



Universidade de Brasília
Departamento de Estatística

Modelos mistos aplicados à recomendação de cultivares no contexto da
ambientômica

César Augusto Galvão

Relatório apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

Brasília
2024

César Augusto Galvão

**Modelos mistos aplicados à recomendação de cultivares no contexto da
ambientômica**

Orientador(a): Prof. Dr. Leandro T. Correia
Coorientador(a): Prof. Dr. Rafael T. Tassinari

Relatório apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília
2024**

Dedico este trabalho (pelo esforço e reflexão, apesar de qualquer qualidade) ao meu pai, cientista e professor que me ensina sobre tudo a todo momento, mesmo que não planeje.

Agradecimentos

- À Universidade de Brasília, a qual insisto em ocupar desde 2007, e ao Prof. Dr. Leandro Correia por encarar o desafio de trabalhar comigo neste TCC;
- Ao Laboratório de Melhoramento de Precisão da Universidade Federal do Goiás;
- À Escola Nacional de Administração Pública, sem o apoio da qual eu não teria concluído o curso;
- À Embrapa Arroz e Feijão e em especial ao Dr. Flávio Breseghello por ceder o banco de dados para nossas análises;
- Aos funcionários do restaurante universitário, o qual frequentei incontáveis vezes, em especial às trabalhadoras Viviane e Luciana (das quais eu infelizmente não conheço o sobrenome);
- Ao meu pai, Dr. Carlos Galvão, por sempre encorajar minhas ideias megalomaniacas e por me mostrar desde criança que ciência é algo belo, nobre e digno de esforço humano;
- À Gabriela Carneiro, que me acompanhou em risos e misérias nos últimos anos e se tornou uma pessoa essencial;
- Aos meus amigos que fiz durante o curso, que me aguentaram e me deram apoio inúmeras vezes: Marco Bahia, João Vasconcelos, e demais, sem perda de generalidade;
- A mim mesmo por não ter desistido.

Resumo

O estudo conduzido teve como objetivo comparar metodologias de redução de dimensionalidade para ajuste de modelos mistos aplicados ao contexto da ambientômica visando a recomendação de cultivares utilizando a base de dados *Embrapa Rice Breeding Dataset*. Na metodologia foram utilizadas duas formas de redução de dimensionalidade, geração de marcadores ambientômicos via algoritmo com modelo linear e análise de componentes principais. No ajuste dos modelos, uma quantidade inferior à desejada foi utilizada devido a limitações computacionais e de identificabilidade dos modelos. Os resultados indicam uma similaridade de desempenho dos modelos ajustados. Em conclusão, o estudo apontou a necessidade de técnicas mais refinadas de redução de dimensionalidade das covariáveis ambientais e de inclusão de estruturas de dependência dos dados para gerar modelos com maior capacidade preditiva.

Palavras-chaves: modelos mistos, ambientômica, redução de dimensionalidade.

Lista de Tabelas

1	Medidas resumo de produtividade por região	22
2	Coincidência de genótipos por região	23
3	Aproveitamento das simulações de geração de marcador ambientômico	24
4	Cargas de rotação máxima e mínima para as primeiras 10 componentes principais	26
5	Interceptos e correlação intraclasse para os modelos com floresta aleatória e com componentes principais	27
6	Percentual de cobertura dos intervalos preditivos com modelos de componentes principais e de marcadores ambientômicos para os genótipos da etapa de validação, discriminado por macrorregião	34
7	Cobertura média e amplitude média dos intervalos de previsão para os modelos PCA e de Marcadores	35
8	Razão entre MSE médio da validação cruzada <i>5-fold</i> e o MSE da amostra completa para os modelos PCA e de Marcadores considerando validações a nível de genótipo e a nível de unidade observacional	36
9	Dicionário de variáveis do ERBD relacionadas aos ensaios	41
10	Dicionário de variáveis do ERBD das unidades experimentais	42
11	Dicionário de variáveis ambientais	43

Lista de Figuras

1	Mapa do Brasil com indicação das localidades com experimento da base ERBD após limpeza dos dados	21
2	Histograma da variável resposta - Rendimento de grãos (GY) em kg ha^{-1} .	22
3	Médias e $\text{IC}(\alpha = 0,05)$ da produtividade por genótipo	23
4	Frequência das covariáveis em todas as simulações	24
5	Frequência das covariáveis em simulações com $r_{gg} > 0,5$	24
6	Barras de frequência relativa de marcadores ambientais aceitos em todas as simulações	25
7	Proporção acumulada da variância explicada pelas componentes principais	25
8	Derivada da curva de variância explicada pelas componentes principais . .	25
9	Histograma das cargas do PCA	26
10	Carga de rotação máxima das covariáveis ambientais para as componentes principais em valor absoluto	27
11	Estimadores pontuais e $\text{IC}(\gamma = 0,95)$ dos coeficientes das componentes principais 1, 6, 9 e 10	28
12	Estimadores pontuais e $\text{IC}(\gamma = 0,95)$ dos coeficientes dos marcadores ambientômicos 1, 6, 9 e 10	29
13	Correlograma para marcadores ambientômicos utilizados no modelo	32
14	Valores reais e ajustados para o modelo misto ajustado com PCA	33
15	Valores reais e ajustados para o modelo misto ajustado com marcadores ambientômicos	33
16	Valores ajustados e resíduos studentizados para o modelo misto ajustado com PCA	33
17	Valores ajustados e resíduos studentizados para o modelo misto ajustado com marcadores ambientômicos	33
18	Mapa do Brasil com indicação das localidades com experimento no ano de 2022 da base ERBD após limpeza dos dados	34

Sumário

1 Introdução	8
1.1 Objetivos	9
2 Referencial Teórico	10
2.1 Redução de dimensionalidade	10
2.2 Modelos Lineares	12
2.3 Modelos Lineares Hierárquicos	13
3 Metodologia	17
3.1 Conjuntos de dados	17
3.2 Software	18
3.3 Geração de Marcadores Ambientômicos	18
3.4 Redução de dimensionalidade	19
3.5 Modelos Mistos	20
3.6 Validação cruzada	20
4 Resultados	21
4.1 Análise Exploratória	21
4.2 Geração de marcadores ambientômicos	24
4.3 Redução de dimensionalidade	25
4.4 Modelos mistos	27
5 Considerações finais	37
Referências	39
Apêndice	41

1 Introdução

A domesticação de espécies silvestres de plantas para a agricultura é uma prática antiga e passou por diversas revoluções até os dias atuais, em que a genética biométrica e o melhoramento de precisão protagonizam a criação de cultivares e seleção de características de interesse (RESENDE; BRONDANI; CHAVES, 2023). Além disso, pressões como crescimento populacional (HICKEY et al., 2019), redução de recursos naturais disponíveis, aquecimento global e uma variedade de consequências desses fatores (JORASCH, 2019) aumentam a necessidade de se produzir alimentos e outros recursos vegetais de forma incrementalmente eficiente. Uma das soluções para isso é justamente o melhoramento de precisão.

Neste contexto, o desenvolvimento e seleção de cultivares é associado a identificação de grupos ambientais (*Target Population of Environments* ou TPE), permitindo que se aproveite ao máximo a característica de interesse (CHENU, 2015). De fato, em posse da informação de que o ambiente em que a planta se desenvolve interfere em seu fenótipo (a característica de interesse, que é uma expressão gênica), cabe estudar a interação de genótipos em seus ambientes ($G \times E$).

O estudo desse tipo de relação é potencializado com o uso de técnicas de Sistemas de Informações Geográficas – SIG, como sensoriamento remoto, entre outros (RESENDE; BRONDANI; CHAVES, 2023). A disponibilização pública de dados coletados via satélite com diversos graus de granularidade permite a inclusão de mais covariáveis ambientais como área cultivada, cobertura vegetal, temperatura, entre outros dados geofísicos¹.

A proposta de Resende et al. (2021), que será usada de estudo de caso, é expandir o uso de TPE para um estudo ômico do ambiente, daí *ambientômica*. Os autores propõem o uso de modelos hierárquicos, e o conceito de ambientipagem, resultante de agrupamentos ambientais, para predição de performance de genótipos não observados. Isto permite, por exemplo, recomendar o melhor genótipo de arroz para uma região em que jamais foi cultivado e assim tornar a região produtiva.

De forma semelhante, Demidenko (2013) propõe modelos mistos como uma categoria adequada para descrever variedades individuais dentro de uma categoria biológica. Essa abordagem assume a existência de um *cluster* com observações correlacionadas – ou seja, parâmetros específicos de um grupo – além dos parâmetros médios populacionais. No contexto deste trabalho, isso se refere ao comportamento médio do arroz cultivado

¹Por exemplo, o serviço Google Earth Engine disponibiliza seu catálogo em <https://developers.google.com/earth-engine/datasets/>

e suas variações individuais, ou seja, de cada genótipo. Nesse paralelo, o genótipo atua como a variável definidora de cada *cluster*.

Pretende-se, portanto, estudar a produtividade de genótipos de arroz expostos a condições ambientais diversas. Serão utilizados dados reais tanto de experimentos agronômicos quanto de condições ambientais e a relação entre os genótipos e o ambiente serão exploradas mediante o uso de modelos lineares mistos.

1.1 Objetivos

O objetivo geral deste trabalho de conclusão de curso é investigar o uso de modelos lineares hierárquicos (ou mistos) na recomendação de genótipos de um cultivar específico em uma região delimitada e ambientalmente caracterizada, ou seja, com a inclusão de dados sobre o maior número possível de características ambientais. Busca-se revisar metodologicamente o estudo de Resende et al. (2021), detalhando o processo de modelagem e sua adequação, além de comparar computacionalmente diferentes variações do modelo empregado. O foco principal deste trabalho é a exploração metodológica para a construção do modelo misto, que, em uma etapa posterior, resultará na recomendação de genótipos para uma determinada região. No entanto, essa recomendação não está incluída no escopo deste trabalho.

Os objetivos específicos são:

- Explorar a técnica de modelagem estatística via modelos lineares hierárquicos incluindo efeitos aleatórios;
- Explorar os conceitos necessários para aplicação do modelo ao contexto de melhoramento de plantas e ambientômica;
- Comparar a adequação do modelo original de Resende et al. (2021) com um modelo que faça composição de marcadores ambientômicos utilizando técnicas de redução de dimensionalidade.

2 Referencial Teórico

2.1 Redução de dimensionalidade

Uma forma geométrica de ilustrar dificuldades de se trabalhar matematicamente em altas dimensões é comparar volumes de objetos, como o volume de uma hipersfera ($V_e(r)$) inscrita em um hipercubo ($V_c(2r)$), ambos em \mathbb{R}^d , com raio r e arestas $2r$. Como se pode notar pelas equações a seguir, a proporção do volume do hipercubo fora da hipersfera tende a 1 quando se toma o limite $d \rightarrow \infty$:

$$\frac{V_c(2r) - V_e(r)}{V_c(2r)} = \frac{(2r)^d - \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}r^d}{(2r)^d} = 1 - \frac{\pi^{d/2}}{2^d \Gamma(\frac{d}{2} + 1)}$$

e

$$\lim_{d \rightarrow +\infty} \frac{\pi^{d/2}}{2^d \Gamma(\frac{d}{2} + 1)} = 0.$$

Esta dificuldade recebeu o nome de “*curse of dimensionality*”, termo cunhado por Bellman e Kalaba (1959 apud CHEN, 2009) e descreve o crescimento de observações necessárias para se estimar com acurácia uma função arbitrária dado um aumento da dimensionalidade da função.

De acordo com Morettin e Singer (2022), técnicas de redução de dimensionalidade são utilizadas quando há um grande número de covariáveis e se deseja estudar as unidades observadas com base em sua estrutura de dependência multivariada. Dentre as técnicas disponíveis, foi escolhida a Análise de Componentes Principais (*PCA — Principal Component Analysis*), que busca expressar as covariáveis originais em termos de outras variáveis não correlacionadas e que resumam as informações contidas em um conjunto de variáveis. Em outras palavras, “as p covariáveis originais (X_1, \dots, X_p) são transformadas em p componentes principais não correlacionadas (Y_1, \dots, Y_p) de modo que Y_1 é aquela que explica a maior parcela da variabilidade total dos dados originais, Y_2 explica a segunda maior parcela e assim por diante” (ARTES; BARROSO, 2023).

Considerando \mathbf{x} o vetor de covariáveis originais $\mathbf{x}^\top = (X_1, \dots, X_p)$ e $\text{Cov}(x) = \Sigma$, as componentes principais são obtidas pela decomposição espectral da matriz de co-

variâncias: $\Sigma = \Gamma \Lambda \Gamma^\top$. Dessa forma Λ é uma matriz diagonal dos autovalores e Γ é uma matriz ortogonal (portanto $\Gamma \Gamma^\top = \mathbf{I}$) dos autovetores de Σ . Obtém-se daí os pares de autovalores e autovetores ortogonais normalizados $(\lambda_i, \boldsymbol{\alpha}_i)$, $i = 1, \dots, p$, ordenados de modo que $\lambda_1 \geq \dots \geq \lambda_p \geq 0$. Dessa forma, pode-se expressar a i -ésima componente principal por

$$Y_i = \boldsymbol{\alpha}_i^\top \mathbf{x} = \alpha_{i1}X_1 + \dots + \alpha_{ip}X_p, \quad i = 1, \dots, p. \quad (2.1.1)$$

Caso as covariáveis tenham variância ou valores em escalas muito discrepantes, é possível normalizá-las e obter as componentes principais a partir da matriz de correlações $\boldsymbol{\rho}$ e seus pares de autovalores e autovetores $(\gamma_i, \boldsymbol{\epsilon}_i)$. As covariáveis são padronizadas utilizando

$$\mathbf{z} = (\mathbf{V}^{1/2})^{-1} (\mathbf{x} - \boldsymbol{\mu}), \quad (2.1.2)$$

em que $\mathbf{V}^{1/2} = \text{diag}(\sigma_1, \dots, \sigma_p)$ e $\boldsymbol{\mu}$ é o vetor de médias. Agora tem-se $\text{Cov}(\mathbf{z}) = \boldsymbol{\rho}$, a matriz de correlações para as covariáveis padronizadas. Finalmente, as componentes são dadas por

$$Y_i = \boldsymbol{\epsilon}_i^\top (\mathbf{V}^{1/2})^{-1} (\mathbf{x} - \boldsymbol{\mu}). \quad (2.1.3)$$

A obtenção de componentes principais para o sistema garante algumas propriedades interessantes como

$$\sum_{j=1}^p \text{Var}(X_j) = \text{tr}(\Sigma) = \text{tr}(\Gamma \Lambda \Gamma^\top) = \text{tr}(\Lambda) = \sum_{i=1}^p \text{Var}(Y_i). \quad (2.1.4)$$

Ou seja, a variância total do sistema é mantida após obtenção das componentes principais, o que também vale quando se utiliza a matriz de correlações.

Efetivamente o que será utilizado serão os estimadores amostrais para as componentes principais, para as matrizes de covariância e correlação e seus pares de autovalores e autovetores; \hat{Y}_i , \mathbf{S} , \mathbf{R} e $(\hat{\gamma}_i, \hat{\boldsymbol{\epsilon}}_i)$ respectivamente.

Quando o intuito da obtenção das componentes principais é reduzir a dimensio-

nalidade do sistema e assim manter uma quantidade menor de variáveis que explique uma parcela razoável da variância do sistema, deve-se adotar um critério de determinação do número de componentes principais a serem retidas. Artes e Barroso (2023) apresentam:

- Critério de Kaiser, segundo a qual se mantêm as componentes com autovalores superiores a 1 — ou seja, componentes que explicariam mais variância do que uma covariável original individual;
- Reter componentes que acumulem ao menos uma certa percentagem de variância explicada do total;
- Reter componentes que acumulem ao menos uma certa percentagem da variância explicada de cada variável original.

Morettin e Singer (2022) apresentam ainda a seleção baseada no fator de aceleração do teste do cotovelo (utilização do *screeplot*). Considerando $af(i) = f''(i) = f(i+1) - 2f(i) - f(i-1)$, $i = 1, \dots, p-1$ o fator de aceleração, o número de fatores retidos corresponde à posição anterior em que $af(i)$ é máximo.

2.2 Modelos Lineares

Modelos Lineares apresentam uma relação estocástica entre duas ou mais variáveis. Sua forma simples com efeitos fixos pode ser representada da forma

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (2.2.1)$$

em que \mathbf{y} é o vetor da variável resposta, $\boldsymbol{\beta}$ é o vetor de p coeficientes associados às covariáveis $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_p)$ e $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}\sigma^2)$ é o vetor de erros estocásticos associado às observações. Ao final do processo de modelagem obtém-se um modelo da forma

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}, \quad (2.2.2)$$

em que $\hat{\boldsymbol{\beta}}$ são estimadores obtidos pelo método de mínimos quadrados ordinários ou máxima verossimilhança para $\boldsymbol{\beta}$ (KUTNER et al., 2005).

Este modelo de regressão, também conhecido como *modelo de efeitos fixos*, requer uma série de suposições sobre a componente aleatória, que são avaliadas durante a etapa de diagnóstico da modelagem. Entre essas suposições, destacam-se a heterocedasticidade, independência, posto completo da matriz \mathbf{X} e distribuição Normal. Quando uma ou mais dessas suposições não são atendidas, ou se houver colinearidade entre as covariáveis, pontos de alavancagem elevados, ou outros problemas que comprometam o modelo, são adotadas medidas corretivas, como transformações de dados ou redução de dimensionalidade. Essas abordagens, bem como a inclusão de variáveis indicativas de grupos ou dados categóricos, frequentemente aumentam a complexidade da interpretação do modelo, exigindo tratamentos inferenciais diferenciados (HOX; MOERBEEK; SCHOOT, 2017).

