

## **PRECIFICAÇÃO ATUARIAL PARA SEGURO PARAMÉTRICO CONTRA VERANICOS EXTREMOS EM FASES CRÍTICAS DO CAFÉ**

**Resumo:** O objetivo desta pesquisa foi a construção de um modelo atuarial para o produto de seguro paramétrico agrícola contra veranicos extremos (VE) em municípios produtores de café no estado de Minas Gerais, buscando estabelecer um índice de indenização que reduzisse o risco base neste tipo de contrato em diferentes fases da planta. Para atingir o objetivo proposto, foi utilizada a distribuição generalizada de valores extremos para modelar os VE e a regressão logística para fornecer as probabilidades de perda nas safras considerando as probabilidades dos VE. Como resultado, verificou-se que as probabilidades de excedência dos VE na primeira fase reprodutiva da planta (R1) e também no período que acumula a primeira e segunda fases reprodutivas (R1R2) explicam as perdas observadas na produção final. Além disso, o índice encontrado que reduz o risco base do produto ocorre quando o modelo calcula 44% ou 45% de probabilidade de perda em R1 e 51% ou 52% em R1R2. Desse modo, nesta pesquisa foi sugerida uma modelagem para produtos paramétricos que intentam reduzir os impactos econômicos dos VE em fases críticas do café, sendo que a metodologia desenvolvida pode ser utilizada como base para modelagem de outros produtos relacionados à agricultura.

**Palavras-chave:** precificação atuarial; seguro paramétrico; risco climático; teoria de valores extremos.

## **ACTUARIAL PRICING FOR PARAMETRIC INSURANCE AGAINST EXTREME DRY SPELLS IN CRITICAL PHASES OF COFFEE**

**Abstract:** This research aimed to build an actuarial model to the agricultural parametric insurance product against extreme dry spells (EDS) in producing cities of coffee in Minas Gerais state, seeking to establish an indemnity index that would reduce the base risk in this kind of contract in different phases of the plant. To reach the proposed objective, the generalized distribution of extreme values was used to model the EDS and the logistic regression was used to provide the probabilities of loss in the harvests considering the EDS probabilities. As a result, it was found that the probability of EDS in the first reproductive phase of the plant (R1) and also in the period that accumulates the first and second reproductive phases (R1R2) explain the losses observed in the final production. Beside, the found index that reduces the base risk of this product occurs when the model calculates 44%

or 45% probability of loss in R1 and 51% or 52% in R1R2. So, this research suggested a modeling for parametric products that intend to reduce the economic impacts of EDS in critical phases of coffee and the developed methodology can be used as a basis for modeling other products related to agriculture.

**Key-words:** actuarial pricing; climate risk; extreme value theory, parametric insurance.

## 1. INTRODUÇÃO

A produtividade agrícola é altamente dependente das condições climáticas (HEIMFARTH; MUSSHOF, 2011). Dessa forma, tendo em vista que estas condições fogem ao controle dos produtores, existe a necessidade da adoção de estratégias que minimizem o impacto de perdas na produção em decorrência de condições climáticas desfavoráveis. No contexto brasileiro a adoção de estratégias de mitigação de riscos agrícolas é bastante relevante, uma vez que cerca de  $\frac{1}{4}$  de toda a riqueza produzida no país é relacionada ao agronegócio (CEPEA, 2022).

Uma das estratégias para a mitigação do risco climático sobre a produção agrícola é a contratação de seguros agrícolas que são ferramentas por meio da qual os produtores são indenizados quando da ocorrência de eventos adversos que prejudiquem a produção e causem, conseqüentemente, perdas financeiras. Desse modo, o seguro agrícola é fundamental para o desenvolvimento do agronegócio, uma vez que possibilita a modernização e a competitividade do setor, tendo em vista que instituições de crédito terão maior segurança em fornecê-lo aos produtores (FORNAZIER, SOUZA, PONCIANO, 2012).

Nessa conjuntura, tem-se a recente emergência de um produto de seguro, chamado “seguro paramétrico” que surge com grande potencial de aplicação para produtos agrícolas (CARTER; CHIU, 2018). Nesta modalidade, diferentemente do que é aplicado em produtos tradicionais nos quais é necessária a constatação de um prejuízo físico para o pagamento da indenização, os pagamentos ocorrem no instante em que surge uma expectativa do prejuízo, sendo esta expectativa verificada quando um índice preestabelecido no contrato é alcançado ou superado (MILLER, 2020).

Dessa forma, tem-se que o seguro paramétrico supera em alguns aspectos a modalidade tradicional quando consideramos as vantagens que esse produto traz como, por exemplo, a maior agilidade no processo de pagamento das indenizações, a solução de problemas de risco moral, a possibilidade de redução da seleção adversa nas carteiras e também tende a ser mais barato do que produtos tradicionais, viabilizando a sua contratação

para pequenos produtores (HEIMFARTH; MUSSHOFF, 2011 & RADU; ALEXANDRU, 2022).

Entretanto, a despeito dos benefícios citados, de acordo com o World Bank (2017), o produto paramétrico ainda tem baixa demanda e enfrenta dificuldades na sua aceitação por parte das seguradoras, que carecem de conhecimento e capacidade técnica para o desenvolvimento de produtos sustentáveis e rentáveis, e por parte dos segurados, que não entendem o mecanismo de funcionamento do produto e, segundo Johnson (2021), em alguns casos o tratam com desconfiança.

Desse modo, a desconfiança e a dificuldade de aceitação de seguradoras e segurados em relação a produtos paramétricos perpassam pela existência do risco base nos contratos. O risco base corresponde à descompatibilização entre o índice indenizatório e as condições vivenciadas no terreno pelo segurado. Segundo Afshar (2021), o risco base faz com que os índices indenizatórios se correlacionem equivocadamente com a ocorrência de perdas na produção e isso pode ocasionar o pagamento de indenizações indevidas ou a omissão de indenizações quando um sinistro ocorrer, sendo ambas as ocorrências problemáticas.

Assim, supõe-se que o desenvolvimento de mecanismos de controle e de redução de risco base em seguros paramétricos trariam como benefício a maior aceitação do produto no mercado brasileiro. Dessa forma, o objetivo central desta pesquisa foi a modelagem de um produto de seguro agrícola paramétrico contra veranicos extremos em cidades produtoras de café de Minas Gerais, buscando o estabelecimento de um índice indenizatório que aplicado em contratos dessa natureza incorra no menor risco base.

Para alcançar o objetivo proposto foi necessário também atender objetivos secundários sendo eles a modelagem dos veranicos extremos, que foi realizada por meio da distribuição generalizada de valores extremos, e a identificação das probabilidades de ocorrência de perdas de produção, observadas as probabilidades de ocorrência dos veranicos extremos, que foi realizada via regressão logística. A partir do modelo logístico ajustado, tornou-se possível confrontar para diferentes limiares de classificação as perdas estimadas com as perdas observadas e, a partir disto, foi possível definir qual limiar de seria o índice indenizatório mais eficiente em compatibilizar as perdas com os pagamentos das indenizações.

A opção pelo estudo de caso voltado à cultura de café se deu pela grande importância econômica do produto para o Brasil, que é o maior produtor e exportador do grão. Minas Gerais foi eleito como foco do estudo pelo fato de, segundo CONAB (2022), este ser o estado com a maior produção nacional e, portanto, o produto modelado nesta pesquisa pode fornecer resultados interessantes a respeito do comportamento do café em períodos de veranicos

extremos para o Estado. Ademais, tendo em vista que a planta do café tem necessidades hídricas diferentes em cada fase fenológica, a pesquisa avaliou o impacto dos veranicos extremos de forma isolada em cada fase.

Por fim, ressalta-se que a metodologia desta pesquisa foi embasada no trabalho de Figueiredo et al. (2018), que buscou modelar contratos de seguro paramétrico contra inundações na Jamaica. Desde modo, a contribuição científica desta pesquisa é a aplicação do método para seguros paramétricos agrícolas, bem como a utilização de dados brasileiros e a inclusão da teoria de valores extremos (TVE) no processo de modelagem.

## **2. REFERENCIAL TEÓRICO**

### **2.1. SEGURO AGRÍCOLA PARAMÉTRICO**

De acordo com o World Bank (2017) um produto de seguro paramétrico pode ser definido como um contrato no qual há o “pagamento de benefícios com base em um índice pré-determinado para a perda de ativos e investimentos que ocorram como resultado do clima ou outros eventos catastróficos”.

No contrato de seguro paramétrico considera-se que o índice para o pagamento das indenizações foi acionado quando algum limite preestabelecido em contrato é alcançado ou superado (ZURICH, 2021). Esse índice que aciona os pagamentos é chamado de gatilho da indenização. No caso de seguros agrícolas, por exemplo, o gatilho pode ser: déficit ou excesso de chuvas, secas, temperaturas extremas, entre outros eventos que possam prejudicar o desenvolvimento das culturas (SHIRSATH et al., 2019).

Para Afshar et al. (2021), o seguro paramétrico pode ser bastante útil para agricultores como forma de adaptação e gestão de riscos à produção advindos de eventos climáticos extremos. A menção a eventos extremos como um foco da utilização destes produtos, segundo os autores, advém do fato de que esses eventos, a despeito de serem menos frequentes, causam prejuízos consideráveis aos agricultores.

De acordo com Turner (2020) a estruturação de um seguro paramétrico ocorre com a determinação de três fatores: 1) O índice ou métrica que se relaciona diretamente com os custos e despesas do segurado; 2) O gatilho relacionado a este índice que cause acionamento da indenização da apólice; e 3) O valor da indenização. Destarte, de acordo com Jonhson (2021, p. 122), “Os contratos paramétricos especificam os dados e os algoritmos a serem usados, as áreas geográficas e os intervalos de tempo para avaliação e os resultados que irão ou não acionar um pagamento, ostensivamente tornando a tomada de decisão mais eficiente e auditável”.

Desse modo, ainda segundo Jonhson (2021) e Heimfarth e Musshoff (2011), a mudança na forma de estruturação de um contrato paramétrico em relação aos seguros tradicionais, bem como a mudança nos cálculos que determinam os pagamentos das indenizações nesta modalidade, resolveria os problemas de informação assimétrica nos contratos, já que os eventos segurados são passíveis de controle pelos segurados.

Além disso, de acordo com Miller (2020), Turner (2020) e Radu e Alexandru (2022) os contratos de seguro paramétrico também trazem vantagens econômicas, uma vez que a seguradora dispensa mão de obra especializada para quantificar perdas e o segurado paga somente pela cobertura que deseja. Também há vantagens de tempo e de flexibilidade, uma vez que as indenizações são pagas automaticamente quando o gatilho é acionado e o produto é totalmente customizável às necessidades do cliente.

Desse modo, coloca-se que a modalidade de seguro paramétrico climático pode ser de grande valia na gestão de riscos ligados à agricultura, tendo em vista que caso ocorra um fator climático adverso que acione o gatilho durante algum período crítico da safra, a indenização é paga instantaneamente sem que seja necessário esperar a colheita para realizar a constatação do prejuízo.

Destarte, tem-se que o seguro paramétrico no Brasil é um produto inovador que está iniciando seu processo de introdução no mercado. Apenas em 2021 foi realizado o primeiro contrato no país, que foi para a proteção de uma lavoura de cacau contra a seca no período da florada, no estado da Bahia (WALENDORFF, 2021). Além disso, foi também somente em 2021 que esta modalidade entrou no catálogo do Governo Federal como produto que participa dos programas públicos de incentivo à contratação de seguro agrícola (NEWE SEGUROS, 2020).

Nesse contexto, apesar das qualidades e potencialidades do seguro paramétrico, um dos entraves à sua aceitação por parte dos segurados e das seguradoras é a gestão de riscos dos contratos, uma vez que este tipo de seguro é suscetível ao chamado “risco base” (AFSHAR, 2021, SUN et al., 2020). O risco base é descrito como incompatibilidade entre o gatilho escolhido e a real ocorrência de perdas na produção, que pode resultar em pagamentos de indenização quando não há a perda real (risco base positiva) ou então na omissão dos pagamentos de indenização quando a perda existe (risco base negativo) (FIGUEIREDO et al., 2018 & TURNER, 2020).

Seguindo, tem-se que a definição equívoca do real índice causador de perdas é uma constante nos novos contratos paramétricos, sendo uma grande preocupação para as seguradoras. Nesse contexto, tem-se que quando o risco base se materializa, por exemplo,

muitas vezes a seguradora acaba tendo de agir fora dos termos dos contratos para atenuar as falhas nos seus dispositivos de definição de gatilhos (JONHSON, 2021). Dessa forma, a correta compatibilização entre os eventos de perda e de indenização não são tarefa fácil, demonstrando que “Embora seja verdade que eles [produtos paramétricos] são a simplicidade personificada uma vez executados e após o recebimento do pagamento, isso não é verdade para sua criação” (Turner, 2020, p. 2).

