Diante dos resultados obtidos, conclui-se que a gravidez na adolescência em Moçambique apresenta um padrão espaço-temporal desigual, com maior incidência sobretudo em regiões caracterizadas por vulnerabilidades socioeconômicas e práticas culturais que dificultam o acesso à educação e aos serviços de saúde sexual e reprodutiva. A presença de uma interação espaço-temporal significativa reforça a importância de se adotarem estratégias mais específicas, sensíveis ao contexto local e voltadas para os territórios mais afectados.

Deste modo, os resultados deste estudo, ampliam o conhecimento actual sobre a gravidez na adolescência em Moçambique, ao evidenciar padrões de persistência e expansão deste evento em regiões específicas, alinhando-se a estudos realizados em países como Etiópia e Nigéria, contribuindo não apenas para o entendimento mais aprofundado da distribuição da gravidez na adolescência em Moçambique, mas também fornecendo subsídios para o desenho de políticas públicas mais direccionadas e mais eficazes.

Diante do exposto, tecem-se as seguintes recomendações:

- Implementar programas de prevenção da gravidez na adolescência de maior alcance, com foco na região norte devido à sua maior incidência para este evento;
- Para estudos futuros, sugere-se considerar dados colectados durante o período de estudo ou dados de registo de rotina de unidades de saúde ou inquéritos anuais desenhados para a colecta desse tipo de informação;
- No modelo bayesiano incluir algumas co-variáveis que possibilitem explicar as razões da existência dos padrões de incidência verificados.

7 Limitações

- A principal limitação deste estudo reside na natureza dos dados utilizados, que foram coletados em um único momento (dados transversais), o que exigiu a distribuição retrospectiva dos dados pelos anos do período em estudo.
- Devido à instabilidade político-social e à não representatividade a nível distrital dos dados, alguns distritos não foram incluídos na recolha dos dados do IDS-2023. Estas áreas foram consideradas como áreas sem informação e, portanto, estão em branco nos mapas apresentados.

8 Referências

- Aransiola JO, Asa S, Obinjuwa P, Olarewaju O, Ojo OO, Fatusi AO. (2013) Teachers' perspectives on sexual and reproductive health interventions for in school adolescents in Nigeria. *Afr J Reprod Health*.
- Assuncao, R. M. (2001) *Estatística Espacial com Aplicações em Epidemiologia, Economia e Sociologia*. Associação Brasileira de Estatística;
- Banco Mundial (2016). Moçambique: Relatório Analisa Pobreza e Suas Causas e Avança Recomendações, <https://www.worldbank.org/pt/news/press-release/2016/12/21/mozambique-report-discusses-poverty-trends-and-recommends-way-forward>
- Baumont, C. (2004). Spatial effects in housing price models: do house prices capitalize urban development policies in the agglomeration Dijon. In: Regional Group Seminar of the Federal Reserve Bank of Chicago;
- Bernardinelli, L., Clayton, D., Pascutto, C., Montomoli, C., Ghislandi, M., e Songini, M. (1995). Bayesian analysis of space—time variation in disease risk. *Statistics in medicine*, 14(21-22), 2433-2443.
- Bernadinelli, L.; Montomli C. (1992). Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Statistics in Medicine*, v.11, p.983-1007, 1992;
- Bolarinwa, O. A., Tessema, Z. T., Frimpong, J. B., Babalola, T. O., Ahinkorah, B. O., e Seidu, A. A. (2022). Spatial distribution and factors associated with adolescent pregnancy in Nigeria: a multi-level analysis. *Archives of Public Health*
- Burns, P. (2002). Robustness of the Ljung-Box test and its rank equivalent. Available at SSRN 443560.
- Câmara *et al.* (2002), *Análise Espacial de Áreas*, Instituto Nacional de pesquisas espaciais-INPE, São José dos Campos-SP.
- Carvalho, C. C. de. (2013). *Gravidez na adolescência: Principais causas e taxas* (p. 27).
- Caumo, R. B. (2006). *Estatística Espacial em Dados de Área: Uma Modelagem Inteiramente Bayesiana para o Mapeamento de Doenças Aplicada à Dados Relacionados com a Natalidade em Mulheres Jovens de Porto Alegre- Monografia de bacharelato em Estatística (UFRGS)*. Porto Alegre;
- Chodota, V., Mbonile, M., e Mwamfupe, A. (2021). Factors That Influence Maternal High-risk Births in Njombe District, Tanzania. *Journal of the Geographical Association of Tanzania*, 40(1), 156–176. <https://doi.org/10.56279/jgat.v40i1.24>
- Da Silva, S. (2009). Mapeamento da mortalidade infantil no Rio Grande Do Sul: Uma Comparação Entre As Abordagens Empírica Bayesiana E Totalmente Bayesiana.