2.3 Modelos Lineares Hierárquicos

Pesquisas em diversos domínios do conhecimento frequentemente investigam fenômenos cujas unidades de análise são agregadas em categorias distintas (ADEWALE et al., 2007; MCMAHON; DIEZ, 2007). Esses diferentes níveis de análise, sejam indivíduos ou grupos, apresentam características que requerem formas específicas de representação e técnicas de inferência capazes de lidar adequadamente com as estruturas de covariância envolvidas.

Modelos hierárquicos, também conhecidos como modelos mistos, substituem duas práticas comuns na aplicação de regressões lineares: a transformação de variáveis categóricas em variáveis binárias (*dummy*) e a planificação do nível de análise, ou seja, o uso direto de medidas de grupos como descritores da unidade de análise. A adoção de um modelo misto permite a construção de estimadores que evitam essas práticas, proporcionando uma melhor representação dos indivíduos dentro de seus agrupamentos e no contexto de suas características (HOX; MOERBEEK; SCHOOT, 2017; GELMAN; HILL, 2006). Esse tipo de modelo linear pode ser representado da forma

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}, \quad (2.3.1)$$

em que \mathbf{y} é o vetor de N observações, $\mathbf{X}_{N \times p}$ é a matriz de incidência de efeitos fixos conhecida, $\boldsymbol{\beta}_{p \times 1}$ é o vetor de efeitos fixos, $\mathbf{Z}_{N \times k}$ é a matriz de incidência de efeitos aleatórios conhecida, $\mathbf{u}_{k \times 1}$ é o vetor de efeitos aleatórios e $\boldsymbol{\varepsilon}_{N \times 1}$ é o vetor de erros aleatórios (MARTINS et al., 1993b).

O modelo pode ser mais facilmente comparado à regressão linear se for reescrito com apenas um termo estocástico (DEMIDENKO, 2013),

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\eta}, \quad (2.3.2)$$

onde

$$\boldsymbol{\eta} = \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}. \quad (2.3.3)$$

Pressupõe-se as seguintes distribuições para \mathbf{y} , \mathbf{u} e $\boldsymbol{\varepsilon}$:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} \sim \mathcal{N} \left\{ \begin{bmatrix} \mathbf{X}\boldsymbol{\beta} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{Z}\mathbf{U}\mathbf{Z}^\top + \mathbf{R} & \mathbf{Z}\mathbf{U} & \mathbf{R} \\ \mathbf{U}\mathbf{Z}^\top & \mathbf{U} & \mathbf{0} \\ \mathbf{R} & \mathbf{0} & \mathbf{R} \end{bmatrix} \right\}, \quad (2.3.4)$$

em que $\mathbf{0}$ são matrizes nulas, \mathbf{U} é a matriz de covariâncias dos efeitos aleatórios \mathbf{u} e \mathbf{R} é a matriz de covariâncias residual.

A matriz \mathbf{U} pode ser ainda decomposta em termos da matriz de correlação $N \times N$ entre os efeitos aleatórios \mathbf{A} e a matriz \mathbf{U}_0 , a matriz de covariância $k \times k$ entre os efeitos aleatórios das q covariáveis de efeito aleatório (MARTINS et al., 1993a). A matriz portanto pode ser expressa como $\mathbf{U} = \mathbf{A} \otimes \mathbf{U}_0$ e, como não serão consideradas parentalidades neste estudo, $\mathbf{U} = \mathbf{I} \otimes \mathbf{U}_0$. Da mesma forma, a matriz de \mathbf{R} é decomposta em $\mathbf{R} = \mathbf{I} \otimes \mathbf{R}_0$.

Soluções de estimação para o modelo 2.3.4 podem ser obtidas a partir da função densidade de probabilidade conjunta

$$f(\mathbf{y}, \mathbf{u}) = \frac{1}{(2\pi)^{n/2} |\mathbf{R}|^{1/2}} \exp \{ (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u})^\top \mathbf{R}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) \}, \quad (2.3.5)$$

e tomando as soluções homogêneas para as derivadas $\frac{\partial \ell}{\partial \boldsymbol{\beta}}$ e $\frac{\partial \ell}{\partial \mathbf{u}}$ da log-verossimilhança ℓ . Obtém-se assim o sistema de Equações de Modelos Mistos, dadas por

$$\begin{bmatrix} \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{X} & \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{Z} \\ \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{X} & \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Z} + \mathbf{U}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{y} \end{bmatrix}. \quad (2.3.6)$$

A solução para $\boldsymbol{\beta}$ é obtida via Mínimos Quadrados Generalizados (GLS) ignorando \mathbf{u} , obtendo-se

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X})^{-} \mathbf{X}^\top \mathbf{V}^{-1} \mathbf{y}, \quad (2.3.7)$$

sendo \mathbf{X}^{-} a inversa generalizada de \mathbf{X} . A variância dos estimadores de efeitos fixos é dada por

$$\text{Var}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X})^{-} \mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X})^{-}, \quad (2.3.8)$$

que depende de $\text{Var}(\mathbf{y}) = \mathbf{V} = \mathbf{Z} \mathbf{U} \mathbf{Z}^\top + \mathbf{R}$, que é uma matriz particionada, de modo que sua inversa é dada por

$$\mathbf{V}^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1} \mathbf{Z} (\mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Z} + \mathbf{U}^{-1})^{-1} \mathbf{Z}^\top \mathbf{R}^{-1}. \quad (2.3.9)$$

Considerando que $E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$, temos os estimadores de efeitos fixos completamente caracterizados. Como consequência de 2.3.7,

$$\begin{aligned} \hat{\mathbf{u}} &= (\mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Z} + \mathbf{U}^{-1})^{-1} \mathbf{Z}^\top \mathbf{R}^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}) \\ &= \mathbf{U} \mathbf{Z}^\top \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}), \end{aligned} \quad (2.3.10)$$

que é o melhor preditor linear não-viesado (BLUP) para os efeitos aleatórios. A variância desse preditor é dada por

$$\begin{aligned}
\text{Var}(\hat{\mathbf{u}}) &= \text{Var} \left[\mathbf{UZ}^\top \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \right] \\
&= \mathbf{UZ}^\top \mathbf{V}^{-1} \left[\text{Var}(\mathbf{y}) - 2\text{Cov}(\mathbf{y}, \hat{\boldsymbol{\beta}}^\top \mathbf{X}^\top) + \text{Var}(\mathbf{X}\hat{\boldsymbol{\beta}}) \right] \mathbf{V}^{-1} \mathbf{ZU}^\top \\
&= \mathbf{UZ}^\top \left[\mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{V}^{-1} \right] \mathbf{ZU}^\top. \tag{2.3.11}
\end{aligned}$$

A recomendação de um genótipo para uma unidade de área ambientipada é baseada no escore

$$\mathbb{S} = \sum_{k=1}^K \hat{\mathbf{u}}_k. \tag{2.3.12}$$

A correlação intraclasse, estatística que pode ser interpretada como a proporção da variância explicada pela clusterização, é dada por

$$\text{ICC} = \frac{\sigma_0^2}{\sigma_0^2 + \sigma}, \tag{2.3.13}$$

onde σ representa a variância residual e σ_0^2 a soma dos demais componentes de variância.

3 Metodologia

3.1 Conjuntos de dados

O conjunto de dados experimentais utilizado foi o *Embrapa Rice Breeding Dataset* (ERBD), desenvolvido e cedido pela Empresa Brasileira de Pesquisa Agropecuária (Embrapa) para pesquisas com cultivares de arroz (*Oryza sativa L.*). A base de dados cedida compreende coleta de dados desde 1982, mas os dados utilizados neste trabalho compreendem apenas os anos de 1995 a 2022 por recomendação dos pesquisadores da Empresa. A base é extensamente documentada em Breseghello et al. (2011) e Breseghello et al. (2021) e os dicionários desses dados estão disponíveis no Apêndice, Tabelas 9 e 10.

Os pontos de experimento do ERBD são identificados geograficamente por duas variáveis, que indicam a unidade da federação e a localidade em que foram realizados. No entanto, diversas localidades possuem nomes que não correspondem a municípios brasileiros. Em posse das coordenadas de latitude e longitude do experimento, quando disponibilizadas, é possível aproximar os pontos para o centro do município mais próximo.

Para o conjunto de dados ambientais foram coletados dados de 393 covariáveis, das quais 19 são oriundas do repositório WorldClim, 130 do NasaPower e 244 do SoilGrids. Todos os dados ambientais advêm de medições via satélite e foram obtidos mediante consulta às API (*Application Programming Interface*) disponibilizadas pelas organizações. O dicionário de dados contendo a fonte e definição de todas as 393 covariáveis está disponível no Apêndice, tabela 11. As três fontes são descritas a seguir.

A *SoilGrids*² é um projeto da *International Soil Reference and Information Centre* e é um sistema para mapeamento digital do solo que faz previsões para as suas distribuições de forma global. Suas variáveis são dadas em termos de distribuições para seis horizontes de profundidade no solo e alguns de seus quantis.

*NASAPOWER (Prediction Of Worldwide Energy Resources)*³ é um projeto que foi desenvolvido para melhorar ou criar dados de sistemas de satélites e tem três comunidades: energia renovável, construções sustentáveis e agroclimatologia. Os dados utilizados são aqueles destinados ao último público.

Finalmente, *WorldClim*⁴ é uma base de dados com dados de tempo e clima de alta

²<https://www.isric.org/explore/soilgrids>

³<https://power.larc.nasa.gov/>

⁴<https://www.worldclim.org/data/index.html>

resolução. As variáveis são derivadas de valores mensais de precipitação e temperatura, seja uma tendência anual ou de fatores ambientais limitantes ou extremos (em termos de temperatura, por exemplo).

Os dados de satélite dessas três fontes foram coletados em 87.155 pontos em todo o país, mas não correspondem exatamente aos pontos de experimento. Embora existam métodos adequados para estimar toda a superfície do país e, subsequentemente, utilizar essas estimativas nos pontos de experimento (JOURNEL; JOURNEL, 1989), esse não é o foco do estudo. Portanto, o erro de estimação das covariáveis para os pontos de experimento será desconsiderado. Para interpolar as covariáveis nesses pontos, optou-se por calcular a média das covariáveis dos três pontos mais próximos, localizados a menos de meio grau de distância, tanto em latitude quanto em longitude.

3.2 Software

Para todas as análises e construção de gráficos foi utilizada a linguagem R 4.4.0 no ambiente de desenvolvimento RStudio (“Chocolate Cosmos” Release (e4392fc9, 2024-06-05) for Ubuntu Jammy) e os seguintes pacotes principais:

- tidyverse 2.0.0 (WICKHAM, 2016);
- lme4 1.1-35.3 (BATES et al., 2003);
- cv 2.0.0 (FOX; MONETTE, 2023).

Em todos os casos em que foi necessário realizar simulações, a semente para geração de números aleatórios utilizada foi 071989.

3.3 Geração de Marcadores Ambientômicos

Para o ajuste de modelo misto proposto em Resende et al. (2021), é necessário inicialmente gerar marcadores ambientômicos, ou seja, variáveis derivadas das covariáveis ambientais originais que serão usadas como variáveis explicativas. Isso é feito via a técnica de floresta aleatória como uma forma de gerar uma árvore de decisões a partir de amostras de dados da base de dados inicial.

O algoritmo proposto pelos autores foi executado 10.000 vezes para se avaliar a distribuição das covariáveis selecionadas para o modelo, utilizando o critério $r_{gg} = \text{cor}(\hat{\mathbf{y}}, \mathbf{y}) > 0,5$ como a decisão de se utilizar ou não o resultado como marcador. O

algoritmo é executando conforme o Algoritmo 1. Para uma execução em tempo adequado, sugere-se que as simulações sejam paralelizadas.

Algorithm 1 Construção paralelizada de marcadores ambientais

```

1: Inicializar lista vazia lista_marcadores de comprimento 10000
2: Ajustar semente para reprodutibilidade
3: Inicializar laço paralelo sobre  $j = 1$  a 10.000:
4: for  $j = 1$  to 10.000 em paralelo do
5:   Initialize lists fitness and vars
6:   Laço interno de construção dos marcadores:
7:   for  $i = 1$  a 150 do
8:     Amostrar número de genótipos e covariáveis ambientais  $n_{\text{samp}}$  de  $[2, 10]$ 
9:      $boots \leftarrow$  observações dos  $n_{\text{samp}}$  genótipos e  $n_{\text{samp}}$  covariáveis
10:    Dividir dados em conjunto de treinamento e validação:
11:     $poptr \leftarrow$  amostra aleatória simples com metade das observações de  $boots$ 
12:     $popvl \leftarrow$  metade remanescente de  $boots$ 
13:    Ajuste de modelo linear em  $poptr$  com covariáveis ambientais sorteadas
14:    Calcular  $r_{gg} = \text{cor}(\hat{\mathbf{y}}, \mathbf{y})$  em  $popvl$ 
15:     $\text{fitness}[i] \leftarrow r_{gg}$ 
16:     $\text{vars}[i] \leftarrow$  nome das covariáveis ambientais usadas
17:   end for
18:   lista_marcadores[ $j$ ]  $\leftarrow$  {fitness, vars}
19: end for

```

3.4 Redução de dimensionalidade

A redução de dimensionalidade foi executada conforme descrito na seção 2.1, utilizando a função *prcomp* do pacote *stats*, com argumentos definidos para padronização das covariáveis antes de realizar as decomposições de matriz. Dessa forma, elimina-se problemas advindo de ordens de grandeza diferentes entre as variáveis no momento de se avaliar as importâncias dos componentes principais gerados.

A seleção de componentes principais a serem utilizadas no modelo foi feita considerando três aspectos: quantidade de variância explicada, capacidade computacional e identificabilidade do modelo. Enquanto se desejou

3.5 Modelos Mistos

Os modelos mistos foram ajustados utilizando a função *lmer* do pacote *lme4* com fórmulas no formato

$$GY \sim 1 + (1 + \text{variáveis}|GEN), \quad (3.5.1)$$

em que GY representa a variável de interesse correspondente à produtividade (*grain yield*) e GEN representa a variável de agrupamento, os genótipos. Segundo Bates et al. (2015), a fórmula em (3.5.1) ajusta um modelo linear com uma média fixa e intercepto e coeficientes aleatórios, os quais se justificam aleatórios por serem efeitos $G \times E$ independentes, condicionados ao genótipo. Dessa forma, serão estimados os u_i da Equação (2.3.4), o critério de seleção dos genótipos para um determinado local. Este é considerado o efeito genético aditivo, que é o que se deseja estimar no contexto de melhoramento (MARTINS et al., 1993a).

Os modelos lineares mistos ajustados pelo pacote *lme4* têm seus estimadores aproximados por um algoritmo de mínimos quadrados penalizados, balizado por critérios de máxima verossimilhança restrita (BATES et al., 2015).

3.6 Validação cruzada

Cross-validation é uma técnica de avaliação de performance de modelos por reamostragem (SCRUCCA et al., 2023). O esquema de validação mais comum e usado neste trabalho é o *V-fold*, caso em que a base de dados de que se deseja realizar reamostragens para avaliar o modelo é dividida em V partições. Em seguida $V - 1$ partições são utilizadas para treinar o modelo, enquanto a partição restante é utilizada como base de validação.

A medida de perda utilizada para avaliar o ajuste do modelo é aferida na partição de validação. Frequentemente é utilizado o erro quadrático médio em cada rodada de validação e o resultado final da validação pode ser a média dos erros.

4 Resultados

4.1 Análise Exploratória

Para a utilização da base de dados ERBD, algumas variáveis tiveram que ser tratadas ou descartadas. O tratamento inicial envolveu, conforme a seção anterior, a inclusão apenas de arroz cultivado em terras altas (“*upland*”) a partir do ano de 1995. Além disso, por uma questão de uniformidade das unidades de análise, foram selecionados apenas os experimentos que foram realizados sob desenho de bloco completamente aleatorizado – na base de dados, “*RCBD* ou *Randomized Complete Block Design*”. Além disso, casos em que a produção, unidade da federação ou local do experimento exibiam dados faltantes, foram removidos. Variáveis ambientais com medidas todas iguais a zero também foram removidas. Finalmente, três genótipos foram eliminados da base de dados por apresentarem apenas uma observação.

A base de dados limpa e considerada para análise consistiu em 73.724 observações, 690 genótipos e 387 covariáveis ambientais. Essas observações ocorreram em 150 locais diferentes em 18 estados brasileiros distribuídos em todas as macrorregiões, mas poucas ocorrências na faixa leste Nordeste. A Figura 1 a seguir ilustra as ocorrências de experimento:

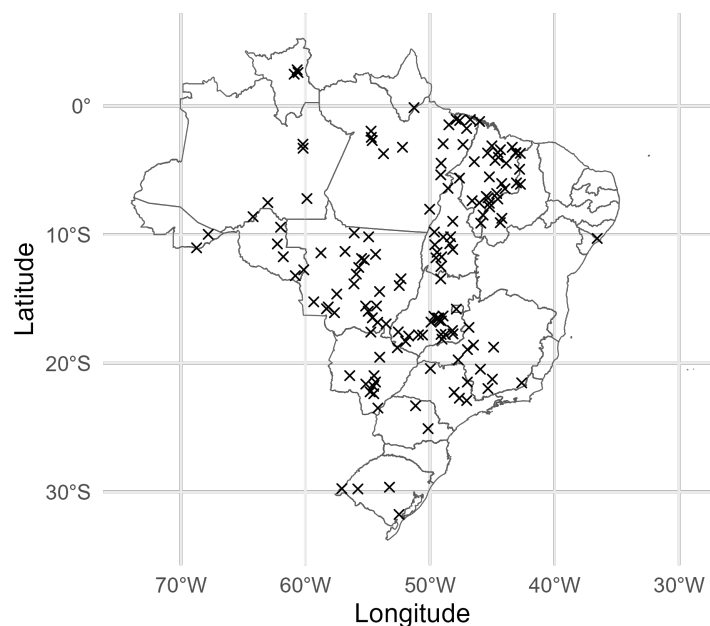


Figura 1: Mapa do Brasil com indicação das localidades com experimento da base ERBD após limpeza dos dados.