Segundo Sun et al. (2020) e Jonhson (2021), o risco base é um problema bastante preocupante quando o gatilho do seguro paramétrico é o clima, como no caso de produtos paramétricos agrícolas. Isso acontece porque, para este caso, existem três riscos que devem ser controlados visando a minimização do risco base. São eles o risco espacial, o risco de projeto e o risco temporal. De acordo com os autores, o risco espacial acontece quando existe grande distância entre a área da plantação e a estação meteorológica, fazendo com que as condições meteorológicas da área segurada não sejam condizentes com os dados obtidos pela estação.

O risco de projeto ocorre quando é feita uma relação equivocada entre o clima e o rendimento das culturas ou quando o gatilho climático escolhido para se relacionar com a produtividade não é adequado. Já o risco temporal surge quando se considera períodos de risco equivocados para a cobertura do contrato, uma vez que o índice dos contratos é, em geral, definido para períodos específicos do calendário (SUN et al. 2020 & JONHSON, 2021).

Neste enquadramento, Carter e Chiu (2018) colocam que a minimização do risco base é possível em contratos bem projetados nos quais as estimativas paramétricas se correlacionam de fato com as perdas dos segurados. Dessa forma, alguns autores realizaram pesquisas buscando projetar contratos de seguros paramétricos que conseguissem relacionar de forma eficiente a ocorrência de perdas, ou variações na produção, com o pagamento das indenizações, visando o estabelecimento de contratos de maior segurança para os envolvidos.

Afshar et al. (2021) buscaram a redução do risco base em produtos paramétricos para segurar lavouras de arroz na Índia contra os riscos da variabilidade das chuvas de monções, sendo a chuva (mm) o gatilho do contrato estudado. Os autores utilizaram na pesquisa modelos de simulação de cultivos e monitoramento da fenologia dos cultivos via satélites. Como resultado, foi obtido um modelo capaz de explicar 54% da variabilidade da produção de arroz considerando a variação de chuvas nas diferentes fases do grão. Segundo os autores propõem, o conhecimento desta relação entre a ocorrência de chuvas em diferentes estágios

da planta e a produção resultante podem auxiliar no projeto e implementação do seguro de índice com a redução de risco base.

Prokopchuk et al. (2020) desenvolveram um produto de seguro paramétrico para lavouras de milho na Ucrânia. Para isso, os autores trabalharam com um gatilho de indenização que combinou fatores de precipitação, temperatura e umidade relativa do ar nas fases de floração e enchimento dos grãos, que são as fases mais sensíveis a estas questões climáticas, para verificar o impacto desses fatores na produção resultante. Como resultado, os autores chegaram em um gatilho combinado capaz de calcular a variação da produção a cada combinação de fatores meteorológicos que houvesse nas fases fenológicas críticas, auxiliando a definição de gatilhos de contratos mais eficientes.

Dalhaus, Musshoff e Finger (2018) os autores realizaram a estruturação de 3 modelos paramétricos para proteger lavouras de trigo alemãs durante a fase de floração. Como resultado, os autores afirmaram que os modelos ajustados foram capazes de reduzir a exposição ao risco dos segurados em relação a não ter seguro, bem como 2 dos modelos também foram capazes de aumentar a utilidade esperada do produto de seguro para os envolvidos. Ademais, os pesquisadores enfatizaram que a consideração da fase fenológica das culturas na estruturação do produto paramétrico reduz o risco base temporal.

Desse modo, como realizado para as culturas de arroz, milho e trigo nos artigos citados, tem-se que o café é uma planta que tem características que também podem ser enquadradas na estruturação de um seguro paramétrico.

## **2.2. CULTURA DO CAFÉ**

No que tange a produção cafeeira mundial, o Brasil tem papel de destaque como o país que mais produz e exporta esse grão no mundo (CONAB, 2021), fazendo com que a cultura cafeeira se configure como um importante alicerce econômico para o país, bem como fator de geração de renda para milhões de brasileiros em centenas de municípios, que atuam de forma direta ou indireta nesta atividade (CEPEA, 2021).

Deste modo, a despeito de sua relevância para a sociedade brasileira, a atividade cafeeira é amplamente arriscada, uma vez que está sujeita às variações climáticas que não são controláveis pelos produtores. Segundo Ozaki (2007), para o desenvolvimento pleno de uma cultura são necessárias combinações de fatores pluviométricos e de temperatura que, caso sejam adversos, podem levar a perdas na produção e, conseqüentemente, a grandes prejuízos econômicos que podem extrapolar as cercas das fazendas e se multiplicar por toda a economia.

A combinação adequada dos fatores meteorológicos para que a safra alcance os níveis de produção desejados, varia a cada fase fenológica do cafeeiro (CAMARGO; CAMARGO, 2001 & CAMARGO, 2010). Desse modo, torna-se necessário o conhecimento sobre a fenologia do cafeeiro para que se possa quantificar o impacto das condições climáticas observadas em cada fase fenológica. De acordo com Camargo e Camargo (2001), o ciclo fenológico completo do café arábica tem duração de dois anos e se divide em seis ciclos fenológicos, apresentados no Quadro 1.

**Quadro 1** - Esquemática das seis fases fenológicas do cafeeiro arábica, durante 24 meses, nas condições climáticas tropicais do Brasil.

1º Ano Fenológico (A1)													
1ª Fase (V1)						2ª Fase (V2)							
Vegetação e Formação das Gemas Foliare						Indução e Maturação das Gemas Foliare							
Set	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago		
2º Ano Fenológico (A2)													
3ª Fase (R1)				4ª Fase (R2)				5ª Fase (R3)				6ª Fase (R4)	
Florada, Chumbinho e Expansão dos Frutos				Granação dos Frutos				Maturação dos Frutos				Repouso e Senescência dos Ramos	
Set	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago		

Fonte: Adaptado de Camargo e Camargo (2001)

A esquematização do ciclo fenológico, segundo os autores, permite que, dentre outras questões, seja possível identificar quais fases fenológicas são mais dependentes de recursos hídricos e em quais o estresse hídrico causa o favorecimento da produção. Segundo Arruda et al. (2000, p.782) “a identificação de fatores limitantes à produção e de épocas críticas de ocorrência de deficiência hídrica durante o ciclo fenológico do cafeeiro são metas importantes no planejamento e manejo da cultura”.

Nesse contexto, apesar de a disponibilidade hídrica durante o desenvolvimento da safra não ser o único fator de influência na produção, esta se configura como principal fator que afeta a produtividade do cafeeiro (PICINI et al., 1999 & MARTINS et al., 2015 &



APARECIDO; ROLIM; SILVA, 2015). Portanto, a pesquisa aqui desenvolvida foi centrada na análise deste fator meteorológico.

Como apresentado no Quadro 1, no primeiro ano fenológico do cafeeiro (A1) ocorrem duas fases vegetativas. A primeira fase vegetativa (V1) acontece entre os meses de setembro e março. Nesta fase ocorre a vegetação e a formação das gemas foliares da planta. A seca neste período prejudica a formação das gemas e a produção do ano seguinte sendo, portanto, um período que exige maior disponibilidade hídrica. A segunda fase vegetativa (V2) ocorre entre os meses de abril e agosto. Nesta fase, ocorre a indução e a maturação das gemas foliares. Aqui a planta não apresenta grande necessidade hídrica, sendo que a ocorrência do estresse hídrico nesta fase pode inclusive trazer benefícios para a produção resultante (EMBRAPA, 2009 & RUIZ-CÁRDENAS et al., 2015).

O segundo ano fenológico possui quatro fases reprodutivas. Na primeira fase reprodutiva (R1), que acontece entre os meses de setembro e outubro, ocorre a florada do café, formação dos chumbinhos e expansão/crescimento dos frutos. Esta fase depende expressivamente de um aumento do potencial hídrico da planta, uma vez que estresse hídrico aqui prejudica o crescimento dos frutos. A segunda fase reprodutiva (R2) que ocorre entre janeiro e março, é a fase de granação dos frutos. Esta fase também tem alta exigência hídrica, sendo que nela estiagens severas prejudicam a qualidade dos grãos, incorrendo em chochamento. (EMBRAPA, 2009 & CAMARGO; CAMARGO, 2001 & RUIZ-CÁRDENAS et al., 2015).

A terceira fase reprodutiva do cafeeiro (R3) ocorre nos meses de abril a junho, fase em que ocorre a maturação dos frutos e, conseqüentemente, a finalização do período de produção. Nesta fase fenológica, a deficiência hídrica é benéfica e favorece a qualidade do produto final. Por fim, a quarta e última fase reprodutiva (R4) ocorre nos meses de julho e agosto e refere-se ao período de repouso e senescência dos ramos, conhecido como período de autopoda, no qual os ramos produtivo não-primários secam e morrem. Excesso de umidade nesta fase também é prejudicial ao cafeeiro (CAMARGO; CAMARGO, 2001 & RUIZ-CÁRDENAS et al., 2015).

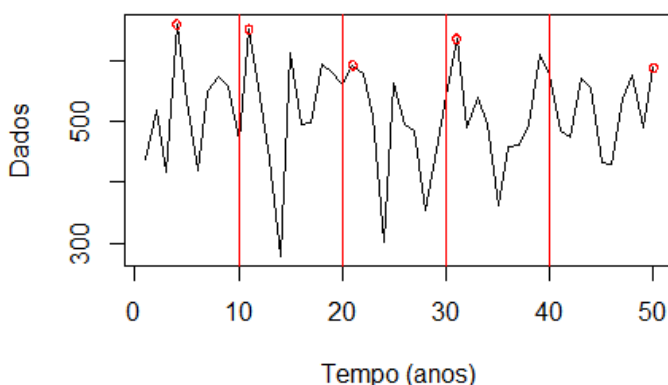
Desse modo, verifica-se que a necessidade hídrica para o bom desenvolvimento de uma safra ocorre nos ciclos V1, R1 e R2 sendo, portanto, coerente que produtos de seguro paramétrico que visem proteção das lavouras contra períodos de veranicos extremos, ou seja, períodos de estiagens severas, sejam modelados para cobrir os períodos de tempo correspondentes a estas fases fenológicas.

### 2.3. ANÁLISE DE VERANICOS EXTREMOS VIA TVE

Segundo Ceballos, Martinez-Fernández e Luengo-Ugidos (2004), considera-se como veranico um período de  $n$  dias consecutivos sem a observação de uma quantidade de precipitação relevante. Desse modo, considerou-se na pesquisa aqui desenvolvida que “veranicos extremos” correspondem às maiores sequências de dias sem chuvas relevantes dentro de cada ciclo fenológico do café. Nesta pesquisa, as séries de veranicos extremos foram obtidas via blocos de máximos.

Segundo Coles (2001), para a obtenção de uma série de blocos de máximos, primeiramente, os dados são divididos em blocos, que em geral são períodos de tempo, e, posteriormente, são extraídos os valores máximos de cada bloco, sendo estes utilizados para compor a série de extremos. Ressalta-se que a criação dos blocos de máximos, segundo este autor, pressupõe que as variáveis sejam independentes e identicamente distribuídas.

Dessa maneira, de acordo com o interesse do pesquisador e com a disponibilidade dos dados, a série de máximos via blocos pode ser obtida apenas pela seleção do maior valor de cada variável aleatória (ou bloco), ou também pode ser obtida, conforme Da Silva e Do Nascimento (2019), pela seleção dos  $k$  maiores valores de cada bloco, sendo  $k$  a ordem da série. A Figura 1 ilustra a forma de extração de dados via blocos de máximos, considerando uma série de extremos construída apenas com o maior valor de cada bloco ( $k=1$ ), sendo os valores da série criada representados pelos círculos vermelhos.



**Figura 1** - Extração de dados para a obtenção de uma série de máximos, via blocos de máximos, considerando apenas o maior valor de cada bloco.

Assim, dadas  $n$  variáveis aleatórias independentes e identicamente distribuídas ( $Z_1, Z_2, \dots, Z_n$ ), ou  $n$  blocos, tem-se que a série de máximos  $M_n^k$ , considerando as  $k$  maiores observações registradas em cada bloco é definida por

$$M_n^k = \text{máximo}\{Z_1^k, Z_2^k, \dots, Z_n^k\}$$

em que  $Z_n^k$  representa os  $k$  valores máximos extraídos da variável  $n$ .

Conforme Rocco (2014), eventos extremos (máximos ou mínimos), a despeito de apresentarem baixa probabilidade de ocorrência (baixa frequência), incorrem em grandes impactos (alta severidade). Desse modo, como abordado por De Oliveira e Carvalho (2020), a estimação das probabilidades de ocorrência de eventos extremos possibilita a quantificação dos riscos e a elaboração de estratégias de gestão de riscos que sejam mais eficientes nas atividades suscetíveis aos seus impactos.

