

Universidade Federal do Paraná
Setor de Ciências Exatas
Departamento de Estatística

Gabriel Agostini Orso

Modelos geoestatísticos multivariados aplicados a dados florestais

**Curitiba
2024**

Gabriel Agostini Orso

Modelos geoestatísticos multivariados aplicados a dados florestais

Trabalho de Conclusão de Curso apresentado à disciplina Laboratório B do Curso de Graduação em Estatística da Universidade Federal do Paraná, como exigência parcial para obtenção do grau de Bacharel em Estatística.

Orientador(a): Paulo Justiniano Ribeiro Junior

Curitiba
2024

Dedico esse trabalho à minha família e amigos.

Agradecimentos

Agradeço aos colegas durante esses anos de estudo. Agradeço aos meus pais por sempre me incentivarem a continuar em frente. Também agradeço aos professores que sempre nos inspiram.

Todo começo é involuntário.

Deus é o agente.

O heroe a si assiste, vário

E inconsciente.

À espada em tuas mãos achada

Teu olhar desce.

Que farei eu com esta espada?

Ergueste-a, e fez-se.

O Conde D. Henrique (Fernando Pessoa)

Resumo

A produção de florestas é uma das principais informações utilizadas para o planejamento da produção. É por meio de estimativas utilizando inventários florestais que são feitas as extrapolações para toda a área plantada. Embora confiáveis, tais estimativas assumem unidades amostrais independentes entre si, ou seja, os valores amostrados em uma parcela de inventário não carregam informação de outras parcelas. Desde o desenvolvimento da Teoria das Variáveis Regionalizadas, diversos trabalhos apontam a existência de correlação entre as parcelas de inventário florestal, o que possibilita o uso de geoestatística para a estimativa da produção. Além disso, é comum nas operações florestais a quantificação indireta do volume utilizando variáveis de mais fácil medição, como o diâmetro das árvores. Nesse contexto, se torna oportuna a utilização de um modelo geoestatístico multivariado, onde se possa contabilizar não só a correlação entre unidades amostrais, bem como a correlação entre variáveis, como diâmetro e volume. Sob essa premissa foram desenvolvidos e avaliados três modelos bivariados com diferentes estruturas de covariância entre as variáveis adotadas. Os dados para o estudo são de plantios de Teca do interior do Mato Grosso. Ao todo foram utilizadas 46 parcelas de inventário florestal, onde se coletaram informações como diâmetro e volume. Foram testadas três especificações de covariância para um modelo geoestatístico bivariado. A primeira especificação de covariância assumiu um processo espacial para cada uma das variáveis, além de um processo espacial comum às duas. A segunda especificação assumiu dependência espacial apenas para a primeira variável, porém com correlação entre variáveis na mesma unidade amostral. A terceira especificação previu apenas covariâncias marginais, em que a covariância cruzada foi derivada implicitamente, sem especificação explícita. As três especificações se ajustaram bem aos dados. Apesar de bom ajuste, uma das especificações não gerou mapas de krigagem adequados, o que reforça a escolha de estruturas de covariância propícias para esse fim. A terceira especificação permitiu a geração de mapas de krigagem e foi parcimoniosa quanto à qualidade da predição e número de parâmetros. O uso de modelos geoestatísticos bivariados permite utilizar os dados de levantamentos florestais, fornecendo melhores ajustes e predições localizadas, auxiliando na quantificação da produção em volume de plantios florestais.

Palavras-chave: Produto de kronecker. Estrutura de covariância. Função de correlação. Covariância cruzada.

Sumário

1	INTRODUÇÃO	7
1.1	Objetivo	8
2	REVISÃO DE LITERATURA	9
3	MATERIAL E MÉTODOS	10
3.1	Fonte dos dados	10
3.2	Modelos espaciais	10
3.3	Modelos testados	12
3.4	Resumo dos modelos testados	15
4	RESULTADOS E DISCUSSÃO	16
4.1	Análise Descritiva	16
4.2	Modelos univariados	17
4.3	Modelos Multivariados	18
4.4	Krigagem	22
5	CONSIDERAÇÕES FINAIS	26
	REFERÊNCIAS	27
	APÊNDICES	30

1 Introdução

A estimativa da produção de plantios florestais é uma das principais informações utilizadas para o planejamento operacional e futuro das empresas de base florestal (BOGNOLA et al., 2008). É por meio dessas estimativas, majoritariamente realizadas por inventários, que são feitas as extrapolações para toda a área plantada, bem como são construídas projeções para idades futuras.

Ainda que confiáveis, essas estimativas comumente assumem unidades amostrais independentes entre si (ROVEDA et al., 2016), *i.e.*, os valores amostrados em uma parcela de inventário não carregam informação de outras parcelas. Entretanto, desde o desenvolvimento da Teoria das Variáveis Regionalizadas (MATHERON, 1971), diversas áreas de estudo consideraram a possibilidade de que alguns processos ou variáveis podem ser correlacionadas ou similares em distâncias mais próximas. Tal constatação é muito bem observada para atributos do solo (JOHN et al., 2021), e logicamente a produção de plantios florestais é influenciada por esse processo (MELLO, 2004; MELLO et al., 2005; PELISSARI, 2015; PELISSARI et al., 2017).

O ramo da estatística que considera a dependência espacial de uma variável é comumente chamado de geoestatística (DIGGLE; RIBEIRO Jr, 2007), em que ampla parte da aplicação desses modelos busca a extrapolação ou espacialização de uma variável em uma determinada região. Tais aplicações são recorrentes na área florestal e ambiental (SALES et al., 2007; KARAHAN; ERŞAHİN, 2018).

No entanto, a estimativa da produção florestal ainda sofre com a dificuldade em medir o volume das árvores na floresta (*e.g.* $m^3.ha^{-1}$), dado o custo e demora para derrubada e medição do volume de cada árvore. Para contornar isso é comum a tomada apenas de variáveis de fácil medição, como diâmetro e altura das árvores, que em seguida são utilizadas para estimativa do volume (MACHADO; FIGUEIREDO FILHO, 2003).

De tal situação surge a necessidade de modelos capazes de relacionar espacialmente não apenas uma variável, mas sim mais de uma variável no espaço, sendo essas correlacionadas entre si (DIGGLE; RIBEIRO Jr, 2007). Nesse contexto a geoestatística multivariada é uma boa alternativa, pois se preocupa em modelar a distribuição espacial conjunta de variáveis em uma região do espaço (RIBEIRO; RIBEIRO Jr; BONAT, 2022). Outro fator importante para o modelo é a estrutura de covariância utilizada entre as variáveis, em que a literatura é diversa quanto à maneira de construção dessa estrutura (DIGGLE; RIBEIRO Jr, 2007; BOGNOLA et al., 2008; RIBEIRO; RIBEIRO Jr; BONAT, 2022).

O emprego da geoestatística multivariada é oportuno para a estimativa da produção florestal, uma vez que é possível relacionar e espacializar variáveis de difícil medição por meio de variáveis rotineiramente mensuradas em inventários florestais.

1.1 Objetivo

Propor, implementar computacionalmente e avaliar três modelos bivariados de dependência espacial para um conjunto de dados de Teca (*Tectona grandis* L.f.) de inventário florestal.

2 Revisão de Literatura

A geoestatística multivariada é uma ferramenta essencial em várias disciplinas, incluindo a ciência florestal. Ao integrar múltiplas variáveis e correlações espaciais, a geoestatística oferece uma abordagem robusta para a análise de padrões espaciais complexos (DIGGLE; RIBEIRO Jr, 2007). Elas permitem modelar a dependência espacial conjunta entre organismos e seu ambiente, proporcionando *insights* sobre as interações ecossistêmicas (ROSSI et al., 1992). Além disso, a geoestatística multivariada pode prever valores em locais não amostrados utilizando correlações espaciais entre observações, o que é de interesse para o monitoramento da qualidade e da produção de plantios florestais (SCOLFORO et al., 2015).

Na ciência florestal, a geoestatística multivariada ajuda a entender a variabilidade espacial das propriedades do solo e sua relação com a produção (MELLO, 2004; BOGNOLA et al., 2008; Castrignanò et al., 2023). Ao combinar técnicas geoestatísticas com análise multivariada, os pesquisadores podem mapear diferentes fontes de variação espacial, oferecendo informações para a gestão florestal.

Além disso, modelos geoestatísticos multivariados possuem um histórico de aplicações na ciência florestal, em grande parte para modelar variáveis dendrométricas em função de outros atributos da floresta (NANOS et al., 2004; PELISSARI; FIGUEIREDO FILHO; CALDEIRA, 2014) ou do solo e clima (SALES et al., 2007; SILVEIRA et al., 2019).

Uma das técnicas mais empregadas é a co-krigagem, que pode ser usada para mapear, por exemplo, o carbono em plantios e florestas naturais (SALES et al., 2007; PFUTZ et al., 2021). A geoestatística mostrou-se eficaz também em conjunto com imagens de sensoriamento remoto, permitindo aprimorar a estimativa de variáveis dendrométricas em plantios florestais a partir de índices de vegetação (PFUTZ et al., 2022).

Na ciência florestal, a aplicação da geoestatística tem se estendido a diversas áreas de pesquisa, incluindo a avaliação da disseminação espacial de doenças em pomares (GOTTWALD et al., 1995) e a análise de emissões de gases de efeito estufa de solos florestais (NISHINA; TAKENAKA; ISHIZUKA, 2009). Essas aplicações destacam a versatilidade da geoestatística em abordar vários desafios na ciência florestal, desde o monitoramento da contaminação ambiental até a previsão de dinâmicas ecológicas em resposta a mudanças nos padrões de uso da terra.

