Technical Sciences 16(4), 2013, 291–307

ESTIMATION OF COVARIANCE PARAMETERS FOR GNSS/LEVELING GEOID DATA BY LEAVE-ONE-OUT VALIDATION

Wojciech Jarmołowski

Department of Satellite Geodesy and Navigation
University of Warmia and Mazury in Olsztyn

Received 2 August 2013; accepted 6 November 2013; avilable on line 12 November 2013

Key words: GNSS/leveling, geoid, least squares collocation, leave-one-out (LOO), covariance, noise.

Abstract

The article describes the estimation of covariance parameters in Least Squares Collocation (LSC) by Leave-One-Out (LOO) validation, which is often considered as a kind of cross validation (CV). Two examples of GNSS/leveling (GNSS/lev) geoid data, characterized by different area extent and resolution are applied in the numerical test. A special attention is focused on the noise, which is not correlated in this case. The noise variance is set to be homogeneous for all points. Two parameters in three covariance models are analyzed via LOO, together with a priori noise standard deviation, which is a third parameter.

The LOO validation finds individual parameters for different applied functions i.e. different correlation lengths and a priori noise standard deviations. Diverse standard deviations of a priori noise found for individual datasets illustrate a relevance of applying LOO in LSC. Two examples of data representing different spatial resolutions require individual noise covariance matrices to obtain optimal LSC results in terms of RMS in LOO validation. The computation of appropriate a priori noise variance is however difficult via typical covariance function fitting, especially in the case of sparse GNSS/leveling geoid data. Therefore LOO validation may be helpful in describing how the a priori noise parameter may affect LSC result and a posteriori error.

List of abbreviations

LSC – least squares collocation
LOO – leave-one-out validation
ECF – empirical covariance function
CV – cross-validation
GNSS/lev – GNSS/leveling geoid height

* Corespondence: Wojciech Jarmołowski, Katedra Geodezji Satelitarnej i Nawigacji, Uniwersytet Warmińsko-Mazurski, ul. Heweliusza 5, 10-724 Olsztyn, Poland, Phone: + 48 89 523 47 64, Fax + 48 89 523 47 23, e-mail: wojciech.jarmolowski@uwm.edu.pl

SSH	– sea surface heights
C_0	– signal variance parameter in covariance model
δn	– a priori noise standard deviation in covariance model
CL	– correlation length in covariance model
CG	– Gaussian covariance model
CGM2	– Gauss-Markov second order covariance model
CGM3	– Gauss-Markov third order covariance model
RMS	– root mean square
RMSL	– root mean square in leave-one-out validation

Introduction

GNSS/leveling (GNSS/lev) data provide an independent source of geoid height information, aside from the gravity data. Both sources may be used to calculate the same functional of disturbing potential with comparable accuracy. However, the spatial resolution of GNSS/lev is usually worse than gravity data sampling. Therefore a special attention is focused in the article on the relation of the horizontal data distribution with the modeling accuracy. GNSS/lev data constitute irreplaceable source of geoid information, used most frequently for an assessment of local gravimetric models (DARBEHESHTI, FEATHERSTONE 2010, ILIFFE et al. 2003, SMITH, MILBERT 1999), global geopotential models (ŁYSZKOWICZ 2009) or also models based on satellite gradiometry (GODAH, KRYŃSKI 2011). GNSS/lev is also frequently used in the combination with gravity or other data in local geoid or quasigeoid modeling (OSADA et al. 2005, TROJANOWICZ 2012). GNSS/lev data provide the information on the local relation between leveling heights and geometric ellipsoidal heights, which is especially useful before the unification of height systems at the global scale. The combination of GNSS with the locally matched geoid is still a common practice in height determination with GNSS (DAWIDOWICZ 2012). The combination of GNSS and leveling heights may be sometimes used alone, as the only source of geoid information, which is compatible with the local height system. This is also applicable in surveying, for dense data and small areas.

Many authors use Least Squares Collocation (LSC) for modeling of the GNSS/lev data or residuals of GNSS/lev and other data (DARBEHESHTI, FEATHERSTONE 2009, DENKER 1998, SMITH, MILBERT 1999). Some authors analyze the efficiency of other methods for GNSS/lev geoid modeling, but use LSC for the validation purposes (KAVZOGLU, SAKA 2005). LSC is also frequently applied in the investigations closely related to geoid or heights in general, i.e.: sea surface heights (SSH) derived from the satellite altimetry

(ANDERSEN, KNUDSEN 1998), vertical crustal movements (EL-FIKY et al. 1997, KOWALCZYK et al. 2010), gravity anomalies (CATALAO, SEVILLA 2009) or deflections of the vertical (ŁYSZKOWICZ 2010a, ŁYSZKOWICZ 2010b). The spectrum of LSC applications is wide, which may be caused by the fact that the covariance in LSC may be modeled in detail and a posteriori error estimates may be calculated. A very frequent method of the covariance estimation is the selection of the analytical model based on the empirical covariance function (ECF) values (HOFMANN-WELLENHOF, MORITZ 2005, MORITZ 1980). There are many practical examples of different numerical techniques of fitting the analytical model into empirical covariance values (ARABELOS, TSCHERNING 2003, DARBEHESHTI, FEATHERSTONE 2009, SMITH, MILBERT 1999). The planar covariance models are often investigated as well, as the spherical models. The standard deviation of a priori noise (δn) is hard to determine using ECF, however YOU, HWANG (2006) have noted its significant role among the covariance parameters. They have also found that different parameters may have an influence on the prediction error in terms of cross-validation (CV) error. The problem of the parameter estimation occurs in the context of combined geoid modeling (FOTOPOULOS et al. 2003) and in the regularization of gravity field from satellite gradiometry (KUSCHE, KLEES 2002). DARBEHESHTI and FEATHERSTONE (2009) apply the non-stationarity for better representation of the local covariance. Among different considerations, the most closely related research has been found in JEKELI, GARCIA (2002), in MARCHENKO et al. (2003) and in MORITZ (1980). They investigate similar problem of optimal covariance matrix by applying Tikhonov regularization parameter. In this paper the problem is treated numerically using a kind of CV. LOO provides some additional observations, which introduces a look from some other side than it is presented in the mentioned works.