Assim, para modelar valores extremos utiliza-se os preceitos da teoria de valores extremos. Dentro da TVE, uma das distribuições mais utilizadas neste tipo de modelagem é a distribuição Generalizada de Valores Extremos (GEV), introduzida por Jenkinson em 1955. De acordo com Butturi-Gomes, Beijo e Avelar (2018) esta distribuição possui bom desempenho ao ser empregada na modelagem de veranicos extremos.

A introdução da GEV simplifica a análise de valores extremos, uma vez que pressupõe que qualquer série de observações obtida via blocos de máximos converge para a distribuição GEV. Esta distribuição é especificada por 3 parâmetros:  $\mu$  que é o parâmetro de posição,  $\sigma$  que é o parâmetro de escala e  $\varepsilon$  que é o parâmetro de forma da distribuição. Na equação (1) está descrita a função da distribuição GEV, conforme Jenkinson (1955):

$$G(x) = \exp\left\{-\left[1 + \varepsilon\left(\frac{x-\mu}{\sigma}\right)\right]\right\}^{-\frac{1}{\varepsilon}} \quad (1)$$

definida em  $1 + \varepsilon\left(\frac{x-\mu}{\sigma}\right) > 0$ , para  $\varepsilon \neq 0$  e  $x \in \mathbb{R}$ , para  $\varepsilon = 0$ .

Nesta pesquisa, os parâmetros  $\mu$ ,  $\sigma$  e  $\varepsilon$  da distribuição GEV foram ajustados via inferência bayesiana, uma vez que de acordo com Coles (2001) a inferência bayesiana apresenta vantagens quando trabalha-se com dados de extremos. Segundo o autor, como em geral dados extremos são escassos e costumam ser oriundos de bases de dados pequenas, a inferência bayesiana auxilia na redução das incertezas no processo inferencial por meio da incorporação de informações *a priori* sobre os parâmetros.

A partir de  $G(x)$  torna-se possível a obtenção da probabilidade de excedência de eventos máximos superiores a determinado quantil  $z$  da variável extrema estudada, sendo que para valores de  $\varepsilon \neq 0$  esta probabilidade é obtida por meio da equação (2).

$$P(X > x) = 1 - G(x) = 1 - \exp\left\{-\left[1 + \varepsilon\left(\frac{x-\mu}{\sigma}\right)\right]\right\}^{-\frac{1}{\varepsilon}} \quad (2)$$

em que definida em  $1 + \varepsilon\left(\frac{x-\mu}{\sigma}\right) > 0$  para  $\varepsilon \neq 0$ .

Já quando  $\varepsilon = 0$ , a probabilidade de excedência de determinado eventos extremo é obtida por meio da equação (3)

$$P(X > x) = 1 - \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\} \quad (3)$$

em que  $-\infty < x < \infty$ .

Ademais, ressalta-se que antes de empregar a distribuição GEV na modelagem de uma série de extremos é necessário verificar algumas características da base de dados. A primeira característica a ser investigada é a independência dos valores na série de máximos, uma vez que a formulação da convergência da série para a distribuição GEV depende deste pressuposto (JENKINSON, 1955). A segunda característica a ser testada é a verificação da presença de tendência na série, uma vez que caso sejam detectadas evidências de tendência é necessário incorporá-la na modelagem.

O teste de Ljung-Box, desenvolvido por Ljung e Box em 1978, é comumente empregado para avaliar a independência de uma série de dados. Trata-se de um teste de hipóteses no qual a hipótese nula ( $H_0$ ) afirma que os dados da série são independentes e a hipótese alternativa ( $H_1$ ) afirma que a série de dados não é independente.

Já para a avaliação de tendência em uma série de dados, um dos testes possível é o teste de Mann-Kendall. Este teste foi desenvolvido por Mann (1945) como um teste não paramétrico empregado na detecção de tendência em séries de dados, sendo complementado por Kendall (1975), que forneceu a distribuição estatística do teste, possibilitando a verificação de existência de tendências não lineares e pontos de virada. De acordo com o teste, a hipótese nula ( $H_0$ ) atesta que a série de dados é estacionária, enquanto que a hipótese alternativa ( $H_1$ ) atesta que a série é não estacionária.

Destarte, o emprego da GEV para modelar os veranicos extremos em cada fase fenológica crítica do café forneceu a probabilidade de excedência de cada quantidade  $x$  de dias de veranicos extremos, sendo que esta informação foi empregada como variável

explicativa no modelo de regressão logística que avaliou a probabilidade de ocorrência de perdas na produção.

#### 2.4. REGRESSÃO LOGÍSTICA

De acordo com Barreto (2011) modelos de regressão consistem em técnicas empregadas tanto para a realização de previsões quanto para a classificação de dados, sendo um método que avalia a variação observada em uma variável (variável dependente) quando outras variáveis (variáveis independentes) se alteram.

Dentre os diversos modelos de regressão existentes, a regressão logística é empregada quando a variável dependente ( $Y$ ) é binária, ou seja, a variável assume apenas dois valores, sendo “0” para eventos de “fracasso” e “1” para eventos de “sucesso”. Assim, como a variável resposta está limitada em 0 ou 1, o modelo logístico ajustado prevê apenas valores que estejam neste intervalo, fazendo com que a regressão logística seja bastante útil na previsão de probabilidades de eventos (FERNANDES et al.,2020).

Na regressão logística a variável dependente é descrita como  $Y \sim Bernoulli(p)$ , em que “ $p$ ” representa a probabilidade desconhecida de sucesso do evento de interesse. Assim, o modelo logístico é obtido por meio da aplicação de uma função de ligação logística construída com o uso de “ $p$ ”, que modela os dados como uma função linear das variáveis independentes ( $x_i$ ). O modelo logístico está representado na equação (4).

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_i x_i \quad (4)$$

em que  $i = 0, \dots, n$ , sendo  $n$  a quantidade de variáveis dependentes do modelo e  $\beta_0$  e  $\beta_i$  os coeficientes do modelo ajustado.

A obtenção das probabilidades de  $Y$  a cada conjunto de variáveis  $x_i$  consideradas é realizada por meio da equação (5)

$$p = \frac{1}{1 + \exp^{-(\beta_0 + \beta_i x_i)}} \quad (5)$$

Por fim, coloca-se que para a realização de tarefas de classificação via regressão logística, deve estabelecer-se um limiar de classificação ( $q$ ). Este limiar representa uma probabilidade na qual eventos que tenham probabilidade de ocorrência maior ou igual a  $q$  são classificados como “1” e eventos com probabilidade menor que  $q$  são classificados como “0”.

### **3. MATERIAL E MÉTODOS**

#### **3.1. Base de dados**

Para a realização desta pesquisa foram utilizados dados de produção estimada e consolidada de café, bem como dados diários de precipitação para as cidades de Guaxupé (coordenadas UTM: E: 322818 N: 7643426, altitude: 870 m), Monte Santo de Minas (coordenadas UTM: E: 295062 - N: 7655267, altitude: 910 m) e Monte Carmelo (coordenadas UTM: E: 237642 - N: 7927660, altitude: 912 m). Todos os dados foram disponibilizados pela Cooperativa dos Cafeicultores de Guaxupé (COOXUPÉ) e correspondem ao período de 2014 a 2021.

A estimativa de produção de café fornecida pela cooperativa foi realizada no meio da fase fenológica V1 (entre novembro e dezembro) de cada safra. Portanto, tendo em vista que a estimativa já foi realizada com o conhecimento das precipitações para este ciclo, as análises aqui desenvolvidas foram realizadas apenas para os veranicos extremos dos ciclos R1 e R2, uma vez que para estes ciclos há o desconhecimento total das condições meteorológicas e isso justifica a estruturação do produto de seguro.

Além disso, as cidades escolhidas para a realização da pesquisa foram selecionadas pelo critério de serem produtoras de café dentro de Minas Gerais, que é o maior produtor de café arábica do País (CONAB, 2022), bem como por possuírem mesma classificação climática Köppen-Geiger, sendo caracterizadas pelo clima tropical de savana com estação seca no inverno (REBOITA; RODRIGUES; ALVES, 2015).

#### **3.2 Construção dos modelos que relacionam as perdas de produção com os veranicos extremos**

Para a criar um modelo que reduza o risco base de um produto de seguro agrícola paramétrico contra veranicos extremos nas fases críticas para lavouras de café, é necessário realizar a adequada compatibilização entre os pagamentos das indenizações e a ocorrência de perdas. Para isso, o primeiro passo foi avaliar dentre dados históricos os períodos nos quais ocorreram perdas na produção de café e verificar o impacto dos veranicos extremos observados nos ciclos R1 e R2 nestas ocorrências. Para tanto, foi realizada a obtenção da variável que descrevesse as safras em que houve ou não perdas, bem como a obtenção das variáveis que descrevessem os veranicos extremos em cada ciclo.

##### **3.2.1 Obtenção da variável perda**

Para cada safra considerada, de 2014 a 2021, foram selecionados os dados de produção esperada e produção realizada por cidade. Assim, quando a produção realizada de uma safra fosse maior ou igual à produção esperada, considerou-se que “não houve perda” ( $L=0$ ). Já para as safras em que a produção realizada foi menor que a produção esperada, considerou-se que “houve perda” ( $L=1$ ). Desse modo, a variável perda ( $L$ ) deste trabalho é uma variável binária que pode ser descrita como

$$L = \begin{cases} 0 & ; \text{ Não houve perda} \\ 1 & ; \text{ Houve perda} \end{cases}$$

Assim, a variável perda de cada ano e cidade amostrada está apresentada na Tabela 1.

**Tabela 1** - Safra estudada, ano da colheita e variável perda ( $L$ ) para as cidades de Guaxupé (G), Monte Carmelo (MC) e Monte Santo de Minas (MSM)

Safra	Ano da Colheita	L		
		Guaxupé	MSM	MC
2012-2014	2014	1	1	0
2013-2015	2015	0	0	1
2014-2016	2016	0	1	0
2015-2017	2017	1	0	1
2016-2018	2018	0	0	0
2017-2019	2019	1	1	1
2018-2020	2020	0	0	0
2019-2021	2021	1	1	1

Para relacionar a perda aos veranicos extremos, o próximo passo foi identificar os veranicos extremos ocorridos em cada ciclo fenológico.

### 3.2.2 Modelagem dos veranicos extremos

A exemplo do que foi feito e descrito em Dalhaus, Musshoff e Finger (2018), nesta pesquisa os ciclos fenológicos foram considerados de forma isolada como uma maneira de reduzir o risco base temporal. Além disso, também foi ajustado um modelo que considerou os dois ciclos conjuntamente. O modelo com esta junção é justificado, uma vez que ambos dependem fortemente de disponibilidade hídrica e ocorrem seguidamente. Deste modo, foram considerados os períodos de análise como dispostos na Tabela 2.

**Tabela 2** - Período de tempo considerado em cada variável fenológica

Ciclo Fenológico	Período Correspondente ao Segundo ano Fenológico
R1	setembro+outubro+novembro+dezembro
R2	janeiro+fevereiro+março
R1R2	setembro+outubro+novembro+dezembro+janeiro+fevereiro+março

Para o ajuste da distribuição GEV que descrevesse os veranicos máximos em cada ciclo fenológico, foram selecionados via blocos de máximos os dois maiores veranicos observados em cada cidade em cada ano da amostra, ou seja, foram selecionadas as duas maiores sequências de dias sem chuvas para cada período em cada ano e cidades amostradas. Ressalta-se que um dia foi considerado “seco” se a precipitação observada neste foi menor do que 3 mm.

Dessa forma, como selecionou-se os dois maiores valores de veranicos da série de dados, considera-se que a série de máximos obtida em cada ciclo, para cada cidade, é de segunda ordem ( $k=2$ ). Logo, a série de máximos para a variável R1 da cidade de Guaxupé, por exemplo, pode ser descrita como

$$\left(R1_{n=8}^{k=2}\right)_{i=1} = \{M_{2014}^2, M_{2015}^2, \dots, M_{2021}^2\}$$

Assim, visto que para cada cidade existem observações para 8 safras, as séries de máximos utilizadas para o ajuste da GEV que contemplaram cada variável fenológica de cada cidade foram compostas por 16 observações cada. Seguindo, para a série de cada ciclo fenológico de cada cidade foram realizados os testes preliminares de Ljung-Box e Mann-Kendall, sendo que foi considerado um nível de significância de 5% para ambos os testes.