A variável resposta – produtividade, cujo nome na base de dados ERBD é GY para *Grain Yield* –, é distribuída conforme a Figura 2 a seguir, com média equivalente a 3408,23 kg/ha. Enquanto a maior parte das observações está concentrada em torno da média, há ainda uma assimetria positiva indicando presença de valores extremos.

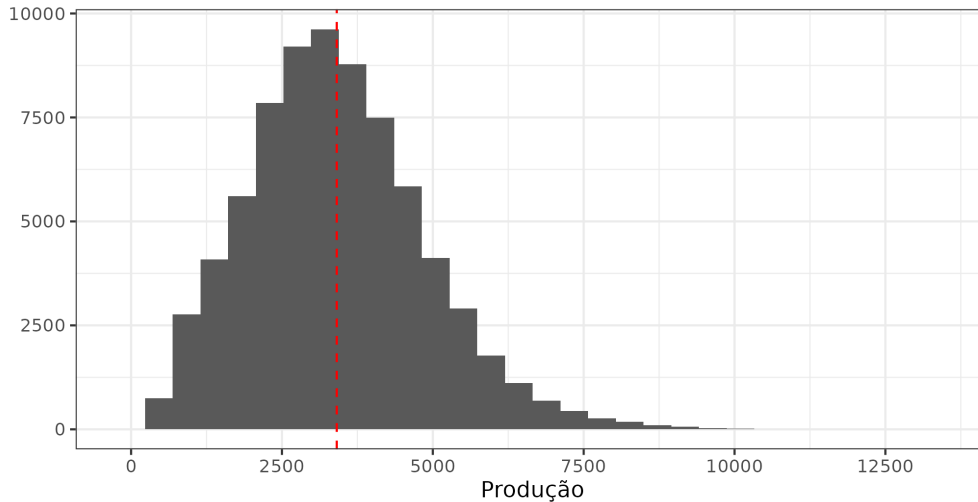


Figura 2: Histograma da variável resposta - Rendimento de grãos (GY) em kg ha^{-1} .

Se observada por região, a produtividade dos genótipos aparenta distribuição similar, exceto pelos máximos das regiões Norte, Nordeste e Centro Oeste. Isso pode ser parcialmente explicado pela quantidade de genótipos testados exclusivamente em cada região: 15, 16, e 84 respectivamente. No entanto, a região Sudoeste também apresenta 16 genótipos exclusivos, enquanto a região Sul apresenta apenas 2.

Tabela 1: Medidas resumo de produtividade por região.

Região	Mín	Q1	Med	Q3	Máx	DP	Média
Norte	261	2.502	3.311	4.263	12.067	1.373,02	3.439,82
Nordeste	302	2.533	3.380	4.265	10.075	1.329	3.448,71
Centro Oeste	200	2.188	3.234	4.334	13.512	1.568,63	3.358,76
Sudeste	270	2620	3.447	4.458	9.562	1.384,12	3.576,19
Sul	355	2500	3.500	4.406	8.250	1.500,20	3.453,87

Pode-se observar ainda que, conforme a Tabela 2, há uma alta coincidência de genótipos experimentados em todas as regiões do país, o que indica que os genótipos de arroz foram expostos a uma variedade de ambientes. Por exemplo, 70% dos genótipos do Sudeste ocorreram também no Nordeste e praticamente todos os genótipos experimentados na região Sul ocorreram em todas as demais regiões. A região Centro Oeste se

destaca com a quantidade de genótipos experimentados, porém o motivo é claro: trata-se da região de maior atuação da Embrapa Arroz e Feijão, executora dos experimentos.

Tabela 2: Coincidência de genótipos por região.

	Centro Oeste	Nordeste	Norte	Sudeste	Sul
Centro Oeste	628 (1,00)				
Nordeste	445 (0,93)	474 (1,00)			
Norte	534 (0,95)	450 (0,80)	562 (1,00)		
Sudeste	101 (0,86)	82 (0,70)	98 (0,83)	117 (1,00)	
Sul	52 (0,96)	51 (0,94)	52 (0,96)	43 (0,79)	54 (1,00)

Finalmente, é interessante verificar se há genótipos com algum comportamento anômalo. Na Figura 3 são projetados os intervalos de confiança de 95% para a produtividade dos genótipos em torno das suas médias e pode-se observar que há casos em que o intervalo não captura a média geral, representada por uma linha pontilhada vermelha. Estes são 42 casos (6%) com 4 ou 3 repetições experimentais e que ocorrem exclusivamente no Centro Oeste ou no Norte ou no Nordeste. Estes genótipos estarão sob maior efeito de *shrinkage*⁵.

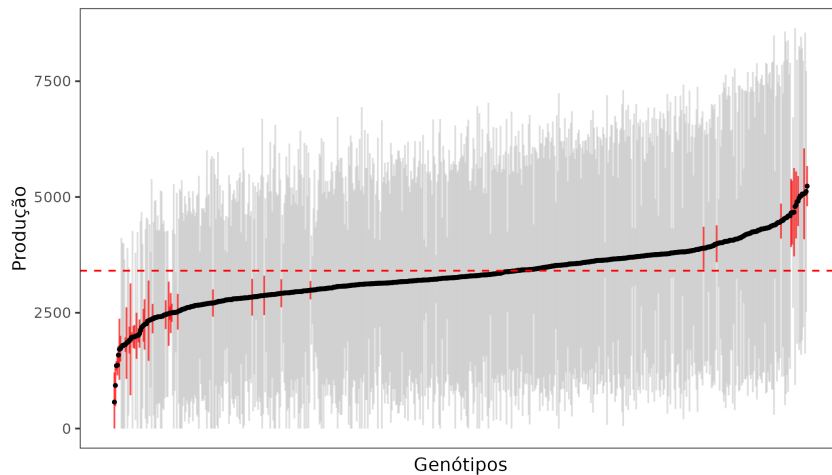


Figura 3: Médias e IC($\alpha = 0,05$) da produtividade por genótipo.

⁵Também conhecido como *partial pooling* (GELMAN; HILL, 2006), clusters com menos observações tendem a se aproximar da média geral, enquanto aqueles com mais observações tendem a se distinguir.

4.2 Geração de marcadores ambientômicos

Foram realizadas 10.000 simulações de geração de marcadores ambientais utilizando o Algoritmo (1) e em seguida foi contada a frequência com que as variáveis ambientais surgem na geração dos marcadores em todos os modelos (Figura 4) e naqueles em que $r_{gg} = \text{cor}(\hat{y}, y) > 0,5$, ou seja, que seriam selecionados para compor covariáveis no modelo misto (Figura 5). Nesta figura, a linha pontilhada vermelha representa a média da frequência com que as covariáveis selecionadas aparecem em marcadores aceitos.

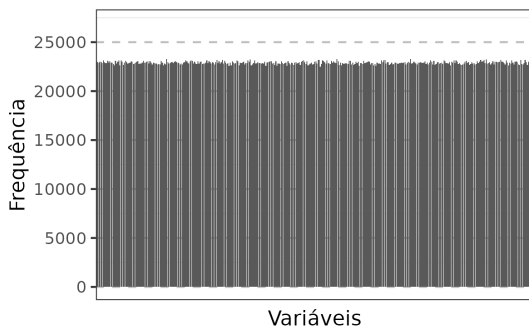


Figura 4: Frequência das covariáveis em todas as simulações.

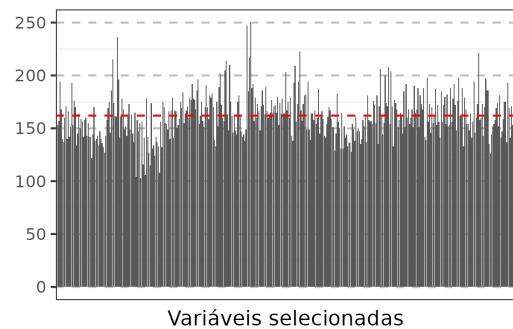


Figura 5: Frequência das covariáveis em simulações com $r_{gg} > 0,5$.

A Tabela 3 fornece uma métrica para avaliação da eficiência do Algoritmo (1). De todas as 1.500.000 simulações – 10.000 ciclos de 150 iterações para cada conjunto de *bootstrap* – houve 1,3% de aproveitamento e uma quantidade relativamente pequena de iterações que não produziram resultados. Após verificação, essas ocorrências se devem a covariáveis altamente correlacionadas presentes na base de dados.

Tabela 3: Aproveitamento das simulações de geração de marcador ambientômico

Simulações	Modelos Aceitos	Sem resultados
1.500.000	20.194 (0.013)	885 (< 0.001)

Além disso, a Figura 6 ilustra a dificuldade de se gerar os marcadores, com mais da metade das simulações gerando um ou dois marcadores úteis.

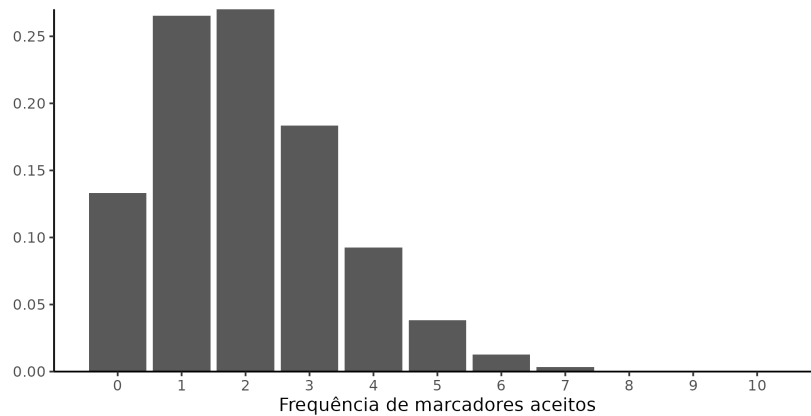


Figura 6: Barras de frequência relativa de marcadores ambientais aceitos em todas as simulações.

Com estas simulações foi possível gerar um máximo de 10 marcadores ambientais em uma iteração, os quais foram utilizados como covariáveis para o modelo misto.

4.3 Redução de dimensionalidade

As Figuras 7 e 8 a seguir exibem a curva de proporção da variância explicada e sua derivada, assim como duas marcas para 10 e 30 componentes. Enquanto seria desejável, pelo critério de Kaiser, utilizar 30 componentes principais, optou-se por utilizar apenas 10. O primeiro motivo para isso foi a dificuldade computacional para o ajuste do modelo misto – foram tentadas opções de computação em nuvem como Google Colab e o computador universitário ao qual o Laboratório de Melhoramento de Precisão (LAMP - UFG) tem acesso, porém sem sucesso de convergência. O segundo motivo é a utilização de um igual número de variáveis entre os dois modelos – com marcadores ambientais e com componentes principais.

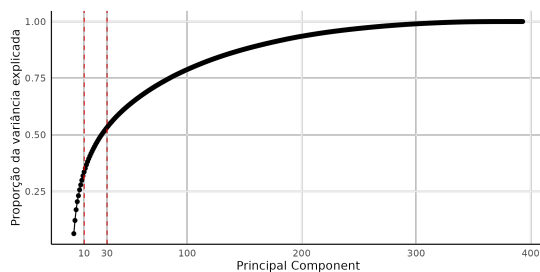


Figura 7: Proporção acumulada da variância explicada pelas componentes principais.

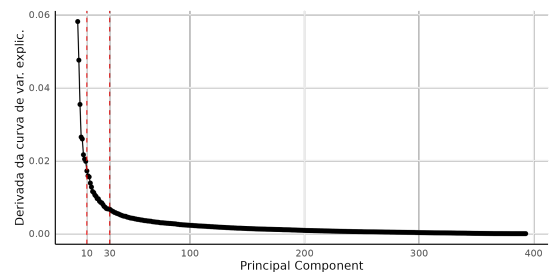


Figura 8: Derivada da curva de variância explicada pelas componentes principais.

Mantendo a escolha de utilizar apenas as primeiras 10 componentes principais, opta-se por componentes cujas cargas de rotação de cada covariável são relativamente

baixas. A Tabela 4 exibe o máximo e o mínimo de contribuição das covariáveis para cada componente principal. Conseqüentemente, tanto as variáveis com maiores frequências exibidas na Figura 5 quanto as covariáveis utilizadas para a construção dos marcadores ambientômicos utilizados no modelo misto têm contribuição dentro dos intervalos a seguir.

Tabela 4: Cargas de rotação máxima e mínima para as primeiras 10 componentes principais.

CP	Máximo	Mínimo
1	0.10	-0.09
2	0.10	-0.08
3	0.06	-0.11
4	0.14	-0.11
5	0.13	-0.11
6	0.14	-0.14
7	0.05	-0.17
8	0.17	-0.14
9	0.13	-0.18
10	0.12	-0.19

De fato, a maior parte das covariáveis têm contribuição muito pequena para as componentes principais, conforme ilustrado na Figura 9.

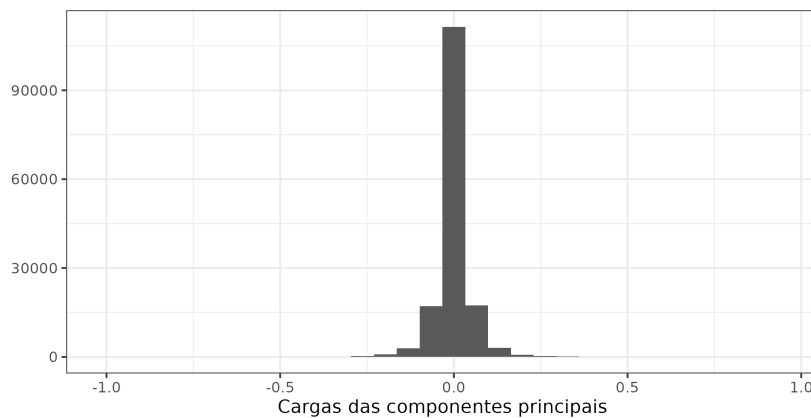


Figura 9: Histograma das cargas do PCA.

Há componentes principais para as quais há contribuições muito maiores, perto dos limites $[-1, 1]$. No entanto, essas componentes estão entre aquelas com menor parcela de variância explicada, conforme a Figura 10. Enquanto há duas componentes atípicas, 27 e 36 com carga máxima próximas a 0,75, parece haver uma gradual troca entre variância explicada e sensibilidade das componentes.

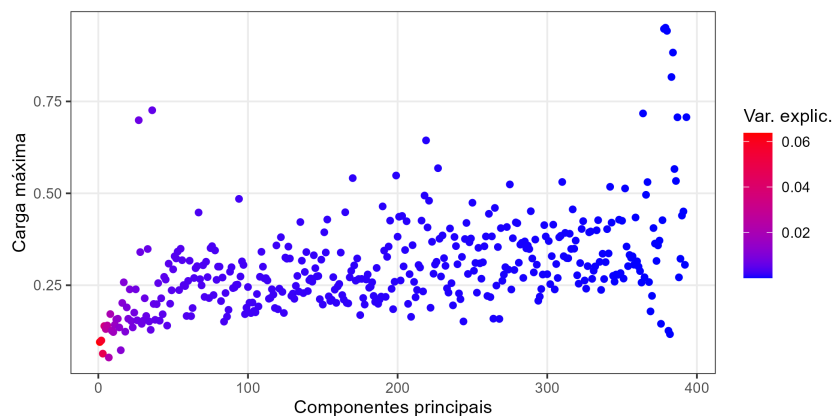


Figura 10: Carga de rotação máxima das covariáveis ambientais para as componentes principais em valor absoluto.

4.4 Modelos mistos

Foram ajustados dois modelos mistos, ambos na forma da Equação (3.5.1), porém utilizando como covariáveis ou marcadores ambientômicos ou componentes principais (PCA) advindos da redução de dimensionalidade.

Os interceptos e correlação intraclasse (ICC) são exibidos na Tabela 5. Enquanto os interceptos, que representam uma média geral de todos os genótipos, estão próximos, os ICC diferem. Esta estatística pode ser interpretada como a proporção da variância explicada pela clusterização, caso em que quase metade da variância do modelo de marcadores ambientômicos seria devida aos agrupamentos. Outra interpretação dessa estatística é a correlação entre observações de um mesmo cluster, de modo que se poderia interpretá-la, no caso do PCA, como 12% de correlação entre genótipos de um mesmo cluster, expostos a condições ambientais diversas.

Tabela 5: Interceptos e correlação intraclasse para os modelos com floresta aleatória e com componentes principais.

Modelo	Intercepto	ICC
PCA	3.270,2	0,12
Marcadores	3.482,3	0,44

Na Figura 11 são exibidos os comportamentos dos coeficientes aleatórios para quatro componentes principais. Esses quatro coeficientes foram escolhidos para ilustrar que estimadores para componentes principais com menor contribuição sobre a variância das variáveis originais parecem não apenas ter variância maior como parecem divergir mais

do comportamento esperado. A linha tracejada vermelha é utilizada como referência para a média teórica dos estimadores. O que se pode observar nos quadros para as componentes 9 e 10 são intervalos de confiança mais largos do que nos demais e pontos nos valores extremos dos estimadores pontuais cada vez mais distantes da média.

