

SOBRE A ESTIMAÇÃO E MODELAGEM DE FUNÇÕES COVARIÂNCIAS NA COLOCAÇÃO POR MÍNIMOS QUADRADOS

On estimation and modelling of the covariance functions in least squares collocation

Sergio Florencio de Souza¹
Nelsi Côgo de Sá²

¹Universidade Federal do Rio Grande do Sul
Instituto de Geociências – Departamento de Geodésia
Av. Bento Gonçalves, 9500 - Porto Alegre - RS - Brasil
CEP: 91501-970 / Caixa Postal: 15001
e-mail: sergio.florencio@ufrgs.br

²Universidade de São Paulo
Instituto de Astronomia, Geofísica e Ciências Atmosféricas
Rua do Matão, 1226 – Cidade Universitária – São Paulo – SP - Brasil
e-mail. nelsi@iag.usp.br

RESUMO

Nestas últimas décadas tem aumentado o interesse na determinação de geóides gravimétricos, um dos principais motivos, é a sua utilização na determinação de altitudes ortométricas a partir das altitudes geométricas, obtidas no posicionamento por satélite (GPS). Vários são os métodos utilizados para a determinação do geóide, como a integral de Stokes e suas variantes, colocação por mínimos quadrado, colocação rápida (*fast collocation*), entre outros. A colocação por mínimos quadrados permite a estimação eficiente do campo de gravidade, mediante a utilização das características estatísticas de dados heterogêneos sobre a forma de funções covariâncias. Portanto, a interpretação estatística da colocação, é baseada no pressuposto de que as observações do campo de gravidade são uma realização de um processo estocástico e que, quer as observações, ou os sinais a serem estimados, são variáveis centradas. A função covariância é o elemento central desse processo, toda a informação relativa ao comportamento do campo de gravidade, sua variabilidade, distancia de correlação e anisotropias, bem como a relação funcional entre os vários elementos do campo estão condicionados a eficiência na sua determinação. O presente trabalho tem por objetivo a discussão da modelagem das funções covariâncias, a partir da utilização de dados com distribuição regular e irregular. A distribuição dos dados tem implicações diretas com a estimação dos parâmetros da função covariância. Sendo assim, serão vistos os fundamentos teóricos dos métodos de estimação das funções covariâncias, seguida da aplicação.

Palavras chaves: Colocação por mínimos quadrados, Funções covariâncias, Geóide.

ABSTRACT

In the last decades there has been an increase of interest in the gravimetric geoids determination. One of the main reasons for this is its use in the determination of orthometric altitudes from the geometric altitudes, obtained through the Global Positioning System (GPS). There are several methods for the geoid determination, such as the Stokes' integral and its variations, least square collocation, fast collocation, and others. The least square collocation allows efficient estimation of the gravitational field, under the utilization of heterogeneous data's statistic characteristics on covariance functions forms. Therefore, the statistic interpretation of the collocation is based on the conjecture that the gravity field observations are the realization of a stochastic process and that both the observations and the signals to be estimated are centered variants. The covariance function is the central element of this process. All information related to the behavior of the gravitational field, its variability, correlation distance and anisotropies, as well as the functional relation between the many elements of the field are conditioned to efficiency in its determination. This work aims the discussion of the covariances functions tracing, from data utilization with regular and irregular distribution. Data distribution has straight

implications on the estimation of the covariance function parameters. Therefore it will be seen the theoretical fundamentals of the covariance function estimation methods, followed by application.

Keywords: Least Square Collocation, Covariance function, Geoid.

1. INTRODUÇÃO

A colocação por mínimos quadrados (C.M.Q) envolve as covariâncias de todos os elementos do campo de gravidade que são utilizados na aproximação. Isto decorre da interpretação do campo de gravidade anômalo como um processo estocástico estacionário. Do ponto de vista estatístico, a função covariância caracteriza a correlação estatística de duas quantidades do campo de gravidade em dois pontos distintos, ou seja, a tendência que apresentam em ter a mesma magnitude e o mesmo sinal. A determinação da função covariância é sempre necessária como indicador do comportamento do campo gravífico numa determinada região, seja o caso em que se pretende estimar novos valores para a anomalia gravimétrica ou no cálculo do geóide. A função covariância representa a estrutura de campo de gravidade anômalo na CMQ. Isto introduz dois requisitos desejáveis nos modelos adotados; o de reproduzir as propriedades estatísticas contidas nos dados e de admitir uma representação analítica simples. No primeiro caso, pode ser verificado através de comparações entre os resultados obtidos na aproximação e na observação de alguns elementos. O segundo é possível para covariâncias homogêneas e isotrópicas. Na prática, o fator dominante na escolha de um modelo é a quantidade e a distribuição geográfica dos dados. Mesmo nos dias de hoje, com milhões de dados gravimétricos, as informações estatísticas sobre alguns parâmetros ainda são precárias. Isto significa que ainda é necessário distinguir-se os modelos globais dos locais. Nos modelos globais, as covariâncias são obtidas de valores médios representativos de blocos cujas dimensões indicam a resolução do campo e das variâncias de baixo grau que resultam do rastreamento de satélites artificiais. Nos modelos locais, as covariâncias são obtidas de dados pontuais e da variância de seus gradientes. Neste trabalho, serão utilizadas na modelagem de covariâncias amostrais, as anomalias gravimétricas, que ainda são em maior número no Brasil. O objetivo principal é estudar o comportamento das funções covariâncias, numa pequena região da Terra, a partir de dados distribuídos regularmente e irregular

2. COLOCAÇÃO POR MÍNIMOS QUADRADOS DOS

A C.M.Q é uma técnica matemática que permite a determinação do campo de gravidade pela combinação de dados diferentes em tipo e precisão. Dado o potencial anômalo (T) e um conjunto de observações relacionadas com o campo de gravidade da Terra (I), pode-se expressar através de um funcional linear (L), tais observações como:

$$I = L(T) \quad (1)$$

Sendo que as observações podem ser representadas por um modelo matemático da forma (Moritz, 1980)

$$I = AX + s + n \quad (2)$$

onde I é o vetor das observações, X é o vetor dos parâmetros que representam a componente determinística, s é o vetor dos sinais do campo anômalo expresso por funcionais aplicados ao potencial anômalo e n é o vetor dos erros observacionais.