The most probable covariance parameters are investigated in this work and a special emphasis is placed on spatially sparse data distribution, which is common in e.g. vertical movements data (EL-FIKY et al. 1997) or deflections of the vertical (ŁYSZKOWICZ 2010a, ŁYSZKOWICZ 2010b). Two datasets of GNSS/lev are used in the numerical test: sparse regional data and significantly denser local data. Three typical, planar covariance models are applied to compare estimates of the same parameters applied in different functions. Two covariance function parameters: correlated signal variance (C_0) and correlation length (CL) are analyzed in the work, together with δn parameter. The method used for the empirical assessment of the parameters is a frequently used form of CV, named leave-one-out validation (LOO) (ARLOT, CELISSE 2010, KOHAVI 1995, KUSCHE, KLEES 2002).

LOO validation applied in LSC

The scalar quantity distributed in 2D space may be represented by the addition of deterministic part and residuals of the signal (MORITZ 1980, RAO, TOUTENBURG 1995). In our case we have:

$$\mathbf{N} = \mathbf{X}\boldsymbol{\beta} + \mathbf{N}^{\mathrm{r}} \tag{1}$$

where **N** is the vector of observed geoid heights, **β** is the vector of unknown trend parameters and **X** is the design matrix of the trend. Different forms of the deterministic part of the signal are often used in practice (GREBENITCHARSKY et al. 2005, OSADA et al. 2005). Deterministic part of the signal is commonly called trend in many papers. This part of the signal approximates large-scale features of the stochastic field, commonly known as long-wavelength part of the signal in the spectral analysis. The trend removal extracts local properties of the data, occuring at higher spatial resolutions. The matrix **X** used for detrending of the data in the case of current test reads:

$$\mathbf{X} = \begin{bmatrix} 1 & \varphi_1 & \lambda_1 & \varphi_1{}^2 & \lambda_1{}^2 & \varphi_1\lambda_1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \varphi_n & \lambda_n & \varphi_n{}^2 & \lambda_n{}^2 & \varphi_n\lambda_n \end{bmatrix} \tag{2}$$

where n is the number of the observations. Such parametric form of the trend is based on the data distribution and is sufficient to obtain the residuals, which have expected value close to zero and can be modeled efficiently by the covariance function. The so-called projection matrix **P** is used for data detrending (RAO, TOUTENBURG 1995):

$$\mathbf{P} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^T\,\mathbf{X})^{-1}\,\mathbf{X}^T \tag{3}$$

Some authors have reported that a linear dependency in the matrix **P** exists for the number of rows equal to the rank of **X** (KITANIDIS 1983, RAO, TOUTENBURG 1995). If we denote the rank of the matrix **X** by p, we may remove p rows from **P** matrix. Since p number has no significance in relation to the whole set here, we compute **Λ** matrix by removing last p rows from **P**. The residuals can be computed, as follows:

$$\mathbf{N}^{\mathrm{r}} = \boldsymbol{\Lambda}\mathbf{N} \tag{4}$$

where $\boldsymbol{\Lambda} = \mathbf{P}_{(n-p)\times n}$. The LSC equation for detrended GNSS/lev data reads:

$$\tilde{\mathbf{N}}^r = \mathbf{C}_P^T \cdot (\mathbf{C} + \mathbf{D})^{-1} \cdot \mathbf{N}^r \tag{5}$$

where $\tilde{\mathbf{N}}^r$ is the vector of predicted, residual values. **C** is the covariance matrix of the data signal, $\mathbf{C}_P$ is the covariance matrix between predicted points and the data and **D** represents the noise covariance matrix. In this case, when we investigate average estimates of a priori noise, the noise is assumed to be homogeneous and uncorrelated, i.e.:

$$\mathbf{D} = \delta n^2 \cdot \mathbf{I}_n \tag{6}$$

where δn represents a priori noise standard deviation, homogeneous for all points in the example test. The studies can be extended to individual data noise values, however this implies weighting and therefore more detail information on the data noise is needed.

Three planar covariance models are applied in the numerical experiment. The choice of the models was arbitrary in some sense, but on the other hand, these models appear in the literature in the same context or even with the same or similar data related to the height system problems. The models are: Gaussian model (DARBEHESHTI, FEATHERSTONE 2009, YOU, HWANG 2006), Gauss-Markov second order model (ANDERSEN, KNUDSEN 1998, ILIFFE et al. 2003, STRYKOWSKI, FORSBERG 1998) and Gauss-Markov third order model (GREBENITCHARSKY et al. 2005, KAVZOGLU, SAKA 2005). The functional covariance models are described by Eq. (7) – Gaussian (CG), Eq. (8) – Gauss-Markov second order (CGM2) and Eq. (9) – Gauss-Markov third order model (CGM3).

$$\mathrm{CG}(C_0, CL, s) = C_0 \cdot \exp\left(\frac{-s^2}{CL^2}\right) \tag{7}$$

$$\mathrm{CGM2}(C_0, CL, s) = C_0 \left(1 + \frac{s}{CL}\right) \cdot \exp\left(\frac{-s}{CL}\right) \tag{8}$$

$$\mathrm{CGM3}(C_0, CL, s) = C_0 \left(1 + \frac{s}{CL} + \frac{s^2}{3 \cdot CL^2}\right) \cdot \exp\left(\frac{-s}{CL}\right) \tag{9}$$

Although spherical distance is a typical variable in the spherical covariance models e.g. Tscherning-Rapp model (ARABELOS, TSCHERNING 2003, HOFMANN-WELLENHOF, MORITZ 2005), it is adopted here to work as a variable in the planar models (Eqs 7–9). This choice is based on the assumption of no advantage coming from cartographic projection, since the area of the regional data is large and the distortion can be significant. The variable distance s is therefore calculated using the spherical distance formula (Eq. 12), also in the ECF and $s = \psi$ in the article.