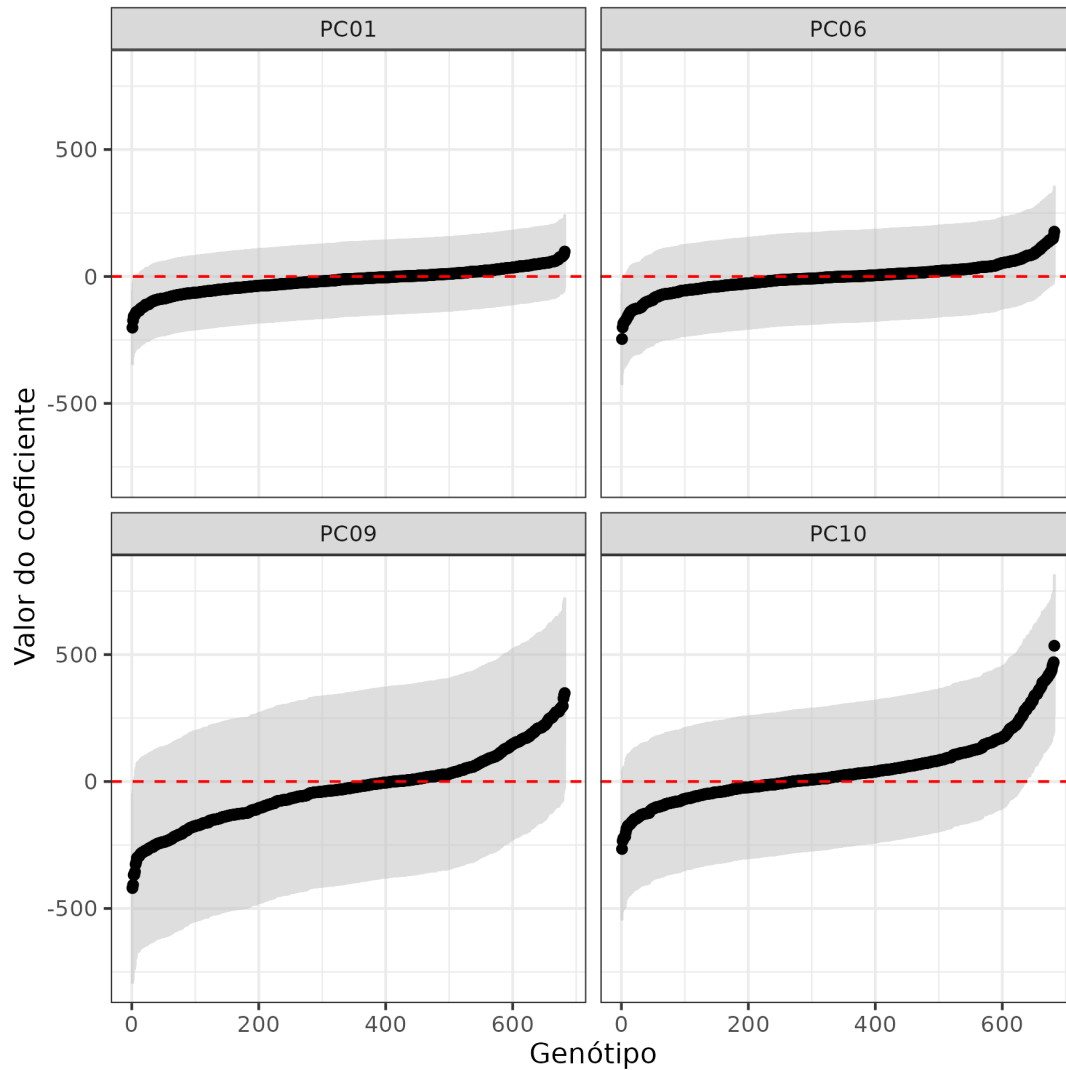


Figura 11: Estimadores pontuais e $IC(\alpha = 0,05)$ dos coeficientes das componentes principais 1, 6, 9 e 10.

A Figura 12 ilustra a mesma informação, porém referente ao modelo ajustado utilizando marcadores ambientômicos. Em contraste ao modelo anterior, seus estimadores não exibem um padrão de desvio progressivo em relação à distribuição teórica. De fato, não há motivo para isto visto que a construção dos marcadores não tem um comportamento bem definido. Enquanto os coeficientes para todos os marcadores parecem, em média, estar melhor comportados que os coeficientes para as componentes principais com menor autovalor, a escala em que variam é superior à da Figura 11.

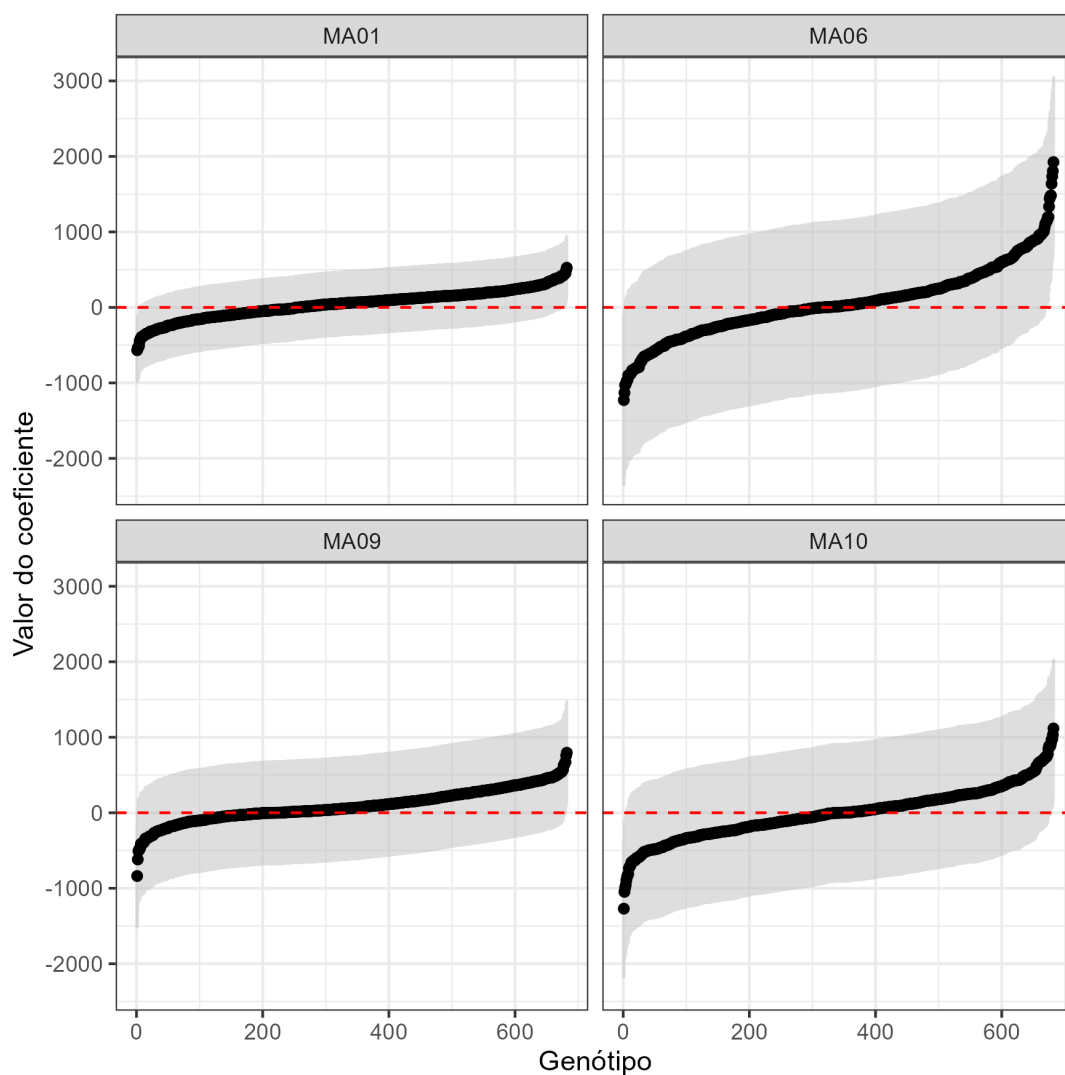


Figura 12: Estimadores pontuais e $IC(\alpha = 0,05)$ dos coeficientes dos marcadores ambientais 1, 6, 9 e 10.

Ao total, 34 variáveis ambientais compuseram os marcadores ambientais, quais sejam:

1. MA 1

- Temperatura máxima média a 2 metros acima do solo durante setembro para cada ano de 2010 a 2019;
- Proporção de partículas de areia (> 0.05 mm) em fração do solo g/kg para centímetros 30 a 60 da superfície – predição para quantil 0,5;

2. MA 2

- Fração volumétrica de fragmentos grosseiros (> 2 mm) em cm^3/dm^3 ($\text{vol}\%$) para centímetros 100 a 200 da superfície – predição para quantil 0,95;

- Proporção de partículas de argila ($< 0,002$ mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – predição para quantil 0,5;

3. MA 3

- Proporção de partículas de silte ($> 0,05$ mm) na fração fina do solo em g/kg para centímetros 0 a 5 da superfície – predição para quantil 0,05;
- Precipitação Anual;
- Nitrogênio Total (N) cg/kg para centímetros 30 a 60 da superfície – predição para quantil 0,5;

4. MA 4

- Temperatura média do ponto de orvalho a 2 metros durante julho para os anos de 2010 a 2019. A temperatura do ponto de orvalho é uma medida de umidade atmosférica;
- Densidade de carbono orgânico em hg/m^3 para centímetros 30 a 60 da superfície – predição para quantil 0,05;
- Umidade relativa média a 2 metros acima do solo durante setembro ao longo do período de 2010 a 2019;

5. MA 5

- Proporção de partículas de silte ($> 0,05$ mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – média;
- Precipitação total corrigida que ocorreu durante agosto de cada ano de 2010 a 2019. "Corrigida" implica ajustes para anomalias ou erros de medição;

6. MA 6

- Proporção de partículas de areia ($> 0,05$ mm) na fração do solo em g/kg para centímetros 30 a 60 da superfície – predição para quantil 0,5;
- Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 0 a 5 da superfície – predição para quantil 0,05;
- pH do solo em $\text{pH} \times 10$ para centímetros 15 a 30 da superfície – predição para quantil 0,5;
- Capacidade de Troca de Cátions do solo em $\text{mmol(c)}/\text{kg}$ para centímetros 5 a 15 da superfície – predição para quantil 0,5;

- Temperatura máxima média a 2 metros acima do solo durante fevereiro para os anos de 2010 a 2019;
- Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 100 a 200 da superfície – predição para quantil 0,95;
- Estoques de carbono orgânico em t/ha para centímetros 0 a 30 da superfície – predição para quantil 0,95;
- Fração volumétrica de fragmentos grosseiros (> 2 mm) em cm^3/dm^3 (vol%) para centímetros 100 a 200 da superfície – predição para quantil 0,95;

7. MA 7

- Proporção de partículas de areia (> 0.05 mm) em fração do solo em g/kg para centímetros 60 a 100 da superfície – predição para quantil 0,05;
- Proporção de partículas de areia (> 0.05 mm) em fração do solo em g/kg para centímetros 15 a 30 da superfície – média;
- Capacidade de Troca de Cátions do solo em $\text{mmol}(c)/\text{kg}$ para centímetros 30 a 60 da superfície – predição para quantil 0,95;
- Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 0 a 5 da superfície – predição para quantil 0,95;
- Proporção de partículas de areia (> 0.05 mm) em fração do solo em g/kg para centímetros 5 a 15 da superfície – predição para quantil 0,5;

8. MA 8

- Nitrogênio Total (N) cg/kg para centímetros 5 a 15 da superfície – predição para quantil 0,05;
- Proporção de partículas de silte ($> 0,05$ mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – média;
- Média anual da velocidade do vento a 2 metros acima do solo para o período de 2010 a 2019;

9. MA 9

- Densidade de carbono orgânico em hg/m^3 para centímetros 60 a 100 da superfície – predição para quantil 0,95;
- Temperatura média a 2 metros acima do solo durante o mês de março para cada ano de 2010 a 2019;

- Estoques de carbono orgânico em t/ha para centímetros 0 a 30 da superfície – predição para quantil 0,5;

10. MA 10

- Proporção de partículas de argila (< 0,002 mm) na fração fina do solo em g/kg para centímetros 100 a 200 da superfície – predição para quantil 0,5;
- Radiação descendente de onda longa na superfície sob todas as condições de céu para julho ao longo dos anos de 2010 a 2019. Reflete a radiação de onda longa que alcança a superfície terrestre;
- Precipitação total corrigida registrada durante fevereiro de cada ano de 2010 a 2019. "Corrigida" implica que os dados foram ajustados para corrigir erros ou vieses de medição;

Além disso, os marcadores ambientômicos via de regra são altamente correlacionados, como explicitado na Figura 13. Isto é problemático para a utilização de modelos lineares, pois é indicador de uma matriz design de posto incompleto.

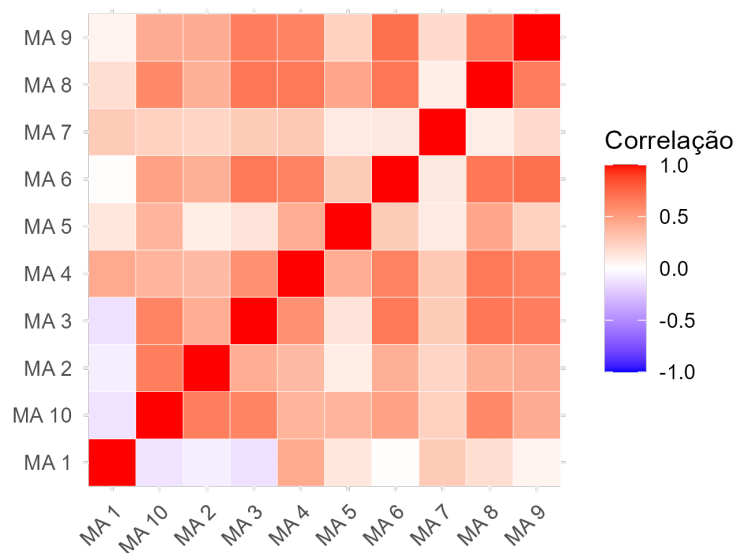


Figura 13: Correlograma para marcadores ambientômicos utilizados no modelo.

Avalia-se também o comportamento dos valores ajustados de ambos os modelos em relação aos valores reais, utilizando uma linha vermelha como referencial para $\hat{y} = y$. Nota-se nas Figuras 14 e 15 que ambos os modelos têm um comportamento similar e apresentam problemas de ajuste.

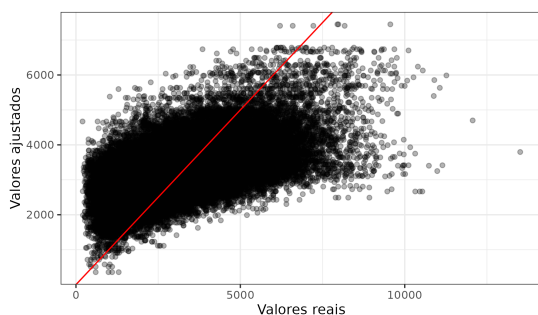


Figura 14: Valores reais e ajustados para o modelo misto ajustado com PCA.

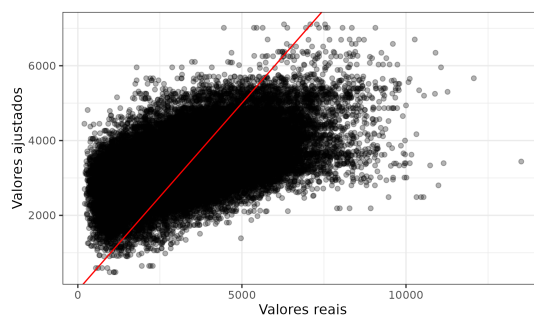


Figura 15: Valores reais e ajustados para o modelo misto ajustado com marcadores ambientômicos.

A avaliação dos resíduos studentizados, Figuras 16 e 17, também não fornece muita informação, exceto por uma aparente heteroscedasticidade e média igual a zero, com maiores variâncias ocorrendo para valores ajustados intermediários.

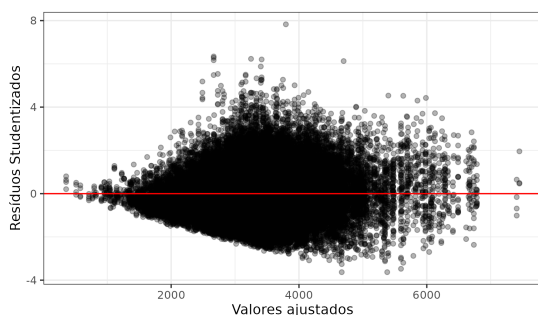


Figura 16: Valores ajustados e resíduos studentizados para o modelo misto ajustado com PCA.

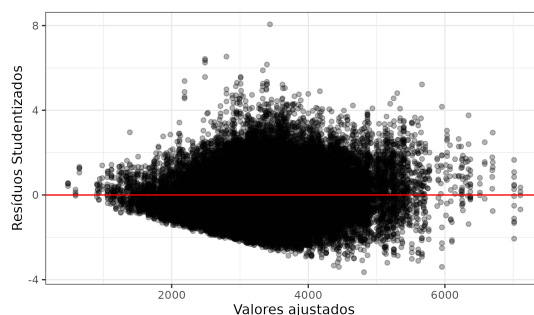


Figura 17: Valores ajustados e resíduos studentizados para o modelo misto ajustado com marcadores ambientômicos.