3 Material e Métodos

3.1 Fonte dos dados

Os dados deste estudo provêm de inventários florestais realizados em povoamentos de Teca (*Tectona grandis* L. f.) localizados no estado do Mato Grosso, com idade de 17 anos. Ao todo, foram medidas 46 parcelas quadradas de tamanhos 30m x 30m (Figura 1) em uma área de 222,45 hectares. As parcelas possuem um padrão sistemático de localização, e a menor, maior e distância média entre parcelas são de 69, 1014 e 2816 m, respectivamente. De cada parcela foram medidos os diâmetros a 1,3 m acima do solo (DAP) em centímetros e os volumes de madeira (V) em metros cúbicos a partir da utilização de fatores de forma (MACHADO; FIGUEIREDO FILHO, 2003). A obtenção do volume foi realizada via cubagem, que consiste na derrubada e seccionamento de árvores para medição do volume *in loco*.

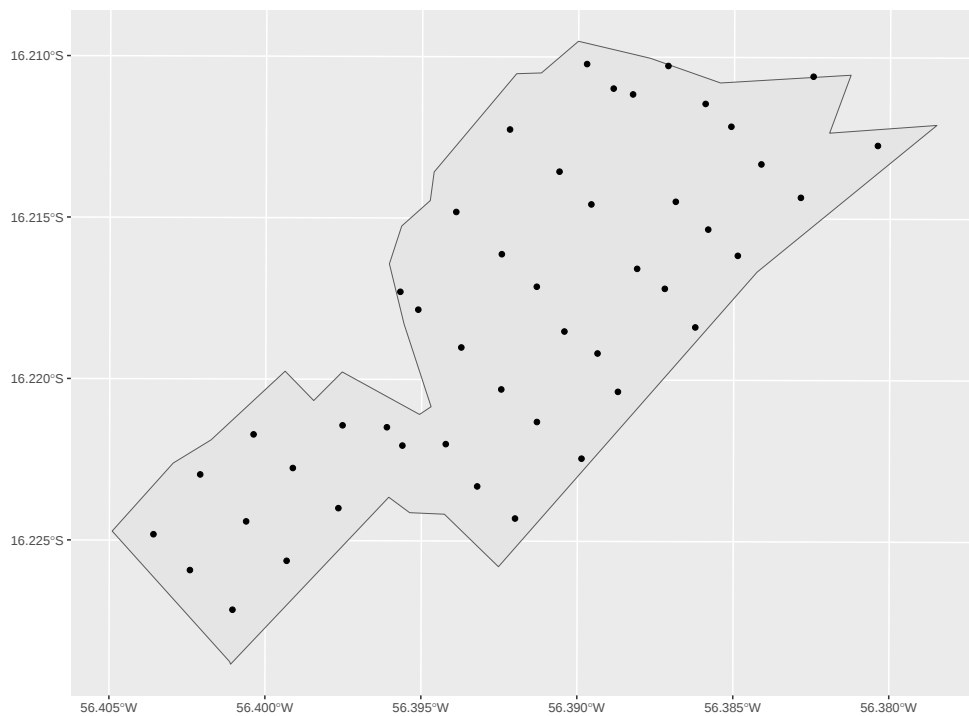


Figura 1 – Croqui de localização das parcelas de inventário.

3.2 Modelos espaciais

Modelos espaciais podem ser utilizados quando há motivação de estudar o comportamento de algumas variáveis em uma região do espaço (RIBEIRO; RIBEIRO Jr; BONAT, 2022). Seja uma amostra de tamanho n composta por unidades s_k , tal que $k = 1, 2, \dots, n$, e $s_k \in \mathbb{R}^d$, onde d é a dimensão onde as unidades estão inseridas. De cada unidade amostral s_k são medidas alguma variável y . Uma vez que cada medida é apenas uma realização de

infinitas possibilidades, tal realização pode ser descrita por um processo estocástico $Y(s_k)$, ou de forma mais geral por $Y(s)$ para qualquer $s \in \mathbb{R}^d$.

O valor esperado de tal processo é definido por $E[Y(s)]$, e a covariância de Y avaliado em dois locais diferentes s_1 e s_2 é dada por

$$Cov(Y(s_1), Y(s_2)) = E[Y(s_1) - E[Y(s_1)]] [Y(s_2) - E[Y(s_2)]]$$

O processo é ainda chamado de estacionário de segunda ordem se o primeiro e segundo momentos são estacionários, e a covariância depende apenas da distância $h \in \mathbb{R}^d$ que separa dois pontos.

$$\begin{aligned} E[Y(s)] &= E[Y(s+h)] \\ Cov[Y(s), Y(s+h)] &= \Sigma(h) \end{aligned}$$

Um processo estocástico é dito gaussiano se, para uma coleção finita de locais $s_1, s_2, \dots, s_k$, com $s_k \in \mathbb{R}^d$, a distribuição de probabilidade de $Y(s_1), Y(s_2), \dots, Y(s_k)$ é uma distribuição normal multivariada (CHILES; DELFINER, 2012). Neste trabalho, a especificação do vetor de médias e matriz de covariâncias dessa distribuição foi feita de duas formas. A primeira é ‘induzida’ a partir da combinação de termos de efeitos aleatórios espaciais e não espaciais, por exemplo

$$Y = \mu + S + \epsilon \quad (3.1)$$

, em que μ é a média da distribuição, S é um termo que descreve o processo espacial subjacente, e ϵ é um termo de erro associado a um componente de variância não estruturada e independente entre observações, $\epsilon \sim N(0, \tau^2)$. Nesse exemplo, S pode ser escrita como $S = \sigma^2 R$, em que σ^2 é um componente de variância estruturada e R é o valor padronizado do processo S .

O processo R é descrito por meio de uma função de correlação ρ , tal que $R = \rho(h, \phi)$, que depende do vetor da distância h e outros parâmetros específicos da função. O papel da função de correlação é retornar a correlação da variável entre dois pontos distintos no espaço. Existem diversos modelos propostos na literatura, em que aqui optou-se pelo uso da função de correlação exponencial dada sua simplicidade, amplo uso e generalidade. Para o modelo exponencial, *phi* maiores valores para o parâmetro suavizam o decaimento da correlação com o aumento da distância.

$$\rho(h, \phi) = e^{\frac{-h}{\phi}} \quad (3.2)$$

A partir do modelo especificado anteriormente (equação (3.1)), a distribuição normal multivariada induzida fica definida como

$$Y \sim N \left(\begin{bmatrix} \mu \\ \vdots \\ \mu \end{bmatrix}, \begin{bmatrix} \sigma^2 + \tau^2 & \sigma^2 \rho(h, \phi) & \cdots & \sigma^2 \rho(h, \phi) \\ \sigma^2 \rho(h, \phi) & \sigma^2 + \tau^2 & \cdots & \sigma^2 \rho(h, \phi) \\ \sigma^2 \rho(h, \phi) & \vdots & \ddots & \vdots \\ \sigma^2 \rho(h, \phi) & \sigma^2 \rho(h, \phi) & \cdots & \sigma^2 + \tau^2 \end{bmatrix} \right) \quad (3.3)$$

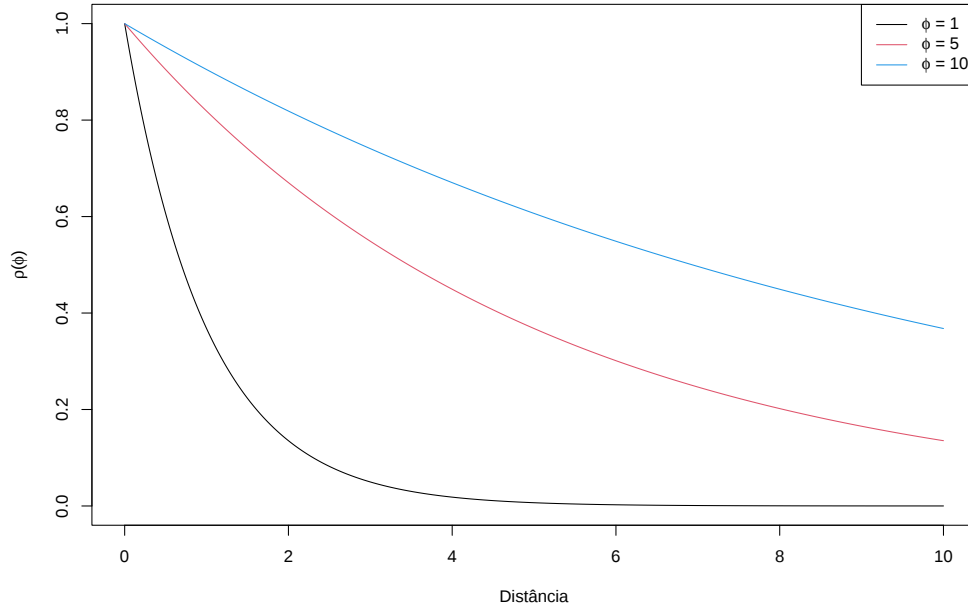


Figura 2 – Exemplos da função de correlação exponencial utilizada no estudo, para diferentes valores do parâmetro ϕ .

Tanto o vetor de médias quanto a matriz de variância-covariância (equações (3.2) e (3.3)) são consequência da especificação proposta para o modelo.