The collocation formula (Eq. 5) is applied in CV test by LOO (Eq. 10). This method is quite frequently applied in the literature (DARBEHESHTI, FEATHERSTONE 2009, KUSCHE, KLEES 2002), but other, often similar kind of CV may be also efficient. The formula of root mean square (RMS) in LOO (RMSL) may be written as:

$$\mathrm{RMSL}((C_0,\ CL,\ \delta n)\,|\,(\tilde{\mathbf{N}}^r_{n\times 1},\ \mathbf{N}^r_{n\times 1})) = \sqrt{\frac{\sum\limits_{i=1}^{n}(\tilde{N}^r_i - N^r_i)^2}{n}}\ \Big|\, N^r_i \notin \mathbf{N}^r_{(n\text{-}1)\times 1} \qquad (10)$$

The vector $\tilde{\mathbf{N}}^r$ estimated using Eq. (5) is compared to $\mathbf{N}^r$ in terms of RMS and this is repeated for every set of covariance parameters. The vector of the residuals $\mathbf{N}^r$ is replaced by the vector $\mathbf{N}^r_{(n\text{-}1)\times 1}$ in the Eq.(5), i.e. the analyzed point i is omitted in the vector as well as in the matrices $\mathbf{C_P}$, $\mathbf{C}$ and $\mathbf{D}$. RMSL is a measure of prediction precision with variable parameters. Additionally the estimates of a posteriori error are provided in the numerical part of the article. Assuming now that $\mathbf{C_P}$ is the vector of covariances limited to one point only, the error of the prediction is (HOFMANN-WELLENHOF, MORITZ 2005):

$$m_P^2 = C_0 - \mathbf{C}_P^T \cdot (\mathbf{C} + \mathbf{D})^{-1} \cdot \mathbf{C}_P \qquad (11)$$

The estimation of covariance parameters by LOO validation is preceded and supported by ECF estimation. ECF provides an initial assessment of the residual data and an approximation of C_0 and CL. These initial values support a search of the parameters in LOO estimation and may be posteriorly compared with LOO results. ECF may be calculated by well-known formula (HOFMANN-WELLENHOF, MORITZ 2005):

$$\forall(i,j)\ \Big|\cos\psi = \cos\theta_i \cos\theta_j + \sin\theta_i \sin\theta_j \cos(\lambda_i - \lambda_j)$$
$$\mathrm{EC}(\psi\ |\mathbf{N}^r) = \frac{\sum\limits_{i,j}^{k} N^r_i N^r_j}{k} \qquad (12)$$

The spherical distance is used as a variable and $\theta_i = \pi/2 - \varphi_i$. The products of residuals are grouped using rings of constant width, which is determined by the sampling interval of the ECF. Average products form the ECF, which is

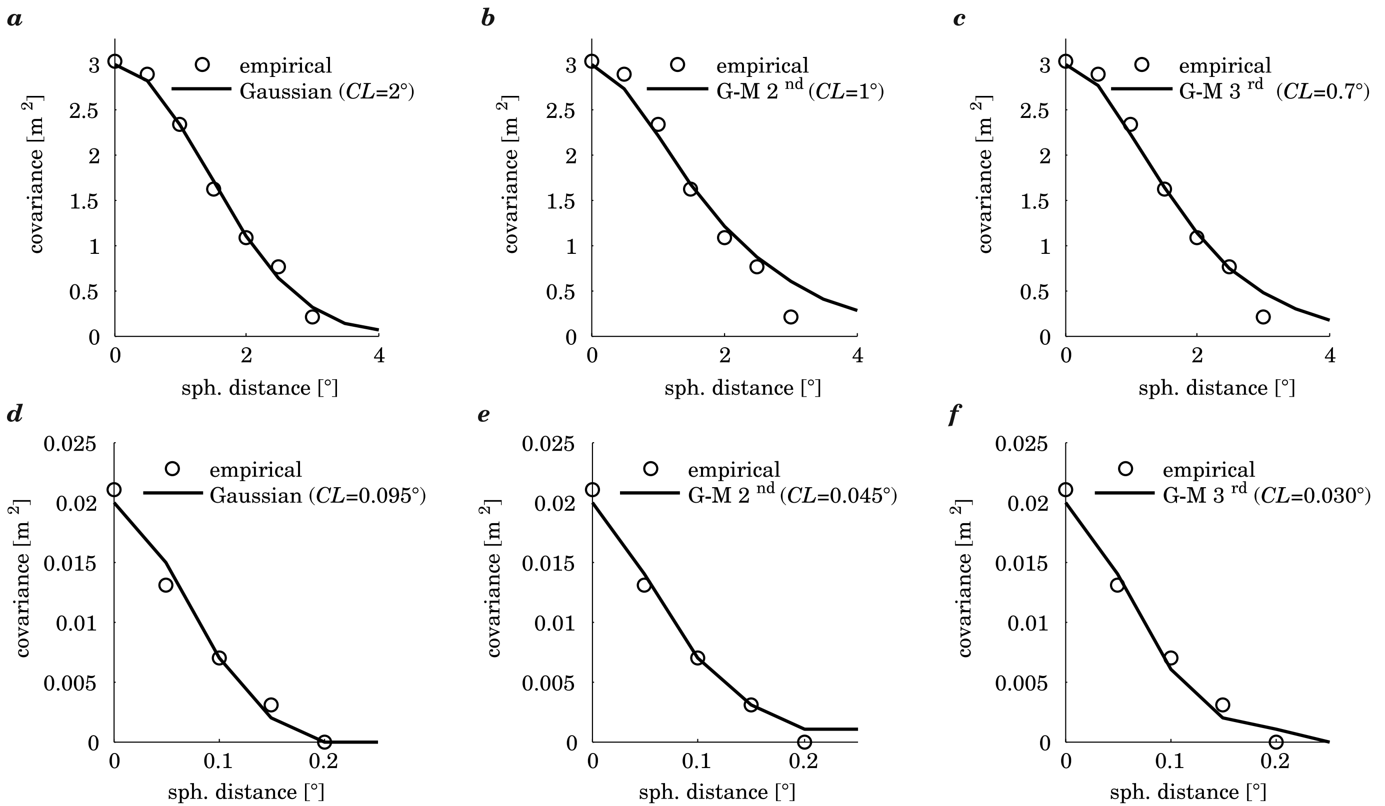

Fig. 1. ECFs of residuals and analytical covariance models fitted graphically disregarding δn: *a* – regional CG, *b* – regional CGM2, *c* – regional CGM3, *d* – local CG, *e* – local CGM2, *f* – local CGM3

shown in Figure 1. Analytical covariance models (Eqs 7–9) are approximately fitted to ECF by graphical manipulation of C_0 and CL, which are roughly determined this way.