- Fernandes, C. M., Conceição, G. M. de S., da Silva, Z. P., Nampo, F. K., e Neto, F. C. (2024). Socioeconomic factors increase the risk of teenage pregnancy: spatial and temporal analysis in a Brazilian municipality. *Revista Brasileira de Epidemiologia*, 27, 1–9. <https://doi.org/10.1590/1980-549720240040>
- Fundo das Nações Unidas para a População [UNFPA]. (2013). Programa Rapariga Biz: Empoderamento de raparigas e redução de casamentos prematuros e gravidez na adolescência em Moçambique. Maputo: UNFPA Moçambique.
- Ganchimeg T, Ota E, Morisaki N, Laopaiboon M, Lumbiganon P, Zhang J, et al. (2014). Pregnancy and childbirth outcomes among adolescent mothers: a World Health Organization multi-country study. *BJOG*.
- Gamerman, D. (1997). Markov Chain Monte Carlo: stochastic simulation for Bayesian inference. Chapman and Hall, London, 245 p.;
- Gelman, A., Carlin, J. B., Stern, H. S., e Rubin, D. B. (1995). Bayesian data analysis. Chapman and Hall/CRC.
- Gelman, A., e Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472.
- Getis, A.; ORD, J.K.. (1992) "The analysis of spatial association by use of distance statistics." *Geographical Analysis*, 24:-206
- Greene, W. (2002). *Econometric Analysis*. 7 Edition, Prentice Hall, Upper Saddle River.
- Griffith, D. A.; Paelinck, J. H. P. (2011). Non-standard spatial statistics and spatial econometrics. *Advances in Geographic Information Science*, v. 1;
- Kelsall, J.E.; Wakefield, J.C. (1999). Bayesian models for spatially correlated disease and exposure data. London: Oxford University p. 151;
- KING, M.L. (1981). The Alternative Durbin-Watson Test: An Assessment of Durbin and Watson's Choice of Statistic. *Journal of Econometrics*, 17, 51-66;
- Knorr-Held, L. (2000). Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in medicine*, 19(17-18), 2555-2567.
- Lawson A. et al (2003). *Disease mapping with WinBUGS and MLwiN*. England: John Wiley e Sons;
- Lima, W. M. S. F. de, Gallas, M., Marrone, L. C. P., Schröder, N. T., e Silveira, E. F. da. (2024). Tendência Temporal E Fatores Associados À Gravidez Na Adolescência. *Revista de Estudos Interdisciplinares*, 6(1), 01–18. <https://doi.org/10.56579/rei.v6i1.865>
- Luzardo, A. J. R., Castañeda Filho, R. M., e Rubim, I. B. (2017). Análise espacial exploratória com o emprego do índice de Moran. *GEOgraphia*, 19(40), 161-79.