Outra maneira de especificar os parâmetros da distribuição é de forma ‘marginal’, isto é, construindo diretamente a matriz de covariâncias. A especificação de um modelo, tal qual na equação (3.1), sempre produz uma respectiva matriz de covariâncias, mas nem sempre o caminho contrário é válido, e a especificação direta da matriz de covariâncias não garante uma única expressão matemática possível, tal qual a equação (3.1).

Para o caso bivariado, além da especificação do vetor de médias e matriz de covariâncias de cada variável, é necessário também especificar uma estrutura de covariância cruzada entre variáveis (equação (3.4)). Para esse caso também são válidas as abordagens ‘induzida’ e ‘marginal’ para especificação da distribuição.

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right) \quad (3.4)$$

, em que Σ_{11} e Σ_{22} são respectivamente as matrizes de variância-covariância das variáveis Y_1 e Y_2 , também chamadas de covariâncias marginais, e Σ_{12} e Σ_{21} são as matrizes da covariância entre variáveis, também chamadas de covariâncias cruzadas (RIBEIRO; RIBEIRO Jr; BONAT, 2022).

3.3 Modelos testados

Foram utilizadas três especificações para o modelo multivariado proposto. A primeira especificação assumiu a existência de um processo espacial único para cada uma das variáveis ($S_i = \sigma_i^2 \rho_i(h, \phi_i)$), para $i = 1, 2$, e outro comum às duas variáveis testadas

($S_0 = \sigma_0^2 \rho_0(\phi_0)$) (Equação (3.5)), ou seja,

$$\begin{aligned} Y_1 &= \mu_1 + S_1 + S_0 + \epsilon_1 \\ Y_2 &= \mu_2 + S_2 + S_0 + \epsilon_2 \end{aligned} \quad (3.5)$$

, em que $Y_i \sim N(\mu_i, \Sigma_i + I\tau_i^2)$, $S_i \sim N(0, \Sigma_i)$ e $\epsilon_i \sim N(0, I\tau_i^2)$

O modelo bivariado assumido fica então com a seguinte distribuição bivariada

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \Sigma \right) \quad (3.6)$$

, em que Σ é a matriz de covariâncias, apresentada no apêndice A do presente trabalho.

A segunda especificação assume correlação entre as variáveis, porém sem dependência espacial da variável Y_2 , e com correlação entre variáveis medidas na mesma parcela, dada pelo processo F_0 (equação (3.7)).

$$\begin{aligned} Y_1 &= \mu_1 + S_1 + F_0 + \epsilon_1 \\ Y_2 &= \mu_2 + F_0 + \epsilon_2 \end{aligned} \quad (3.7)$$

Em que $F_0 = \tau_1 \tau_2 \rho_{12}$, e ρ_{12} é a correlação entre as variáveis. Dessa forma, existe uma estrutura de correlação entre as variáveis, porém não espacializada.

$$\Sigma = \left[\begin{array}{cccc|cccc} \sigma_1^2 + \tau_1^2 & \sigma_1^2 \rho_1(\phi_1) & \cdots & \sigma_1^2 \rho_1(\phi_1) & \tau_1 \tau_2 \rho_{12} & 0 & \cdots & 0 \\ \sigma_1^2 \rho_1(\phi_1) & \sigma_1^2 + \tau_1^2 & \cdots & \sigma_1^2 \rho_1(\phi_1) & 0 & \tau_1 \tau_2 \rho_{12} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_1^2 \rho_1(\phi_1) & \sigma_1^2 \rho_1(\phi_1) & \cdots & \sigma_1^2 + \tau_1^2 & 0 & \cdots & \cdots & \tau_1 \tau_2 \rho_{12} \\ \hline \tau_1 \tau_2 \rho_{12} & 0 & \cdots & 0 & \tau_2^2 & 0 & \cdots & 0 \\ 0 & \tau_1 \tau_2 \rho_{12} & \cdots & 0 & 0 & \tau_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \tau_1 \tau_2 \rho_{12} & 0 & \cdots & \cdots & \tau_2^2 \end{array} \right] \quad (3.8)$$

O terceiro modelo é especificado pela construção direta das matrizes de covariância marginal (RIBEIRO; RIBEIRO Jr; BONAT, 2022), isto é, sem especificar as covariâncias cruzadas. Ainda assim, essas podem ser obtidas indiretamente (RIBEIRO; RIBEIRO Jr; BONAT, 2022). O modelo bivariado tem distribuição na forma

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \Sigma \right) \quad (3.9)$$

Para o cálculo da matriz de covariâncias, seja Σ a matriz de covariâncias do modelo (Equação (3.10))

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \quad (3.10)$$

Em que Σ_{11} e Σ_{22} são as matrizes de covariâncias marginais (Equação (3.11)), e Σ_{12} e Σ_{21} são as matrizes de covariâncias cruzadas.

$$\Sigma_{ii} = \begin{bmatrix} \sigma_i^2 + \tau_i^2 & \sigma_i^2 \rho_i(\phi_i) & \cdots & \sigma_i^2 \rho_i(\phi_i) \\ \sigma_i^2 \rho_i(\phi_i) & \sigma_i^2 + \tau_i^2 & \cdots & \sigma_i^2 \rho_i(\phi_i) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_i^2 \rho_i(\phi_i) & \sigma_i^2 \rho_i(\phi_i) & \cdots & \sigma_i^2 + \tau_i^2 \end{bmatrix} \quad (3.11)$$

Na abordagem de Ribeiro, Ribeiro Jr e Bonat (2022) a matriz de covariâncias Σ pode ser obtida por

$$\Sigma = B \text{diag}(\tilde{\Sigma}_{11}, \tilde{\Sigma}_{22})(\Sigma_b \otimes I) B \text{diag}(\tilde{\Sigma}_{11}^\top, \tilde{\Sigma}_{22}^\top) \quad (3.12)$$

Em que $\tilde{\Sigma}_{ii}$ é a matriz triangular inferior da decomposição de Cholesky da matriz Σ_{ii} , $B \text{diag}$ indica uma matriz diagonal de blocos, Σ_b é a matriz de correlação entre as variáveis, $\otimes$ é o produto de Kronecker e I é uma matriz identidade.

$$\Sigma_b = \begin{bmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{bmatrix}$$

Dessa forma, é possível especificar uma estrutura de covariância multivariada definindo apenas as covariâncias marginais.

Feita a especificação do vetor de médias e matriz de covariância da distribuição bivariada, os modelos foram ajustados via máxima verossimilhança utilizando a função *optim* da linguagem R (R Core Team, 2023). Embora seja possível utilizar a verossimilhança da normal multivariada via pacotes como *mvtnorm* (GENZ; BRETZ, 2009), a função de custo utilizada foi otimizada para diminuir o esforço computacional, de forma que $\Sigma = LL^\top$, e L é a matriz triangular inferior da decomposição de choleski de Σ . Portanto

$$\begin{aligned} L(Y; \mu, \Sigma) &= (2\pi)^{-k/2} |\Sigma|^{-1/2} \exp \left(-\frac{1}{2} (y - \mu)^\top \Sigma^{-1} (y - \mu) \right) \\ &= (2\pi)^{-k/2} |LL^\top|^{-1/2} \exp \left(-\frac{1}{2} (y - \mu)^\top (LL^\top)^{-1} (y - \mu) \right) \\ &= (2\pi)^{-k/2} |L|^{-1} \exp \left(-\frac{1}{2} (y - \mu)^\top L^{-\top} L^{-1} (y - \mu) \right) \end{aligned}$$

Uma vez que $|LL^\top| = |L||L^\top| = |L||L| = |L|^2$

$$L(Y; \mu, \Sigma) = (2\pi)^{-k/2} (\prod l_{i,i})^{-1} \exp \left(-\frac{1}{2} (L^{-1}(y - \mu))^\top (L^{-1}(y - \mu)) \right)$$

Uma vez que $|L|$ é o produto da diagonal de L , e $(y - \mu)^\top L^{-\top} = (L^{-1}(y - \mu))^\top$

$$\log(L(Y; \mu, \Sigma)) \propto -\log(\prod l_{i,i}) - \frac{1}{2} (A^\top A)$$

Em que $A = L^{-1}(y - \mu)$, um sistema com solução facilmente encontrada via função *solve*.

Os ajustes não sofreram nenhuma transformação prévia. Tentativas anteriores de transformação de variável utilizando o método Box-Cox se mostraram insatisfatórias, com

amplitudes grandes para λ , e falhas de convergência numérica durante a busca do melhor parâmetro de transformação.

Os ajustes foram comparados via log-verossimilhança maximizada, critérios de informação de Akaike e Bayesiano, além da visualização dos resíduos do modelo com o pacote *ggplot2* (WICKHAM, 2016). Os valores iniciais para a otimização numérica vieram de ajustes de modelos geoespaciais para cada uma das variáveis utilizando o pacote *geoR* (RIBEIRO Jr et al., 2022). Tais ajustes foram realizados via máxima verossimilhança, com a mesma função de correlação e mesma estrutura de covariância marginal quando possível.

De cada modelo foi realizada a krigagem simples das variáveis a partir das propriedades da distribuição condicional da normal multivariada

$$E[Y_P|Y_O] = \mu_P + \Sigma_{PO}\Sigma_{OO}^{-1}(y - \mu_O)$$

$$Var[Y_P|Y_O] = \Sigma_{PP} - \Sigma_{PO}\Sigma_{OO}^{-1}\Sigma_{OP}$$

Em que O e P são os índices para valores observados e preditos, respectivamente.