Para testar a capacidade preditiva dos modelos, o ano de 2022 foi separado como partição de validação da base de dados. No entanto nem todos os pontos puderam ser usados, pois só é possível realizar predições para genótipos que já estavam presentes no momento de ajuste do modelo. Dessa forma, a base de dados de validação contou com 815 observações de 16 genótipos, com ocorrência conforme a Figura 18.

Complementarmente, a Tabela 7 resume informações a respeito de cobertura de predição e amplitude média dos intervalos de predição. O modelo baseado em indicadores ambientômicos apresenta tanto cobertura média quanto amplitude média dos intervalos de predição um pouco maiores.

Tabela 7: Cobertura média e amplitude média dos intervalos de previsão para os modelos PCA e de Marcadores.

	Modelo PCA			Modelo Marcadores		
	Centro Oeste	Norte	Nordeste	Centro Oeste	Norte	Nordeste
Amplitude média	4.926,81	4.914,6	4.898,53	4.951,65	4.944,01	4.947,02
Cobertura média	0,79	0,56	0,97	0,86	0,62	0,98

Para avaliar a consistência dos modelos, foram realizadas validações cruzadas (*k-fold cross-validation*) utilizando 5 “dobras” ou pastas. Além disso, dois níveis de validação cruzada foram testados: nível de cluster, em que cada cluster corresponde a um genótipo, e nível de unidade observacional, em que as partições da base de dados são feitas mediante amostras aleatórias simples da base de dados completa. O conjunto de dados utilizados foi aquele destinado ao treinamento e a medida utilizada para avaliar a consistência nas validações foi o erro quadrático médio (*MSE*), que foi avaliada pela razão entre o MSE médio entre as validações e o MSE da amostra completa.

A Tabela 8 resume a avaliação das validações cruzadas. Para os casos em que a razão está próxima de 1, há indícios de que o modelo original foi ajustado de forma consistente com a validação cruzada. Os casos em que a razão é menor que 1 sugerem que houve um melhor desempenho do modelo na validação quando comparado com o modelo original. Isso é especialmente notável para a validação a nível de cluster do modelo de marcadores ambientômicos. Mantida a complexidade do modelo e considerando que há considerável desbalanceamento entre a quantidade de observações para os genótipos, uma possível interpretação da diferença de desempenho na validação é que os modelo de marcadores ambientômicos são sensíveis à quantidade de genótipos disponíveis para ajuste do modelo.

Tabela 8: Razão entre MSE médio da validação cruzada *5-fold* e o MSE da amostra completa para os modelos PCA e de Marcadores considerando validações a nível de genótipo e a nível de unidade observacional.

Modelo	Nível de cluster	Nível de unidade
PCA	1,01	0,82
Marcadores	0,51	1,09

5 Considerações finais

A proposta ambientômica para modelagem de comportamento de cultivares no contexto de interação $G \times E$ presume um estudo ômico, ou seja, considerando a maior quantidade de variáveis possível. Essa abordagem implica em dificuldades metodológicas para inferência estatística como o crescimento da dimensionalidade do problema avaliado, colinearidades entre covariáveis e consequentemente matrizes de postos incompletos, entre outros. Nesse sentido, Resende et al. (2021) propõe a criação de marcadores ambientômicos como uma forma de aumentar a capacidade de representação do modelo misto. Com vistas à avaliação metodológica dos autores, foi proposta a comparação do modelo dos autores com outro, utilizando como covariáveis também variáveis derivadas, mas advindas de análise de componentes principais.

O algoritmo para geração dos marcadores ambientais tem baixo aproveitamento das simulações e em muitas delas gera-se uma pequena quantidade de marcadores. Enquanto seria possível agregar uma grande quantidade de marcadores ambientais caso fossem utilizadas todas as simulações, não foi possível identificar fundamentação teórica para avaliação das propriedades de marcadores gerados de diferentes formas.

Por outro lado, a decomposição da matriz *design* em componentes principais produz covariáveis com propriedades estatísticas conhecidas. No entanto, a interpretação das componentes principais ainda é dificultosa devido à grande quantidade de covariáveis iniciais e devido à distribuição da variância explicada mediante a inclusão de mais componentes. Espera-se que seja possível tomar uma conclusão a respeito de quantas componentes devem ser utilizadas e em seguida interpretar as componentes, mas isso não pode ser feito visto que o incremento na variância explicada mediante a inclusão de mais componentes era pequeno e não apresentava pontos de corte claros.

O uso de modelos mistos para propósitos agrônômicos é amplamente indicado na literatura, com evidentes vantagens sobre modelos sem uma estrutura de hierarquia entre parâmetros. As ferramentas disponíveis para um trabalho aprofundado com essa classe de modelos fora da perspectiva bayesiana, entretanto, são especializadas e com exemplos de caso limitados. Além disso, material sobre diagnósticos desses modelos e funções em R predefinidas que ajam sobre os resultados também são escassos.

A maior dificuldade encontrada durante o processo de modelagem foi a ponderação entre a quantidade de covariáveis a serem utilizadas e a identificabilidade do modelo misto. Diversas combinações de quantidade de covariáveis e quantidade de genótipos

foram testadas, indicando que o limitador computacional era a quantidade de covariáveis. Enquanto era desejável incluir 30 componentes principais no modelo, essa dificuldade levou à inclusão de apenas 10 – correspondendo a aproximadamente 30% de variância explicada –, mesma quantidade de marcadores ambientômicos.

Comparar e avaliar os modelos gerados também envolveu uma série de dificuldades, quando comparados a modelos mais simples. Análises de resíduos e do comportamento dos coeficientes carecem de uma conclusão e são vistas como características finais do modelo, como heteroscedasticidade dos resíduos. Além disso, como os modelos não utilizaram as mesmas variáveis explicativas não era possível utilizar medidas mais tradicionais como critérios de informação e outros métodos que utilizam a verossimilhança. Decidiu-se por uma opção mais simples, e arbitrária, de comparar a capacidade preditiva dos modelos sobre o último ano de experimentos disponível.

O desempenho dos modelos a princípio não é diferenciável. A Tabela 7 ilustra isso na medida em que intervalos de predição de maior amplitude têm uma cobertura maior das respostas na base de dados de validação. Análises de comportamento dos resíduos e de comparação entre valores reais e valores ajustados da variável resposta sugerem o mesmo.

Considerando as dificuldades computacionais encontradas bem como a limitação de se estar lidando com relações evidentemente não lineares entre genótipo e ambiente, pesquisas futuras podem aprofundar o estudo do desempenho de técnicas como

- Utilização de dados de parentalidade e técnicas mais refinadas de geração de marcadores ambientômicos mais explicativos, como em Resende et al. (2024);
- Inclusão de variáveis quadráticas;
- Inclusão de interação entre variáveis;
- Utilização de medidas resumo de subregiões do Brasil como covariáveis;
- Utilização de variáveis mais simples como covariáveis, por exemplo: região, bioma, pluviometria média, etc;
- Redução de covariáveis àquelas com uma mesma “temática”, tais como: característica de solo, pluviometria, temperatura, etc.

Referências

- ADEWALE, A. J. et al. Understanding hierarchical linear models: applications in nursing research. *Nursing Research*, LWW, v. 56, n. 4, p. S40–S46, 2007.
- ARTES, R.; BARROSO, L. P. *Métodos Multivariados de Análise Estatística*. [S.l.]: Blucher, 2023.
- BATES, D. et al. *lme4: Linear Mixed-Effects Models using “Eigen” and S4*. The R Foundation, 2003. Disponível em: <http://dx.doi.org/10.32614/CRAN.package.lme4>.
- BATES, D. et al. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, v. 67, n. 1, p. 1–48, 2015. Disponível em: <https://www.jstatsoft.org/index.php/jss/article/view/v067i01>.
- BELLMAN, R.; KALABA, R. A mathematical theory of adaptive control processes. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 45, n. 8, p. 1288–1290, 1959.
- BRESEGHELLO, F. et al. Building the embrapa rice breeding dataset for efficient data reuse. *Crop Science*, v. 61, n. 5, p. 3445–3457, 06 2021. Disponível em: <http://dx.doi.org/10.1002/csc2.20550>.
- BRESEGHELLO, F. et al. Results of 25 years of upland rice breeding in brazil. *Crop Science*, v. 51, n. 3, p. 914–923, 05 2011. Disponível em: <http://dx.doi.org/10.2135/cropsci2010.06.0325>.
- CHEN, L. Curse of dimensionality. In: _____. *Encyclopedia of Database Systems*. Boston, MA: Springer US, 2009. p. 545–546. ISBN 978-0-387-39940-9. Disponível em: https://doi.org/10.1007/978-0-387-39940-9_133.
- CHENU, K. Characterizing the crop environment – nature, significance and applications. *Crop physiology*, Elsevier, p. 321–348, 2015.
- DEMIDENKO, E. *Mixed models: theory and applications with R*. [S.l.]: John Wiley & Sons, 2013.
- FOX, J.; MONETTE, G. *cv: Cross-Validating Regression Models*. The R Foundation, 2023. Disponível em: <http://dx.doi.org/10.32614/CRAN.package.cv>.
- GELMAN, A.; HILL, J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. [S.l.]: Cambridge University Press, 2006. (Analytical Methods for Social Research).
- HICKEY, L. T. et al. Breeding crops to feed 10 billion. *Nature biotechnology*, Nature Publishing Group US New York, v. 37, n. 7, p. 744–754, 2019.
- HOX, J.; MOERBEEK, M.; SCHOOT, R. Van de. *Multilevel analysis: Techniques and applications*. [S.l.]: Routledge, 2017.
- JORASCH, P. The global need for plant breeding innovation. *Transgenic Research*, Springer, v. 28, n. Suppl 2, p. 81–86, 2019.

- JOURNEL, A. G.; JOURNEL, A. G. *Fundamentals of geostatistics in five lessons*. [S.l.]: American Geophysical Union Washington, DC, 1989. v. 8.
- KUTNER, M. H. et al. *Applied linear statistical models*. [S.l.]: McGraw-hill, 2005.
- MARTINS, E. N. et al. *Estimação de Componentes de Variância*. [S.l.]: Editora UFV, 1993. v. 39. (Cadernos Didáticos, v. 39).
- MARTINS, E. N. et al. *Modelo Linear Misto*. [S.l.]: Editora UFV, 1993. v. 38. (Cadernos Didáticos, v. 38).
- MCMAHON, S. M.; DIEZ, J. M. Scales of association: hierarchical linear models and the measurement of ecological systems. *Ecology letters*, Wiley Online Library, v. 10, n. 6, p. 437–452, 2007.
- MORETTIN, P. A.; SINGER, J. d. M. *Estatística e Ciência de Dados*. [S.l.]: LTC, 2022.
- RESENDE, R. T.; BRONDANI, C.; CHAVES, L. J. O melhoramento na era de agricultura de precisão. In: RESENDE, R. T.; BRONDANI, C. (Ed.). *Melhoramento de Precisão*. Santo Antônio de Goiás, GO: Embrapa Arroz e Feijão, 2023. cap. 1, p. 13–40.
- RESENDE, R. T. et al. Enviromics in breeding: applications and perspectives on envirotypic-assisted selection. *Theoretical and Applied Genetics*, Springer, v. 134, p. 95–112, 2021.
- RESENDE, R. T. et al. Gis-based $g \times e$ modeling of maize hybrids through enviromic markers engineering. *New Phytologist*, 2024.
- SCRUCCA, L. et al. *Model-based clustering, classification, and density estimation using mclust in R*. [S.l.]: Chapman and Hall/CRC, 2023.
- WICKHAM, H. *tidyverse: Easily Install and Load the “Tidyverse”*. The R Foundation, 2016. Disponível em: <http://dx.doi.org/10.32614/CRAN.package.tidyverse>.

Tabela 9: Dicionário de variáveis do ERBD relacionadas aos ensaios

Variável	Nome	Detalhes
TRIAL	Código do ensaio	String único que identifica o ensaio
SYST	Sistema de cultivo	Indica tanto o subprograma de melhoramento quanto o ambiente do ensaio. Níveis: Irrigado ou de Sequeiro
YEAR	Ano do ensaio	Ano de preparação do ensaio. Ex: 2005: temporada 2005/2006
DATE	Data de plantio	Dia de plantio de sementes secas. Formato DD/MM/AAAA
ST	Estado do Brasil	Estado do Brasil onde o ensaio foi conduzido
LOCATION	Local de plantio	Nome do município onde o ensaio foi conduzido
LOC	Local de plantio	Termo abreviado que indica o município
TYPE	Tipo de ensaio	Tipo de ensaio. ER: Ensaios Regionais de Rendimento; VCU: Valor de Cultivo e Uso (Ensaios Avançados de Rendimento)
DESIGN	Desenho Experimental	O desenho estatístico do ensaio. RCB: delineamento de blocos completos ao acaso; LAT: delineamento em látice
MEAN	Média de rendimento de grãos	Média geral do ensaio do rendimento de grãos (kg ha^{-1})
H ²	Hereditariedade	Hereditariedade de sentido amplo do rendimento de grãos
CV	Coefficiente de Variação	Coefficiente de variação experimental para rendimento de grãos (%)

Fonte: Breseghello et al. (2021)

Tabela 10: Dicionário de variáveis do ERBD das unidades experimentais

Nome na base	Nome da variável	Tipo	Detalhes
TRIAL	Código do ensaio	Link para Metadados	String único que identifica o ensaio
REP	Número da Repetição	Fator de Design	Inteiro que indica a repetição dentro do ensaio
BLO	Número do Bloco	Fator de Design	Inteiro que indica o bloco dentro da repetição (apenas em delineamento em látice)
GEN	Nome do Genótipo	Fator Experimental	Identificação do germoplasma (linhagem pura, variedade local ou cultivar)
GY	Rendimento de grãos	Numérico	Peso do arroz em casca com 13% de umidade, em kg ha^{-1}
PHT	Altura da Planta	Numérico	Altura da planta do solo até a ponta da panícula primária, no estágio pré-colheita, em cm
DTF	Dias até a Floração	Numérico	Número de dias desde o plantio de sementes secas até 50% das plantas estarem floridas
LOD	Tombamento	Escores de 1 a 9 ⁶	Nível de tombamento da copa da parcela, avaliado no estágio pré-colheita
LBL	Blast na Folha	Escores de 1 a 9	Severidade da doença de brusone do arroz, avaliada em folhas no estágio vegetativo
PBL	Blast na Panícula	Escores de 1 a 9	Severidade da doença de brusone do arroz, avaliada em panículas no estágio pré-colheita
BSP	Mancha Parda	Escores de 1 a 9	Severidade da doença causada por <i>Bipolaris oryzae</i> , avaliada em folhas no estágio pré-colheita
LSC	Escaldadura da Folha	Escores de 1 a 9	Severidade da doença causada por <i>Monographella albescens</i> , avaliada em folhas no estágio pré-colheita
GDS	Descoloração do Grão	Escores de 1 a 9	Severidade do escurecimento ou manchas nos grãos, causada por vários fungos, avaliada nas glumas no estágio pré-colheita

Fonte: Breseghello et al. (2021)

⁶Escores mais altos indicam níveis crescentes de tombamento ou doença.