- Llop-Gironés, A., et al. (2019). Inequalities in the access to and quality of healthcare in Mozambique: Evidence from the household budget survey. *International Journal for Quality in Health Care*, 31(10), 813-820. <https://doi.org/10.1093/intqhc/mzy171>
- Martins, M. D. G., Santos, G. H. N. D., Sousa, M. D. S., Costa, J. E. F. B. D., e Simões, V. M. F. (2011). Associação de gravidez na adolescência e prematuridade. *Revista Brasileira de Ginecologia e Obstetrícia*, 33, 354-360.
- Martinez, W. L., Martinez, A. R., e Solka, J. (2017). *Exploratory data analysis with MATLAB*. Chapman and Hall/CRC.
- Ministério da Educação e Desenvolvimento Humano [MINEDH]. (2020). *Plano Curricular do Ensino Secundário Geral: Educação Sexual e Saúde Reprodutiva*. Maputo: MINEDH.
- Ministério da Saúde [MISAU]. (2024). *Moçambique inquérito demográfico e de saúde 2022-23*. Instituto Nacional de Estatística (INE). www.ine.gov.mz
- Ministério da Saúde [MISAU]. (2015). *Plano estratégico do programa Geração Biz 2015–2020*. Maputo: Ministério da Saúde.
- Morris, M., Wheeler-Martin, K., Simpson, D., Mooney, S. J., Gelman, A., e DiMaggio, C. (2019). Bayesian hierarchical spatial models: Implementing the Besag York Mollié model in stan. *Spatial and spatio-temporal epidemiology*, 31, 100301.
- Natário, I. (2004). *Hierarchical Bayesian Models for Epidemiological Analysis for Rare Events*. Dissertação de tese de Doutoramento não publicada, Faculdade de Ciências da Universidade Nova de Lisboa;
- Neto, F. R. (2007). Gravidez na adolescência: motivos e percepções de adolescentes. *Revista Brasileira de enfermagem*
- Nhampoca, J. M., e Maritz, J. (2022). Early marriage and adolescent pregnancy in Mozambique. *African Journal of Reproductive Health*, 26(3), 114-123.
- Nogueira, D.A. (2004). *Proposta e Avaliação de Critérios de Convergência para o Método de Monte Carlo via Cadeias de Markov: Caso uni e Multivariado- Dissertação de Mestrado em Agronomia (área de concentração em Estatística e Experimentação Agrária) - Minas Gerais*;
- O'Hagan A. (1994) *Kendall's Advanced Theory of Statistics: Bayesian Inference*, vol. 2B. Oxford University Press: Oxford. Springer;
- Olsen, S., et al. (1996). Cluster analysis and disease mapping- Why, when, how? A step by step guide, *British Medical Journal*, 313, 863-865;
- Organização Mundial da Saúde. (2009). *Child and adolescent health and development*. Genebra: Organização Mundial da Saúde. <https://www.who.int/child-adolescent-health/>
- Organização Mundial da Saúde. (2014). *Adolescência: gravidez precoce*. Genebra: <https://www.who.int/media>

- Pinheiro, L. C. (2009). Alguns aspectos de modelos espaço-temporais.
- Pacca, G. P. (2015). Aplicações de modelos markovianos ocultos em modelos dinâmicos não-lineares e não gaussianos.
- Roy, V. (2020). Convergence diagnostics for markov chain monte carlo. *Annual Review of Statistics and Its Application*, 7(1), 387-412.
- Salhan, S. (2007). High-Risk Pregnancies. *Textbook of Obstetrics*, 394–394. <https://doi.org/10.5005/jp/books/1>
- Silva, M. W. da, Franco, E. C. D., Gadelha, A. K. O. A., Costa, C. C., e Sousa, C. F. de. (2021). Adolescência e Saúde: significados atribuídos por adolescentes. *Research, Society and Development*, 10(2), e27510212482. <https://doi.org/10.33448/rsd-v10i2.12482>
- Studenmund, A.H. (2001). *Using Econometrics—A Practical Guide*. Addison-Wesley Longman, London;
- Tigabu, S., Liyew, A. M., e Geremew, B. M. (2021). Modeling spatial determinants of teenage pregnancy in Ethiopia: Geographically weighted regression. *BMC Women's Health*, 21, 316. <https://doi.org/10.1186/s12905-021-01400-7>
- UNFPA. (2013). *Gravidez na adolescência: Desafios e respostas de Moçambique – Suplemento do relatório sobre a situação da população mundial – Moçambique*. Rádio Moçambique.
- UNICEF. (2021). *A situação das crianças em Moçambique 2021*. UNICEF Moçambique.
- UNICEF e UNFPA. (2018). *Casamento prematuro e gravidez na adolescência em Moçambique: Causas e impacto*. Maputo: UNICEF Moçambique.
- União Africana. (2015). *Primeira sessão do comitê técnico especializado sobre o desenvolvimento social, trabalho e desemprego (STC-SDLE-1), Adis Abeba, Etiópia, 20-24 de abril de 2015: Tema: "Proteção social para o desenvolvimento inclusivo" Campanha para eliminação do casa*.
- Vicente, L.S. (2011). *Modelos Bayesianos Hierárquicos no Planeamento de Recursos Humanos*. Doutoramento em Estatística e Investigação Operacional- Universidade de Lisboa. Lisboa;
- Vieira, R. S. (2009). *Crescimento econômico no estado de São Paulo: uma análise espacial* (Publicação online). Editora UNESP; Cultura Acadêmica. <http://books.scielo.org>