Valendo-se das propriedades de que $\Sigma_{OP} = \Sigma_{PO}^\top$ e que Σ_{OO} é simétrica, então $Var[Y_P|Y_O]$ pode ser reescrito como

$$\begin{aligned} V[Y_P|Y_O] &= \Sigma_{PP} - \Sigma_{PO}(LL^\top)^{-1}\Sigma_{PO}^\top \\ &= \Sigma_{PP} - \Sigma_{PO}L^{-\top}L^{-1}\Sigma_{PO}^\top \\ &= \Sigma_{PP} - (L^{-1}\Sigma_{PO}^\top)^\top(L^{-1}\Sigma_{PO}^\top) \end{aligned}$$

3.4 Resumo dos modelos testados

- Modelo 1: Há um processo espacial específico para cada uma das variáveis (S_1 e S_2), além de um processo espacial conjunto (S_0). Para cada variável há também um termo de erro independente entre observações.
- Modelo 2: Processo espacial apenas para a variável Volume. Porém, agora os termos de erro são correlacionados entre variáveis para um mesmo ponto amostral.
- Modelo 3: Há um processo espacial específico para cada variável. Entretanto, da formulação da matriz de covariância resulta uma estrutura de dependência espacial entre variáveis.

4 Resultados e Discussão

4.1 Análise Descritiva

Pela Tabela 1 é perceptível a maior variabilidade do volume em relação ao DAP. Esse último é uma das medidas mais simples e de fácil medição da área florestal (MACHADO; FIGUEIREDO FILHO, 2003), enquanto o primeiro possui outras componentes de variação, como altura e forma da árvore. Ainda assim, ambos possuem baixa variabilidade.

A distribuição das variáveis (Figura 3) mostra que o diâmetro médio da parcela possui comportamento aproximadamente simétrico, enquanto a variável volume tem sua média levemente deslocada do centro da distribuição. O diagrama de dispersão entre as variáveis mostra a forte relação positiva entre elas. A correlação entre volume e diâmetro foi de 0,84.

Tabela 1 – Estatísticas descritivas do DAP e volume médios das parcelas de inventário.

Medida	DAP (cm)	Volume (m ³)
Média	29,97	360,58
Desvio Padrão	2,43	59,46
Máximo	34,81	458,50
Mínimo	24,35	209,84
CV (%)	8,12	16,49

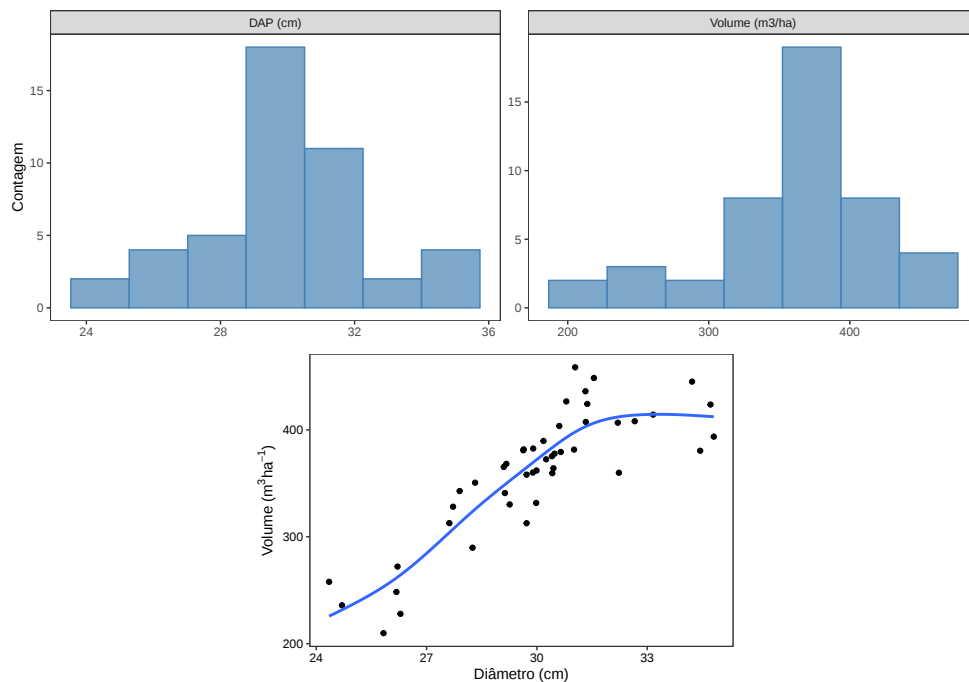


Figura 3 – Histogramas e diagrama de dispersão das variáveis diâmetro e volume médio das parcelas de inventário.

A distribuição espacial na Figura 4 demonstra a ocorrência de regiões de maior e menor valor para cada uma das variáveis em regiões coincidentes.

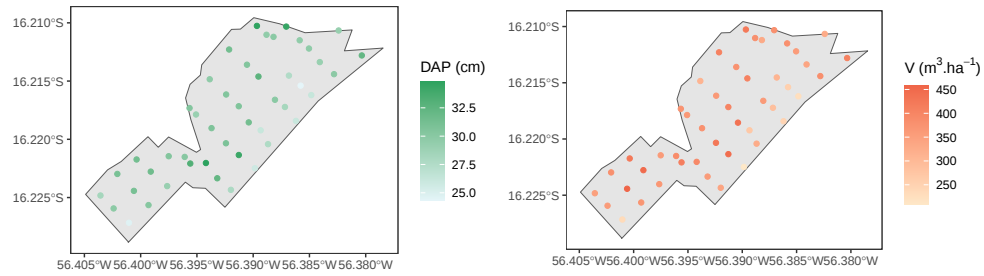


Figura 4 – Distribuição espacial das variáveis diâmetro e volume médio das parcelas de inventário.

4.2 Modelos univariados

Os resultados dos semivariogramas ajustados via máxima verossimilhança (Figura 5) indicam presença de dependência espacial para ambas as variáveis, mas mais acentuada para a variável volume. De acordo com os parâmetros estimados (Tabela 2), as médias foram de $354,32 \text{ m}^3.\text{ha}^{-1}$ para volume e $29,89 \text{ cm}$ para o diâmetro. Os parâmetros de alcance foram $210,09$ e $181,18$, respectivamente, com alcance assintótico de 629m e 543m para volume e diâmetro, respectivamente.

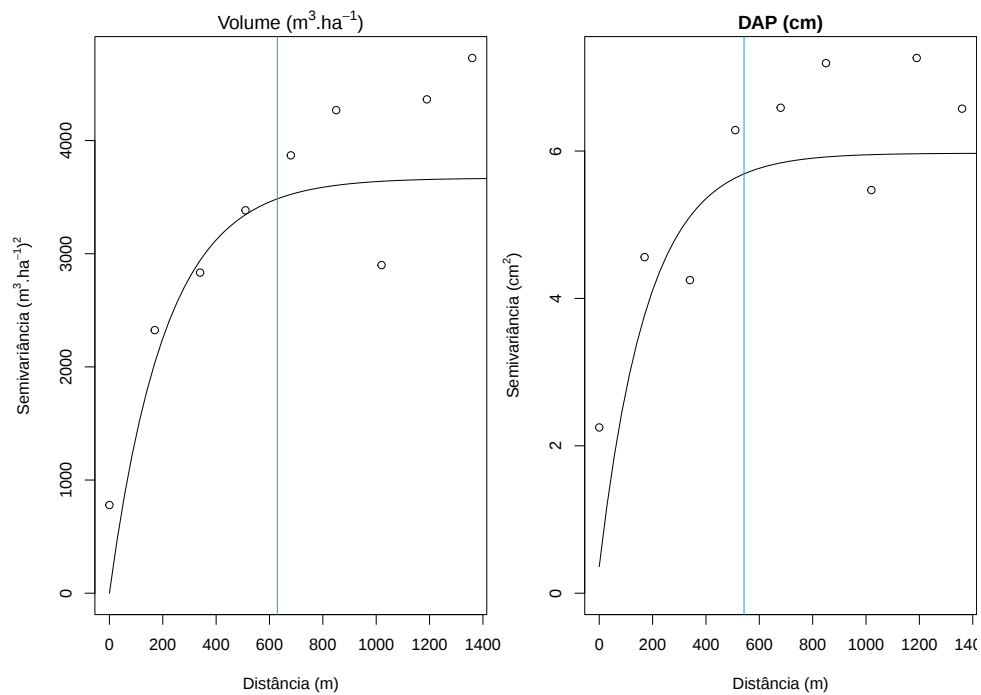


Figura 5 – Semivariogramas univariados das variáveis volume de DAP, ajustados via máxima verossimilhança. A linha azul define o alcance prático para cada um dos modelos ajustados.

Tabela 2 – Parâmetros estimados e log-verossimilhança maximizada para os modelos univariados ajustados via máxima verossimilhança.

Variável	$\hat{\beta}$	$\hat{\sigma}$	$\hat{\phi}$	$\hat{\tau}$	L
Volume	354,32	60,57	210,09	0,00	-248,47
Diâmetro	29,89	2,37	181,18	0,60	-102,66

Apesar de os valores de alcance dependerem tanto de fatores endógenos (como espécie) quanto fatores ambientais, os ajustes univariados indicam a existência de correlação entre as parcelas de inventário, porém em menor grau para o diâmetro. O maior alcance do volume em relação ao diâmetro já foi observado em estudos da área (BOGNOLA et al., 2008; ROVEDA et al., 2016; LIN et al., 2021; OLIVEIRA et al., 2023). Tais constatações podem ser proveitosas no planejamento florestal, uma vez que a intensidade amostral pode ser reduzida ao fazer uso da geoestatística para espacialização das variáveis de interesse (RAIMUNDO et al., 2017).