Tabela 11: Dicionário de variáveis ambientais

Nome na base	Nome da variável	Descrição
WorldClim	bio_1	Temperatura Média Anual
WorldClim	bio_2	Amplitude Diurna Média (Média das diferenças mensais (temp máx - temp mín))
WorldClim	bio_3	Isotermia (BIO2/BIO7) ($\times 100$)
WorldClim	bio_4	Sazonalidade da Temperatura (desvio padrão $\times 100$)
WorldClim	bio_5	Temperatura Máxima do Mês Mais Quente
WorldClim	bio_6	Temperatura Mínima do Mês Mais Frio
WorldClim	bio_7	Amplitude Térmica Anual (BIO5-BIO6)
WorldClim	bio_8	Temperatura Média do Trimestre Mais Chuvoso
WorldClim	bio_9	Temperatura Média do Trimestre Mais Seco
WorldClim	bio_10	Temperatura Média do Trimestre Mais Quente
WorldClim	bio_11	Temperatura Média do Trimestre Mais Frio
WorldClim	bio_12	Precipitação Anual
WorldClim	bio_13	Precipitação do Mês Mais Chuvoso
WorldClim	bio_14	Precipitação do Mês Mais Seco
WorldClim	bio_15	Sazonalidade da Precipitação (Coeficiente de Variação)
WorldClim	bio_16	Precipitação do Trimestre Mais Chuvoso
WorldClim	bio_17	Precipitação do Trimestre Mais Seco
WorldClim	bio_18	Precipitação do Trimestre Mais Quente
WorldClim	bio_19	Precipitação do Trimestre Mais Frio
SoilGrids	bdod_0-5cm_mean	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 0 a 5 da superfície – média
SoilGrids	bdod_0-5cm_Q0p05	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 0 a 5 da superfície – predição para quantil 0,05
SoilGrids	bdod_0-5cm_Q0p5	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 0 a 5 da superfície – predição para quantil 0,5

SoilGrids	bdod_0-5cm_Q0p95	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 0 a 5 da superfície – predição para quantil 0,95
SoilGrids	bdod_5-15cm_mean	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 5 a 15 da superfície – média
SoilGrids	bdod_5-15cm_Q0p05	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 5 a 15 da superfície – predição para quantil 0,05
SoilGrids	bdod_5-15cm_Q0p5	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 5 a 15 da superfície – predição para quantil 0,5
SoilGrids	bdod_5-15cm_Q0p95	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 5 a 15 da superfície – predição para quantil 0,95
SoilGrids	bdod_15-30cm_mean	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 15 a 30 da superfície – média
SoilGrids	bdod_15-30cm_Q0p05	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 15 a 30 da superfície – predição para quantil 0,05
SoilGrids	bdod_15-30cm_Q0p5	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 15 a 30 da superfície – predição para quantil 0,5
SoilGrids	bdod_15-30cm_Q0p95	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 15 a 30 da superfície – predição para quantil 0,95
SoilGrids	bdod_30-60cm_mean	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 15 a 30 da superfície – média
SoilGrids	bdod_30-60cm_Q0p05	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 30 a 60 da superfície – predição para quantil 0,05
SoilGrids	bdod_30-60cm_Q0p5	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 30 a 60 da superfície – predição para quantil 0,5

SoilGrids	bdod_30-60cm_Q0p95	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 30 a 60 da superfície – predição para quantil 0,95
SoilGrids	bdod_60-100cm_mean	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 60 a 100 da superfície – média
SoilGrids	bdod_60-100cm_Q0p05	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 60 a 100 da superfície – predição para quantil 0,05
SoilGrids	bdod_60-100cm_Q0p5	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 60 a 100 da superfície – predição para quantil 0,5
SoilGrids	bdod_60-100cm_Q0p95	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 60 a 100 da superfície – predição para quantil 0,95
SoilGrids	bdod_100-200cm_mean	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 100 a 200 da superfície – média
SoilGrids	bdod_100-200cm_Q0p05	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 100 a 200 da superfície – predição para quantil 0,05
SoilGrids	bdod_100-200cm_Q0p5	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 100 a 200 da superfície – predição para quantil 0,5
SoilGrids	bdod_100-200cm_Q0p95	Densidade aparente da fração fina do solo em cg/cm^3 para centímetros 100 a 200 da superfície – predição para quantil 0,95
SoilGrids	cec_0-5cm_mean	Capacidade de Troca de Cátions do solo em $\text{mmol}(c)/\text{kg}$ para centímetros 0 a 5 da superfície – média
SoilGrids	cec_0-5cm_Q0p05	Capacidade de Troca de Cátions do solo em $\text{mmol}(c)/\text{kg}$ para centímetros 0 a 5 da superfície – predição para quantil 0,05
SoilGrids	cec_0-5cm_Q0p5	Capacidade de Troca de Cátions do solo em $\text{mmol}(c)/\text{kg}$ para centímetros 0 a 5 da superfície – predição para quantil 0,5

SoilGrids	cec_0-5cm_Q0p95	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 0 a 5 da superfície – predição para quantil 0,95
SoilGrids	cec_5-15cm_mean	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 5 a 15 da superfície – média
SoilGrids	cec_5-15cm_Q0p05	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 5 a 15 da superfície – predição para quantil 0,05
SoilGrids	cec_5-15cm_Q0p5	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 5 a 15 da superfície – predição para quantil 0,5
SoilGrids	cec_5-15cm_Q0p95	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 5 a 15 da superfície – predição para quantil 0,95
SoilGrids	cec_15-30cm_mean	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 15 a 30 da superfície – média
SoilGrids	cec_15-30cm_Q0p05	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 15 a 30 da superfície – predição para quantil 0,05
SoilGrids	cec_15-30cm_Q0p5	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 15 a 30 da superfície – predição para quantil 0,5
SoilGrids	cec_15-30cm_Q0p95	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 15 a 30 da superfície – predição para quantil 0,95
SoilGrids	cec_30-60cm_mean	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 15 a 30 da superfície – média
SoilGrids	cec_30-60cm_Q0p05	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 30 a 60 da superfície – predição para quantil 0,05
SoilGrids	cec_30-60cm_Q0p5	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 30 a 60 da superfície – predição para quantil 0,5

SoilGrids	cec_30-60cm_Q0p95	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 30 a 60 da superfície – predição para quantil 0,95
SoilGrids	cec_60-100cm_mean	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 60 a 100 da superfície – média
SoilGrids	cec_60-100cm_Q0p05	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 60 a 100 da superfície – predição para quantil 0,05
SoilGrids	cec_60-100cm_Q0p5	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 60 a 100 da superfície – predição para quantil 0,5
SoilGrids	cec_60-100cm_Q0p95	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 60 a 100 da superfície – predição para quantil 0,95
SoilGrids	cec_100-200cm_mean	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 100 a 200 da superfície – média
SoilGrids	cec_100-200cm_Q0p05	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 100 a 200 da superfície – predição para quantil 0,05
SoilGrids	cec_100-200cm_Q0p5	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 100 a 200 da superfície – predição para quantil 0,5
SoilGrids	cec_100-200cm_Q0p95	Capacidade de Troca de Cátions do solo em mmol(c)/kg para centímetros 100 a 200 da superfície – predição para quantil 0,95
SoilGrids	cfvo_0-5cm_mean	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 0 a 5 da superfície - média
SoilGrids	cfvo_0-5cm_Q0p05	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 0 a 5 da superfície – predição para quantil 0,05
SoilGrids	cfvo_0-5cm_Q0p5	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 0 a 5 da superfície – predição para quantil 0,5

SoilGrids	cfvo_0-5cm_Q0p95	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 0 a 5 da superfície – predição para quantil 0,95
SoilGrids	cfvo_5-15cm_mean	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 5 a 15 da superfície – média
SoilGrids	cfvo_5-15cm_Q0p05	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 5 a 15 da superfície – predição para quantil 0,05
SoilGrids	cfvo_5-15cm_Q0p5	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 5 a 15 da superfície – predição para quantil 0,5
SoilGrids	cfvo_5-15cm_Q0p95	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 5 a 15 da superfície – predição para quantil 0,95
SoilGrids	cfvo_15-30cm_mean	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 15 a 30 da superfície – média
SoilGrids	cfvo_15-30cm_Q0p05	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 15 a 30 da superfície – predição para quantil 0,05
SoilGrids	cfvo_15-30cm_Q0p5	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 15 a 30 da superfície – predição para quantil 0,5
SoilGrids	cfvo_15-30cm_Q0p95	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 15 a 30 da superfície – predição para quantil 0,95
SoilGrids	cfvo_30-60cm_mean	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 15 a 30 da superfície – média

SoilGrids	cfvo_30-60cm_Q0p05	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 30 a 60 da superfície – predição para quantil 0,05
SoilGrids	cfvo_30-60cm_Q0p5	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 30 a 60 da superfície – predição para quantil 0,5
SoilGrids	cfvo_30-60cm_Q0p95	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 30 a 60 da superfície – predição para quantil 0,95
SoilGrids	cfvo_60-100cm_mean	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 60 a 100 da superfície – média
SoilGrids	cfvo_60-100cm_Q0p05	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 60 a 100 da superfície – predição para quantil 0,05
SoilGrids	cfvo_60-100cm_Q0p5	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 60 a 100 da superfície – predição para quantil 0,5
SoilGrids	cfvo_60-100cm_Q0p95	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 60 a 100 da superfície – predição para quantil 0,95
SoilGrids	cfvo_100-200cm_mean	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 100 a 200 da superfície – média
SoilGrids	cfvo_100-200cm_Q0p05	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 100 a 200 da superfície – predição para quantil 0,05
SoilGrids	cfvo_100-200cm_Q0p5	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 100 a 200 da superfície – predição para quantil 0,5

SoilGrids	cfvo_100- 200cm_Q0p95	Fração volumétrica de fragmentos grosseiros (>2 mm) em cm ³ /dm ³ (vol%) para centímetros 100 a 200 da superfície – predição para quantil 0,95
SoilGrids	clay_0-5cm_mean	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 0 a 5 da superfície - média
SoilGrids	clay_0-5cm_Q0p05	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 0 a 5 da superfície – predição para quantil 0,05
SoilGrids	clay_0-5cm_Q0p5	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 0 a 5 da superfície – predição para quantil 0,5
SoilGrids	clay_0-5cm_Q0p95	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 0 a 5 da superfície – predição para quantil 0,95
SoilGrids	clay_5-15cm_mean	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 5 a 15 da superfície – média
SoilGrids	clay_5-15cm_Q0p05	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 5 a 15 da superfície – predição para quantil 0,05
SoilGrids	clay_5-15cm_Q0p5	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 5 a 15 da superfície – predição para quantil 0,5
SoilGrids	clay_5-15cm_Q0p95	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 5 a 15 da superfície – predição para quantil 0,95
SoilGrids	clay_15-30cm_mean	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – média

SoilGrids	clay_15-30cm_Q0p05	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – predição para quantil 0,05
SoilGrids	clay_15-30cm_Q0p5	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – predição para quantil 0,5
SoilGrids	clay_15-30cm_Q0p95	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – predição para quantil 0,95
SoilGrids	clay_30-60cm_mean	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – média
SoilGrids	clay_30-60cm_Q0p05	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 30 a 60 da superfície – predição para quantil 0,05
SoilGrids	clay_30-60cm_Q0p5	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 30 a 60 da superfície – predição para quantil 0,5
SoilGrids	clay_30-60cm_Q0p95	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 30 a 60 da superfície – predição para quantil 0,95
SoilGrids	clay_60-100cm_mean	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 60 a 100 da superfície – média
SoilGrids	clay_60-100cm_Q0p05	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 60 a 100 da superfície – predição para quantil 0,05
SoilGrids	clay_60-100cm_Q0p5	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 60 a 100 da superfície – predição para quantil 0,5

SoilGrids	clay_60-100cm_Q0p95	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 60 a 100 da superfície – predição para quantil 0,95
SoilGrids	clay_100-200cm_mean	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 100 a 200 da superfície – média
SoilGrids	clay_100-200cm_Q0p05	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 100 a 200 da superfície – predição para quantil 0,05
SoilGrids	clay_100-200cm_Q0p5	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 100 a 200 da superfície – predição para quantil 0,5
SoilGrids	clay_100-200cm_Q0p95	Proporção de partículas de argila (<0,002 mm) na fração fina do solo em g/kg para centímetros 100 a 200 da superfície – predição para quantil 0,95
SoilGrids	nitrogen_0-5cm_mean	Nitrogênio Total (N) cg/kg para centímetros 0 a 5 da superfície - média
SoilGrids	nitrogen_0-5cm_Q0p05	Nitrogênio Total (N) cg/kg para centímetros 0 a 5 da superfície – predição para quantil 0,05
SoilGrids	nitrogen_0-5cm_Q0p5	Nitrogênio Total (N) cg/kg para centímetros 0 a 5 da superfície – predição para quantil 0,5
SoilGrids	nitrogen_0-5cm_Q0p95	Nitrogênio Total (N) cg/kg para centímetros 0 a 5 da superfície – predição para quantil 0,95
SoilGrids	nitrogen_5-15cm_mean	Nitrogênio Total (N) cg/kg para centímetros 5 a 15 da superfície – média
SoilGrids	nitrogen_5-15cm_Q0p05	Nitrogênio Total (N) cg/kg para centímetros 5 a 15 da superfície – predição para quantil 0,05
SoilGrids	nitrogen_5-15cm_Q0p5	Nitrogênio Total (N) cg/kg para centímetros 5 a 15 da superfície – predição para quantil 0,5

SoilGrids	nitrogen_5-15cm_Q0p95	Nitrogênio Total (N) cg/kg para centímetros 5 a 15 da superfície – predição para quantil 0,95
SoilGrids	nitrogen_15-30cm_mean	Nitrogênio Total (N) cg/kg para centímetros 15 a 30 da superfície – média
SoilGrids	nitrogen_15-30cm_Q0p05	Nitrogênio Total (N) cg/kg para centímetros 15 a 30 da superfície – predição para quantil 0,05
SoilGrids	nitrogen_15-30cm_Q0p5	Nitrogênio Total (N) cg/kg para centímetros 15 a 30 da superfície – predição para quantil 0,5
SoilGrids	nitrogen_15-30cm_Q0p95	Nitrogênio Total (N) cg/kg para centímetros 15 a 30 da superfície – predição para quantil 0,95
SoilGrids	nitrogen_30-60cm_mean	Nitrogênio Total (N) cg/kg para centímetros 15 a 30 da superfície – média
SoilGrids	nitrogen_30-60cm_Q0p05	Nitrogênio Total (N) cg/kg para centímetros 30 a 60 da superfície – predição para quantil 0,05
SoilGrids	nitrogen_30-60cm_Q0p5	Nitrogênio Total (N) cg/kg para centímetros 30 a 60 da superfície – predição para quantil 0,5
SoilGrids	nitrogen_30-60cm_Q0p95	Nitrogênio Total (N) cg/kg para centímetros 30 a 60 da superfície – predição para quantil 0,95
SoilGrids	nitrogen_60-100cm_mean	Nitrogênio Total (N) cg/kg para centímetros 60 a 100 da superfície – média
SoilGrids	nitrogen_60-100cm_Q0p05	Nitrogênio Total (N) cg/kg para centímetros 60 a 100 da superfície – predição para quantil 0,05
SoilGrids	nitrogen_60-100cm_Q0p5	Nitrogênio Total (N) cg/kg para centímetros 60 a 100 da superfície – predição para quantil 0,5
SoilGrids	nitrogen_60-100cm_Q0p95	Nitrogênio Total (N) cg/kg para centímetros 60 a 100 da superfície – predição para quantil 0,95
SoilGrids	nitrogen_100-200cm_mean	Nitrogênio Total (N) cg/kg para centímetros 100 a 200 da superfície – média

SoilGrids	nitrogen_100-200cm_Q0p05	Nitrogênio Total (N) cg/kg para centímetros 100 a 200 da superfície – predição para quantil 0,05
SoilGrids	nitrogen_100-200cm_Q0p5	Nitrogênio Total (N) cg/kg para centímetros 100 a 200 da superfície – predição para quantil 0,5
SoilGrids	nitrogen_100-200cm_Q0p95	Nitrogênio Total (N) cg/kg para centímetros 100 a 200 da superfície – predição para quantil 0,95
SoilGrids	ocd_0-5cm_mean	Densidade de carbono orgânico em hg/m ³ para centímetros 0 a 5 da superfície - média
SoilGrids	ocd_0-5cm_Q0p05	Densidade de carbono orgânico em hg/m ³ para centímetros 0 a 5 da superfície – predição para quantil 0,05
SoilGrids	ocd_0-5cm_Q0p5	Densidade de carbono orgânico em hg/m ³ para centímetros 0 a 5 da superfície – predição para quantil 0,5
SoilGrids	ocd_0-5cm_Q0p95	Densidade de carbono orgânico em hg/m ³ para centímetros 0 a 5 da superfície – predição para quantil 0,95
SoilGrids	ocd_5-15cm_mean	Densidade de carbono orgânico em hg/m ³ para centímetros 5 a 15 da superfície – média
SoilGrids	ocd_5-15cm_Q0p05	Densidade de carbono orgânico em hg/m ³ para centímetros 5 a 15 da superfície – predição para quantil 0,05
SoilGrids	ocd_5-15cm_Q0p5	Densidade de carbono orgânico em hg/m ³ para centímetros 5 a 15 da superfície – predição para quantil 0,5
SoilGrids	ocd_5-15cm_Q0p95	Densidade de carbono orgânico em hg/m ³ para centímetros 5 a 15 da superfície – predição para quantil 0,95
SoilGrids	ocd_15-30cm_mean	Densidade de carbono orgânico em hg/m ³ para centímetros 15 a 30 da superfície – média
SoilGrids	ocd_15-30cm_Q0p05	Densidade de carbono orgânico em hg/m ³ para centímetros 15 a 30 da superfície – predição para quantil 0,05

SoilGrids	ocd_15-30cm_Q0p5	Densidade de carbono orgânico em hg/m^3 para centímetros 15 a 30 da superfície – predição para quantil 0,5
SoilGrids	ocd_15-30cm_Q0p95	Densidade de carbono orgânico em hg/m^3 para centímetros 15 a 30 da superfície – predição para quantil 0,95
SoilGrids	ocd_30-60cm_mean	Densidade de carbono orgânico em hg/m^3 para centímetros 15 a 30 da superfície – média
SoilGrids	ocd_30-60cm_Q0p05	Densidade de carbono orgânico em hg/m^3 para centímetros 30 a 60 da superfície – predição para quantil 0,05
SoilGrids	ocd_30-60cm_Q0p5	Densidade de carbono orgânico em hg/m^3 para centímetros 30 a 60 da superfície – predição para quantil 0,5
SoilGrids	ocd_30-60cm_Q0p95	Densidade de carbono orgânico em hg/m^3 para centímetros 30 a 60 da superfície – predição para quantil 0,95
SoilGrids	ocd_60-100cm_mean	Densidade de carbono orgânico em hg/m^3 para centímetros 60 a 100 da superfície – média
SoilGrids	ocd_60-100cm_Q0p05	Densidade de carbono orgânico em hg/m^3 para centímetros 60 a 100 da superfície – predição para quantil 0,05
SoilGrids	ocd_60-100cm_Q0p5	Densidade de carbono orgânico em hg/m^3 para centímetros 60 a 100 da superfície – predição para quantil 0,5
SoilGrids	ocd_60-100cm_Q0p95	Densidade de carbono orgânico em hg/m^3 para centímetros 60 a 100 da superfície – predição para quantil 0,95
SoilGrids	ocd_100-200cm_mean	Densidade de carbono orgânico em hg/m^3 para centímetros 100 a 200 da superfície – média
SoilGrids	ocd_100-200cm_Q0p05	Densidade de carbono orgânico em hg/m^3 para centímetros 100 a 200 da superfície – predição para quantil 0,05