Data assessment and numerical experiment

The data are acquired from the National Geodetic Survey website. GNSS/lev geoid heights are located in the area of the North America. Two subsets are selected from the data and the first set is slightly reduced to obtain homogeneous horizontal distribution (Fig. 2*c*). The first set of GNSS/lev geoid heights (Fig. 2*a*) covers a large area in the western part of the continent and therefore will be named regional dataset in further considerations. The second dataset (Fig. 2*b*) represents much smaller area in Texas and consequently

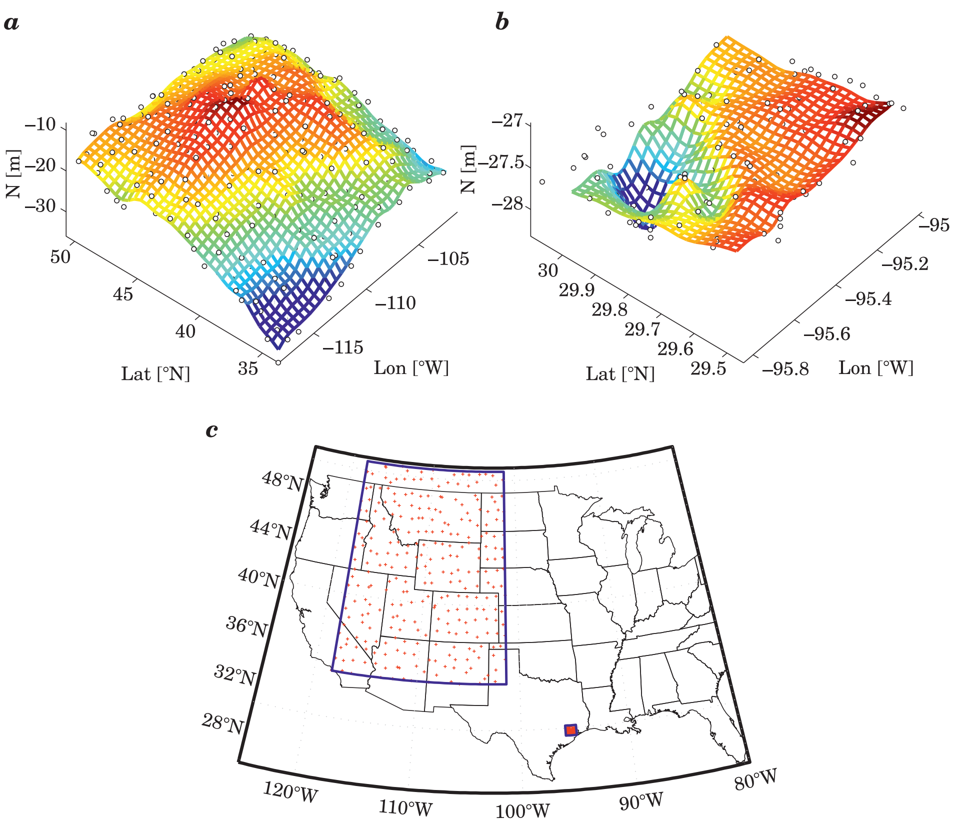

Fig. 2. Selected GNSS/lev samples representing regional and local GNSS/leveling geoid data: *a* – regional geoid (meters), *b* – local geoid (meters)

Table 1

Basic statistics of regional and local data samples (meters)

	Regional (287 points)	Reg. residual	Local (139 points)	Loc. residual
Min.	-36.01	-3.71	-28.32	-0.66
Max.	-8.14	5.04	-26.95	0.27
Mean	-18.60	0.00	-27.41	0.00
Median	18.02	-0.04	-27.35	0.01
RMS	19.26	1.74	27.41	0.14
Std dev.	5.01	1.74	0.26	0.14

is named local dataset. The regional data are reduced to have very sparse distribution with average data spacing about 1° in latitude and similarly in longitude. The local data are denser, i.e. around 0.05° in latitude and longitude. Moreover, the local data are slightly more irregularly spaced than the regional data. There are no evident outlying observations in the datasets, therefore larger observed errors are assumed as random ones.

Numerical calculations were performed using ellipsoidal coordinates, starting from the data detrending (Eq. 2–4). Both datasets were reduced by the polynomial trend of the second order (Eq. 2). ECF is computed for both residual datasets, according to the Eq. (12). The sampling interval of ECF represented by the interval of ψ cannot be significantly smaller than average minimum distance between the data points. Otherwise, we would search for the covariances at the distances that occur only occasionally between data points and the estimation accuracy will be poor for ECF at small distances. Three functional models are graphically compared to ECF values (Eq. 12) and different CL parameter is found for each covariance model (Fig. 1). At this moment C_0 is assumed to be equal to the variance of the residuals and constant for all covariance models. ECF estimation is an additional, separate tool for the covariance parameters estimation and it is used for the further comparisons with LOO estimation. It should be noted, that δn is not estimated by the fitting of analytical model. The values of the empirical covariance may strongly depend on the sampling interval. Therefore one should be careful in assessing δn by fitting the analytical model into the empirical covariance samples, which have arbitrarily chosen interval. Some authors find a priori noise empirically, e.g. iteratively searching for the consistency between a priori noise variance and RMS in CV (SMITH, MILBERT 1999). Some others use a priori noise based on the observational accuracy (DENKER 1998). The proposed method is based on the minimum RMSL in the space of three covariance parameters.