4.3 Modelos Multivariados

O modelo ajustado com a primeira especificação da estrutura de covariância teve parâmetros estimados (Tabela 3) similares aos modelos univariados, à exceção do termo adicional $\hat{\sigma}_0$, que acrescenta um componente de variância e simultaneamente diminui a estimativa de ϕ . Os resíduos foram bem distribuídos ao longo dos valores estimados, porém ambas as variáveis tiveram resíduos com distribuição assimétrica (Figura 6).

Tabela 3 – Parâmetros do primeiro modelo bivariado ajustado para diâmetro e volume.

Variável	$\hat{\beta}$	$\hat{\sigma}$	$\hat{\phi}$	$\hat{\tau}$
Volume	354,55	58,39	206,25	0,00
Diâmetro	29,89	0,00	180,00	0,00
-		2,44	168,60	

Para a segunda especificação, as médias estimadas $\hat{\beta}$ foram similares às anteriores (Tabela 4), porém agora τ_i retém maior variabilidade dos dados, uma vez que acumula toda a variância de Y_2 , bem como grande parte da covariância entre as variáveis. O parâmetro estimado $\hat{\sigma}$, que captura a variabilidade do processo espacial subjacente da variável volume (Y_1), teve um menor valor que a estimativa anterior, *i.e.*, essa especificação sugere um processo espacial mais ameno para o volume, e que pode ser parcialmente explicado pela estrutura $\tau_1\tau_2\rho_{12}$ adotada para o modelo. Em outras palavras, uma maior parte da variabilidade do volume está sendo explicada pelo diâmetro medido nas mesmas parcelas. A predição para essa especificação também apresentou resíduos levemente assimétricos para o volume (Figura 7), e mais assimétricos para o diâmetro.

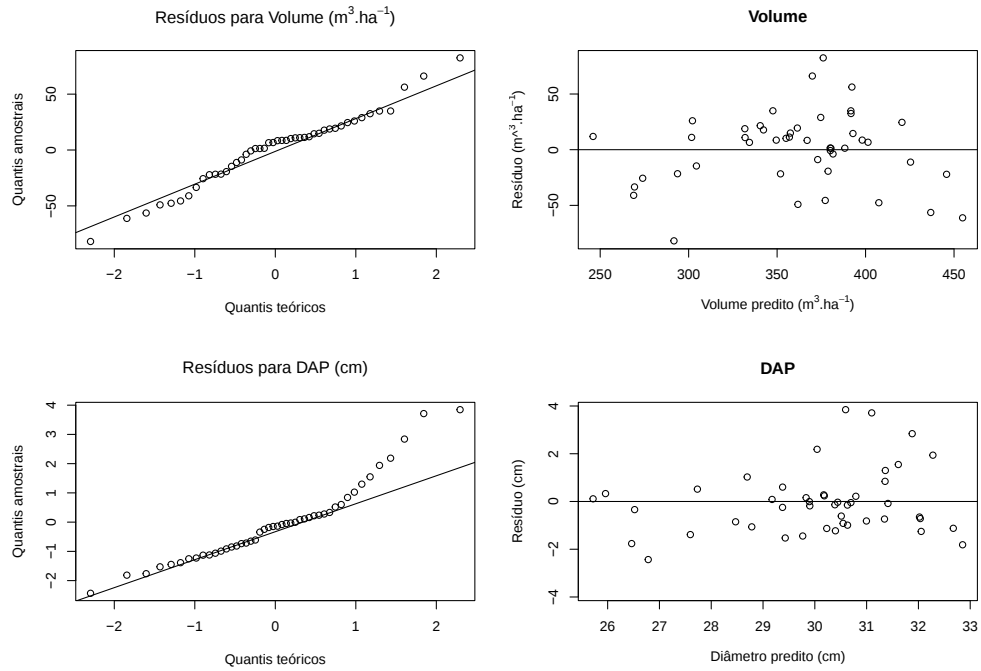


Figura 6 – Gráficos quantil-quantil e dispersão dos resíduos do primeiro modelo ajustado.

Tabela 4 – Parâmetros do segundo modelo bivariado ajustado para diâmetro e volume.

Variável	$\hat{\beta}$	$\hat{\sigma}$	$\hat{\phi}$	$\hat{\tau}$	$\hat{\rho}$
Volume	354,38	56,93	213,50	5,32	0,85
Diâmetro	29,97			2,33	

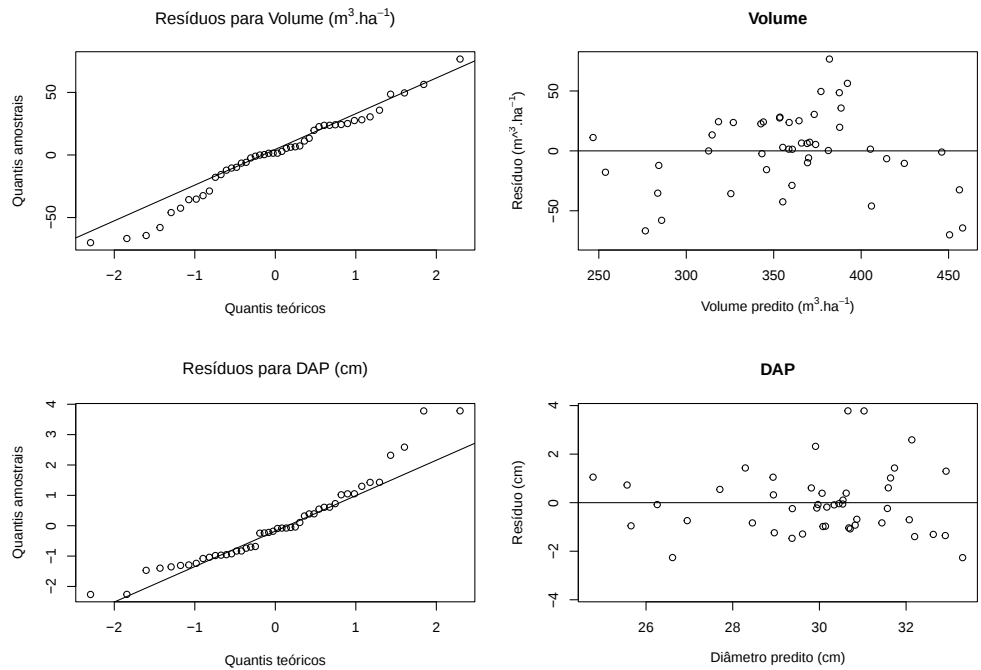


Figura 7 – Gráficos quantil-quantil e dispersão dos resíduos do segundo modelo ajustado.

A terceira especificação da matriz de covariância, baseada em Ribeiro, Ribeiro Jr e Bonat (2022), também resultou em parâmetros similares à primeira especificação, porém

com um parâmetro a menos (Tabela 5), além de maior valor do parâmetro de alcance ϕ para o diâmetro e menor para volume.

Tabela 5 – Parâmetros do terceiro modelo bivariado ajustado para diâmetro e volume.

Variável	$\hat{\beta}$	$\hat{\sigma}$	$\hat{\phi}$	$\hat{\tau}$	$\hat{\rho}$
Volume	358,20	56,25	114,09	7,49	0,81
DAP	29,91	1,70	242,49	1,65	

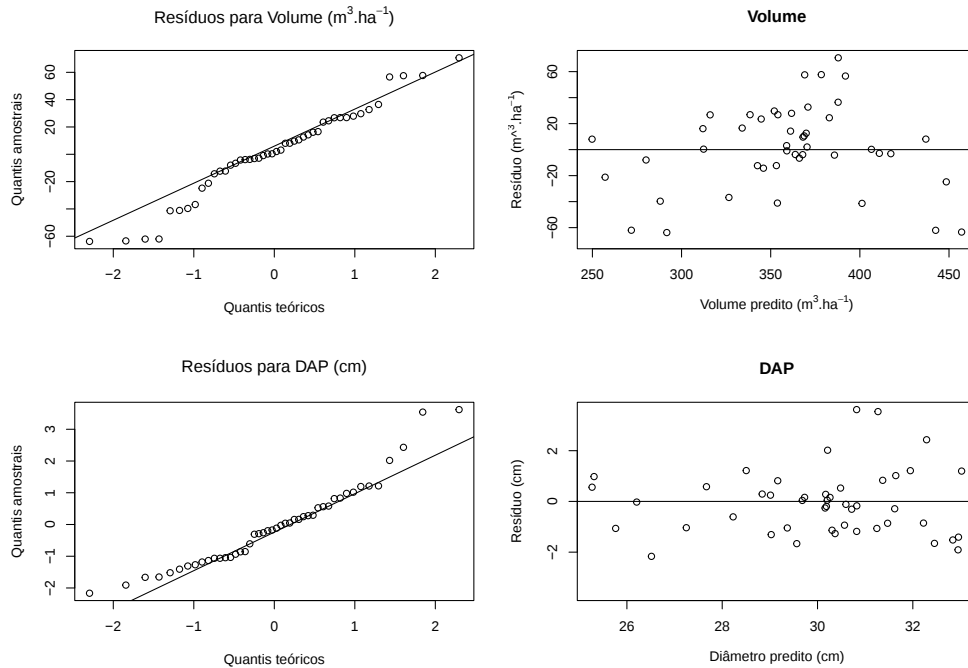


Figura 8 – Gráficos quantil-quantil e dispersão dos resíduos do terceiro modelo ajustado.