Atendido o pressuposto de independência e avaliada a tendência na série de dados, foi realizado o ajuste dos parâmetros da distribuição GEV para cada cidade e ciclo fenológico. Portanto, foram ajustadas 9 distribuições GEV, uma vez que são três ciclos avaliados para três cidades diferentes. Os parâmetros  $\mu$ ,  $\sigma$  e  $\varepsilon$  da distribuição GEV foram ajustados via inferência bayesiana, sendo a estimativa a posteriori dos parâmetros é descrita pela equação 6

$$p(\mu, \sigma, \varepsilon|x) \propto p(\mu, \sigma, \varepsilon) \cdot L(\mu, \sigma, \varepsilon|x) \quad (6)$$



em que  $p(\mu, \sigma, \varepsilon|x)$  representa a distribuição à *posteriori* conjunta de  $\mu, \sigma, \varepsilon$ ,  $p(\mu, \sigma, \varepsilon)$  é a distribuição à *priori* conjunta fornecida para  $\mu, \sigma, \varepsilon$  e  $L(\mu, \sigma, \varepsilon|x)$  é a função de verossimilhança conjunta da distribuição GEV.

Destarte, considerando que para realização da pesquisa não se dispunha de informações à *priori* que fossem suficientes para favorecer determinado conjunto de valores, utilizou-se *prioris* não informativas para os parâmetros. Assim, para descrever o parâmetro  $\mu$  utilizou-se a distribuição normal com média 0 e variância  $10^{-4}$  ( $N(M_1 = 0; S_1^2 = 10^{-4})$ ), para o parâmetro  $\sigma$  utilizou-se a distribuição log-normal com média 0 e variância 9,21 ( $(Ln(M_2 = 0; S_2^2 = 9,21))$ ) e para o parâmetro  $\varepsilon$  utilizou-se a distribuição normal com média 0 e variância  $10^{-2}$  ( $N(M_3 = 0; S_3^2 = 10^{-2})$ ).

A escolha da distribuição das *prioris* foi baseada no trabalho de Coles e Paul (1996). Destarte, seguindo Coles (2001), a *priori* utilizada nesta pesquisa pode ser descrita como na equação 7.

$$p(\mu, \sigma, \varepsilon) = \exp\left(-\frac{1}{2}\left(\frac{\mu-M_1}{S_1}\right)^2\right) \cdot \exp\left(-\frac{1}{2}\left(\frac{\ln(\sigma)-M_2}{S_2}\right)^2\right) \cdot \exp\left(-\frac{1}{2}\left(\frac{\varepsilon-M_3}{S_3}\right)^2\right) \quad (7)$$

Já a função de verossimilhança da GEV pode ser descrita de acordo com a equação 8.

$$L(\mu, \sigma, \varepsilon) = \frac{1}{\sigma^n} \prod_{i=1}^n \left\{ \left[ 1 + \varepsilon \left( \frac{x_i - \mu}{\sigma} \right) \right]^{-\frac{1+\varepsilon}{\varepsilon}} \right\} \exp \left\{ \sum_{i=1}^n \left\{ - \left[ 1 + \varepsilon \left( \frac{x_i - \mu}{\sigma} \right) \right]^{-\frac{1}{\varepsilon}} \right\} \right\} \quad (8)$$

em que  $x_i$  são os valores observados e  $n$  é o tamanho da amostra.

Dessa forma, após obtidas a distribuição *posteriori* conjunta dos parâmetros foi necessário obter-se a marginal da distribuição a *posteriori* para cada um. Como a solução analítica para a marginal não é tarefa fácil empregou-se para a tarefa o algoritmo Metropolis-Hastings via MCMC. Desse modo, para verificar a convergência das cadeias *posteriori* dos parâmetros adotou-se a avaliação dos critérios de Raftery e Lewis, Geweke e Heidelberger e Welch.

Segundo Nogueira, Safadi e Ferreira (2004), de acordo com estes critérios existem evidências de convergência das cadeias a *posteriori* quando no critério de Raftery e Lewis (1992) os valores do fator de dependência são próximos de 1, quando o módulo da estatística de Geweke (1992)  $|GW| < 1,96$  e quando o critério de Heidelberger e Welch (1983) apresenta valor-p  $> 0,05$ .

Continuando a estruturação dos modelos, o próximo passo consistiu na obtenção de uma nova série de máximos de primeira ordem ( $k=1$ ) composta apenas pelos maiores veranicos extremos de cada fase fenológica e cidade. A critério de exemplificação, coloca-se que a série de quantis máximos para a variável R1 da cidade de Guaxupé pode ser descrita como

$$\left( R1_{n=8}^{k=1} \right)_{i=1} = \{ M_{2014}^1, M_{2015}^1, \dots, M_{2021}^1 \}.$$

Assim, foram obtidas 9 séries de veranicos extremos de tamanho 8, sendo uma série para cada cidade e variável fenológica analisada. Os quantis obtidos na série de máximos de primeira ordem foram empregados nas distribuições GEV ajustadas correspondentes fornecendo, conforme a equação (3), as probabilidades de excedência dos  $x$  dias de veranicos extremos observados em cada variável fenológica e cidade ( $p(VE)$ ).

Desse modo, para cada variável fenológica e cidade obteve-se as probabilidades de ocorrência para os 8 quantis extremos de veranicos. Como são 3 cidades, considerou-se então que para cada variável fenológica se dispunha de 24 dados que representavam probabilidades de ocorrência de veranicos extremos, sendo este, portanto, o tamanho da série  $p(VE)$ .

### 3.2.3 Ajuste dos modelos previsores de perdas

O próximo passo consistiu no ajuste de três modelos de regressão logística para que fosse possível verificar qual a probabilidade de se observar perda na produção final, dada a probabilidade de excedência de  $x$  dias de veranico extremo em cada fase fenológica. Dessa forma, foi ajustado um modelo logístico para a fase R1, um para a fase R2 e um modelo para o período correspondente a R1+R2, em que a variável perda ( $L$ ) era a variável resposta do modelo e as probabilidades  $p(VE)$  formavam a variável explicativa do modelo.

Para exemplificação, tem-se que a probabilidade de perda em R1 ( $L_{R1} = 1$ ) fornecida pelo modelo logístico observada a probabilidade de excedência de  $x$  dias de veranico extremo foi obtida conforme a equação 5 por

$$p(L_{R1} = 1) = \frac{1}{1 + \exp^{-(\beta_0 + \beta_1 p(VE_2))}}$$

em que utilizou-se o método dos mínimos quadrados ordinários para a estimação dos coeficientes  $\beta_0$  e  $\beta_1$  das regressões.

### 3.3 Avaliação dos modelos ajustados

A decisão em relação ao gatilho mais eficiente foi feita pela avaliação conjunta da qualidade estatística que o gatilho apresenta, bem como pela avaliação do benefício financeiro que cada gatilho resulta para a seguradora e segurado. Desse modo, para verificar o desempenho do modelo a cada gatilho  $q$  considerado, o primeiro procedimento foi verificar quantos acertos de eventos de perda ( $A(q)$ ), alarmes falsos de perda ( $B(q)$ ), falhas na previsão de perdas ( $C(q)$ ) e acertos de eventos de não perda ( $D(q)$ ) cada gatilho teve.

Para isso, considerou-se que a cada  $q$  considerado, todos os eventos que apresentassem probabilidade de perda maiores ou iguais a  $q$  (limiar da classificação), seriam indenizados ( $L_{ciclo} = 1$ ) e todos os eventos com probabilidade de perda menor do que o  $q$  considerado, não seriam indenizados ( $L_{ciclo} = 0$ ). A partir disto, tornou-se possível confrontar o indicativo de perda dos modelos a cada  $q$  ( $L_{ciclo}$ ), com a variável perda observada ( $L$ ), obtida em 3.2.1, e verificar para qual gatilho os modelos ajustados compatibilizaram melhor as ocorrências de perda com o pagamento das indenizações, ou seja, para qual  $q$  ocorreu mais vezes de o  $L_{ciclo} = L$ .

#### 3.3.1 Avaliação dos modelos pela qualidade do ajuste

Para a avaliação da eficiência estatística dos modelos propostos para cada ciclo a cada gatilho indenizatório considerado, foram analisadas para cada  $q$  as métricas de acurácia, viés, sensibilidade e especificidade. A acurácia corresponde a proporção de eventos corretamente classificados, sendo que foi obtida por

$$Acurácia = \frac{A(q)+D(q)}{A(q)+B(q)+C(q)+D(q)}$$

O viés é uma métrica que apresenta a razão entre o número de previsões do evento de interesse ( $L_{ciclo} = 1$ ) e o número de eventos de ocorrência do evento de interesse ( $L = 1$ ), sendo fornecido por

$$Viés = \frac{A(q)+B(q)}{A(q)+C(q)}$$

em geral, deseja-se que o modelo apresente o viés mais próximo possível de 1, pois isto indica que os eventos são previstos em proporção aproximada ao que é de fato observado.

Ademais, a métrica de sensibilidade mede a quantidade de perdas corretamente previstas dentre todas as perdas previstas e a métrica de especificidade mede a quantidade de eventos de não perda previstos dentre todas as ocorrências de não perda. Estas métricas são obtidas por

$$sensibilidade = \frac{A(q)}{A(q)+C(q)}$$

$$especificidade = \frac{D(q)}{D(q)+B(q)}$$

### 3.3.2 Avaliação dos modelos pelo benefício financeiro

A avaliação dos modelos pelo benefício financeiro, ou avaliação da utilidade esperada dos modelos, será realizada por meio da avaliação dos custos que a seguradora e o segurado terão com os modelos ajustados a cada  $q$  considerado, em cada ciclo para o qual foi ajustado o modelo. Para isso, considerou-se que os modelos ajustados trazem como custo total

$$Custo = \frac{A(q)}{n}C_A + \frac{B(q)}{n}C_B + \frac{C(q)}{n}C_C + \frac{D(q)}{n}C_D$$

em que  $C_A, C_B, C_C, C_D$  são os custos atrelados a ocorrência dos eventos  $A(q), B(q), C(q), D(q)$  e  $n$  é a quantidade de eventos previstos pelo modelo.

A despeito disto, é preciso considerar que a estrutura de custo de um produto é diferente para a seguradora e para o segurado. Portanto, assim como em Figueiredo et al. (2018), foi avaliado o custo do modelo de cada variável fenológica de forma isolada para ambos. A estrutura de custo utilizada nesta pesquisa foi a mesma sugerida pelos autores e está apresentada na Tabela 3.

**Tabela 3** - Estrutura de custo para a seguradora e segurado em relação aos modelos de contrato ajustado

Agente	$L_{ciclo}$	$L_{ij}$	
		1	0
Seguradora	1	$C_A = E_O$	$C_B = E_O + E_{AF}$
	0	$C_C = E_F$	$C_D = 0$
Segurado	1	$C_A = E_P(q)$	$C_B = E_P(q) - Indenização$
	0	$C_C = E_P(q) + E_{PJ}$	$C_D = E_P(q)$

Fonte: Adaptado de Figueiredo et al. (2018)

De acordo com a Tabela 4, para eventos  $A(q)$ , a seguradora incorre somente em custos operacionais relacionados ao pagamento da indenização ( $E_o$ ), uma vez que a perda ocorrida já estava prevista. Já para eventos  $B(q)$ , a seguradora além dos custos operacionais relacionados ao pagamento das indenizações, ainda terá de recorrer a custos extras para a recalibração do modelo, além de incorrer também em custos de imagem relacionados ao alarme falso incorrido ( $E_{AF}$ ). Para eventos  $C(q)$ , a seguradora terá de arcar também com a recalibração do modelo, bem como com o custo de imagem relacionado a omissão no pagamento de indenizações quando esta deveria ter sido paga ( $E_p$ ). Por fim, a seguradora incorre em custo zero quando não prevê um evento e ele de fato não ocorre.

Já em relação ao segurado, todos os custos variam de acordo com o valor do prêmio ( $E_p(q)$ ), sendo que para eventos em que o modelo acerta eventos de perda  $A(q)$  e de não perda  $D(q)$ , o custo do segurado é somente o valor despendido com o prêmio. Para eventos de falha no modelo  $C(q)$ , além do prêmio pago o segurado ainda terá o custo para arcar com o prejuízo do sinistro ( $E_{pj}$ ), uma vez que não receberá a indenização. Por fim, para eventos de alarme falso  $B(q)$ , o segurado incorrerá em um custo negativo, uma vez que vai receber a indenização sem ter sofrido perdas e, conseqüentemente, terá um “ganho”.

Desse modo, ressalta-se que o valor do prêmio cobrado do segurado vai depender do gatilho considerado. Isso ocorre porque o valor cobrado pelo prêmio leva em conta o risco de pagamento de determinada indenização e isto vai variar a cada  $q$ . Dessa forma, ainda segundo Figueiredo et al. (2018), o cálculo do prêmio é dado pela equação

$$E_p(q) = \frac{A(q)+B(q)}{A(q)+B(q)+C(q)+D(q)} * Indenização * m$$

em que o valor da indenização é o valor máximo definido em contrato e  $m$  representa 1+ %carregamentos.

Ademais, para avaliar o desempenho do modelo foram calculadas duas métricas, sendo uma que apresenta qual seria o custo para a seguradora e para o segurado quando o modelo fosse perfeito e previsse todas as ocorrências de perda e de não perda corretamente ( $C_{perfeito}$ ) e o custo para os agentes caso ficassem na aleatoriedade, sem utilizar um modelo para prever as possíveis perdas ( $C_{aleatório}$ ). A Tabela 4 apresenta os valores empregados para os cálculos destas métricas.