O problema da representação do campo de gravidade, a partir de dados heterogêneos, expresso em (2), é solucionado através da condição de unicidade, dada por

$$\alpha \langle T, T \rangle + \beta n^T D^{-1} n = \text{mínimo} \quad (3)$$

que tem solução analítica igual a (Moritz, 1980)

$$\hat{X} = [A^T (\beta C + \alpha D)^{-1} A]^{-1} A^T (\beta C + \alpha D)^{-1} I \quad (4)$$

$$\hat{T} = (\beta BK)^T (\beta C + \alpha D)^{-1} (I - AX) \quad (5)$$

onde α e β são parâmetros numéricos, K é o núcleo do espaço de Hilbert, C e D são matrizes simétricas definidas positivas. Dependendo do que se adota para os parâmetros, o núcleo e as matrizes, a solução geral assume particulares e admite interpretações distintas. A colocação por mínimos quadrados constitui o caso particular em que T é considerado um processo estocástico. Isto proporciona interpretações estatísticas para as normas do potencial anômalo e do erro observacional.

A C.M.Q é aquela que minimiza as variâncias dos parâmetros σ_x^2 e do potencial anômalo σ_T^2 , isto é, (Sá, 1988)

$$\sigma_x^2 = \text{mínimo} \quad (6)$$

$$\sigma_T^2 = \text{mínimo} \quad (7)$$

Estas condições de mínimo são satisfeitas por (4) e (5) através de (2) e (3) se (Moritz, 1980):

- os parâmetros forem unitários

$$\alpha = \beta = 1 \quad (8)$$

- a função que representa o núcleo é igual à função variância do sinal

$$K(P, Q) = C(T_p, T_q) = C_u \quad (9)$$

- a matriz peso é igual a matriz das covariâncias do ruído

$$C(n_p, n_q) = C_m \quad (10)$$

Assim, com (9) e (10) pode-se escrever

$$C(T_p, T_q) + C(n_p, n_q) = C_u + C_m = C \quad (11)$$

$$BK = C_{st} \quad (12)$$

$$t = BT \quad (13)$$

e

$$s \begin{bmatrix} t \\ u \end{bmatrix} \quad (14)$$

onde t e u representam os vetores dos sinais conhecidos e desconhecidos, respectivamente. Nestas condições, o modelo matemático da colocação por mínimos quadrados tem forma geral

$$l = AX + s + n \quad (15)$$

Como a componente determinística (AX), neste caso, é representada pelo modelo EGM96, a expressão (15) se reduz ao modelo não paramétrico ($X = \theta$), expresso por

$$l = s + n \quad (16)$$

As soluções para o vetor dos sinais e a respectiva variância são dadas por (Moritz, 1980)

$$\hat{s} = C_{st} C^{-1} (l) \quad (17)$$

$$\sigma_s^2 = C_{ss} - C_{st} C^{-1} C_{ts} \quad (18)$$

onde as matrizes covariâncias são obtidas através de funções ajustadas às covariâncias amostrais.

Para ilustrar a aplicação destas expressões, faz-se aqui um exemplo teórico da aplicação do método de colocação por mínimos quadrados. Supondo que se deseja determinar anomalias de altitude (ζ) a partir de anomalias gravimétricas (Δg), e que os dados estejam referidos ao geóide. Dada a expressão (17), e

considerando f como sendo a quantidade total de dados utilizados, tem-se

$$\hat{s} = C_{st} C^{-1} (l)$$

lembrando que, $C = C_{tt} + C_{nn}$

e

$$\hat{s} = \zeta(P) \quad (19)$$

onde a expressão (19) representa a anomalia de altitude no ponto P , o vetor das observações (l) é dado por

$$l = \begin{bmatrix} \Delta g_1 \\ \vdots \\ \Delta g_f \end{bmatrix} \quad (20)$$

Admitindo que se conheçam os erros dos dados gravimétricos e que todas as observações são não-correlacionadas, tem-se a matriz covariância do ruído C_{nn} igual a,

$$C_{nn} = \begin{bmatrix} \sigma_{\Delta g_1}^2 & & \\ & \ddots & \\ & & \sigma_{\Delta g_f}^2 \end{bmatrix} \quad (21)$$

a matriz covariância cruzada e a matriz covariância do sinal são dadas por,

$$C_{st} = \begin{bmatrix} C_{P_1}^{\Delta g} & \dots & C_{P_f}^{\Delta g} \end{bmatrix} \quad (22)$$

$$C_{tt} = \begin{bmatrix} C_{\Delta g_1, \Delta g_1} & \dots & C_{\Delta g_1, \Delta g_f} \\ \vdots & & \vdots \\ C_{\Delta g_f, \Delta g_1} & \dots & C_{\Delta g_f, \Delta g_f} \end{bmatrix} \quad (23)$$

As matrizes C_{tt} e C_{st} são modeladas a partir de funções covariâncias, e os procedimentos envolvidos na sua determinação estão descritos no item seguinte

3. FUNÇÕES DE COVARIÂNCIAS

A C.M.Q requer as covariâncias de todos os dados envolvidos na aproximação. Em geral, a função covariância caracteriza a correlação reflete o comportamento do campo gravífico descrevendo a magnitude da sua variação e da rugosidade. Do ponto de vista estatístico, a função covariância caracteriza a correlação estatística de duas quantidades do campo de gravidade. Teoricamente, qualquer tipo de dado do campo de gravidade pode ser usado para a obtenção das covariâncias. Mas, como a quantidade e distribuição geográfica dos dados são importantes nos processos de amostragem, as anomalias gravimétricas são geralmente usadas por constituírem a maior população e a distribuição mais homogênea dos dados atualmente disponíveis.

Assim, representando a anomalia observada (Δg) através das componentes determinística (Δg^d) e estocásticas (Δg^r), tem-se:

$$\Delta g = \Delta g^d + \Delta g^r \quad (24)$$

A componente determinística é representada pelo modelo geopotencial, a componente estocástica, na realidade, é a combinação do sinal (Δg^s) e do ruído (Δg^n) resultante dos erros observacionais e dos erros de aproximação das correções adotadas as observações

$$\Delta g^r = \Delta g^s + \Delta g^n \quad (25)$$

Para distribuições geográficas apropriadas, as covariâncias amostrais podem ser calculadas pelo método direto. Assim, as covariâncias entre n elementos amostrais da componente estocástica

$$\Delta g^r = \{\Delta g_1^r, \Delta g_2^r, \dots, \Delta g_n^r\} \quad (26)$$

são dadas pelo operador média

$$C(\psi_k) = \text{cov}(\Delta g_i^r, \Delta g_j^r) = M\{\Delta g_i^r \Delta g_j^r\} \quad (27)$$

e obtidas numericamente por

$$C(\psi_k) = \frac{1}{n_k} \sum_{i=1}^n \sum_{j=1}^n \Delta g_i^r \Delta g_j^r \quad (28)$$

onde k é o índice de classificação segundo as distâncias. Assim, $C(\psi_k)$ representa a média dos produtos dos elementos existentes em pontos que estão à distância $\psi_{i,j}$ tal que

$$\psi_k - \frac{\Delta\psi}{2} \leq \psi_{i,j} \leq \psi_k + \frac{\Delta\psi}{2} \quad (29)$$

onde $\Delta\psi$ é o intervalo de distância das classes.