A Tabela 6 contém a log-verossimilhança de cada modelo, além da comparação com os ajustes independentes (univariados) de cada variável. Os três modelos bivariados tiveram melhor desempenho que os ajustes independentes. Dada a grande similaridade na estrutura dos modelos, era esperado a obtenção de valores próximos nos valores apresentados. Ainda assim, o modelo 1 teve o melhor desempenho, algo esperado dada a maior quantidade de parâmetros. O segundo melhor modelo foi o número 3, onde há a especificação apenas das covariâncias marginais. Tanto o segundo quanto o terceiro modelo tiveram desempenho similar, intercalando posições entre AIC e BIC. Embora similares em desempenho, o segundo modelo não prevê dependência espacial para a segunda variável, o que inviabiliza a predição para novas observações. O modelo 3, além de maior verossimilhança, é capaz de realizar a predição para ambas as variáveis.

Os semivariogramas marginais (Figura 9) de cada especificação mostram diferenças para cada uma das variáveis em relação ao ajuste univariado, especialmente para o diâmetro. Para a segunda especificação, o parâmetro $\hat{\tau}_1$ acumulou boa parte da variabilidade, resultando em um menor valor estimado de $\hat{\sigma}_1$ (56,93, linha verde do painel esquerdo). No terceiro modelo, o alcance prático para o volume é consideravelmente menor, o que

Tabela 6 – Estatísticas de ajuste dos três modelos testados e comparação com o ajuste univariado (modelo 0). LL é a log-verossimilhança maximizada, AIC e BIC são os critérios de informação de Akaike e Bayesiano, e npar é o número de parâmetros estimados.

Modelo	LL	AIC	BIC	npar
1	-312,15	644,31	669,52	10
2	-330,79	675,59	693,24	7
3	-327,94	673,88	696,57	9
0	-351,13	718,26	738,43	8

implica em estimativas com maior incerteza na krigagem conforme se afasta de um ponto observado.

Os semivariogramas construídos para o diâmetro não representaram adequadamente as semivariâncias empíricas calculadas para a primeira e terceira especificação (9). Comparando com os semivariogramas ajustados para os modelos univariados, há uma clara subestimativa da semivariância para a variável diâmetro. A segunda especificação não prevê dependência espacial, por isso o semivariograma é uma constante que passa pela variância do diâmetro.

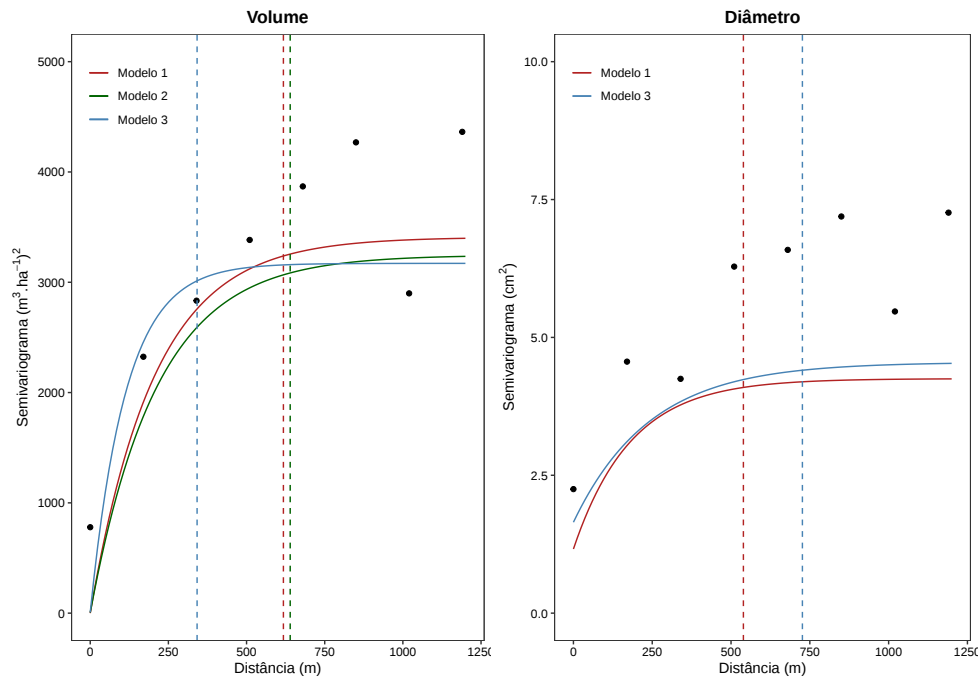


Figura 9 – Semivariogramas univariados do volume e diâmetro para os três modelos testados. A reta vertical indica o alcance prático.

4.4 Krigagem

A krigagem realizada para 10000 novos pontos igualmente espaçados foi similar entre os três modelos (Figuras 10, 11 e 12), os quais apresentaram padrão similar ao dos pontos amostrais (Figura 4). Os modelos 1 e 2 resultaram em mapas de krigagem muito semelhantes para o volume. O primeiro modelo apresentou padrão espacial similar tanto para o volume quanto diâmetro.

A variância estimada para cada ponto da Figura 10 acompanha a localização das parcelas de campo, uma vez que foram alocadas de forma sistemática na região, em que ambas as variáveis são medidas na mesma parcela. O padrão circular é consequência do aumento da incerteza conforme se afasta de um ponto observado na região. A krigagem do segundo modelo para o diâmetro foi constante para toda a região.

Outra característica notável é a completa dissimilaridade entre os mapas de média e variância, reflexo direto da ortogonalidade dos parâmetros de média e variância da distribuição normal.

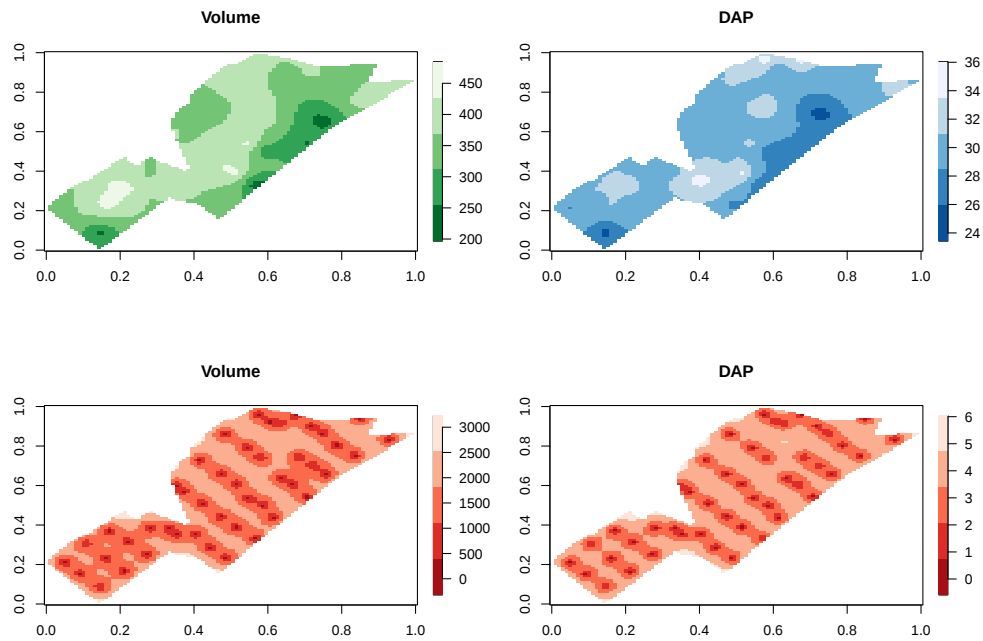


Figura 10 – Krigagem ordinária do volume e diâmetro utilizando o primeiro modelo. Na parte superior a predição das variáveis, e na parte inferior a variância das estimativas.

A krigagem para a terceira especificação resultou em mapas similares às anteriores para média, porém com mapas de variância predita de menor alcance para volume e maior para diâmetro, em comparação com a krigagem da primeira especificação. Tal contraste também é perceptível ao comparar os parâmetros estimados $\hat{\phi}$ de cada especificação. Para o terceiro modelo, o volume teve uma menor estimativa do parâmetro ϕ_1 , que implica em um decaimento mais acentuado da correlação, e consequentemente um maior aumento de incerteza conforme se afasta de um ponto observado na região. Por sua vez, o contrário

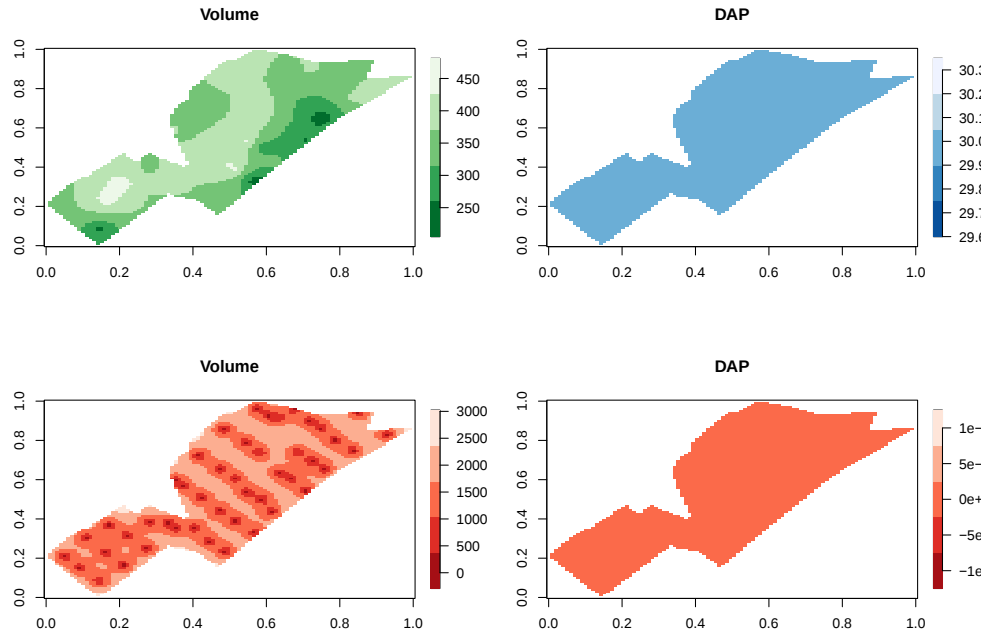


Figura 11 – Krigagem ordinária do volume e diâmetro utilizando o segundo modelo. Na parte superior a predição das variáveis, e na parte inferior a variância das estimativas.

ocorre com a variável diâmetro, pois o parâmetro estimado ϕ_2 para a terceira especificação foi maior que para a primeira.