Besides the ECF computation for initial assessment of the parameters, an additional data are prepared for the comparisons and discussion. Geoid height residuals are computed using the harmonic expansion of EGM2008 geopotential model. The two applied terrestrial datasets represent significantly different area sizes and spatial resolutions. The spatial resolution of the local data is around 5–10 km, which may be approximately comparable with the resolution of EGM2008 geopotential model with its maximum harmonic expansion degree (PAVLIS et al. 2012). The spatial resolution of regional data is around 100 km, which corresponds approximately to 180 degree and order of the harmonic expansion. Therefore the residual geoid signal is calculated using EGM2008 coefficients for the area of regional data (Fig. 3). These residuals represent the spectrum of the geoid between 180 and 2190 degree and order and show significant variance in the area of regional data. The measurement accuracy of GNSS/lev data is usually at the centimeter level and therefore is insignificant in relation to the signal in Figure 3, which can be lost in LSC due to limited resolution. The corresponding residuals for the local data are not computed,

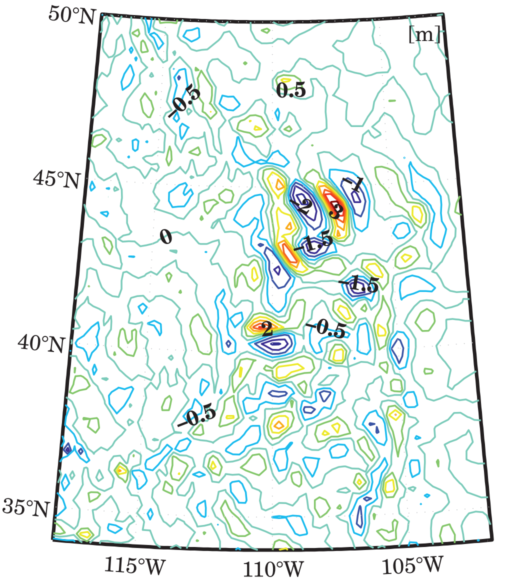

Fig. 3. Residual geoid from EGM2008 model equal $N_{2190} - N_{180}$ (area of regional dataset)

because there are no degrees in EGM2008 supplying more than local data resolution. On the other hand, geoid height residuals at the frequency higher than 0.05° may have the variance comparable to or less than GNSS/lev measurement error (RAPP 1973).

Two sets of residual data are interpolated with use of three above mentioned covariance models. LOO is performed as an iterative LSC process, which

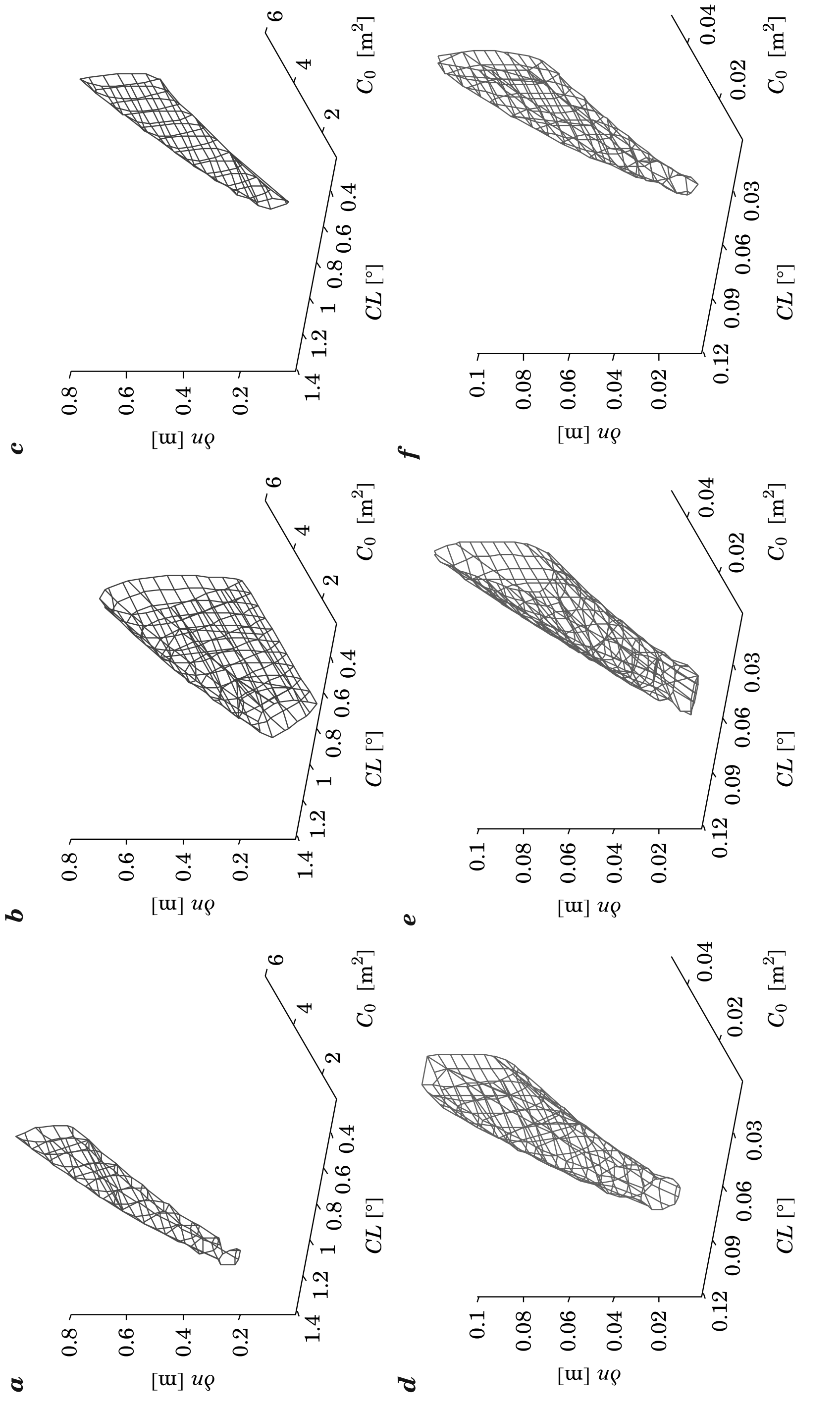

Fig. 4. Surfaces of equal RMSL in space of three covariance parameters: *a* – regional CG, RMSL = 0.784 m, *b* – regional CGM2, RMSL = 0.754 m, *c* – regional CGM3, RMSL = 0.761 m, *d* – local CG, RMSL = 0.089 m, *e* – local CGM2, RMSL = 0.086 m, *f* – local CGM3, RMSL = 0.087 m

uses variable covariance parameters. The estimate of geoid height for each point excludes this point from the dataset (Eq. 10). The LSC prediction is made with 3° distance limit for each regional $\tilde{N}_i^r$. Respective distance limit for the local data is 0.2°. These distances are assessed from Figure 1, by finding the maximum distances of the correlated residuals. LOO process is applied and the predictions are compared to the measured data by RMSL calculation. The analysis of three covariance parameters is performed in their 3D space (Fig. 4). The minimum RMSL is a measure of optimum prediction possible for the chosen data and covariance model. This minimum RMSL indicates covariance parameters enabling the prediction that fits best the data in the least squares sense. Figure 4 describes regions in 3D space of three parameters, which have RMSL smaller than individually specified values. These values are computed by adding 0.001 m to the global RMSL minimum in every case. One may note that these regions have elongated shapes. This means that various parameter sets may provide similar LSC results if properly combined. *CL* parameter has a tendency to be more constant than δn, when C_0 is increasing. More specifically, δn rises with the increase of C_0, which may indicate a correlation between the parameters.