A expressão (28) proporciona covariâncias em função apenas da distância relativa ψ , o que significa covariâncias isotrópicas. Na prática, estas são a de maior interesse devido à simplicidade de cálculo. As funções covariâncias são caracterizadas por três parâmetros (Figura 1): a variância (C_0), a distância de correlação (ξ_1) e o parâmetro de curvatura (χ). A variância é o valor que a função covariância assume para ψ_0 (Moritz, 1980)

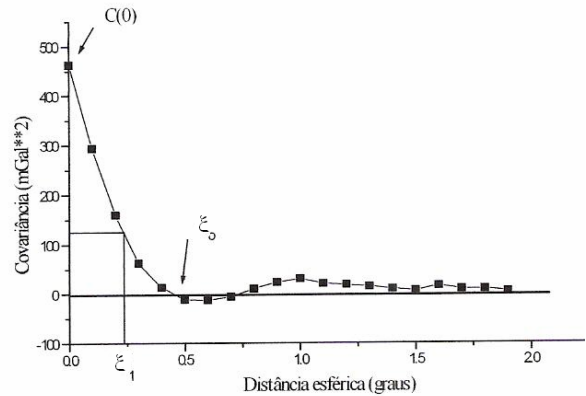


Figura 1 – Parâmetros essenciais da função covariância (adaptada de Fernandes (1999))

A variância C_0 é o valor da função covariância $C(\psi)$ para o argumento $\psi = 0$.

$$C_0 = C(0) \quad (30)$$

A distância de correlação é o valor do argumento para o qual a covariância é igual a metade da variância.

$$C(\xi_1) = \frac{C_0}{2} \quad (31)$$

O parâmetro de curvatura é relacionado com a curvatura da função em $\psi=0$ e é dado por

$$\chi = \xi_1^2 \frac{G_0}{C_0} \quad (32)$$

onde G_0 é a variância do gradiente horizontal da anomalia gravimétrica. Estes parâmetros são chamados de essenciais.

3.1 Covariâncias esféricas

Dados dois pontos $P(\theta, \lambda)$ e $Q(\theta', \lambda')$ no espaço, e o potencial perturbador $T(P)$ e $T(Q)$ nesses pontos, define-se a função covariância do potencial perturbador $K(P, Q)$ como

$$K(P, Q) = M\{T(P)T(Q)\} \quad (33)$$

onde $M\{\bullet\}$ representando o estimador da média do sinal do campo de gravidade anômalo. Admitindo-se a inexistência o termo de ordem zero no potencial anômalo, tem-se

$$M\{T\} = 0 \quad (34)$$

característica que implica na anulação dos harmônicos de ordens zero e um do potencial perturbador. Sobre a esfera, o operador $M\{\bullet\}$ é considerado homogêneo

(independe de translação) e isotrópico (independe de direção). Deste modo, a função dependerá somente da distância esférica (ψ). De fato:

$$K(P, Q) = K(\psi) = M\{T(P)T(Q)\} = \frac{1}{8\pi^2} \int_{\lambda=0}^{2\pi} \int_{\theta=0}^{\pi} \int_{\alpha=0}^{2\pi} T(\theta, \lambda) T(\theta', \lambda') \sin\theta d\theta d\lambda d\alpha \quad (35)$$

onde a distância esférica pode ser expressa como

$$\cos\psi = \cos\theta \cos\theta' + \sin\theta \sin\theta' \cos(\lambda - \lambda') \quad (36)$$

Considerando uma esfera de raio unitário e o fato da função ser harmônica sobre a esfera, justifica-se a sua representação em harmônicos esféricos a série

$$K(P, Q) = \sum_{n=2}^{\infty} \sum_{m=0}^n (\bar{a}_{nm} \cos m\lambda + \bar{b}_{nm} \sin m\lambda) P_{nm}(\cos\psi) \quad (37)$$

sendo $\bar{a}_{nm}, \bar{b}_{nm}$, coeficientes dos harmônicos esféricos associados à função covariância. Com o caráter isotrópico da função covariância a expressão (37) se reduz a:

$$K(\psi) = \sum_{n=2}^{\infty} k_n P_n(\cos\psi) \quad (38)$$

onde $P_n(\cos\psi)$ são os polinômios de Legendre. Dado que k_n são obtidos por (Heiskanen & Moritz, 1967):

$$k_n = \sum_{m=0}^n (\bar{a}_{nm}^2 + \bar{b}_{nm}^2) \quad (39)$$

onde $\bar{a}_{nm}, \bar{b}_{nm}$ coeficientes plenamente normalizados, relacionados ao potencial perturbador.

Para o exterior da esfera, a função covariância do potencial anômalo entre os pontos P e Q , situados a uma distância (ψ), será

$$K(P, Q) = \sum_{n=2}^{\infty} k_n \left(\frac{R_B}{r_P r_Q} \right)^{n+1} P_n(\cos\psi) \quad (40)$$

sendo que r_P, r_Q são os raios vetores dos pontos P e Q , R_B é o raio da esfera de *Bjerhammar*, respectivamente. Em princípio a função covariância expressa em (40) é por definição uma função covariância global. A sua determinação é feita com recurso ao cálculo de infinitos coeficientes com base em observações distribuídas por toda a Terra. Tscherning and Rapp (1974) desenvolveram alguns modelos da função covariância determinando uma expressão para a função covariância

empírica global. Nesse estudo, os dados estão distribuídos numa pequena região, pelo qual deveria ser introduzida alguma simplificação na determinação da função covariância, designando-se neste caso por função covariância local. A função covariância local foi definida por Goad et al.(1984) apud Fernandes (1999), que introduziu esta noção como “*um caso especial da função covariância global em que a informação de comprimento de onda superior à extensão da área local é removida e é assumido que a informação no exterior, mas junto à área, tem uma variação similar no seu interior*”. Estes sugerem que os modelos de Tscherning and Rapp (1974) devem ser utilizados de modo a ajustarem uma expressão aos valores empíricos, da mesma forma que no caso global.