**Tabela 4** - Custo para a seguradora e para o segurado considerando um modelo de previsão de perdas perfeito e a utilização de nenhum modelo

Agente	$C_{perfeito}$	$C_{aleatório}$
Seguradora	$sE_O$	$sE_F$
Segurado	$E_p(q)$	$sE_{PJ} + E_p(q)$

Fonte: Conceitos adaptados de Figueiredo et al. (2018)

Pela Tabela 4, tem-se que o custo de um modelo perfeito para a seguradora será relacionado apenas à proporção de eventos de perdas observadas dentre todos os eventos ocorridos ( $s$ ) multiplicada pelo custo operacional que ela terá para realizar os pagamentos dos eventos de perda. Já em relação ao segurado, um modelo perfeito trará como custo apenas o valor do prêmio. Destarte, considerando o não uso do modelo, o custo da seguradora será a proporção  $s$  multiplicada pelo custo de uma falha, bem como para o segurado será a proporção  $s$  multiplicada pelo prejuízo incorrido a cada perda não prevista, somado ao valor despendido com o prêmio.

Por fim, para avaliar a economia que a seguradora e o segurado tiveram com os modelos a cada gatilho considerado, deve-se realizar a subtração do custo que se tem com a não utilização o modelo a cada  $q$  do custo que a utilização do modelo traz neste  $q$ , sendo

$$economia = C_{aleatório}(q) - C_{modelo}(q).$$

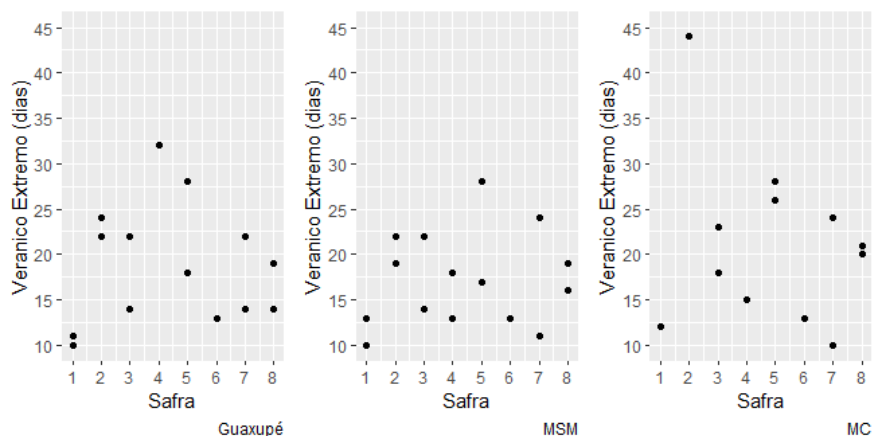
Por fim, ressalta-se que todas as análises estatísticas foram desenvolvidas por meio do *software* estatístico *R core team (2017)* pelos pacotes *evd*, *R2OpenBUGS*, *kendall*, *coda*, *rstanarm*, *ggplot2* e *hnp*.

## 4. RESULTADOS E DISCUSSÃO

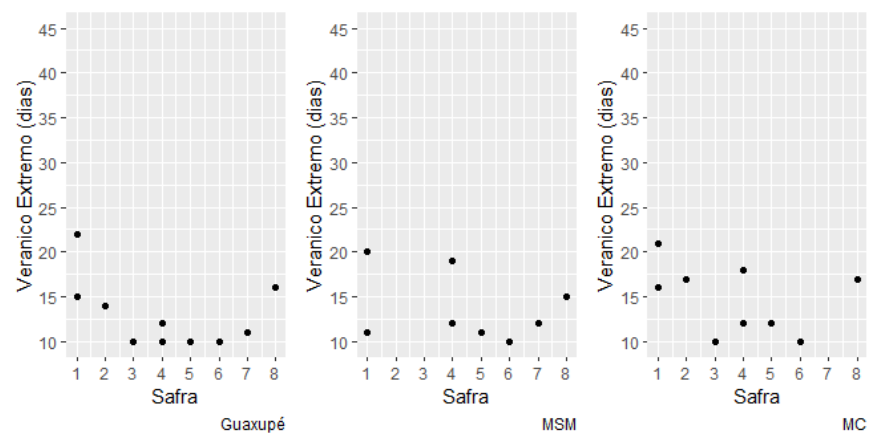
### 4.1 Análise descritiva dos veranicos máximos

Os valores das séries de máximos de segunda ordem obtidos para os ajustes das distribuições GEV para a modelagem dos veranicos extremos de cada variável fenológica de cada cidade amostrada estão dispostos nos gráficos de dispersão apresentados na Figura 2.

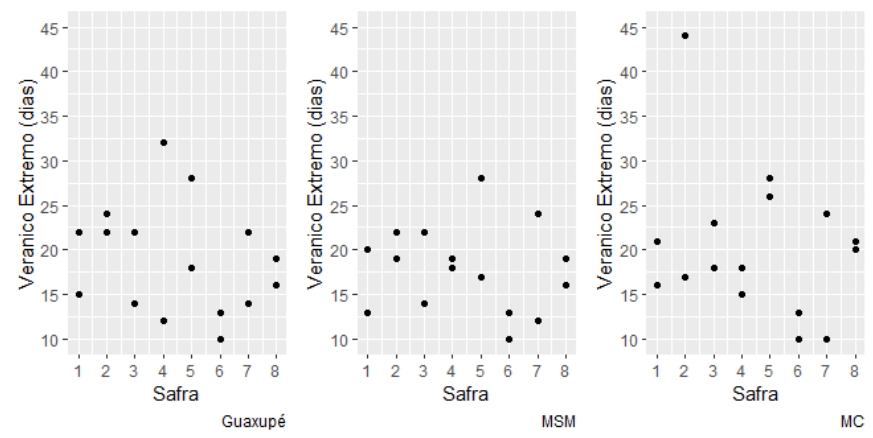
### Variável R1



### Variável R2



### Variável R1R2



**Figura 2** - Valores das séries de máximos de segunda ordem dos veranicos extremos para as variáveis temporais R1, R2, R1R2 nas cidades de Guaxupé, Monte Santo de Minas (MSM) e Monte Carmelo (MC)

De acordo com a Figura 2 é possível constatar que o tamanho dos veranicos extremos observados na fase R1, em geral, é mais expressivo do que o tamanho dos veranicos observados em R2, por exemplo, os dois maiores veranicos observados em R1 são de 44 e 33 dias em Monte Carmelo e Guaxupé, respectivamente, enquanto que os dois maiores veranicos observados em R2 são de 22 e 21 dias em Guaxupé e Monte Carmelo, respectivamente. Tal observação já era esperada, uma vez que o ciclo R1 apresenta uma parte de seu período de tempo situada no período seco no ano (abril a setembro), enquanto que o ciclo R2 se situa por completo dentro período chuvoso (outubro a março).

Entretanto, para alguns anos aconteceram exceções como, por exemplo, para a safra de 2012-2014 (Safra 1) em Guaxupé, na qual os dois maiores veranicos observados foram provenientes de R2. Assim, em relação a série R1R2, como foram considerados os dois maiores veranicos extremos de todo o período, em algumas safras os dois veranicos extremos vieram do ciclo R1, em outras o primeiro veranico extremo veio da fase R1 e o segundo da fase R2 e em alguns casos ambos foram oriundos da fase R2.

#### 4.2 Determinação da probabilidade de excedência dos veranicos máximos

Considerando as séries de veranicos máximos obtidas para cada cidade e em cada período apresentado na Tabela 2, foram realizados os testes de Ljung-Box e Mann-Kendall. Os resultados dos valores-p obtidos nos testes estão apresentados na Tabela 5.

**Tabela 5** - Valores-p obtidos nos testes de Ljung-Box (LB) e Mann-Kendall (MK) para as séries de veranicos máximos para cada variável temporal e cidade consideradas.

Variável fenológica	Cidades					
	Guaxupé		Monte Carmelo		Monte Santo de Minas	
	LB	MK	LB	MK	LB	MK
R1	0,231	0,964	0,098	0,964	0,162	0,928
R2	0,248	0,386	0,054	0,172	0,769	0,927
R1R2	0,290	0,363	0,731	0,527	0,414	0,415

Considerando o nível de significância de 5% para ambos os testes, obteve-se que os testes de Ljung-Box e Mann-Kendall apresentaram resultados não significativos (valor-p >



5%) para todas as variáveis temporais analisadas de todas as cidades. Desse modo, a pressuposição de independência foi atendida em todas as 9 séries de extremos analisadas, bem como não foi apontada a necessidade de modelar tendência em nenhuma delas. Destarte, o próximo passo consistiu em ajustar a distribuição GEV para modelar os veranicos extremos de cada cidade em cada variável fenológica analisada.

Na Tabela 6 são apresentadas as estimativas médias dos parâmetros da distribuição GEV para as diferentes estruturas cidades e variáveis fenológicas e os resultados dos critérios de convergência das cadeias de distribuição *a posteriori* dos parâmetros. Ressalta-se que para produzir a distribuição *a posteriori* via MCMC considerou-se a geração de 1 cadeia, 45 mil iterações, queima de 3 mil e salto igual a 4.

**Tabela 6** - Estimativas médias dos parâmetros para cada distribuição GEV, de acordo com a cidade e variável fenológica, e os resultados dos critérios para análise de convergência das cadeias da distribuição *a posteriori* dos parâmetros, sendo eles Raftery e Lewis (R&L), Geweke (GW) e Heidelberger-Welch (H&W).

Cidade	Variável Fenológica	Parâmetro	Estimativa	R&L	GW	H&L
Guaxupé	R1	$\mu$	13,90	1,01	0,05	0,830
		$\sigma$	6,43	0,99	0,08	0,731
		$\varepsilon$	0,02	1,00	0,59	0,619
	R2	$\mu$	8,54	0,99	0,90	0,115
		$\sigma$	2,69	1,02	0,94	0,119
		$\varepsilon$	0,34	0,98	0,64	0,799
	R1R2	$\mu$	15,97	1,04	1,05	0,063
		$\sigma$	5,28	0,99	0,49	0,130
		$\varepsilon$	0,04	1,00	1,42	0,212
Monte Santo de Minas	R1	$\mu$	14,23	1,00	0,78	0,648
		$\sigma$	5,09	1,01	0,18	0,939
		$\varepsilon$	-0,05	1,00	1,92	0,574
	R2	$\mu$	7,83	1,01	0,26	0,370
		$\sigma$	3,70	1,00	1,43	0,185
		$\varepsilon$	0,14	1,01	0,66	0,782
	R1R2	$\mu$	15,72	0,99	1,34	0,678
		$\sigma$	4,61	1,00	0,28	0,821

		$\varepsilon$	-0,08	1,01	1,00	0,605
		$\mu$	11,76	1,16	0,20	0,968
	R1	$\sigma$	4,19	0,99	0,40	0,906
		$\varepsilon$	1,05	1,01	0,35	0,974
		$\mu$	9,52	1,02	0,76	0,807
Monte Carmelo	R2	$\sigma$	3,40	0,98	1,19	0,315
		$\varepsilon$	0,19	1,01	0,40	0,775
		$\mu$	16,24	1,00	0,99	0,847
	R1R2	$\sigma$	6,11	1,00	0,12	0,052
		$\varepsilon$	0,16	1,02	0,83	0,577

A análise da Tabela 6 permite constatar que todas as cadeias de distribuição *posteriori* dos parâmetros atenderam as condições propostas nos três critérios de convergência avaliados, sendo este um indicativo de que as cadeias convergiram em todos os parâmetros para todos os modelos ajustados. Em relação ao critério de Raftery e Lewis, todos foram próximo a 1, o que indica independência entre as iterações. Pelo critério de Geweke tem-se que  $|Gw| < 1,96$ , indicando que não há indícios de ausência de convergência. Ademais, os valores-p do critério de Heidelberger e Welch foram não significativos ( $p > 0,05$ ), atestando que a série é estacionária.

Dessa forma, prosseguiu-se o estudo para a obtenção da variável explicativa do modelo previsor de perdas.

Utilizando a distribuição GEV ajustada para cada cidade e em cada variável fenológica foi possível obter as probabilidades de excedência para todas as quantidades de dias de veranicos extremos obtidos na série de máximos de primeira ordem, como descrito em 3.2.2.3. Estas probabilidades foram usadas como variável explicativa do modelo de previsão de perdas proposto nesta pesquisa. A critério de exemplificação, a Tabela 7 mostra quais as probabilidades de excedência de cada quantidade de dias de veranico extremo para a variável R2 na cidade de Monte Carmelo.