SoilGrids	ocd_100-200cm_Q0p5	Densidade de carbono orgânico em hg/m^3 para centímetros 100 a 200 da superfície – predição para quantil 0,5
SoilGrids	ocd_100-200cm_Q0p95	Densidade de carbono orgânico em hg/m^3 para centímetros 100 a 200 da superfície – predição para quantil 0,95
SoilGrids	ocs_0-30cm_mean	Estoques de carbono orgânico em t/ha para centímetros 0 a 30 da superfície - média
SoilGrids	ocs_0-30cm_Q0p05	Estoques de carbono orgânico em t/ha para centímetros 0 a 30 da superfície – predição para quantil 0,05
SoilGrids	ocs_0-30cm_Q0p5	Estoques de carbono orgânico em t/ha para centímetros 0 a 30 da superfície – predição para quantil 0,5
SoilGrids	ocs_0-30cm_Q0p95	Estoques de carbono orgânico em t/ha para centímetros 0 a 30 da superfície – predição para quantil 0,95
SoilGrids	phh2o_0-5cm_mean	pH do solo em $\text{pH} \times 10$ para centímetros 0 a 5 da superfície - média
SoilGrids	phh2o_0-5cm_Q0p05	pH do solo em $\text{pH} \times 10$ para centímetros 0 a 5 da superfície – predição para quantil 0,05
SoilGrids	phh2o_0-5cm_Q0p5	pH do solo em $\text{pH} \times 10$ para centímetros 0 a 5 da superfície – predição para quantil 0,5
SoilGrids	phh2o_0-5cm_Q0p95	pH do solo em $\text{pH} \times 10$ para centímetros 0 a 5 da superfície – predição para quantil 0,95
SoilGrids	phh2o_5-15cm_mean	pH do solo em $\text{pH} \times 10$ para centímetros 5 a 15 da superfície – média
SoilGrids	phh2o_5-15cm_Q0p05	pH do solo em $\text{pH} \times 10$ para centímetros 5 a 15 da superfície – predição para quantil 0,05
SoilGrids	phh2o_5-15cm_Q0p5	pH do solo em $\text{pH} \times 10$ para centímetros 5 a 15 da superfície – predição para quantil 0,5
SoilGrids	phh2o_5-15cm_Q0p95	pH do solo em $\text{pH} \times 10$ para centímetros 5 a 15 da superfície – predição para quantil 0,95
SoilGrids	phh2o_15-30cm_mean	pH do solo em $\text{pH} \times 10$ para centímetros 15 a 30 da superfície – média
SoilGrids	phh2o_15-30cm_Q0p05	pH do solo em $\text{pH} \times 10$ para centímetros 15 a 30 da superfície – predição para quantil 0,05

SoilGrids	phh2o_15-30cm_Q0p5	pH do solo em pHx10 para centímetros 15 a 30 da superfície – predição para quantil 0,5
SoilGrids	phh2o_15-30cm_Q0p95	pH do solo em pHx10 para centímetros 15 a 30 da superfície – predição para quantil 0,95
SoilGrids	phh2o_30-60cm_mean	pH do solo em pHx10 para centímetros 15 a 30 da superfície – média
SoilGrids	phh2o_30-60cm_Q0p05	pH do solo em pHx10 para centímetros 30 a 60 da superfície – predição para quantil 0,05
SoilGrids	phh2o_30-60cm_Q0p5	pH do solo em pHx10 para centímetros 30 a 60 da superfície – predição para quantil 0,5
SoilGrids	phh2o_30-60cm_Q0p95	pH do solo em pHx10 para centímetros 30 a 60 da superfície – predição para quantil 0,95
SoilGrids	phh2o_60-100cm_mean	pH do solo em pHx10 para centímetros 60 a 100 da superfície – média
SoilGrids	phh2o_60-100cm_Q0p05	pH do solo em pHx10 para centímetros 60 a 100 da superfície – predição para quantil 0,05
SoilGrids	phh2o_60-100cm_Q0p5	pH do solo em pHx10 para centímetros 60 a 100 da superfície – predição para quantil 0,5
SoilGrids	phh2o_60-100cm_Q0p95	pH do solo em pHx10 para centímetros 60 a 100 da superfície – predição para quantil 0,95
SoilGrids	phh2o_100-200cm_mean	pH do solo em pHx10 para centímetros 100 a 200 da superfície – média
SoilGrids	phh2o_100-200cm_Q0p05	pH do solo em pHx10 para centímetros 100 a 200 da superfície – predição para quantil 0,05
SoilGrids	phh2o_100-200cm_Q0p5	pH do solo em pHx10 para centímetros 100 a 200 da superfície – predição para quantil 0,5
SoilGrids	phh2o_100-200cm_Q0p95	pH do solo em pHx10 para centímetros 100 a 200 da superfície – predição para quantil 0,95
SoilGrids	sand_0-5cm_mean	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 0 a 5 da superfície - média
SoilGrids	sand_0-5cm_Q0p05	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 0 a 5 da superfície – predição para quantil 0,05

SoilGrids	sand_0-5cm_Q0p5	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 0 a 5 da superfície – predição para quantil 0,5
SoilGrids	sand_0-5cm_Q0p95	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 0 a 5 da superfície – predição para quantil 0,95
SoilGrids	sand_5-15cm_mean	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 5 a 15 da superfície – média
SoilGrids	sand_5-15cm_Q0p05	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 5 a 15 da superfície – predição para quantil 0,05
SoilGrids	sand_5-15cm_Q0p5	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 5 a 15 da superfície – predição para quantil 0,5
SoilGrids	sand_5-15cm_Q0p95	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 5 a 15 da superfície – predição para quantil 0,95
SoilGrids	sand_15-30cm_mean	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 15 a 30 da superfície – média
SoilGrids	sand_15-30cm_Q0p05	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 15 a 30 da superfície – predição para quantil 0,05
SoilGrids	sand_15-30cm_Q0p5	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 15 a 30 da superfície – predição para quantil 0,5
SoilGrids	sand_15-30cm_Q0p95	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 15 a 30 da superfície – predição para quantil 0,95

SoilGrids	sand_30-60cm_mean	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 15 a 30 da superfície – média
SoilGrids	sand_30-60cm_Q0p05	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 30 a 60 da superfície – predição para quantil 0,05
SoilGrids	sand_30-60cm_Q0p5	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 30 a 60 da superfície – predição para quantil 0,5
SoilGrids	sand_30-60cm_Q0p95	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 30 a 60 da superfície – predição para quantil 0,95
SoilGrids	sand_60-100cm_mean	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 60 a 100 da superfície – média
SoilGrids	sand_60-100cm_Q0p05	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 60 a 100 da superfície – predição para quantil 0,05
SoilGrids	sand_60-100cm_Q0p5	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 60 a 100 da superfície – predição para quantil 0,5
SoilGrids	sand_60-100cm_Q0p95	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 60 a 100 da superfície – predição para quantil 0,95
SoilGrids	sand_100-200cm_mean	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 100 a 200 da superfície – média
SoilGrids	sand_100-200cm_Q0p05	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 100 a 200 da superfície – predição para quantil 0,05

SoilGrids	sand_100-200cm_Q0p5	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 100 a 200 da superfície – predição para quantil 0,5
SoilGrids	sand_100-200cm_Q0p95	Proportion of sand particles (>0.05 mm) in the fine earth fraction g/kg para centímetros 100 a 200 da superfície – predição para quantil 0,95
SoilGrids	silt_0-5cm_mean	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 0 a 5 da superfície - média
SoilGrids	silt_0-5cm_Q0p05	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 0 a 5 da superfície – predição para quantil 0,05
SoilGrids	silt_0-5cm_Q0p5	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 0 a 5 da superfície – predição para quantil 0,5
SoilGrids	silt_0-5cm_Q0p95	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 0 a 5 da superfície – predição para quantil 0,95
SoilGrids	silt_5-15cm_mean	Proporção de partículas de areia (>0,05 mm) na fração fina do solo em g/kg para centímetros 5 a 15 da superfície – média
SoilGrids	silt_5-15cm_Q0p05	Proporção de partículas de areia (>0,05 mm) na fração fina do solo em g/kg para centímetros 5 a 15 da superfície – predição para quantil 0,05
SoilGrids	silt_5-15cm_Q0p5	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 5 a 15 da superfície – predição para quantil 0,5
SoilGrids	silt_5-15cm_Q0p95	Proporção de partículas de areia (>0,05 mm) na fração fina do solo em g/kg para centímetros 5 a 15 da superfície – predição para quantil 0,95

SoilGrids	silt_15-30cm_mean	Proporção de partículas de areia (>0,05 mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – média
SoilGrids	silt_15-30cm_Q0p05	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – predição para quantil 0,05
SoilGrids	silt_15-30cm_Q0p5	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – predição para quantil 0,5
SoilGrids	silt_15-30cm_Q0p95	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – predição para quantil 0,95
SoilGrids	silt_30-60cm_mean	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 15 a 30 da superfície – média
SoilGrids	silt_30-60cm_Q0p05	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 30 a 60 da superfície – predição para quantil 0,05
SoilGrids	silt_30-60cm_Q0p5	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 30 a 60 da superfície – predição para quantil 0,5
SoilGrids	silt_30-60cm_Q0p95	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 30 a 60 da superfície – predição para quantil 0,95
SoilGrids	silt_60-100cm_mean	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 60 a 100 da superfície – média
SoilGrids	silt_60-100cm_Q0p05	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 60 a 100 da superfície – predição para quantil 0,05

SoilGrids	silt_60-100cm_Q0p5	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 60 a 100 da superfície – predição para quantil 0,5
SoilGrids	silt_60-100cm_Q0p95	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 60 a 100 da superfície – predição para quantil 0,95
SoilGrids	silt_100-200cm_mean	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 100 a 200 da superfície – média
SoilGrids	silt_100-200cm_Q0p05	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 100 a 200 da superfície – predição para quantil 0,05
SoilGrids	silt_100-200cm_Q0p5	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 100 a 200 da superfície – predição para quantil 0,5
SoilGrids	silt_100-200cm_Q0p95	Proporção de partículas de silte (>0,05 mm) na fração fina do solo em g/kg para centímetros 100 a 200 da superfície – predição para quantil 0,95
SoilGrids	soc_0-5cm_mean	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 0 a 5 da superfície - média
SoilGrids	soc_0-5cm_Q0p05	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 0 a 5 da superfície – predição para quantil 0,05
SoilGrids	soc_0-5cm_Q0p5	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 0 a 5 da superfície – predição para quantil 0,5
SoilGrids	soc_0-5cm_Q0p95	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 0 a 5 da superfície – predição para quantil 0,95

SoilGrids	soc_5-15cm_mean	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 5 a 15 da superfície – média
SoilGrids	soc_5-15cm_Q0p05	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 5 a 15 da superfície – predição para quantil 0,05
SoilGrids	soc_5-15cm_Q0p5	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 5 a 15 da superfície – predição para quantil 0,5
SoilGrids	soc_5-15cm_Q0p95	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 5 a 15 da superfície – predição para quantil 0,95
SoilGrids	soc_15-30cm_mean	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 15 a 30 da superfície – média
SoilGrids	soc_15-30cm_Q0p05	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 15 a 30 da superfície – predição para quantil 0,05
SoilGrids	soc_15-30cm_Q0p5	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 15 a 30 da superfície – predição para quantil 0,5
SoilGrids	soc_15-30cm_Q0p95	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 15 a 30 da superfície – predição para quantil 0,95
SoilGrids	soc_30-60cm_mean	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 15 a 30 da superfície – média
SoilGrids	soc_30-60cm_Q0p05	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 30 a 60 da superfície – predição para quantil 0,05

SoilGrids	soc_30-60cm_Q0p5	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 30 a 60 da superfície – predição para quantil 0,5
SoilGrids	soc_30-60cm_Q0p95	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 30 a 60 da superfície – predição para quantil 0,95
SoilGrids	soc_60-100cm_mean	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 60 a 100 da superfície – média
SoilGrids	soc_60-100cm_Q0p05	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 60 a 100 da superfície – predição para quantil 0,05
SoilGrids	soc_60-100cm_Q0p5	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 60 a 100 da superfície – predição para quantil 0,5
SoilGrids	soc_60-100cm_Q0p95	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 60 a 100 da superfície – predição para quantil 0,95
SoilGrids	soc_100-200cm_mean	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 100 a 200 da superfície – média
SoilGrids	soc_100-200cm_Q0p05	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 100 a 200 da superfície – predição para quantil 0,05
SoilGrids	soc_100-200cm_Q0p5	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 100 a 200 da superfície – predição para quantil 0,5
SoilGrids	soc_100-200cm_Q0p95	Conteúdo de carbono orgânico do solo na fração fina do solo em dg/kg para centímetros 100 a 200 da superfície – predição para quantil 0,95

NasaPower	T2M 2010a2019ANN	Média anual da temperatura média a 2 metros acima do solo para o período de 2010 a 2019.
NasaPower	RH2M 2010a2019ANN	Média anual da umidade relativa a 2 metros acima do solo para o período de 2010 a 2019.
NasaPower	WS2M 2010a2019ANN	Média anual da velocidade do vento a 2 metros acima do solo para o período de 2010 a 2019.
NasaPower	T2MDEW 2010a2019ANN	Média anual da temperatura do ponto de orvalho a 2 metros acima do solo para o período de 2010 a 2019.
NasaPower	T2M MAX 2010a2019ANN	Média anual da temperatura máxima a 2 metros acima do solo para o período de 2010 a 2019.
NasaPower	T2M MIN 2010a2019ANN	Média anual da temperatura mínima a 2 metros acima do solo para o período de 2010 a 2019.
NasaPower	PRECTOTCORR 2010a2019ANN	Precipitação total corrigida para o período de 2010 a 2019. O "CORR" indica que esta medição foi corrigida ou ajustada de alguma forma.
NasaPower	ALLSKY SFC LW DWN 2010a2019ANN	Radiação descendente de onda longa na superfície sob todas as condições de céu para o período de 2010 a 2019. Média ao longo da década.
NasaPower	ALLSKY SFC SW DWN 2010a2019ANN	Radiação descendente de onda curta na superfície sob todas as condições de céu para o período de 2010 a 2019. Média ao longo da década.
NasaPower	alt 2010a2019ANN	Altitude média do local para o qual os dados foram registrados durante o período de 2010 a 2019.
NasaPower	T2M 2010a2019APR	Temperatura média a 2 metros acima do solo durante o mês de abril para cada ano de 2010 a 2019.
NasaPower	RH2M 2010a2019APR	Umidade relativa média a 2 metros acima do solo durante abril para os anos especificados.

NasaPower	WS2M 2010a2019APR		Velocidade média do vento a 2 metros acima do solo durante abril ao longo do período de 2010 a 2019.
NasaPower	T2MDEW 2010a2019APR		Temperatura média do ponto de orvalho a 2 metros durante abril para cada ano no intervalo especificado. O ponto de orvalho é um indicador de umidade atmosférica.
NasaPower	T2M 2010a2019APR	MAX	Temperatura máxima média a 2 metros acima do solo durante abril para os anos de 2010 a 2019.
NasaPower	T2M 2010a2019APR	MIN	Temperatura mínima média a 2 metros acima do solo durante abril durante o mesmo período.
NasaPower	PRECTOTCORR 2010a2019APR		Precipitação total corrigida registrada durante abril de cada ano de 2010 a 2019. As correções normalmente ajustam diversos fatores para garantir precisão.
NasaPower	ALLSKY SFC LW DWN 2010a2019APR		Radiação descendente de onda longa na superfície sob todas as condições de céu durante abril de cada ano na década. Mede a radiação de onda longa que alcança a superfície terrestre.
NasaPower	ALLSKY SFC SW DWN 2010a2019APR		Radiação descendente de onda curta na superfície sob todas as condições de céu durante abril para cada ano de 2010 a 2019. Mede a radiação de onda curta (como a luz solar) que alcança a superfície.
NasaPower	alt 2010a2019APR		Representa a altitude média do(s) local(is) de onde os dados de abril foram coletados ao longo desses anos.
NasaPower	T2M 2010a2019AUG		Temperatura média a 2 metros acima do solo durante o mês de agosto para cada ano de 2010 a 2019.
NasaPower	RH2M 2010a2019AUG		Umidade relativa média a 2 metros acima do solo registrada durante agosto ao longo dos anos especificados.

NasaPower	WS2M 2010a2019AUG		Velocidade média do vento a 2 metros acima do solo medida durante agosto para os anos de 2010 a 2019.
NasaPower	T2MDEW 2010a2019AUG		Temperatura média do ponto de orvalho a 2 metros durante agosto para os anos de 2010 a 2019. O ponto de orvalho dá uma indicação do conteúdo de umidade no ar.
NasaPower	T2M 2010a2019AUG	MAX	Temperatura máxima média a 2 metros acima do solo durante agosto para a década de interesse.
NasaPower	T2M 2010a2019AUG	MIN	Temperatura mínima média a 2 metros acima do solo durante agosto durante o mesmo período.
NasaPower	PRECTOTCORR 2010a2019AUG		Precipitação total corrigida que ocorreu durante agosto de cada ano de 2010 a 2019. "Corrigida" implica ajustes para anomalias ou erros de medição.
NasaPower	ALLSKY SFC LW DWN 2010a2019AUG		Radiação descendente de onda longa na superfície sob todas as condições de céu durante o mês de agosto ao longo dos anos de 2010 a 2019. Reflete a radiação de onda longa que alcança a superfície terrestre.
NasaPower	ALLSKY SFC SW DWN 2010a2019AUG		Radiação descendente de onda curta na superfície sob todas as condições de céu durante agosto para a década em questão.
NasaPower	alt 2010a2019AUG		Representa a altitude média do(s) local(is) de onde os dados de agosto foram coletados ao longo desses anos.
NasaPower	T2M 2010a2019DEC		Temperatura média a 2 metros acima do solo durante o mês de dezembro para cada ano de 2010 a 2019.
NasaPower	RH2M 2010a2019DEC		Umidade relativa média a 2 metros acima do solo durante dezembro para o período especificado.
NasaPower	WS2M 2010a2019DEC		Velocidade média do vento a 2 metros acima do solo durante dezembro ao longo dos anos de 2010 a 2019.