As covariâncias associadas a qualquer elemento do campo de gravidade podem ser expressas em função de $K(P, Q)$. Representando a covariância de um elemento qualquer do campo anômalo por

$$C_{ij} = \text{cov}(l_i, l_j) = M\{l_i, l_j\} \quad (41)$$

Mas, os elementos de l podem ser representados por funcionais lineares de T . Assim, nos pontos P e Q , tem-se as expressões

$$l_i = L_i^P T(P) \quad (42)$$

e

$$l_j = L_j^Q T(Q) \quad (43)$$

que substituindo em (35), proporcionam

$$C_{ij} = M\{L_i^P T(P) L_j^Q T(Q)\} = L_i^P L_j^Q \{T(P)T(Q)\} \quad (44)$$

resultando, assim, a expressão que representa a propagação das covariâncias

$$C_{ij} = L_i^P L_j^Q K(P, Q) = \sum_{n=2}^{\infty} c_n \left(\frac{R_B}{r_P r_Q} \right)^{n+1} P_n(\cos\psi) \quad (45)$$

com $c_n = k_n \frac{(n-1)^2}{r_P r_Q}$

Os coeficientes c_n designam-se variâncias de grau e uma análise harmônica de um conjunto de valores estimados, permitirá determinar as primeiras n variâncias de grau, em que n depende do intervalo de amostragem.

Desta forma, pode-se obter as expressões para as covariâncias entre os elementos do campo de gravidade a partir de $K(P, Q)$.

$$\text{cov}(T_P, T_Q) = K(P, Q) \quad (46)$$

$$\text{cov}(\Delta g_p, \Delta g_q) = C(P, Q) = \frac{\partial^2}{\partial r_p \partial r_q} K(P, Q) + \frac{2}{r_p} \frac{\partial}{\partial r_q} K(P, Q) + \frac{2}{r_q} \frac{\partial}{\partial r_p} K(P, Q) + \frac{4}{r_p r_q} K(P, Q) \quad (47)$$

$$\text{cov}(\Delta g_p, \zeta_q) = -\frac{1}{\gamma_p} \left[\frac{\partial}{\partial r_p} K(P, Q) + \frac{2}{r_p} K(P, Q) \right] \quad (48)$$

$$\text{cov}(\zeta_p, \zeta_q) = \frac{K(P, Q)}{\gamma_p \gamma_q} \quad (49)$$

3.2. Ajuste da função covariância (Método de Knudsen)

A função covariância ajustada é determinada com base na covariância empírica local em que os parâmetros R_B , A e N são estimados por ajuste aos dados empíricos. Os critérios para o ajuste consistem em igualar valores dos parâmetros essenciais das duas funções covariância, ou seja, determinar valores para R_B , A e N de modo que a função covariância modelo tenha os mesmos valores da variância $C(0)$, a correlação (ξ_1) e o parâmetro de curvatura (χ), que representam a função covariância local. Para o ajuste da covariância empírica à covariância ajustada, definida pela expressão (50), Knudsen (1987) apresentou um método baseado na inversão por mínimos quadrados. O ajuste do modelo nessa expressão é conseguido por ajuste dos parâmetros R_B e A . Adicionalmente, é introduzido um terceiro parâmetro α , com este parâmetro, as variâncias de grau dos erros dos coeficientes do potencial são escalonadas, de modo que estes representem a qualidade da aproximação dos coeficientes do potencial na região de estudo e o sistema de referencia.

$$C(\psi) = \alpha \sum_{n=2}^N \sigma_n \left[\frac{R_B^2}{r_p r_q} \right]^{n+2} P_n(\cos \psi) + \sum_{n=N+1}^{\infty} \frac{A(n-1)}{(n-2)(n+B)} \left[\frac{R_B^2}{r_p r_q} \right]^{n+2} P_n(\cos \psi) \quad (50)$$

onde σ_n representa a variância de grau do erro contido no modelo do geopotencial, adotado na representação da componente determinística. O segundo termo da expressão (50), representa o modelo de função covariância proposto por Tscherning & Rapp (1974). Onde A (uma constante em unidades de (m/s^4)) e R_B (raio da esfera de *Bjehammar*) são determinados através de um ajustamento não linear (Knudsen, 1987). O número N é um valor fixo, que é igual ao grau de truncamento do modelo geopotencial, neste caso

($N=180$). A variável B , que descreve a estrutura da variância de grau, não pode ser obtida a partir de dados locais. Assim, geralmente usa-se o valor numérico ($B=24$), determinado por Tscherning & Rapp (1974), a partir de dados globais.

A determinação dos parâmetros para cada interação pode ser expressa (Knudsen, 1987)

$$\mathbf{x} - \mathbf{x}_0 = \left(A^T C_y^{-1} A + C_x^{-1} \right)^{-1} A^T C_y^{-1} (\mathbf{y} - \mathbf{y}_0) \quad (51)$$

onde \mathbf{x} é o vetor de parâmetros ajustados a partir dos valores aproximados do vetor \mathbf{x}_0 ; \mathbf{y} contém as covariâncias amostrais; \mathbf{y}_0 contém os valores gerados pelo modelo para $\mathbf{x}=\mathbf{x}_0$; A é a matriz Jacobiana contendo os valores de $\frac{\partial y_i}{\partial x_j}$; C_y é a matriz dos erros de \mathbf{y} , e C_x é a

matriz variância de $(\mathbf{x}-\mathbf{x}_0)$. Adotando os valores iniciais \mathbf{x}_0 para o modelo, aplica-se a expressão (51) e iterativamente determina-se o conjunto de parâmetros. A não linearidade da relação entre o modelo e R_B conduz à implementação de um processo iterativo, no qual, em cada interação, o ajuste dos parâmetros \mathbf{x} é calculado pela expressão (51).

O valor de N , relativo à ordem do modelo do geopotencial, é selecionado por ajuste do primeiro zero da função covariância empírica, uma vez que quanto maior for a ordem mais perto da origem esta o primeiro zero (ξ_0), local onde a função covariância é nula (Fernandes, 1999). A ordem de N é determinada pela relação $N = \frac{90^\circ}{\xi_0}$. O valor R_B , ou mais exatamente,

$(R_B - R_E)$ (*raio terrestre médio*), deverá ser relativamente pequeno, procurando que o ruído (diferença entre a variância empírica e ajustada) seja positivo. O raio da esfera de *Bjehammar* é substituído pela distancia a essa esfera $(R_B - R_E)$, de modo que se consiga obter uma maior precisão no ajuste deste valor face ao seu pequeno valor e o coeficiente A é um termo livre com unidades de mGal^2 .

A precisão do ajuste é dada pela expressão

$$Q^2 = \frac{1}{(n-m)} (\mathbf{y} - \mathbf{y}_0)^T C_y^{-1} (\mathbf{y} - \mathbf{y}_0) \quad (52)$$

Em que n é o numero de dados e m é o numero de parâmetros ($m=3$).