**Tabela 7** - Veranicos extremos observados em cada safra da amostra para a variável fenológica R2 e suas probabilidades de excedência na cidade de Monte Carmelo

Safra	Veranico Extremo Observado (VE) em dias	$p(VE)$ (%)
2012-2014	21	7,10
2013-2015	17	14,71
2014-2016	10	58,12
2015-2017	18	12,18
2016-2018	12	39,67
2017-2019	10	58,12
2018-2020	9	68,89
2019-2021	17	14,71

Como apresentado na Tabela 7, o maior valor de veranico observado no ciclo R2 foi para a safra de 2012-2014, que vivenciou 21 dias seguidos sem precipitação relevante. A probabilidade de que ocorram 21 dias ou mais de veranico na cidade, considerando a distribuição GEV ajustada, é de 7,10%. Desse mesmo modo, foram estimadas as probabilidades de excedência ( $p(VE)$ ) dos quantis obtidos dos veranicos extremos de primeira ordem para cada variável fenológica e cidade, totalizando 24 dados com estas probabilidades.

#### 4.3 Análise da compatibilização de perdas em função da ocorrência de veranicos extremos

Partindo das  $p(VE)$  calculadas de cada quantidade de  $z$  dias de veranicos extremos como variável explicativa da variável perda ( $L$ ), ajustou-se um modelo logístico para cada variável fenológica. A saída obtida nos modelos de regressão estão apresentados na Tabela 8.

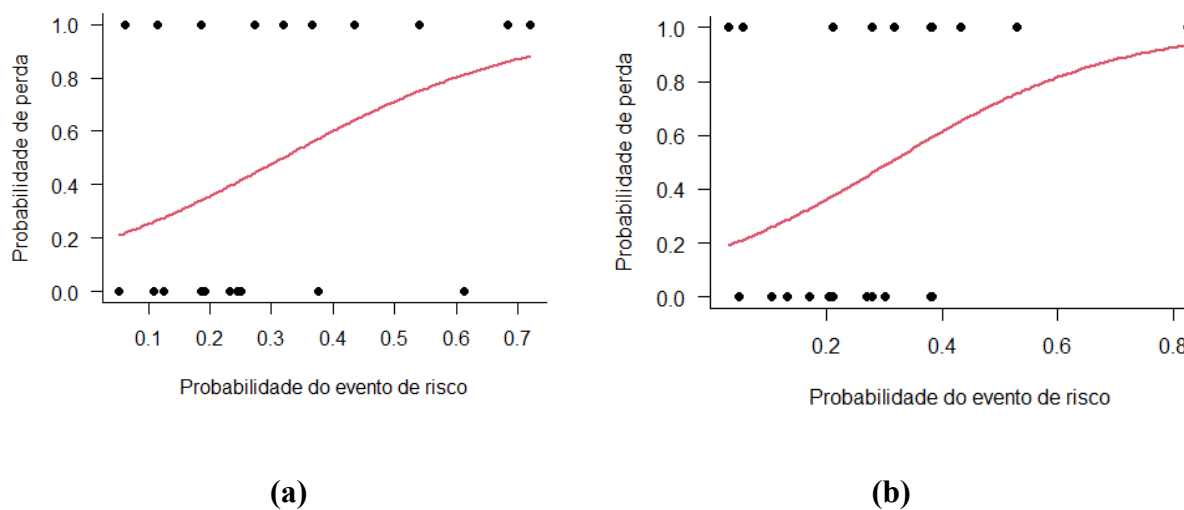
**Tabela 8** - Coeficientes, erro padrão e valor-p dos modelos de regressão logística ajustados para cada variável fenológica

Ciclo	Variáveis	Coeficientes	Erro padrão	valor-p
R1	$\beta_0$	-1,59	0,88	<b>0,071</b>
	$\beta_1$	4,98	2,51	<b>0,047</b>
R2	$\beta_0$	0,948	0,82	0,248

	$\beta_1$	-2,944	2,19	0,180
<b>R1R2</b>	$\beta_0$	-1,60	0,93	<b>0,0867</b>
	$\beta_1$	5,15	2,83	<b>0,0688</b>

Pela Tabela 8 constata-se que, ao nível de 5% de significância, as probabilidades de excedência dos veranicos extremos na fase fenológica R1 influenciam positivamente a ocorrência de perdas na produção final, ou seja, quanto maior a probabilidade de excedência de determinada quantidade  $z$  de veranico extremo, maior a probabilidade de observar-se uma perda na produção final.

Em relação às probabilidades de excedência de veranicos extremos na fase R1R2 constatou-se que, considerando um nível de 10% de significância, essa variável também impacta positivamente a ocorrência de perdas na produção final. Não foi encontrada significância estatística no modelo que correlacionou as probabilidades de ocorrência de veranicos extremos na fase R2 e com as perdas observadas na produção final, portanto, essa variável fenológica foi desconsiderada nas análises subsequentes. Na Figura 3 estão apresentados os gráficos dos modelos de regressão ajustados para R1 e R1R2.



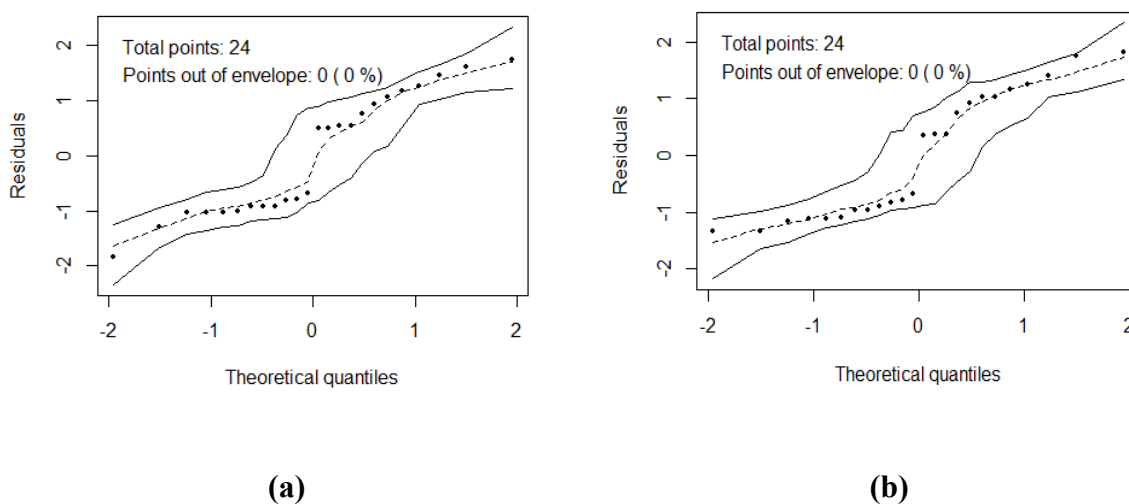
**Figura 3** - Modelos de regressão logística ajustados para os ciclos fenológicos R1 (a) e R1R2 (b)

O ajuste dos modelos de regressão ilustrados na Figura 1 permitem a afirmação de que quanto maior a probabilidade de ocorrência de um evento de risco, no caso o veranico

extremo, maior também é a probabilidade de que perdas sejam observadas em decorrência deste, ou seja, maior a probabilidade de que a seguradora tenha de pagar uma indenização.

Por exemplo, se consideramos 15 dias de veranico, o modelo vai considerar as probabilidades de excedência de 15 dias de veranicos, ou seja, o modelo vai acumular as probabilidades se serem observados 15 ou mais dias de veranico e isso contabiliza a probabilidade de 15, 16, 17, ... até os  $n$  dias possíveis de veranicos que compreendem o tamanho do ciclo fenológico considerado. Portanto, se consideramos um veranico extremo de 45 dias, por exemplo, a probabilidade de excedência associada a ele é menor do que a de 15 dias, pois aquela probabilidade está inclusa nesta, mas o oposto não ocorre.

Assim, ressalta-se que pelo modelo ajustado, quanto maior o veranico considerado em dias, menor é a probabilidade de o evento de risco ser observado. Portanto, a probabilidade de ocorrência de um evento de perda ( $L=1$ ) em função da probabilidade de ocorrência de veranicos muito grandes é menor do que a probabilidade de se ter de pagar indenização nos casos de veranicos menores. Na Figura 4 estão apresentados os gráficos de envelope simulado dos modelos logísticos ajustados.



**Figura 4** - Gráfico de envelope simulado dos modelos de regressão logística ajustados para os ciclos fenológicos R1 (a) e R2 (b)

Pelos gráficos apresentados na Figura 4 é possível verificar que os modelos propostos se ajustaram bem aos dados, uma vez que nenhum ponto foi observado fora do envelope nos dois ciclos analisados. Desse modo, por meio dos modelos logísticos ajustados foi possível estimar a probabilidade de ser observada perda na produção de café ( $L=1$ ) a cada probabilidade de veranicos extremos observados nos ciclos R1 e R1R2.

O modelo logístico ajustado para o ciclo R1 estimou probabilidades de perda que variaram de 21% a 88%. Já o modelo logístico ajustado para o ciclo R1R2 estimou probabilidades de perda que variaram de 19% a 94%. Nessa linha, considera-se que o gatilho mais assertivo a ser empregado em um seguro paramétrico contra veranicos extremos para lavouras de café em Minas Gerais em cada ciclo analisado está no intervalo das probabilidades previstas.

#### **4.4 Estruturação do contrato de seguro paramétrico**

Considerando que para cada variável fenológica existem diversos gatilhos indicativos de perda que podem ser considerados, deve-se atentar para que o gatilho escolhido no contrato do seguro paramétrico seja aquele que tenha maior probabilidade de compatibilizar adequadamente os pagamentos das indenizações e as ocorrências de perda. Então, definiu-se o gatilho ótimo para os contratos propostos como aqueles que apresentaram as melhores métricas de ajuste associadas ao maior benefício financeiro para a seguradora e ao melhor custo-benefício para o segurado, levando em conta o custo e a segurança do produto.

##### **4.4.1 Avaliação pela qualidade do ajuste**

Para contratos de seguro paramétrico contra veranicos extremos no ciclo fenológico R1, verificou-se que os gatilhos que trazem maior acurácia, ou seja, os gatilhos que mais compatibilizaram adequadamente as perdas ocorridas com o pagamento de indenização foram os gatilhos de 44% e 45%, que conseguiram acertar 79,17% das indenizações devidas. Em relação ao viés, obteve-se para este ciclo valores mais próximos de 1 para os gatilhos de 42% a 45%, demonstrando que gatilhos nessa faixa de valores apresentam resultados menos viesados. Desse modo, sugere-se que quando o modelo de R1 apontar para uma probabilidade de perda de 44% ou 45%, os pagamentos de indenização devem ser realizados. Além disso, para essa faixa de valores também foram obtidos bons resultados de precisão, sendo que neles obteve-se 75% de sensibilidade e 83% de especificidade.

Já para contratos de seguro paramétrico contra veranicos extremos no ciclo fenológico R1R2, verificou-se que os gatilhos que resultam em maior acurácia são 51% e 52%, que conseguiram acertar 75% das indenizações devidas. Em relação ao viés, obteve-se para este ciclo valores mais próximos de 1 para os gatilhos de 48% a 52%. Deste modo, para o modelo ajustado para R1R2 sugere-se que os pagamentos de indenização devem ser realizados quando o modelo apontar para uma probabilidade de perda de 51% ou 52%. Ademais, para

essa faixa de valores também foram obtidos resultados razoáveis de precisão, sendo que neles obteve-se 67% de sensibilidade e 83% de especificidade.

Salienta-se que por meio das distribuições GEV ajustadas para cada cidade e pelo modelo logístico ajustado, torna-se possível verificar a quantidade de dias de veranico extremo que correspondem aos gatilhos ótimos encontrados para cada ciclo para cada cidade. Essa quantidade de dias para o ciclo R1 é de 21, 19 e 20 e para o ciclo R1R2 é de 20, 19 e 22, para as cidades de Guaxupé, Monte Santo de Minas e Monte Carmelo, respectivamente.

Estas quantidades de dias significam a extensão do veranico na qual o modelo apresenta bom desempenho estatístico. No caso de o contrato ser realizado considerando veranicos menos extensos do que os valores apresentados como ótimos, as probabilidades de ocorrência destes veranicos serão maiores, resultando na previsão de probabilidades de perda elevadas. Tal fato induzirá à geração de níveis maiores de alarme falsos (B) no modelo e aumentará, conseqüentemente, o risco base positivo.

Já no caso em que o contrato é realizado considerando veranicos mais extensos do que os valores apresentados como ótimos, as probabilidades de ocorrência desses veranicos irão ser menores, resultando na previsão de probabilidades de perda inferiores. Isto induzirá à geração de maiores níveis de falhas (C) no modelo e aumentará, portanto, o risco base negativo.

Ademais, ressalta-se que a elucidação da quantidade de dias de veranico que correspondem aos gatilhos ótimos encontrados torna a negociação e explicação do seguro e das condições da apólice mais compreensíveis para o segurado, o que pode facilitar a aceitação deste em relação ao produto.