9 Anexos

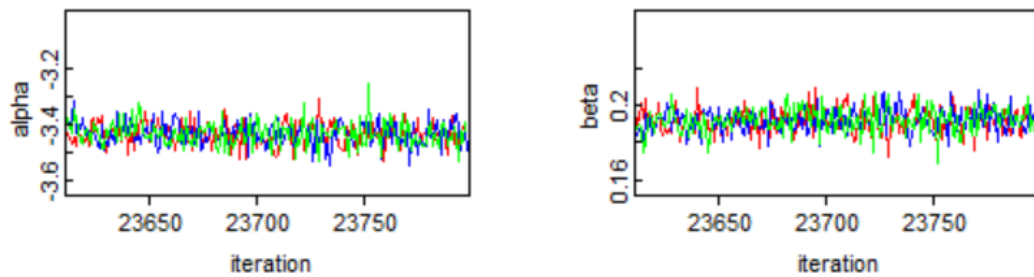


Figura 7: Gráfico de traços para alfa e beta

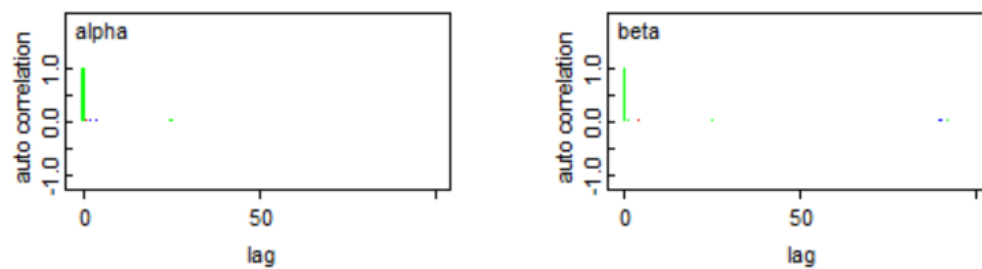


Figura 8: gráfico de correlação para alfa e beta

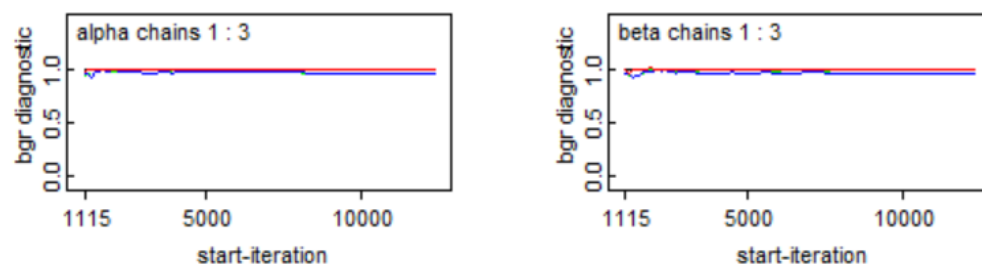


Figura 9: diagnostico de Gelman e Rubin para alfa e beta

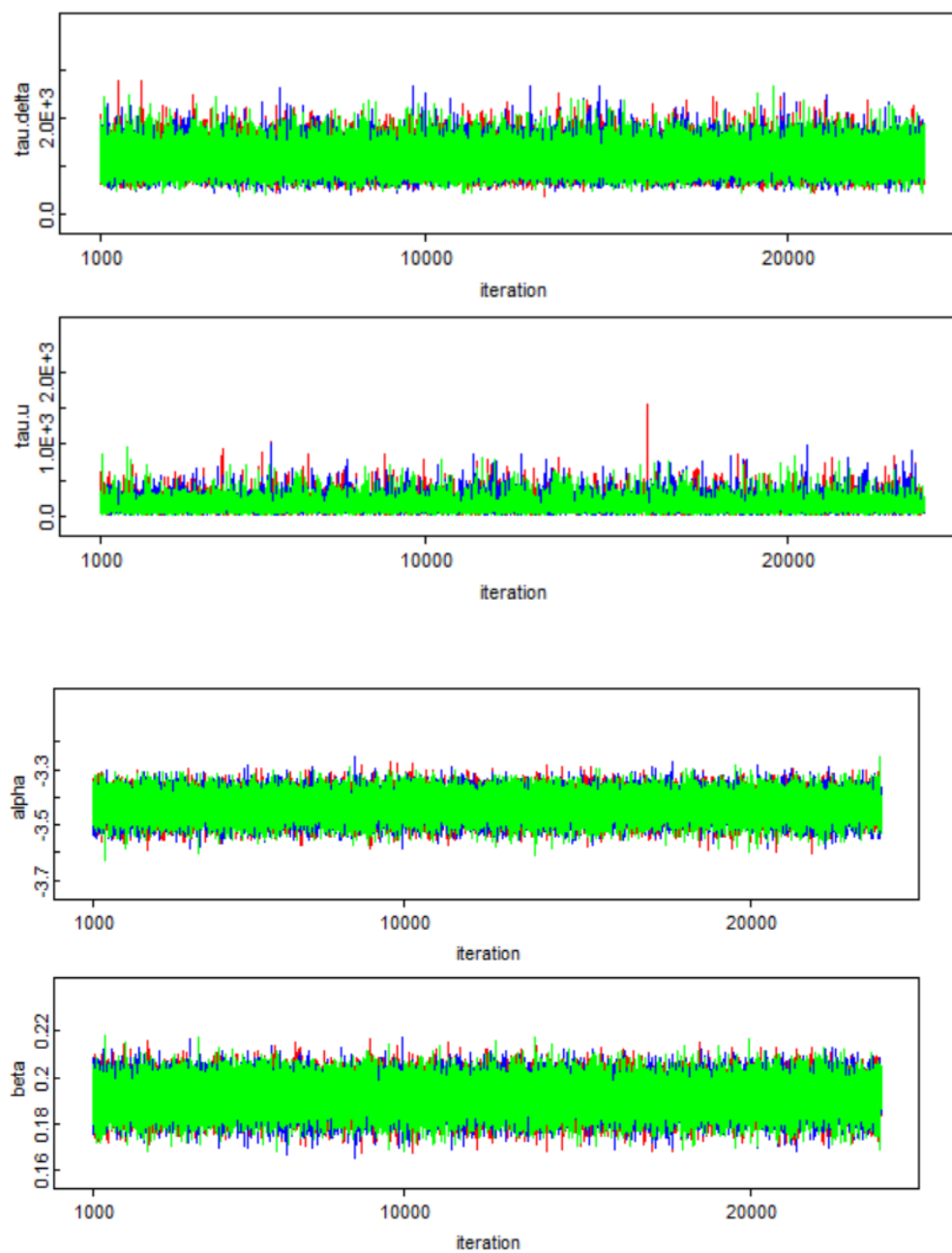


Figura 10: gráfico do histórico das MCMC

9.1 Códigos

9.1.1 Anexo A - Código do Modelo BUGS

```
model {  
  for (i in 1:N) {  
    for (j in 1:T) {  
      o[i,j] ~ dpois(mu[i,j])  
      log(mu[i,j]) <- log(e[i,j]) + alpha + u[i] + (beta + delta[i]) * t[j]  
      RR[i,j] <- exp(alpha + u[i] + (beta + delta[i]) * t[j])  
    }  
  }  
  
  # Modelagem espacial CAR para u e delta  
  u[1:N] ~ car.normal(adj[], weights[], num[], tau.u)  
  delta[1:N] ~ car.normal(adj[], weights[], num[], tau.delta)  
  
  for(k in 1:sumNumNeigh) {  
    weights[k] <- 1  
  }  
  
  # Priors ajustadas  
  alpha ~ dnorm(0, 0.001)  
  beta ~ dnorm(0, 0.001)  
  
  # Hiperpriors ajustados  
  tau.u ~ dgamma(1, 0.01)  
  tau.delta ~ dgamma(1, 0.01)  
}
```

.0.2 Anexo B - Código R para execução do modelo no OpenBUGS

```
modelo_resultado <- bugs(  
  data = dados,  
  inits = list(inits1(), inits2(), inits3()),  
  parameters.to.save = parametros,  
  model.file = "C:/Users/Administrator/Downloads/OpenBUGS323/OpenBUGS323/modelo22.txt",  
  n.chains = 3,  
  n.iter = 3000,  
  n.burnin = 1000,  
  n.thin = 10,  
  debug = TRUE,  
  DIC = TRUE,  
  OpenBUGS.pgm = "C:/Users/Administrator/Downloads/OpenBUGS323/OpenBUGS323/OpenBUGS.exe"  
)
```