Ainda outro reflexo dessa diferença é a amplitude de valores estimados para cada uma das variáveis. Quando há maior variabilidade associada (maior valor do parâmetro σ), maior será a amplitude das estimativas. Isso pode ser percebido comparando as escalas entre a krigagem do diâmetro dos modelos 1 e 3 (Figuras 10 e 12).

Comparando as estimativas de krigagem de sem considerar correlação entre variáveis (Figura 13), os modelos 1 e 3 foram muito similares para a estimativa do volume, enquanto o modelo 3 foi levemente divergente na estimativa. Na estimativa do diâmetro, também na figura é visível a menor amplitude de estimativas para o modelo 3, com estimativas maiores em menores diâmetros, e menores estimativas para maiores diâmetros, comparado com o modelo 1.

Para comparar a estimativa da produção em volume com métodos tradicionalmente utilizados na ciência florestal, foi realizada a estimativa da produção para a área total utilizando amostragem aleatória simples, extrapolando os resultados para toda a área de 222,45 hectares (Tabela 7). Para a estimativa dos modelos, a predição de cada pixel (dada em $m^3.ha^{-1}$) foi extrapolada para a área do pixel, e então foram somados os valores. Possíveis diferenças entre as estimativas dos modelos e a AAS podem surgir principalmente da configuração amostral, uma vez que o modelo geoestatístico compensa regiões com mais amostras, atribuindo ‘pesos’ menores para dados agrupados.

Todas as quatro estimativas foram similares, tanto entre os modelos testados quanto em comparação com a AAS. Tal resultado é esperado, uma vez que a grade

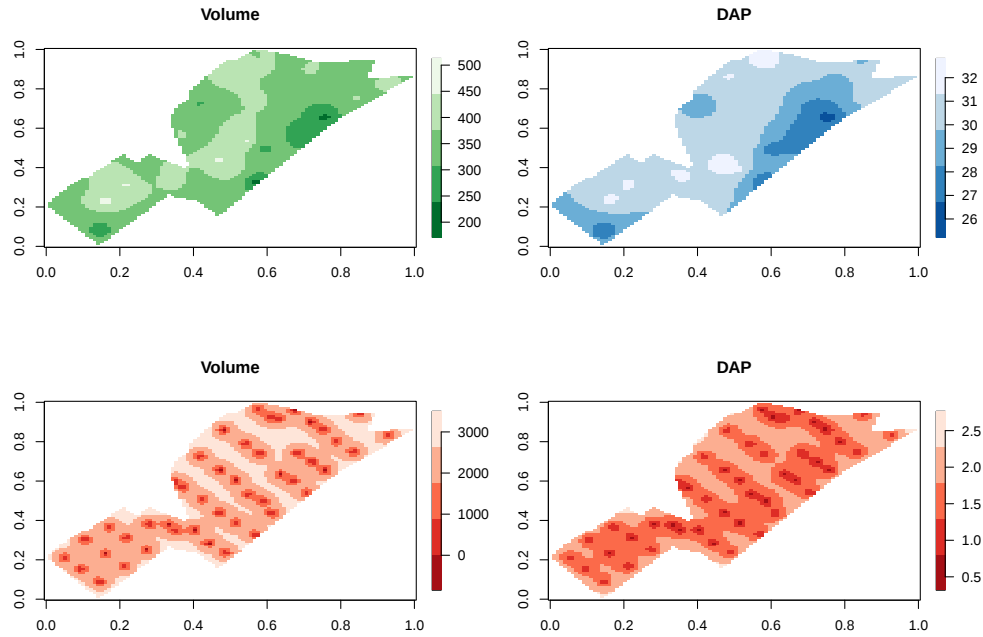


Figura 12 – Krigagem ordinária do volume e diâmetro utilizando o terceiro modelo. Na parte superior a predição das variáveis, e na parte inferior a variância das estimativas.

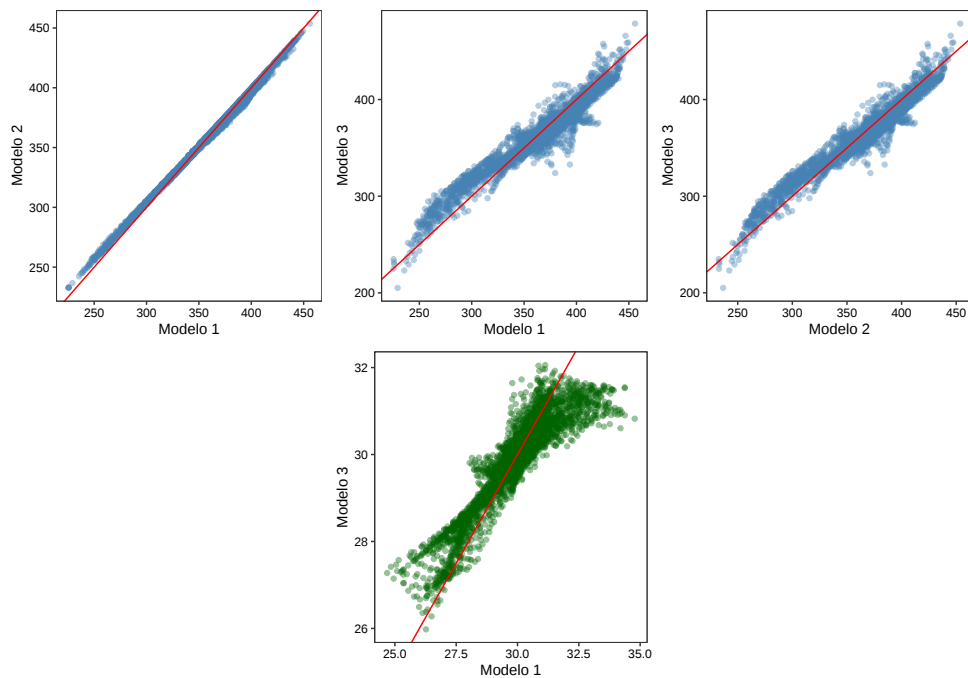


Figura 13 – Comparação de estimativas de predição entre os modelos para as variáveis volume (pontos azuis) e diâmetro (pontos verdes).

amostral é relativamente regular e não há valores muito discrepantes entre as unidades amostrais. Entretanto, esses resultados reforçam a adequabilidade dos modelos propostos para predição da produção, pois além de fornecerem estimativas condizentes com métodos consolidados na área, permitem a espacialização dessa produção.

Ainda assim, o impacto na incerteza da produção total pode ser mais significativo, pois os modelos geoestatísticos estimam não apenas a média geral, bem como realizam

predições espacializadas na área. A variância da média global dos modelos, porém, não foi estimada nesse estudo.

Tabela 7 – Comparação da estimativa do volume total m^3 de cada modelo com a amostragem aleatória simples (AAS).

AAS	Modelo 1	Modelo 2	Modelo 3
80211	80417	80410	80185

5 Considerações Finais

As três especificações para a estrutura de covariância permitiram modelar de forma conjunta o volume e o diâmetro de plantios florestais. Diferentes especificações resultaram em mapas de krigagem com capacidades de predições semelhantes para novos locais.

A primeira estrutura de covariância teve a maior log-verossimilhança, embora seja a com maior número de parâmetros no modelo. A segunda especificação, apesar de ter um bom ajuste, não foi capaz de gerar mapas de krigagem úteis, dada a especificação de covariância adotada. A terceira estrutura foi parcimoniosa entre número de parâmetros necessários e qualidade do mapa de predição das variáveis.

Os modelos propostos possuem uma considerável complexidade de implementação e interpretação, além de um custo computacional elevado, a depender da quantidade de observações e resolução dos mapas de krigagem. O histograma de resíduos também revelou distribuição assimétrica em alguns modelos, especialmente para o diâmetro.

Os modelos geoestatísticos possuem a capacidade de estimativas espacializadas da produção, além da possibilidade de criação de mapas de incerteza de cada variável. Esses modelos ainda permitem a obtenção de outras informações relevantes na área florestal, como a delimitação das regiões com produção acima de um valor especificado, e até mesmo o cálculo da produção na área total, valor esse muito importante para o ramo, que de outra forma é obtido via extrapolação à partir do valor médio das parcelas de inventário.

Por fim, outras alternativas podem ser usadas na modelagem geoestatística, em substituição ou complemento à análise multivariada, como a inserção de variáveis exógenas como preditores lineares das médias da distribuição, ou a construção de semivariogramas direcionais para a determinação de uma direção preferencial de variação do dado avaliado.