#### **4.4.2 Avaliação do modelo pelo benefício financeiro**

Para realizar a avaliação da utilidade dos modelos, foram supostos alguns valores financeiros relacionados à operação de seguros tanto para o segurado, quanto para a seguradora. Desse modo, considerou-se uma indenização máxima de R\$ 10.000, um prejuízo para o segurado em caso de falha do modelo ( $E_F$ ) de R\$ 15.000, carregamentos ( $m$ ) de 5%, o custo operacional para a seguradora em casos de pagamento de indenização para perdas corretamente previstas ( $E_O$ ) de R\$ 500, o custo de imagem da seguradora e de recalibração do modelo em caso de alarme falso ( $E_{AF}$ ) de R\$ 4.000 e o custo de imagem da seguradora e de recalibração do modelo em caso de falha ( $E_F$ ) de R\$ 7.000.

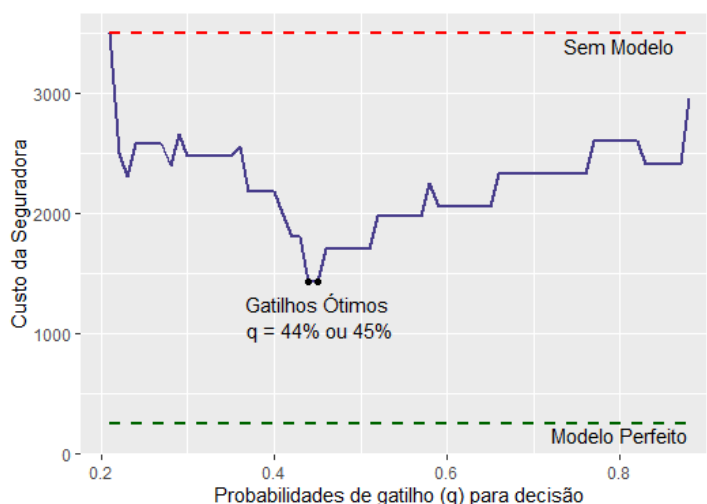
A partir desses valores, foram calculados os custos para a operadora e para o segurado a cada possível gatilho a ser considerado em cada variável fenológica.

#### 4.4.2.1 Custos da seguradora no ciclo R1

A partir dos custos considerados obteve-se que o menor custo para a seguradora para contratos paramétricos contra veranicos extremos ocorre justamente quando se considera os gatilhos que apresentaram o melhor desempenho estatístico, ou seja, para os gatilhos de 44% e 45%. Nesses gatilhos, levando em conta os acertos e erros do modelo proposto, a seguradora teve um custo esperado de R\$ 1.437,5.

Se considerarmos um caso em que não havia modelo de previsão de perdas, ou seja, no caso em que a seguradora trabalha à mercê da aleatoriedade, o custo incorrido pela seguradora seria de R\$ 3.500, havendo, portanto, uma economia de R\$ 2.062,50 pela utilização do modelo sugerido nesta pesquisa. Ademais, considerando um modelo perfeito que previsse corretamente todas as ocorrências de perda e de não perda, ou seja, no caso de um modelo 100% acurado, o custo da seguradora com o contrato seria de apenas R\$ 250,00. Isto sugere que existe um potencial financeiro para este produto paramétrico para seguradoras que investirem no desenvolvimento de modelos preditivos ainda melhores. A Figura 5 ilustra o modelo proposto para identificar o gatilho que traz a o maior benefício financeiro para seguradoras ao ser empregado em contratos de seguro paramétrico contra veranicos extremos no ciclo fenológico R1 em lavouras de café de Minas Gerais.

**Figura 5** - Ilustração do modelo proposto para a identificar gatilho que traz a maior utilidade esperada para seguradoras ao ser empregado em contratos de seguro paramétrico contra veranicos extremos no ciclo fenológico R1 em lavouras de café de Minas Gerais.





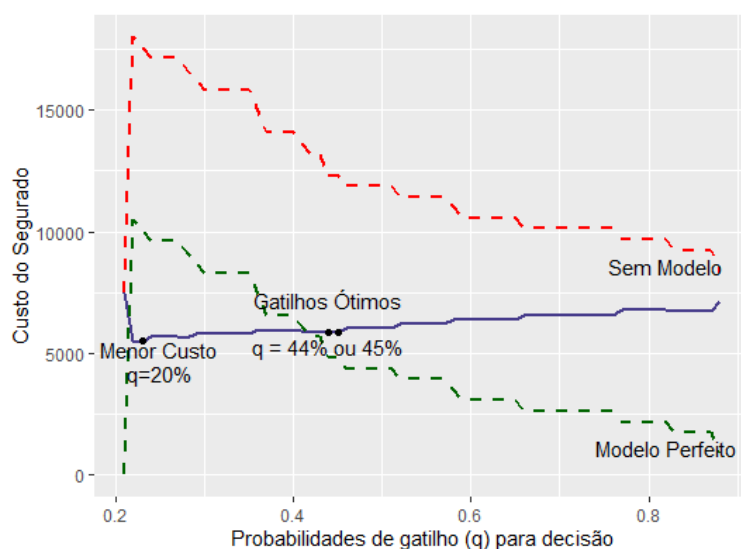
De acordo com a Figura 5, verifica-se que o maior custo para a seguradora ocorreu para os primeiros e últimos gatilhos considerados que foram 21% e 88%, pois nestes gatilhos a seguradora cometeu elevados equívocos de alarmes falsos e falhas, respectivamente.

#### 4.4.2.2 Custos do segurado no ciclo R1

Para os custos supostos como exemplo na modelagem, obteve-se que o menor custo que o segurado tem é para quando considera-se o gatilho de 23%. Entretanto, esse menor custo do segurado ocorre em decorrência de um modelo mal calibrado que incorre em muitos alarmes falsos e, portanto, efetua pagamentos de indenização indevidos aos clientes.

Dessa forma, coloca-se que diferentemente do que ocorre com a seguradora, na qual o modelo ótimo é aquele que coincide com os menores custos, para o segurado o melhor modelo é o que traz um bom custo-benefício, ou seja, um custo justo e ao mesmo tempo maior segurança. Isso ocorre justamente nos gatilhos de 44% e 45%, tendo em vista a maior acurácia do modelo nesses índices. A Figura 6 ilustra o modelo proposto para identificar o gatilho que traz a o maior benefício financeiro para os segurados ao ser empregado em contratos de seguro paramétrico contra veranicos extremos no ciclo fenológico R1 em lavouras de café de Minas Gerais.

**Figura 6** - Ilustração do modelo proposto para a verificação do gatilho que traz a maior utilidade esperada para o segurado ao ser empregado em contratos de seguro paramétrico contra veranicos extremos no ciclo fenológico R1 em lavouras de café de Minas Gerais.



Nesse contexto, obteve-se que nos gatilhos mais acurados na modelagem desse contrato o custo esperado para o segurado foi de R\$ 5.854,17. Além disso, o custo de não

utilizar um modelo para prever as perdas considerando o gatilho ótimo é de R\$ 10.562,50, havendo, portanto, uma economia de R\$ 4.708,33 com o uso do modelo.

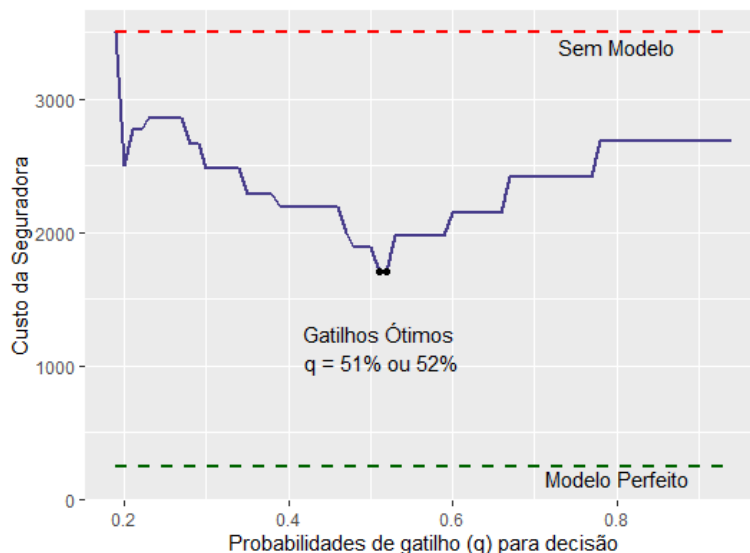
Ademais, um modelo que acertasse todas as previsões faria com que o segurado tivesse como custo exatamente o valor do prêmio pago, independentemente de ocorrência de sinistro ou não. Como o prêmio varia a cada gatilho considerado, considerando-se o valor dos gatilhos ótimos, tem-se que um modelo que acertasse tudo custaria R\$3062,50.

#### 4.4.2.3 Custos da seguradora nos ciclos R1R2

Assim como observado para o modelo de R1, a seguradora obteve o menor custo para os contratos paramétricos nos gatilhos com melhor desempenho estatístico, ou seja, para os gatilhos de 51% e 52%. Nestes gatilhos, levando em conta os acertos e erros do modelo proposto, a seguradora teve um custo de R\$ 1.708,33. Para o caso em que não havia modelo para a previsão de perdas o custo da seguradora seria de R\$ 3.500, havendo, portanto, uma economia de R\$ 1.791,67 com o emprego do modelo sugerido.

Ademais, considerando um modelo perfeito que previsse corretamente todas as ocorrências de perda e de não perda o custo da seguradora com o contrato seria de apenas R\$ 250,00. A Figura 7 ilustra o modelo proposto para identificar o gatilho que traz a maior utilidade esperada para seguradoras ao ser empregado em contratos de seguro paramétrico contra veranicos extremos no ciclo fenológico R1R2 em lavouras de café de Minas Gerais.

**Figura 7** - Ilustração do modelo proposto para a identificar gatilho que traz a maior utilidade esperada para seguradoras ao ser empregado em contratos de seguro paramétrico contra veranicos extremos no ciclo fenológico R1R2 em lavouras de café de Minas Gerais.

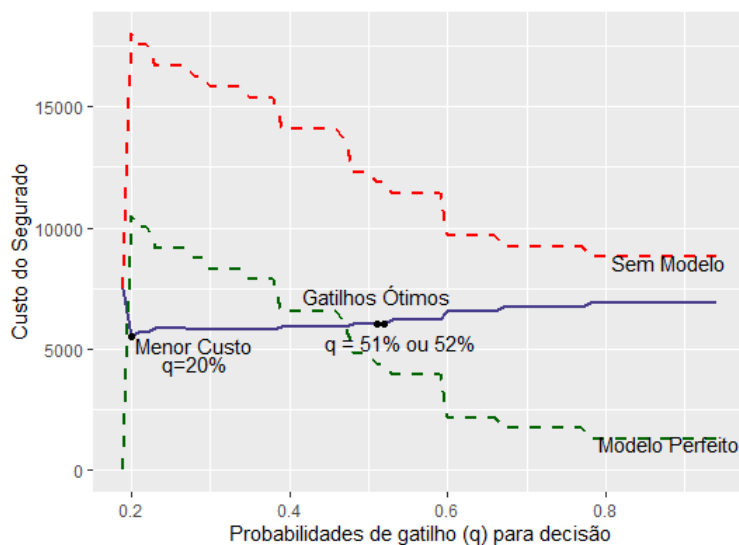


Assim como observado na Figura 5, o maior custo para a seguradora ocorreu nos primeiros e últimos gatilhos considerados que foram 21% e 88% em decorrência do grande número de alarmes falsos e falhas nestes gatilhos, respectivamente.

#### 4.4.2.4 Custos do segurado nos ciclos R1R2

Para este ciclo o menor custo que o segurado tem é para o gatilho de 20%, quando existe alta taxa de alarmes falsos e o recebimento de pagamentos de indenização indevida. Entretanto, o modelo que apresenta maior segurança ao segurado é o mais acurado, onde o gatilho é de 51% ou 52%. Na Figura 8 é apresentada a ilustração do modelo proposto para a verificação do gatilho que traz a maior utilidade esperada para o segurado no contrato firmado para o período R1R2.

**Figura 8** - Ilustração do modelo proposto para a verificação do gatilho que traz a maior utilidade esperada para o segurado ao ser empregado em contratos de seguro paramétrico contra veranicos extremos no ciclo fenológico R1R2 em lavouras de café de Minas Gerais.



Dessa forma, obteve-se que nos gatilhos mais acurados na modelagem desse contrato o custo esperado para o segurado foi de R\$ 6.041,67. Além disso, nestes gatilhos o custo de não utilizar um modelo para prever as perdas considerando o prêmio pago é de R\$ 9.250,50, havendo, portanto, uma economia de R\$ 3.208,33 com o uso do modelo. Ademais, um modelo que acertasse todas as previsões nos gatilhos ótimos teria um custo de R\$ 1.750,00.