Para aumentar a precisão e reduzir o tempo computacional no cálculo das covariâncias, são usadas as *fórmulas finitas* (Sá, 1988), equivalentes a (50). A técnica utilizada para obter as expressões finitas é bastante simples e passa pela separação das expressões (40) e (45) em componentes, que multiplicadas por constantes apropriadas nos reconstruam a função covariância. Estas componentes são

$$F = \sum_{n=0}^{\infty} s^{n+1} P_n(t) = s \sum_{n=0}^{\infty} s^n P_n(t) = \frac{s}{\sqrt{1-2st+s^2}} \quad (53)$$

$$F_i = \sum_{n=0}^{\infty} \frac{1}{n+i} s^{n+1} P_n(t) \quad \text{para } i > 0 \quad (54)$$

$$F_i = \sum_{n=0}^{\infty} \frac{1}{n+i} s^{n+1} P_n(t) \quad \text{para } i \leq 0 \quad (55)$$

O denominador da expressão (53) é uma das quantidades básicas nos próximos desenvolvimentos pelo que foi utilizada a seguinte simplificação:

$$L = \sqrt{1-2st+s^2} \quad M = 1-L-st \\ N = 1+M-st \quad (56)$$

Com estas simplificações, os primeiros elementos das funções F_i , são dados por

$$F_{-2} = s \left(M \left(\frac{3ts+1}{2} \right) + s^2 \left(P_2(t) \ln \frac{2}{N} + \frac{1-t^2}{4} \right) \right) \quad (57)$$

$$F_{-1} = s \left(M + ts \ln \frac{2}{N} \right) \quad (58)$$

$$F_0 = s \ln \frac{2}{N} \quad F_1 = \ln \left(1 + \frac{2s}{1-s+L} \right) \quad (59)$$

Assim a variância de grau pode ser escrita como

$$c_n = \frac{1}{(B+2)(B+1)} \left[\frac{B+1}{n-2} - \frac{B+2}{n-1} + \frac{1}{n+B} \right] \quad (60)$$

Desta forma, a expressão da função covariância do potencial anômalo, que também depende desses parâmetros, tem a forma

$$K(\psi) = \frac{AR_B^2}{(B+2)(B+1)} \\ \left[\sum_{n=3}^{\infty} \frac{B+1}{n-2} s^{n+1} P_n(t) - \sum_{n=3}^{\infty} \frac{B+2}{n-1} s^{n+1} P_n(t) + \sum_{n=3}^{\infty} \frac{1}{n+B} s^{n+1} P_n(t) \right] \\ = \frac{AR_B^2}{(B+2)(B+1)} + \\ + \left[(B+1)F_{-2} - (B+2)(F_{-1} - s^3 P(t)) - \frac{s}{B} - \frac{s^2 t}{B+1} - \frac{s^3 P_2(t)}{B+2} \right] \quad (61)$$

onde

$$s = \left(\frac{R_B^2}{r_p r_q} \right) \quad (62)$$

e

$$t = \cos \psi \quad (63)$$

As funções são funções auxiliares que dependem de s e t (Sá, 1988), a relação entre para os elementos do campo de gravidade envolvidos nesta pesquisa, podem ser expressas (Sá, 1988)

$$\text{cov}(\zeta_p, \zeta_q) = \frac{1}{\gamma_p \gamma_q} \left\{ \frac{AR_B^2}{(B+2)(B+1)} + \left[(B+1)F_{-2} - (B+2)(F_{-1} - s^3 P(t)) \right] \right. \\ \left. + \left[-\frac{s}{B} - \frac{s^2 t}{B+1} - \frac{s^3 P_2(t)}{B+2} \right] \right\} \quad (64)$$

para a anomalias de altitude, e

$$\text{cov}(\Delta g_p, \zeta_q) = \frac{AR_B^2}{r_p \gamma_p (B+2)} + \\ + \left[F_{-2} - F_B - \frac{s}{B} - \frac{s^2 t}{B+1} - \frac{s^3 P_2(t)}{B+2} \right] \quad (65)$$

para as covariâncias entre a anomalia ar-livre e a anomalia de altitude, onde γ é a gravidade teórica no ponto, geralmente substituída pelo valor da aceleração de gravidade média na superfície.

Há, ainda, a possibilidade de estudar a representação das covariâncias amostrais, como funções covariâncias planas, neste caso, poderá ser utilizado o modelo logarítmico desenvolvido por Forsberg (1987). Sá (2005) calculou o geóide gravimétrico para a América do Sul utilizando-se desse modelo, e comparou com outros modelos anteriormente determinados para América do Sul (Barzaghi et al., 2002 *apud* Sá, 2005), demonstrando que a metodologia utilizada é eficiente para a determinação de modelos geoidais de alta resolução espacial e alta precisão. Neste trabalho, será analisada apenas a modelagem de funções covariâncias esféricas.

Outros modelos matemáticos podem ser usados na representação de covariâncias planas, como por exemplo, o modelo *Markov* de segunda ordem. Normalmente, esses modelos são utilizados na interpolação de dados, porque no seu sentido mais amplo, a colocação por mínimos quadrados combina ajustamento, filtragem e interpolação.

4. APLICAÇÃO A DADOS REAIS

Os métodos de aquisição de dados podem ser: por pontos amostrados com espaçamento irregular e regular. A distribuição irregular implica que os dados estão disponibilizados da mesma forma que são coletados no campo, na regular são igualmente espaçados formando uma malha.

Os dados gravimétricos, que foram utilizados neste trabalho, são resultado da integração de três tipos de informações: terrestres, oceânicas e de altimetria por satélite. As informações terrestres foram obtidas em levantamentos realizados por várias instituições (Souza, 2002), com objetivos diversos; as oceânicas foram coletadas por instituições internacionais nos levantamentos de Geofísica marinha, processadas e fornecidas pelo *U. S. National Geophysical Data Center* (NGDC); as de altimetria por satélite são aquelas que resultaram das missões GEOSAT e ERS-1, convertidas em anomalias ar-livre na forma digital (Andersen & Knudsen, 1998). As medidas terrestres foram referidas a *International Gravity Standardization Net 1971* (IGSN71) através da Rede Gravimétrica Fundamental Brasileira (RGFB), implantada a partir de 1975. Os dados gravimétricos terrestres foram acrescidos com os de levantamentos recentes realizados por diversas instituições. Com a junção de todos os dados gravimétricos existentes (terrestres e oceânicos), obtendo um arquivo com 31000 registros (Figura 2). Com esses dados foi gerado um modelo digital, e as observações foram reduzidas para 10285 pontos, e

delimitadas por uma área com as seguintes coordenadas: paralelos (-27° e -17°) e os meridianos (-56° e -42°) e resolução de 5', para maiores detalhes ver Souza (2002). Nos cálculos são utilizados os programas *EMPCOV* (Tscherning, 1997) para a modelagem das covariâncias amostrais, e o *COVFIT* (Knudsen, 1987), para o ajuste dos parâmetros da função covariância esférica. Para melhor entender o comportamento das funções covariâncias, os dados gravimétricos foram divididos em blocos com 6° graus de amplitude em latitude e longitude, tanto para os dados irregulares (Figura 2), como para o modelo digital. Com isso, obteve-se 6 blocos quadrados (Figura 2), que foram diferenciados pelas letras A, B, C, D, E e F. Este procedimento ajudará na análise do comportamento das funções covariâncias obtidas de anomalias gravimétricas em diferentes regiões da área estudada, a partir de dados distribuídos regularmente e irregularmente.