Figure 5 presents the cross-sections of subfigures from Figure 4. The sections are realized for C_0 parameter equal 3 m^2 for regional data and 0.02 m^2 for local geoid heights. These values are variances of the residuals, which are often used to approximate C_0. The sections show the problem in more detail and indicate the regions of the parameters, where the prediction has decreased accuracy. The choice of too small *CL* may strongly affect accuracy of the prediction as well as small δn.

CL parameter is different in particular covariance models, which is a confirmation of the previous observations (Fig. 1). *CL* values in Figure 5 are usually smaller than their respective estimates in Figure 1. This is especially noticeable in the case of regional data (Figs. 5 *a–c*). The parameter δn that represents a priori noise is similar for all covariance models used and large in the case of regional data (Fig. 5 *a–c*). The predictions, where δn is closer to GNSS/lev data accuracy are poor, especially in case of CG (Fig. 5*a*) and CGM3 (Fig. 5*c*).

Figure 3 describes a part of the geoid signal that represents some details of the geoid height between 180 and 2160 degree and order of the harmonic expansion. These spectra may be present in GNSS/lev data due to the centimeter accuracy of GPS and leveling, however, the spatial resolution of the regional data is not sufficient to interpolate them. The information about the correlation of this higher frequency signal may be impossible to gain from data with 100 km spacing. This signal may be treated as a noise, since the sampling is here insufficient to find it as a correlated signal.

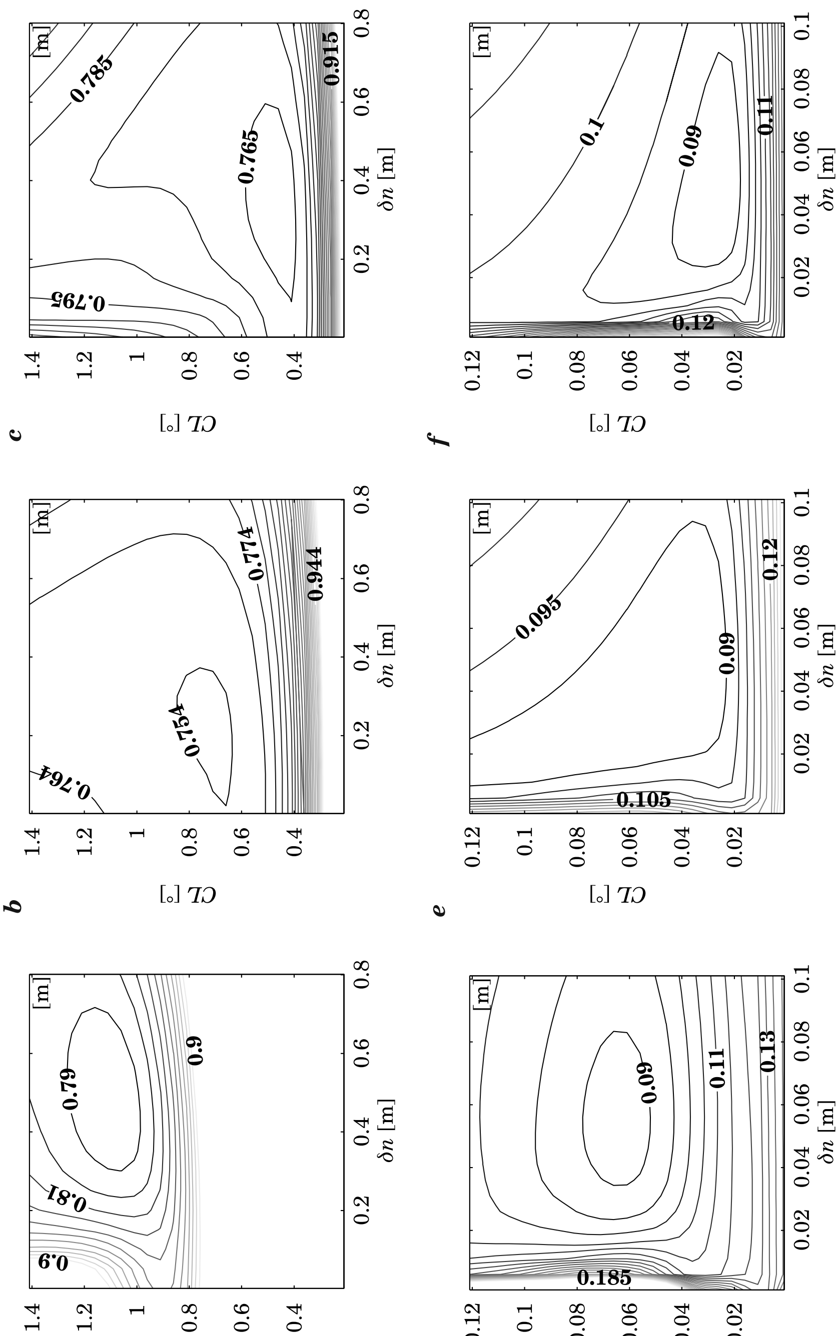

Fig. 5. RMSL for C_0 parameter equal to variance of residuals (i.e. 3 m^2 and 0.02 m^2): *a* – regional, CG, C_0 = 3 m^2, *b* – regional, CGM2, C_0 = 3 m^2, *c* – regional, CGM3, C_0 = 3 m^2, *d* – local, CG, C_0 = 0.02 m^2, *e* – local, CGM2, C_0 = 0.02 m^2, *f* = local, CGM3, C_0 = 0.02 m^2