Referências

- BOGNOLA, I. A. et al. Modelagem uni e bivariada da variabilidade espacial de rendimento de pinus taeda l. *FLORESTA*, Universidade Federal do Parana, v. 38, 6 2008. ISSN 0015-3826.
- Castrignanò, A. et al. Improving the spatial prediction of sand content in forest soils using a multivariate geostatistical analysis of lidar and hyperspectral data. *Remote Sensing*, Multidisciplinary Digital Publishing Institute (MDPI), v. 15, 9 2023. ISSN 20724292.
- CHILES, J.-P.; DELFINER, P. *Geostatistics: modeling spatial uncertainty*. [S.l.]: John Wiley & Sons, 2012. v. 713.
- DIGGLE, P. J.; RIBEIRO Jr, P. J. *Model-based Geostatistics*. 1. ed. [S.l.]: Springer Science, 2007. v. 1. 1-241 p. ISBN 0-387-32907-2.
- GENZ, A.; BRETZ, F. *Computation of Multivariate Normal and t Probabilities*. Heidelberg: Springer-Verlag, 2009. (Lecture Notes in Statistics). ISBN 978-3-642-01688-2.
- GOTTWALD, T. et al. Analysis of the spatial spread of sharka (plum pox virus) in apricot and peach orchards in eastern Spain. *Plant Disease*, v. 79, p. 266, 1995.
- JOHN, K. et al. Mapping soil properties with soil-environmental covariates using geostatistics and multivariate statistics. *International Journal of Environmental Science and Technology*, Springer Science and Business Media Deutschland GmbH, v. 18, p. 3327–3342, 11 2021. ISSN 17352630.
- KARAHAN, G.; ERŞAHİN, S. Geostatistics in characterizing spatial variability of forest ecosystems. *Eurasscience Journals Eurasian Journal of Forest Science*, v. 6, p. 9–22, 2018.
- LIN, C. et al. Generating a baseline map of surface fuel loading using stratified random sampling inventory data through cokriging and multiple linear regression methods. *Remote Sensing*, MDPI AG, v. 13, 4 2021. ISSN 20724292.
- MACHADO, S. do A.; FIGUEIREDO FILHO, A. *Dendrometria*. 1. ed. [S.l.: s.n.], 2003. 1-309 p.
- MATHERON, G. The theory of regionalised variables and its applications. *Les Cahiers du Centre de Morphologie Mathématique*, v. 5, p. 1–212, 1971.
- MELLO, J. M. D. et al. Estudo da dependência espacial de características dendrométricas para eucalyptus grandis. *Cerne*, v. 11, p. 113–126, 2005.
- MELLO, J. M. de. Geoestatística aplicada ao inventário florestal. p. 1–121, 2004.
- NANOS, N. et al. Geostatistical prediction of height/diameter models. *Forest Ecology and Management*, v. 195, p. 221–235, 6 2004. ISSN 03781127.
- NISHINA, K.; TAKENAKA, C.; ISHIZUKA, S. Spatial variations in nitrous oxide and nitric oxide emission potential on a slope of Japanese cedar (Cryptomeria japonica) forest. *Soil Science and Plant Nutrition*, v. 55, p. 179–189, 2009.

OLIVEIRA, R. G. D. et al. Modelagem geoestatística do volume comercial e sortimentos em povoamentos de pinus taeda l. geostatistical modeling of commercial volume and assortments in pinus taeda l. stands. *Adv. For. Sci, Cuiabá*, p. 2031–2041, 2023. Disponível em: <<http://dx.doi.org/10.34062/afs.v10i2.13499>>.

PELISSARI, A. L. Geoestatística aplicada ao manejo de povoamentos de tectona grandis l. f. p. 1–120, 2015.

PELISSARI, A. L.; FIGUEIREDO FILHO, A.; CALDEIRA, S. F. Geostatistic applied to spatial modeling of hypsometric relationships in forest stands. *American International Journal of Contemporary Research*, v. 4, 2014. Disponível em: <www.aijcrnet.com>.

PELISSARI, A. L. et al. Modelagem geoestatística da variabilidade espacial do volume de madeira para o manejo de precisão de tectona grandis l. f. *Cerne*, Federal University of Lavras, v. 23, p. 115–122, 2017. ISSN 01047760.

PFUTZ, I. et al. Geostatistical modelling of below- and above-ground carbon stocks of teak stands at different ages. *Source: Journal of Tropical Forest Science*, v. 33, p. 105–112, 2021.

PFUTZ, I. F. P. et al. Geostatistical modeling and remotely sensed data to improve dendrometric variables prediction in tectona grandis l. f. stand. *Revista Forestal Mesoamericana Kurú*, v. 19, p. 71–83, 2022.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2023. Disponível em: <<https://www.R-project.org/>>.

RAIMUNDO, M. R. et al. Geostatistics applied to growth estimates in continuous forest inventories. *Forest Science*, Society of American Foresters, v. 63, p. 29–38, 2 2017. ISSN 19383738.

RIBEIRO, A. M. T.; RIBEIRO Jr, P. J.; BONAT, W. H. A kronecker-based covariance specification for spatially continuous multivariate data. *Stochastic Environmental Research and Risk Assessment*, Springer Science and Business Media Deutschland GmbH, v. 36, p. 4087–4102, 12 2022. ISSN 14363259.

RIBEIRO, A. M. T.; Ribeiro Jr, P. J.; BONAT, W. H. A kronecker-based covariance specification for spatially continuous multivariate data. *Stochastic Environmental Research and Risk Assessment*, Springer Science and Business Media Deutschland GmbH, v. 36, p. 4087–4102, 12 2022. ISSN 14363259.

RIBEIRO Jr, P. J. et al. *geoR: Analysis of Geostatistical Data*. [S.l.], 2022. R package version 1.9-2. Disponível em: <<https://CRAN.R-project.org/package=geoR>>.

ROSSI, R. et al. Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecological Monographs*, v. 62, p. 277–314, 1992.

ROVEDA, M. et al. Continuidade espacial do número de árvores e da área basal de uma floresta ombrófila mista sob diferentes formas e tamanhos de unidades amostrais. *Cerne*, Federal University of Lavras, v. 22, p. 189–196, 4 2016. ISSN 01047760.

SALES, M. H. et al. Improving spatial distribution estimation of forest biomass with geostatistics: A case study for rondônia, brazil. *Ecological Modelling*, v. 205, p. 221–230, 7 2007. ISSN 03043800.

SCOLFORO, H. F. et al. Spatial distribution of aboveground carbon stock of the arboreal vegetation in brazilian biomes of savanna, atlantic forest and semi-arid woodland. *PLoS ONE*, Public Library of Science, v. 10, 6 2015. ISSN 19326203.

SILVEIRA, E. M. et al. Spatial distribution of wood volume in brazilian savannas. *Anais da Academia Brasileira de Ciencias*, Academia Brasileira de Ciencias, v. 91, 2019. ISSN 16782690.

WICKHAM, H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016. ISBN 978-3-319-24277-4. Disponível em: <<https://ggplot2.tidyverse.org>>.

Apêndices

APÊNDICE A - Matriz de variância-covariância do modelo 1

$$\Sigma = \left[\begin{array}{cccc|cccc} \sigma_1^2 + \sigma_0^2 + \tau_1^2 & \sigma_1^2 \rho_1(\phi_1) + \sigma_0^2 \rho_0(\phi_0) & \cdots & \sigma_1^2 \rho_1(\phi_1) + \sigma_0^2 \rho_0(\phi_0) & \sigma_0 & \sigma_0 \rho_0(\phi_0) & \cdots & \sigma_0 \rho_0(\phi_0) \\ \sigma_1^2 \rho_1(\phi_1) + \sigma_0^2 \rho_0(\phi_0) & \sigma_1^2 + \sigma_0^2 + \tau_1^2 & \cdots & \sigma_1^2 \rho_1(\phi_1) + \sigma_0^2 \rho_0(\phi_0) & \sigma_0 \rho_0(\phi_0) & \sigma_0 & \cdots & \sigma_0 \rho_0(\phi_0) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_1^2 \rho_1(\phi_1) + \sigma_0^2 \rho_0(\phi_0) & \sigma_1^2 \rho_1(\phi_1) + \sigma_0^2 \rho_0(\phi_0) & \cdots & \sigma_1^2 + \sigma_0^2 + \tau_1^2 & \sigma_0 \rho_0(\phi_0) & \cdots & \cdots & \sigma_0 \\ \hline \sigma_0 & \sigma_0 \rho_0(\phi_0) & \cdots & \sigma_0 \rho_0(\phi_0) & \sigma_2^2 + \sigma_0^2 + \tau_2^2 & \sigma_2^2 \rho_2(\phi_2) + \sigma_0^2 \rho_0(\phi_0) & \cdots & \sigma_2^2 \rho_2(\phi_2) + \sigma_0^2 \rho_0(\phi_0) \\ \sigma_0 \rho_0(\phi_0) & \sigma_0 & \cdots & \sigma_0 \rho_0(\phi_0) & \sigma_2^2 \rho_2(\phi_2) + \sigma_0^2 \rho_0(\phi_0) & \sigma_2^2 + \sigma_0^2 + \tau_2^2 & \cdots & \sigma_2^2 \rho_2(\phi_2) \sigma_0^2 \rho_0(\phi_0) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_0 \rho_0(\phi_0) & \cdots & \cdots & \sigma_0 & \sigma_2^2 \rho_2(\phi_2) + \sigma_0^2 \rho_0(\phi_0) & \cdots & \cdots & \sigma_2^2 + \sigma_0^2 + \tau_2^2 \end{array} \right] \quad (1)$$