A pergunta principal é porque não se pode modelar uma única função covariância para a região de estudo.

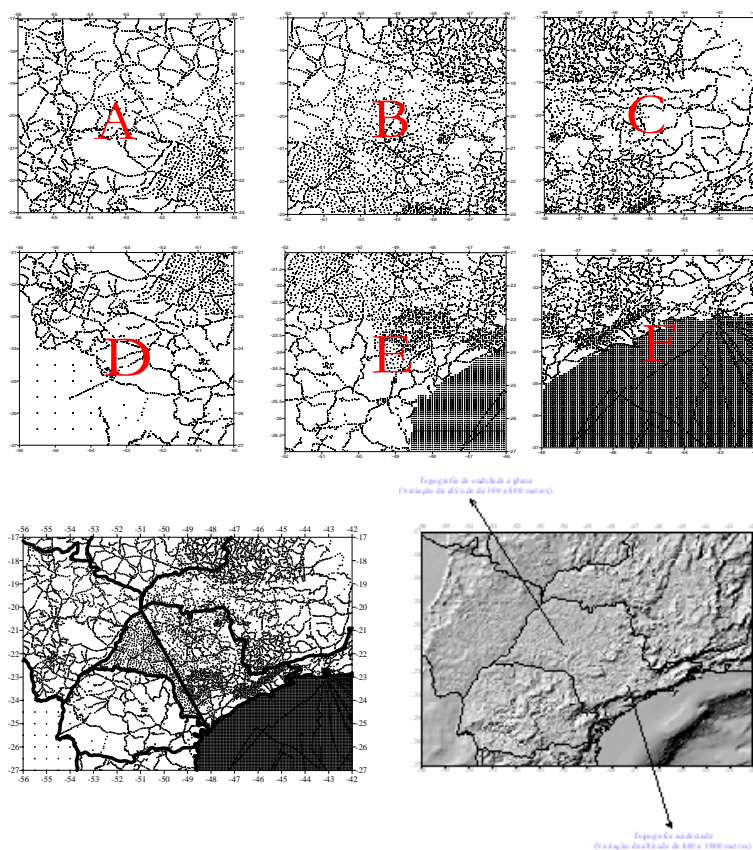


Figura 2 - Divisão da Área em Blocos, Distribuição dos Dados Gravimétricos e Modelo Digital de Terreno

A determinação da covariância local será efetuada para cada Bloco, estimando os coeficientes de um modelo paramétrico, por meio do ajuste aos dados empíricos da função covariância. A primeira decisão consiste na definição da dimensão dos blocos estudados,

que deverão ser suficientemente grandes para dispormos de dados de observação que cumpram a condição do método (média zero), mas por outro lado, devem respeitar a condição de que o campo de gravidade seja homogêneo e isotrópico; respondendo a pergunta levantada no parágrafo anterior.

5. EXPERIMENTOS

A distribuição e amostragem espacial dos dados é um fator importante para a modelagem das funções covariâncias. A amostragem compreende a aquisição de um conjunto de amostras que geralmente representa um determinado elemento do campo de gravidade. Ela não pode ser insuficiente, nem tampouco redundante. A falta de dados pode levar a modelagem de funções covariâncias pobres. A superabundância, ou a excessiva quantidade de dados, sobrecarrega o sistema com o uso excessivo de memória computacional e pode prejudicar os resultados de uma modelagem. Porque nenhuma modelagem, por mais complexa ou sofisticada que seja, pode compensar os efeitos de uma amostragem mal feita.

A distribuição irregular de dados é vantajosa quando se quer evitar os erros provenientes da geração de modelos digitais. A determinação de modelos digitais é relativamente simples, pois existem vários algoritmos disponíveis para isso. Entretanto, a escolha do procedimento para a interpolação dos dados não é uma tarefa fácil. A modelagem de funções covariâncias, neste caso, é diretamente afetada pela distribuição espacial dos dados. A distribuição espacial dos dados no processo de estimação da função covariância é uma condição importante na determinação dos parâmetros, podendo prejudicar os resultados obtidos. Para que um efeito melhor deve-se considerar não somente a quantidade, mas também o seu posicionamento em relação aos demais. A principal dificuldade de utilizar dados com distribuição irregular é verificar a isotropia e homogeneidade. Que na prática são de interesse, devido a simplicidade de cálculo que proporcionam; covariâncias em função apenas da distância relativa entre os pontos.

As covariâncias amostrais foram determinadas a partir da expressão (28), utilizando as anomalias

gravimétricas residuais distribuídas de forma irregular, e utilizando-se do programa *EMPCOV*. Os valores obtidos servem de entrada para o programa *COVFIT*, onde são modelados os parâmetros da função covariância esférica, definidos na expressão (50). Neste caso, não é possível a determinação de covariâncias bidimensionais, para a verificação de isotropia, porque só é possível com os dados distribuídos regularmente. Os valores estimados, para as covariâncias amostrais e ajustadas, para os dados irregulares, são observados na Tabela 1.

TABELA 1 - COVARIÂNCIAS E PARÂMETROS AJUSTADOS PARA OS BLOCOS COM DADOS IRREGULARES

BLOCO	Covariâncias Amostrais (mGal ²)	Covariâncias Ajustada (mGal ²)	ξ (km)	$R_B - R$ (km)
A	138,2919	137,7401	17	-3541,126
B	245,3668	243,5838	19	-3979,994
C	480,2679	478,6879	17	-3695,113
D	207,8470	207,5220	15	-2464,568
E	410,2448	409,7262	25	-4428,017
F	367,4591	367,2670	17	-3692,807

A Figura 3 ilustra graficamente o comportamento das funções covariâncias empíricas e ajustadas exemplificada para o Bloco A. Nota-se que o modelo de função covariância esférica, ajusta-se melhor quando dos dados são distribuídos regularmente, do que com os irregulares. Vê-se com isso que a função covariância, e consequentemente os seus parâmetros, são influenciados pela distribuição dos dados. Por isso, para a região de estudo, não seria conveniente a escolha de uma única função, pois as covariâncias estimadas variam de maneira significativa de uma região para outra.