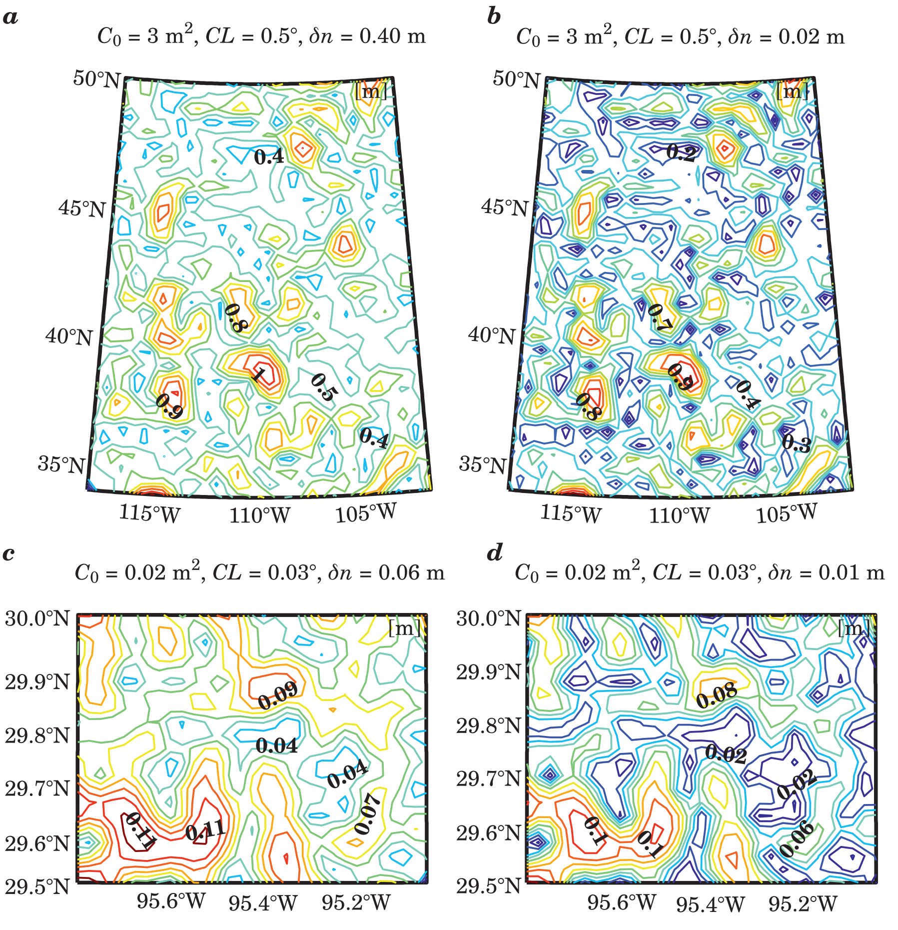

Fig. 6. A posteriori standard deviations for optimal parameters found by LOO and when only noise variances are decreased (CGM3 model): *a* – a posteriori st. dev. for regional geoid, *b* – a posteriori st. dev. for regional geoid, *c* – a posteriori st. dev. for local geoid, *d* – a posteriori st. dev. for local geoid

Additional problem that may be considered here is the a posteriori error, which is usually used in the assessment of LSC accuracy. It is strongly related with a priori noise, however remaining covariance parameters may have also significant influence on LSC error. SANSÓ et al. (1999) and DARBEHESHTI, FEATHERSTONE (2009) report that a posteriori error estimate may be even more affected by the changes of parameters than the LSC result. Figure 6 presents standard deviations of LSC prediction computed from Eq. 11 with optimal parameters (Fig. 5) and with the same C_0 and CL, but smaller δn values used. The underestimated a posteriori standard deviations may be found in Figure 6*b* and 6*d*.

Conclusions

CV methods like e.g. LOO are useful tools for finding covariance parameters that enable optimal prediction with arbitrarily selected covariance model. In some cases, wrong covariance parameters may give significantly worse result and some ranges of the parameters are especially inappropriate. A posteriori LSC error may even stronger depend on the parameters, especially on δn. The parameter δn should represent the noise existing in the observations, therefore applying smaller parameter can provide too optimistic accuracy estimate. It is suspected from analyses that δn depends not only on the measurement error. Limited spatial resolution of the data may exclude higher frequency signal from the correlated data part. Consequently, this part of the signal, which exists in the data due to the high accuracy of the measurement, can be assessed as uncorrelated. Such occurrence is very hard to detect when the data spatial resolution corresponds in a measure to its accuracy. This is quite frequent in practice, however, the case similar to regional data used in this paper may also be found.

Observed LSC properties are essentially consistent with mentioned works, where Tikhonov regularization is applied. In this work, the actual covariance parameters have been estimated instead of the regularization parameter. The paper reveals probable influence of data spatial resolution on the noise covariance matrix. However, the data resolution may be not sole variable influencing a priori noise. Some additional tests have to be performed to explain this in detail.

Acknowledgements

The datasets were downloaded from the National Geodetic Survey website (U.S.).

References

ANDERSEN O.B., KNUDSEN P. 1998. *Global marine gravity field from the ERS-1 and Geosat geodetic mission altimetry*. J. Geophys. Res., 103(C4): 8129–8137.

ARABELOS D., TSCHERNING C.C. 2003. *Globally covering a-priori regional gravity covariance models*. Adv. Geosci., 1: 143–147, DOI:10.5194/adgeo-1-143-2003.

ARLOT S., CELISSE A. 2010. *A survey of cross-validation procedures for model selection*. Stat. Surv., 4: 40–79, DOI: 10.1214/09-SS054.

CATALAO J., SEVILLA M.J. 2009. *Mapping the geoid for Iberia and the Macaronesian Islands using 430 multi-sensor gravity data and the GRACE geopotential model*. J Geod., 48(1): 6–15. DOI: 10.1016/j.jog.2009.03.001.

DARBEHESHTI N., FEATHERSTONE W.E. 2009. *Non-stationary covariance function modelling in 2D least-squares collocation*. J Geod., 83(6): 495–508. DOI:10.1007/s00190-008-0267-0

DARBEHESHTI N., FEATHERSTONE W.E. 2010. *Tuning a gravimetric quasigeoid to GPSlevelling by non-stationary least-squares collocation*. J. Geod., 84 (7): 419–431.