NasaPower	T2MDEW 2010a2019DEC		Temperatura média do ponto de orvalho a 2 metros durante dezembro para os anos de 2010 a 2019. A temperatura do ponto de orvalho é uma medida de umidade atmosférica.
NasaPower	T2M 2010a2019DEC	MAX	Temperatura máxima média a 2 metros acima do solo durante dezembro para a década em questão.
NasaPower	T2M 2010a2019DEC	MIN	Temperatura mínima média a 2 metros acima do solo durante dezembro durante o mesmo período.
NasaPower	PRECTOTCORR 2010a2019DEC		Precipitação total corrigida observada durante dezembro de cada ano de 2010 a 2019. As correções normalmente envolvem ajustes nos dados brutos para eliminar vieses ou erros conhecidos.
NasaPower	ALLSKY SFC LW DWN 2010a2019DEC		Radiação descendente de onda longa na superfície sob todas as condições de céu para dezembro ao longo dos anos de 2010 a 2019. Reflete a radiação de onda longa que alcança a superfície terrestre.
NasaPower	ALLSKY SFC SW DWN 2010a2019DEC		Radiação descendente de onda curta na superfície sob todas as condições de céu durante dezembro para cada um dos anos de 2010 a 2019.
NasaPower	alt 2010a2019DEC		Representa a altitude média do(s) local(is) de onde os dados de dezembro foram coletados ao longo desses anos.
NasaPower	T2M 2010a2019FEB		Temperatura média a 2 metros acima do solo durante o mês de fevereiro para cada ano de 2010 a 2019.
NasaPower	RH2M 2010a2019FEB		Umidade relativa média a 2 metros acima do solo registrada durante fevereiro ao longo dos anos especificados.
NasaPower	WS2M 2010a2019FEB		Velocidade média do vento a 2 metros acima do solo medida durante fevereiro para os anos de 2010 a 2019.

NasaPower	T2MDEW 2010a2019FEB		Temperatura média do ponto de orvalho a 2 metros durante fevereiro para os anos de 2010 a 2019. A temperatura do ponto de orvalho é um indicador importante de umidade atmosférica.
NasaPower	T2M 2010a2019FEB	MAX	Temperatura máxima média a 2 metros acima do solo durante fevereiro para os anos de 2010 a 2019.
NasaPower	T2M 2010a2019FEB	MIN	Temperatura mínima média a 2 metros acima do solo durante fevereiro durante o período especificado.
NasaPower	PRECTOTCORR 2010a2019FEB		Precipitação total corrigida registrada durante fevereiro de cada ano de 2010 a 2019. "Corrigida" implica que os dados foram ajustados para corrigir erros ou vieses de medição.
NasaPower	ALLSKY SFC LW DWN 2010a2019FEB		Radiação descendente de onda longa na superfície sob todas as condições de céu para o mês de fevereiro ao longo dos anos de 2010 a 2019. Reflete a radiação de onda longa que alcança a superfície terrestre.
NasaPower	ALLSKY SFC SW DWN 2010a2019FEB		Radiação descendente de onda curta na superfície sob todas as condições de céu durante fevereiro para cada um dos anos de 2010 a 2019.
NasaPower	alt 2010a2019FEB		Representa a altitude média do(s) local(is) de onde os dados de fevereiro foram coletados ao longo desses anos.
NasaPower	T2M 2010a2019JAN		Temperatura média a 2 metros acima do solo durante o mês de janeiro para cada ano de 2010 a 2019.
NasaPower	RH2M 2010a2019JAN		Umidade relativa média a 2 metros acima do solo durante janeiro para o período especificado.
NasaPower	WS2M 2010a2019JAN		Velocidade média do vento a 2 metros acima do solo medida durante janeiro para os anos de 2010 a 2019.

NasaPower	T2MDEW 2010a2019JAN		Temperatura média do ponto de orvalho a 2 metros durante janeiro para os anos de 2010 a 2019. O ponto de orvalho fornece uma medida da quantidade de umidade no ar.
NasaPower	T2M 2010a2019JAN	MAX	Temperatura máxima média a 2 metros acima do solo durante janeiro para a década de interesse.
NasaPower	T2M 2010a2019JAN	MIN	Temperatura mínima média a 2 metros acima do solo durante janeiro durante o período especificado.
NasaPower	PRECTOTCORR 2010a2019JAN		Precipitação total corrigida observada durante janeiro de cada ano de 2010 a 2019. O aspecto "corrigido" normalmente aborda quaisquer anomalias ou erros de medição.
NasaPower	ALLSKY SFC LW DWN 2010a2019JAN		Radiação descendente de onda longa na superfície sob todas as condições de céu para janeiro ao longo dos anos de 2010 a 2019. Reflete a radiação de onda longa que alcança a superfície terrestre.
NasaPower	ALLSKY SFC SW DWN 2010a2019JAN		Radiação descendente de onda curta na superfície sob todas as condições de céu durante janeiro para cada um dos anos de 2010 a 2019.
NasaPower	alt 2010a2019JAN		Representa a altitude média do(s) local(is) de onde os dados de janeiro foram coletados ao longo desses anos.
NasaPower	T2M 2010a2019JUL		Temperatura média a 2 metros acima do solo durante o mês de julho para cada ano de 2010 a 2019.
NasaPower	RH2M 2010a2019JUL		Umidade relativa média a 2 metros acima do solo durante julho para o período especificado.
NasaPower	WS2M 2010a2019JUL		Velocidade média do vento a 2 metros acima do solo medida durante julho para os anos de 2010 a 2019.

NasaPower	T2MDEW 2010a2019JUL		Temperatura média do ponto de orvalho a 2 metros durante julho para os anos de 2010 a 2019. A temperatura do ponto de orvalho é uma medida de umidade atmosférica.
NasaPower	T2M 2010a2019JUL	MAX	Temperatura máxima média a 2 metros acima do solo durante julho para os anos de 2010 a 2019.
NasaPower	T2M 2010a2019JUL	MIN	Temperatura mínima média a 2 metros acima do solo durante julho durante o período especificado.
NasaPower	PRECTOTCORR 2010a2019JUL		Precipitação total corrigida registrada durante julho de cada ano de 2010 a 2019. O termo "corrigida" sugere que os dados foram ajustados para contabilizar erros ou vieses de medição.
NasaPower	ALLSKY SFC LW DWN 2010a2019JUL		Radiação descendente de onda longa na superfície sob todas as condições de céu para julho ao longo dos anos de 2010 a 2019. Reflete a radiação de onda longa que alcança a superfície terrestre.
NasaPower	ALLSKY SFC SW DWN 2010a2019JUL		Radiação descendente de onda curta na superfície sob todas as condições de céu durante julho para cada um dos anos de 2010 a 2019.
NasaPower	alt 2010a2019JUL		Representa a altitude média do(s) local(is) de onde os dados de julho foram coletados ao longo desses anos.
NasaPower	T2M 2010a2019JUN		Temperatura média a 2 metros acima do solo durante o mês de junho para cada ano de 2010 a 2019.
NasaPower	RH2M 2010a2019JUN		Umidade relativa média a 2 metros acima do solo durante junho para o período especificado.
NasaPower	WS2M 2010a2019JUN		Velocidade média do vento a 2 metros acima do solo medida durante junho para os anos de 2010 a 2019.

NasaPower	T2MDEW 2010a2019JUN		Temperatura média do ponto de orvalho a 2 metros durante junho para os anos de 2010 a 2019. O ponto de orvalho é um indicador da quantidade de umidade no ar.
NasaPower	T2M 2010a2019JUN	MAX	Temperatura máxima média a 2 metros acima do solo durante junho para os anos de 2010 a 2019.
NasaPower	T2M 2010a2019JUN	MIN	Temperatura mínima média a 2 metros acima do solo durante junho durante o período especificado.
NasaPower	PRECTOTCORR 2010a2019JUN		Precipitação total corrigida registrada durante junho de cada ano de 2010 a 2019. O termo "corrigida" indica ajustes feitos para contabilizar possíveis erros ou vieses de medição.
NasaPower	ALLSKY SFC LW DWN 2010a2019JUN		Radiação descendente de onda longa na superfície sob todas as condições de céu para junho ao longo dos anos de 2010 a 2019. Reflete a radiação de onda longa que alcança a superfície terrestre.
NasaPower	ALLSKY SFC SW DWN 2010a2019JUN		Radiação descendente de onda curta na superfície sob todas as condições de céu durante junho para cada um dos anos de 2010 a 2019.
NasaPower	alt 2010a2019JUN		Representa a altitude média do(s) local(is) de onde os dados de junho foram coletados ao longo desses anos.
NasaPower	T2M 2010a2019MAR		Temperatura média a 2 metros acima do solo durante o mês de março para cada ano de 2010 a 2019.
NasaPower	RH2M 2010a2019MAR		Umidade relativa média a 2 metros acima do solo durante março para o período especificado.
NasaPower	WS2M 2010a2019MAR		Velocidade média do vento a 2 metros acima do solo medida durante março para os anos de 2010 a 2019.

NasaPower	T2MDEW 2010a2019MAR		Temperatura média do ponto de orvalho a 2 metros durante março para os anos de 2010 a 2019. O ponto de orvalho é uma medida de quanto de umidade está no ar.
NasaPower	T2M 2010a2019MAR	MAX	Temperatura máxima média a 2 metros acima do solo durante março para os anos de 2010 a 2019.
NasaPower	T2M 2010a2019MAR	MIN	Temperatura mínima média a 2 metros acima do solo durante março durante o período especificado.
NasaPower	PRECTOTCORR 2010a2019MAR		Precipitação total corrigida registrada durante março de cada ano de 2010 a 2019. O termo "corrigida" geralmente se refere a ajustes feitos nos dados para corrigir imprecisões ou vieses conhecidos.
NasaPower	ALLSKY SFC LW DWN 2010a2019MAR		Radiação descendente de onda longa na superfície sob todas as condições de céu para março ao longo dos anos de 2010 a 2019. Reflete a radiação de onda longa que alcança a superfície terrestre.
NasaPower	ALLSKY SFC SW DWN 2010a2019MAR		Radiação descendente de onda curta na superfície sob todas as condições de céu durante março para cada um dos anos de 2010 a 2019.
NasaPower	alt 2010a2019MAR		Representa a altitude média do(s) local(is) de onde os dados de março foram coletados ao longo desses anos.
NasaPower	T2M 2010a2019MAY		Temperatura média a 2 metros acima do solo durante o mês de maio para cada ano de 2010 a 2019.
NasaPower	RH2M 2010a2019MAY		Umidade relativa média a 2 metros acima do solo durante maio para o período especificado.
NasaPower	WS2M 2010a2019MAY		Velocidade média do vento a 2 metros acima do solo medida durante maio para os anos de 2010 a 2019.

NasaPower	T2MDEW 2010a2019MAY		Temperatura média do ponto de orvalho a 2 metros durante maio para os anos de 2010 a 2019. A temperatura do ponto de orvalho é uma medida de umidade atmosférica.
NasaPower	T2M 2010a2019MAY	MAX	Temperatura máxima média a 2 metros acima do solo durante maio para os anos de 2010 a 2019.
NasaPower	T2M 2010a2019MAY	MIN	Temperatura mínima média a 2 metros acima do solo durante maio durante o período especificado.
NasaPower	PRECTOTCORR 2010a2019MAY		Precipitação total corrigida registrada durante maio de cada ano de 2010 a 2019. As correções geralmente abordam imprecisões ou vieses de medição.
NasaPower	ALLSKY SFC LW DWN 2010a2019MAY		Radiação descendente de onda longa na superfície sob todas as condições de céu para maio ao longo dos anos de 2010 a 2019. Reflete a radiação de onda longa que alcança a superfície terrestre.
NasaPower	ALLSKY SFC SW DWN 2010a2019MAY		Radiação descendente de onda curta na superfície sob todas as condições de céu durante maio para cada um dos anos de 2010 a 2019.
NasaPower	alt 2010a2019MAY		Representa a altitude média do(s) local(is) de onde os dados de maio foram coletados ao longo desses anos.
NasaPower	T2M 2010a2019NOV		Temperatura média a 2 metros acima do solo durante o mês de novembro para cada ano de 2010 a 2019.
NasaPower	RH2M 2010a2019NOV		Umidade relativa média a 2 metros acima do solo durante novembro para o período especificado.
NasaPower	WS2M 2010a2019NOV		Velocidade média do vento a 2 metros acima do solo medida durante novembro para os anos de 2010 a 2019.
NasaPower	T2MDEW 2010a2019NOV		Temperatura média do ponto de orvalho a 2 metros durante novembro para os anos de 2010 a 2019. O ponto de orvalho indica a quantidade de umidade no ar.

NasaPower	T2M 2010a2019NOV	MAX	Temperatura máxima média a 2 metros acima do solo durante novembro para os anos de 2010 a 2019.
NasaPower	T2M 2010a2019NOV	MIN	Temperatura mínima média a 2 metros acima do solo durante novembro durante o período especificado.
NasaPower	PRECTOTCORR 2010a2019NOV		Precipitação total corrigida registrada durante novembro de cada ano de 2010 a 2019. O termo "corrigida" geralmente se refere a ajustes feitos nos dados para corrigir imprecisões ou vieses conhecidos.
NasaPower	ALLSKY SFC LW DWN 2010a2019NOV		Radiação descendente de onda longa na superfície sob todas as condições de céu para novembro ao longo dos anos de 2010 a 2019. Reflete a radiação de onda longa que alcança a superfície terrestre.
NasaPower	ALLSKY SFC SW DWN 2010a2019NOV		Radiação descendente de onda curta na superfície sob todas as condições de céu durante novembro para cada um dos anos de 2010 a 2019.
NasaPower	alt 2010a2019NOV		Representa a altitude média do(s) local(is) de onde os dados de novembro foram coletados ao longo desses anos.
NasaPower	T2M 2010a2019OCT		Temperatura média a 2 metros acima do solo durante o mês de outubro para cada ano de 2010 a 2019.
NasaPower	RH2M 2010a2019OCT		Umidade relativa média a 2 metros acima do solo durante outubro para o período especificado.
NasaPower	WS2M 2010a2019OCT		Velocidade média do vento a 2 metros acima do solo medida durante outubro para os anos de 2010 a 2019.
NasaPower	T2MDEW 2010a2019OCT		Temperatura média do ponto de orvalho a 2 metros durante outubro para os anos de 2010 a 2019. O ponto de orvalho é uma medida de umidade atmosférica.

NasaPower	T2M 2010a2019OCT	MAX	Temperatura máxima média a 2 metros acima do solo durante outubro para os anos de 2010 a 2019.
NasaPower	T2M 2010a2019OCT	MIN	Temperatura mínima média a 2 metros acima do solo durante outubro durante o período especificado.
NasaPower	PRECTOTCORR 2010a2019OCT		Precipitação total corrigida registrada durante outubro de cada ano de 2010 a 2019. Essa correção leva em conta imprecisões ou vieses de medição.
NasaPower	ALLSKY SFC LW DWN 2010a2019OCT		Radiação descendente de onda longa na superfície sob todas as condições de céu para outubro ao longo dos anos de 2010 a 2019. Isso mede a radiação de onda longa que alcança a superfície terrestre.
NasaPower	ALLSKY SFC SW DWN 2010a2019OCT		Radiação descendente de onda curta na superfície sob todas as condições de céu durante outubro para cada um dos anos de 2010 a 2019.
NasaPower	alt 2010a2019OCT		Representa a altitude média do(s) local(is) de onde os dados de outubro foram coletados ao longo desses anos.
NasaPower	T2M 2010a2019SEP		Temperatura média medida a 2 metros acima do solo durante o mês de setembro ao longo dos anos de 2010 a 2019.
NasaPower	RH2M 2010a2019SEP		Umidade relativa média a 2 metros acima do solo durante setembro ao longo do período de 2010 a 2019.
NasaPower	WS2M 2010a2019SEP		Velocidade média do vento a 2 metros acima do solo para setembro durante os anos especificados de 2010 a 2019.
NasaPower	T2MDEW 2010a2019SEP		Temperatura média do ponto de orvalho a 2 metros acima do solo em setembro para os anos de 2010 a 2019. A temperatura do ponto de orvalho é uma medida da umidade atmosférica.

NasaPower	T2M 2010a2019SEP	MAX	Temperatura máxima média a 2 metros acima do solo durante setembro para cada ano de 2010 a 2019.
NasaPower	T2M 2010a2019SEP	MIN	Temperatura mínima média a 2 metros acima do solo durante setembro para os anos de 2010 a 2019.
NasaPower	PRECTOTCORR 2010a2019SEP		Precipitação total registrada para setembro a cada ano de 2010 a 2019, corrigida para quaisquer vieses ou imprecisões conhecidos na medição.
NasaPower	ALLSKY SFC LW DWN 2010a2019SEP		Quantidade média de radiação descendente de onda longa na superfície sob todas as condições de céu durante setembro para os anos especificados de 2010 a 2019, indicando a radiação de onda longa que alcança a superfície da Terra.
NasaPower	ALLSKY SFC SW DWN 2010a2019SEP		Esta variável mede a média da radiação descendente de onda curta na superfície sob todas as condições de céu (como a luz solar) durante setembro para os anos de 2010 a 2019.
NasaPower	alt 2010a2019SEP		Representa a altitude média do(s) local(is) de onde os dados de setembro foram coletados ao longo desses anos.

Fonte: elaboração própria.