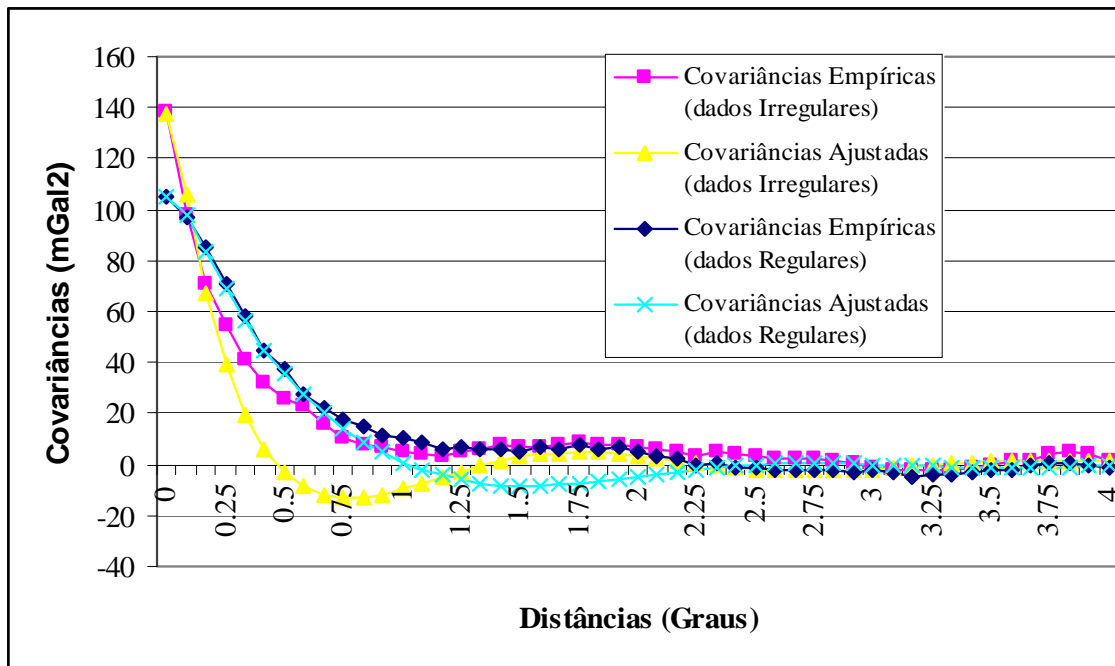


Figura 3 – Representação das Funções Covariâncias Empíricas e Ajustadas para dados com distribuição Regular e Irregular (Exemplo: Bloco A)

Para os dados regulares, os resultados obtidos para as variâncias empíricas e ajustadas são vistos na Tabela 2.

TABELA 2 - COVARIÂNCIAS E PARÂMETROS AJUSTADOS PARA OS BLOCOS COM DADOS REGULARES

BLOCO	Covariâncias Amostrais (mGal ²)	Covariâncias Ajustada (mGal ²)	ξ (km)	$R_B - R$ (km)
A	105,4752	105,3492	38	-6982,451
B	148,5823	148,0662	28	-7955,999
C	301,4680	300,8120	23	-6349,202
D	129,8566	128,8305	24	-8081,289
E	281,2315	279,3523	27	-5026,052
F	293,1245	292,1379	25	-8119,423

No caso de dados distribuídos regularmente, a determinação e modelagem de funções covariâncias são realizadas explorando a distribuição geográfica dos dados. Para isso, determinam-se as chamadas covariâncias bidimensionais (Figura 4), que fazem uso dos modelos gravimétricos digitais residuais.

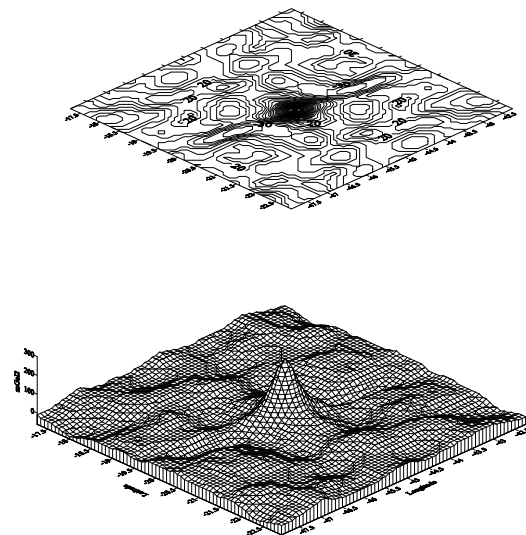


Figura 4 – Ilustração de Covariâncias Bidimensionais

As covariâncias bidimensionais são determinadas em função de duas variáveis r e s . O procedimento de cálculo é uma simples extensão da técnica que se usa na obtenção das unidimensionais, para elementos com distribuição regular, expressa em (28). Na prática, a determinação de covariâncias isotrópicas, é de maior interesse devido a sua simplicidade de cálculo. Contudo, é necessário verificar até que distância pode-se supor isotropia. Portanto, a covariância no ponto de coordenadas r_k e s_l com origem no centro da área, é representada por (Sá, 1988).

$$C(r_i, s_j) = \frac{1}{(N-k)(N-l)} \sum_{i=1}^{N-k} \sum_{j=1}^{N-l} \Delta g_{r_i} \Delta g_{s_j} \quad (66)$$

$$C(r_i, -s_i) = \frac{1}{(N-k)(N-l)} \sum_{i=1}^{N-k} \sum_{j=1}^N \Delta g_{r_i} \Delta g_{s_{i+j}} \quad (67)$$

tem-se, pela propriedade de simetria

$$C(-r_i, -s_i) = C(r_i, s_i) \quad (68)$$

e

$$C(-r_i, s_i) = C(r_i, -s_i) \quad (69)$$

que completam as expressões para a obtenção de covariâncias nos quatro quadrantes.

As covariâncias bidimensionais, calculadas para as anomalias, indicam que a condição de isotropia em geral é válida para distâncias inferiores ao triplo da distância de correlação.