DAWIDOWICZ K. 2012. *GNSS satellite levelling using the ASG-EUPOS system services*, Techn. Sc., 15(1): 35–48.

DENKER H. 1998. *Evaluation and Improvement of the EGG97 Quasigeoid Model for Europe by GPS and Leveling Data*. In: *Proceed. Continental Workshop on the Geoid in Europe*. Eds. M. Vermeer, J. Adam. Budapest, Hungary, March 10–14, Reports of the Finnish Geodetic Institute, 98(4): 53–61, Masala.

EL-FIKY G.S., KATO T., FUJII Y. 1997. *Distribution of vertical crustal movement rates in the Tohoku district*, *Japan, predicted by least-squares collocation*. J. Geod., 71(7): 432–442.

FOTOPOULOS G., KOTSAKIS C., SIDERIS M.G. 2003. *How accurately can we determine orthometric height differences from GPS and geoid data*? J. Surv. Eng., 129: 1–10.

GODAH W., KRYŃSKI J. 2011. *Validation of GOCE gravity field models over Poland using the EGM2008 and GPS/levelling data*. Geoinformation Issues, 3, 1(3): 5–17.

GREBENITCHARSKY R.S., RANGELOVA E.V., SIDERIS M.G. 2005. *Transformation between gravimetric and GPS/levelling-derived geoids using additional gravity information*. J. Geodyn., 39(5): 527–544.

HOFMANN-WELLENHOF B., MORITZ H. 2005. *Physical Geodesy*. Springer, New York.

ILIFFE J.C., ZIEBART M., CROSS P.A., FORSBERG R., STRYKOWSKI G., TSCHERNING C.C. 2003. *OGSM02: A new model for converting GPS-derived heights to local height datums in Great Britain and Ireland*. Surv. Rev., 37(290): 276–293.

JEKELI C., GARCIA R. 2002. *Local geoid determination with in situ geopotential data obtained from satellite-to-satellite tracking data*. In: *Gravity, Geoid and Geodynamics 2000*. Ed. M.G. Sideris. Springer, Berlin, p. 123–128.

KAVZOGLU T., SAKA M.H. 2005. *Modeling local GPS/levelling geoid undulations using artificial neural networks*. J. Geod., 78: 520–527. DOI 10.1007/s00190-004-0420-3.

KITANIDIS P.K. 1983. *Statistical estimation of polynomial generalized covariance functions and hydrologic applications*. Water Resour. Res., 19(4): 909–921.

KOHAVI R. 1995. *A study of cross-validation and bootstrap for accuracy estimation and model selection*. Proceedings of the 14th International Joint Conference on Artificial Intelligence. Montreal, Canada, 2: 1137–1143.

KOWALCZYK K., RAPIŃSKI J., MRÓZ M. 2010. *Analysis of vertical movements modelling through various interpolation techniques*. Acta Geodyn. Geomater., 7, 4(160): 1–11.

KUSCHE J., KLEES R. 2002. *Regularization of gravity field estimation from satellite gravity gradients*. J Geod 76: 359–368.

ŁYSZKOWICZ A. 2010a. *Refined astrogravimetric geoid in Poland*. Part I. Geomatics and Environmental Engineering, 4(1): 57–67.

ŁYSZKOWICZ A. 2010b. *Refined astrogravimetric geoid in Poland*. Part II. Geomatics and Environmental Engineering, 4(2): 63–73.

ŁYSZKOWICZ A. 2009. *Assessment of accuracy of EGM08 model over the area of Poland*. Technical Sciences, 12: 118-134, DOI 10.2478/v10022-009-0011-x.

MARCHENKO A., TARTACHYNSKA Z., YAKIMOVICH A., ZABLOTSKYJ F. 2003. *Gravity anomalies and geoid heights derived from ERS-1,ERS-2, and Topex/Poseidon altimetry in the Antarctic peninsula area*. Proceedings of the 5th International Antarctic Geodesy Symposium AGS'03, September 15-17, Lviv, Ukraine, SCAR Report No. 23, http://www.scar.org/publications/reports/23/

MORITZ H. 1980. *Advanced Physical Geodesy*. Herbert Wichmann Verlag, Karlsruhe.

OSADA E., KRYŃSKI J., OWCZAREK M. 2005. *A robust method of quasigeoid modelling in Poland based on GPS/levelling data with support of gravity data*. Geodesy and Cartography, 54(3): 99–117.

PAVLIS N.K., HOLMES S.A., KENYON S.C., FACTOR J.F. 2012. *The development and evaluation of Earth Gravitational Model (EGM2008)*, J. Geophys. Res., 117, B04406, DOI:10.1029/2011JB008916.

RAO C.R., TOUTENBURG H. 1995. *Linear Models: Least Squares and Alternatives*. New York: Springer--Verlag, pp. 352.

RAPP R.H. 1973. *Geoid information by wavelength*. Bull. Geod., 110(1): 405–411.

SANSÓ F., VENUTI G., TSCHERNING C.C. 1999. *A theorem of insensitivity of the collocation solution to variations of the metric of the interpolation space*. In: *Geodesy Beyond 2000*. Ed. K.P. Schwarz. The Challenges of the First Decade, International Association of Geodesy Symposia, 121: 233–240, Springer, Berlin.

SMITH D.A., MILBERT D.G. 1999. *The GEOID96 high-resolution geoid height model for the United States*. J. Geod., 73(5): 219–236.
STRYKOWSKI G., FORSBERG R. 1998. *Operational merging of satellite airborne and surface gravity data by draping techniques*. In: *Geodesy on the Move*. Eds. R. Forsberg, M. Feissl, R. Dietrich. Springer, Berlin, p. 207–212.
TROJANOWICZ M. 2012. *Local modelling of quasigeoid heights with the use of the gravity inverse method – case study for the area of Poland.* Acta Geodyn. Geomater., 9, 1(165): 5–18.
YOU R.J., HWANG H.W. 2006. *Coordinate transformation between two geodetic datums of Taiwan by least-squares collocation.* J. Surv. Eng.-ASCE, 132(2): 64–70.