Ressalta-se que para o segurado, à medida que o gatilho para o pagamento das indenizações cresce, a vantagem econômica para o segurado em adquirir o produto diminui. Isso acontece porque os gatilhos são mais altos para veranicos menores e, dessa forma, como

os veranicos menores são mais prováveis de acontecer do que os maiores, há uma probabilidade maior de pagamento de indenizações e, portanto, o custo do produto para o segurado tende a crescer na mesma proporção.

## 5. CONSIDERAÇÕES FINAIS

Nesta pesquisa foram estruturados modelos de seguros paramétricos contra veranicos extremos em fases críticas da produção de café. Encontrou-se relação significativa entre as probabilidades de excedência de veranicos extremos e a produção resultante de café para os ciclos fenológicos R1 e para o período de tempo correspondente a junção dos ciclos R1R2.

Dessa forma, os procedimentos metodológicos propostos neste trabalho podem auxiliar na elaboração de contratos para eventos segurados desta natureza uma vez que atuou-se na definição de gatilhos indenizatórios que compatibilizassem da melhor forma possível a ocorrência das perdas com o pagamento das indenizações nos contratos, reduzindo o risco base. Assim, os gatilhos mais adequados para acionarem os pagamentos das indenizações foram estabelecidos para quando o modelo apontar 44% ou 45% de probabilidade perda na fase R1 e 51% ou 52% no período R1R2, pois estes gatilhos causam maior economia no produto para a seguradora e trazem maior segurança para o segurado.

Ademais, coloca-se que para dentre os modelos ajustados, o modelo de contrato apenas para a fase R1 incorre em custos esperados menores, tanto para a seguradora quanto para o segurado. Isto pode ser reflexo de o modelo ajustado para R1 ter maior acurácia e sensibilidade em relação ao modelo para o período R1R2. Desse modo, sugere-se que seja mais vantajoso, nos moldes desta pesquisa, que o segurado opte por adquirir o produto para o ciclo R1, uma vez que para este ciclo o custo de não se ter um modelo também é maior do que seria para R1R2.

Por fim, esse estudo pode ser extrapolado a um conjunto de municípios maior para que a base de dados possa refletir mais insights no contexto nacional. Além disso, sugere-se para pesquisas futuras a expansão do campo de estudo, com outras culturas como soja, milho ou cacau, além de um modelo que contemple um efeito espacial no processo de precificação.

## REFERÊNCIAS BIBLIOGRÁFICAS

AFSHAR, M. H. et al. Improving the Performance of Index Insurance Using Crop Models and Phenological Monitoring. **Remote Sens.** v. 13, n. 5, p. 924-942, 2021.

ARRUDA, F. B. et al. Estudo da influência do clima e do consumo hídrico na produção de cafeeiros (*Coffea arabica* L.) em Pindorama, SP. **Bragantia**, v. 62, n.1, p. 782-785, 2000.

APARECIDO, L. E. O.; ROLIM, G. S.; SOUZA, P. S. Sensitivity of newly transplanted coffee plants to climatic conditions at altitudes of Minas Gerais, Brazil. **Australian Journal of Crop Science**, Queensland, v.9, n. 2, p. 160-167, 2015.

BARRETO, A. S. **Modelos de Regressão: Teoria e Aplicações com o Programa estatístico R**. 1 ed. Brasília: Edição do Autor, 2011.

BUTTURI-GOMES, D.; BEIJO, L. A.; AVELAR, F. G. On modeling the maximum duration of dry spells: a simulation study under a Bayesian approach. **Theoretical and Applied Climatology**, v. 137, n. 1, p. 1337-1346, 2019.

CAMARGO, M. B. P. de. The impact of climatic variability and climate change on arabic coffee crop in Brazil. **Bragantia**, Campinas, v. 69, p. 239-247, 2010.

CAMARGO, A. P.; CAMARGO, M. B. P. Definição e esquematização das fases fenológicas do cafeeiro arábica nas condições tropicais do Brasil. **Bragantia**, Campinas, v. 60, n. 1, p. 65-68, 2001.

CARTER, M.R., CHIU, T. A **Minimum Quality Standard to Ensure Index Insurance Contracts do no Harm**. Feed the Future Innovation Lab for Assets and Market Access, Davis, 2018.

CEBALLOS, A.; MARTINEZ-FERNÁNDEZ, J.; LUENGO-UGIDOS, M. Á. Analysis of rainfall trends and dry periods on a pluviometric gradient representative of Mediterranean climate in the Duero Basin, Spain. **Journal of Arid Environments**, Elsevier, v. 58, n. 2, p. 215–233, 2004.

Centro de Estudos Avançados em Economia Aplicada (Cepea). Mercado de Trabalho/CEPEA: PIB-AGRO/CEPEA: PIB DO AGRO CRESCE 8,36% EM 2021; PARTICIPAÇÃO NO PIB BRASILEIRO CHEGA A 27,4%. 2022. Disponível em: <<https://bityli.com/TUWgAMY>>. Acesso em 11 ago. 2022.

COLES, S. G. **An Introduction to Statistical Modeling of Extreme Values**. Bristol: Springer, 2001.

COLES, S. G.; POWELL, E. A. Bayesian methods in extreme value modeling: a review and new developments. **International Statistical Review**, v. 64, n. 1, p. 119–136, 1996.

Companhia Nacional de Abastecimento (CONAB). Café: Conjuntura semanal - de 2/8/21-6/8/21. CONAB, 2021. Disponível em: <https://www.conab.gov.br/info-agro/analises-do-mercado-agropecuario-e-extrativista/analises-do-mercado/historico-de-conjunturas-de-cafe>. Acesso em: 22 jul. 2022.

Companhia Nacional de Abastecimento (CONAB). Safra de café 2022 pode chegar a 53,4 milhões de sacas, impactada por clima adverso. ConabCast, 2021. Disponível em: <[https://cast.conab.gov.br/post/2022-05-19\\_2\\_lev\\_de\\_cafe/](https://cast.conab.gov.br/post/2022-05-19_2_lev_de_cafe/)>. Acesso em: 22 jul. 2022.

DA SILVA, R. S.; DO NASCIMENTO, F. F. Extreme Value Theory Applied to  $r$  Largest Order Statistics Under the Bayesian Approach. **Revista Colombiana de Estadística**, v. 42, n. 2, p. 143-166, 2019.

DALHAUS, T.; MUSSHOF, O.; FINGER, R. Phenology information contributes to reduce temporal basis risk in agricultural weather index insurance. **Scientific reports**, v. 8, n. 1, p. 1-10, 2018.

DE OLIVEIRA, L. H. A.; CARVALHO, J. V. F. We are Living on the Edge! Gerenciando Sinistros de Extrema Severidade com o Apoio da Teoria de Valores Extremos. In: XX USP International Conference in Accounting. **Anais eletrônicos**, São Paulo, EAC/FEA/USP, 2020. Disponível em: <https://congressousp.fipecafi.org/anais/Anais2020/ArtigosDownload/1902.pdf>. Acesso em: 22 jun. 2022.

EMBRAPA. **Agrometeorologia dos Cultivos: O fator meteorológico na produção agrícola**. Brasília - DF: Instituto Nacional de Meteorologia (INMET), 2009.

EMBRECHTS, P.; RESNICK, S. I.; SAMORODNITSKY, G. Extreme value theory as a risk management tool. **North American Actuarial Journal**, v. 3, n. 2, p. 30-41, 1999.

FIGUEIREDO, R. et al. A probabilistic paradigm for the parametric insurance of natural hazards. **Risk Analysis**, v. 38, n. 11, p. 2400-2414, 2018.

FERNANDES, A. A. T. et al. Leia este artigo se você quiser aprender regressão logística. **Revista de Sociologia e Política**, v. 28, 2021.

FORNAZIER, A.; DE SOUZA, P. M.; PONCIANO, N. J. A Importância do Seguro Rural na Redução de Riscos da Agropecuária. **Revista de Estudos Sociais**, v. 14, n. 28, p. 39-52, 2012.

HEIMFARTH, L. E.; MUSSHOF, O. Weather index-based insurances for farmers in the North China Plain: An analysis of risk reduction potential and basis risk. **Agricultural Finance Review**, v. 71, n. 2, p. 218-239, 2011.

JENKINSON, A. F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. **Quarterly Journal of the Royal Meteorological Society**, v. 81, n. 348, p. 158-171, Apr. 1955.

JOHNSON, L. Paying ex gratia: Parametric insurance after calculative devices fail. **Geoforum**, v. 125, p. 120-131, 2021.

LJUNG, G. M.; BOX, G. E. On a measure of lack of fit in time series models. **Biometrika**, v. 65, n. 2, p. 297-303, 1978.

MANN, H. B. Nonparametric tests against trend. *Econometrica*: **Journal of the Econometric Society**, JSTOR, p. 245-259, 1945.

MARTINS, E. et al. Influência das condições climáticas na produtividade e qualidade do cafeeiro produzido na região do sul de Minas Gerais. **Coffee Science**, Lavras, v. 10, n. 4, p. 499 - 506, 2015.

MILLER. **Parametric Insurance Solutions**: Guaranteed, objective and transparent protection. Londres, 2020. Disponível em: <[https://www.miller-insurance.com/-/media/Bulletins/NA\\_bulletin/Nov-20/Parametric-insurance-solutions-from-Miller.pdf](https://www.miller-insurance.com/-/media/Bulletins/NA_bulletin/Nov-20/Parametric-insurance-solutions-from-Miller.pdf)> . Acesso em 02 ago. 2022.

NEWE Seguros. Seguro paramétrico terá subvenção a partir de 2021 e promete ser mais barato. NEWE Seguros. 2 nov. 2020. Disponível em: <<https://neweseguros.com.br/seguro-parametrico-tera-subvencao-a-partir-de-2021-e-promete-ser-mais-barato/>>. Acesso em: 08 de setembro de 2021.

NOGUEIRA, D.; SÁFADI, T.; FERREIRA, D. Avaliação de critérios de convergência univariados para o método de Monte Carlo via Cadeias de Markov. *Revista Brasileira de Estatística*, v. 65, n. 224, 2004.

OZAKI, V. A. O papel do seguro na gestão do risco agrícola e os empecilhos para o seu desenvolvimento. **Revista Brasileira de Risco e Seguro**, v. 2, n. 4, p. 75-92, 2007.

PICINI, A. G. et al. Desenvolvimento e teste de modelos agrometeorológicos para a estimativa de produtividade do cafeeiro. **Bragantia**, Campinas, v. 58, n. 1, p. 157-170, 1999.

PROKOPCHUK, O. et al. Parametric insurance as innovative development factor of the agricultural sector of economy. **AGRIS on-line Papers in Economics and Informatics**, v. 12, n. 3, p. 69-86, 2020.

R Core Team. R: uma linguagem e ambiente para computação estatística, 2017.

RADU, N.; ALEXANDRU, F. Parametric Insurance: A Possible and Necessary Solution to Insure the Earthquake Risk of Romania. **Risks**, v. 10, n. 3, p. 59-75, 2022.

ROCCO. M. Extreme Value Theory in Finance: A Survey. **Journal of Economic Surveys**, Malden, v. 28, n. 1, p.82-108, 2014.

RUÍZ-CÁRDENAS, R. et al. A cafeicultura e sua relação com o clima. 2015. Disponível em: <https://toolbox.coffeeandclimate.org/wp-content/uploads/Folheto-Cartilha-cafe-yclima-Brasil.pdf>. Acesso em: 25 jul. 2022

SHIRSATH, P. et al. Designing weather index insurance of crops for the increased satisfaction of farmers, industry and the government. **Climate Risk Management**, v. 25, n.1, p. 100189-100201, 2019.

TURNER, S. Parametric insurance Policies: Do they have to be weather-related? Armic: White paper, 2020. Disponível em: <<https://www.airmic.com/sites/default/files/Stuart-Turner-Risk-Leadership-Dissertation-White-Paper.pdf>> . Acesso em 01 ago. 2022.

WALENDORFF, R. Seguro paramétrico inédito protege cacau no sul da Bahia. **Valor Econômico**. 17 ago. 2021. Disponível em: <<https://valor.globo.com/agronegocios/noticia/2021/08/17/seguro-parametrico-inedito-protege-cacau-no-sul-da-bahia.ghtml>>. Acesso em 01 ago. 2022.

WORLD BANK. Index insurance is having a development impact where it's needed most. **World Bank Blogs**. 16 ago. 2017. Disponível em: <<https://blogs.worldbank.org/voices/index-insurance-having-development-impact-where-it-s-needed-most>> . Acesso em 01 ago. 2022.

ZURICH. Zurich Construction Weather Parametric Insurance Protecting your profits from extreme climate events. 2021. Disponível em: <<https://bityli.com/qAzzoW>>. Acesso em 08 de setembro de 2021.