6. CONCLUSÕES

Observando os resultados encontrados para os dados irregulares, na Tabela 1; vê-se que os blocos apresentam valores distintos para as covariâncias amostrais, como era de se esperar. As causas estão relacionadas a diversos fatores; primeiramente, a representação da componente determinística pelo modelo do geopotencial, que pode ser eficiente em algumas áreas e deficiente em outras. É deficiente quando não é possível representar feições locais, a partir dos dados existentes; ou nas áreas onde os dados existentes foram utilizados na determinação dos seus coeficientes. A topografia é outro aspecto que pode influenciar na modelagem de funções covariâncias, visto que as anomalias gravimétricas são altamente correlacionadas com a topografia. Como as anomalias ar-livre são altamente correlacionadas com a topografia, espera-se um comportamento semelhante, ou seja, nas áreas planas uma variação menor do que nos locais de onde a topografia é mais acidentada.

Pode-se notar a ocorrência destes casos nos blocos selecionados. Os blocos A e D apresentam as menores variâncias em relação aos demais blocos; nestas regiões a topografia é plana, e os blocos apresentam poucas observações gravimétricas. Nos blocos restantes, verifica-se que a topografia é mais elevada, e existe uma maior quantidade de dados gravimétricos, incluindo os levantamentos gravimétricos mais recentes. É provável que os coeficientes do modelo do geopotencial não representem de maneira adequada o campo de gravidade nestas regiões, devido a variação topográfica existente e dos dados gravimétricos que não foram incluídos na determinação dos coeficientes do modelo geopotencial.

Comparando os dados da Tabela 2, com a mesma estimada para os dados irregulares, nota-se que houve uma suavização das covariâncias ajustadas. Este efeito é causado na geração de modelos digitais, quanto maior a resolução, mais suavizada será a resposta encontrada pelas funções covariâncias. Uma completa

interpretação do efeito da suavização não é possível; contudo, um tipo de suavização é óbvia; a da utilização do modelo digital, que implica em uma estimativa média. Como no caso anterior, os parâmetros foram estimados para cada um dos blocos, sendo novamente diferentes, reforçando mais uma vez a idéia que estes parâmetros são modificados conforme a característica e a distribuição dos dados locais. As distâncias de correlação estimadas para estas funções covariâncias são maiores do que as determinadas para os dados irregulares. A distribuição regular dos dados, novamente, é responsável por este efeito. No trabalho de Kotsakis (2007), é proposta algumas alternativas para a solução do problema da suavização da função covariância quando são utilizados dados regulares.

O raio da esfera de Bjerhammar é o parâmetro mais afetado no ajuste das funções covariâncias. Por ser um valor pequeno, na verdade o que se ajusta é a diferença ($R_B - R_E$), quaisquer variações nos outros parâmetros, por menor que seja, modifica o seu valor bruscamente; conforme pode ser visto nas tabelas 1 e 2. A exatidão da estimativa é muito baixa, sendo impossível garantir qualquer resultado que se obtenha deste processo de ajuste. A mesma conclusão pode ser observada na tese de doutorado de Fernandes (1999).

Qualquer modelagem mal feita tem implicações diretas na determinação e estimativa do erro do funcional calculado, como pode ser visto nas expressões (17 e 18). Pode-se afirmar que para cada bloco existe a possibilidade da determinação de modelos geoidais diferentes, isto porque a distribuição de dados na área tem um papel fundamental na modelagem de funções covariâncias. Os problemas, levantados neste trabalho, constitui uma estratégia para ser aplicada na determinação do geóide, na obtenção de um modelo de alta resolução utilizando-se da CMQ. Isto porque, essa metodologia vem sendo empregada na determinação de geóides (Genro e Quintas (2007); Souza (2002); Sá (2005, 1988).

AGRADECIMENTOS

Os dados utilizados resultaram de vários projetos que contaram com auxílios da FAPESP, PADCT/FINEP e CNPq. Agradecemos aos programas colocados à nossa disposição por C. C. Tscherning e P. Knudsen.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANDERSEN, O.B. & KNUDSEN, P. –1998- Global Marine gravity Field from the ERS-1 and GEOSAT Geodetic Mission Altimeter, *J. Geoph. Res.* **102**(C4), 8129-8137.
- BARZAGHI, R.; BORGHI, A.; SANSÒ, F.; BLITZKOW, D.; LOBIANCO, M.C.; KENYON, S.C.; BEALE, J. & FACTOR, J. –2002- Quasi-geoid

estimations in South America. **Anais IAG – International Symposium**, Berlin, Alemanha, 8 pp.

FERNANDES, J. C.C. C. -1999- Modelação do Campo Gravítico na Região do Atlântico Oriental (Açores-Portugal Continental). **Tese de Doutorado**. Lisboa, PT. 298p.

FORSBERG, R. -1987- A new covariance model for inertial gravimetry and gradiometry. **Journal of Geophysical Research**, **92(B2)**, 1305-1310.

GENRO, R.; QUINTAS, M.- 2007- Determinação de um Geóide Gravimétrico para o Estado do Paraná utilizando Colocação por Mínimos Quadrados. **Boletim de Ciências Geodésicas**, América do Sul, v. 13, n.1: 186-208p.

GOAD, C.C.; TSCHERNING, C.C.; CHIN, M.M - 1984- Gravity empirical covariance values for the continental United States”. **Journal of Geophysical Research**, **89(B9)**, 7962-7968.

HEISKANEN, W. & MORITZ, H. -1967- **Physical Geodesy**. W.H. Freeman., N. York, 364 pp.

KNUDSEN, P. -1987- Estimation and modelling of the local empirical covariance function using gravity and satellite altimeter data, **Bull. Géodésique**, **61**: 145-160p.

KOTSAKIS C -2007- Least-squares collocation with covariance-matching constraints. **Journal of Geodesy** 81:661–677

MORITZ, H. -1980- **Advanced Physical Geodesy**. Abacus Press, Wichmann, 500 pp.

SÁ, N.C.de -1988- Modelos estatísticos e representação integrada do campo de gravidade no Brasil. **Tese de Doutorado, IAG/USP**, São Paulo, 240 pp.

SÁ, N.C.de -2005- O Geóide Gravimétrico na América do Sul obtido com uso da Colocação por Mínimos Quadrados. In: XXII Congresso Brasileiro de Cartografia. Macaé, RJ. **Anais**. Macaé-RJ: Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto.

SOUZA, S. F. -2002- Contribuição do GPS para Aprimoramento do Geóide no Estado de São Paulo. **Tese de Doutorado, IAG/USP**, São Paulo, 240 pp.

TSCHERNING, C.C. & RAPP, R.H. -1974- Closed covariance expressions for gravity anomalies, geoid undulations and deflections of the vertical implied by anomaly degree-variance models, Ohio State University. **Report. No 208**.

TSCHERNING, C.C. -1997- Geoid Determination by Least-square Collocation Using GRAVSOF. In:

Lectures Notes of the 2nd. International School for the Determination and Use of the Geoid. International Geoid Service, Rio de Janeiro. 135-164.