



# On the drawback of local detrending in universal kriging in conditions of heterogeneously spaced regional TEC data, low-order trends and outlier occurrences

Wojciech Jarmołowski<sup>1</sup> · Paweł Wielgosz<sup>1</sup> · Xiaodong Ren<sup>2</sup> · Anna Krypiak-Gregorczyk<sup>1</sup>

Received: 5 June 2019 / Accepted: 5 October 2020 / Published online: 23 December 2020  
© The Author(s) 2020

## Abstract

The study intercompares three stochastic interpolation methods originating from the same geostatistical family: least-squares collocation (LSC) known from geodesy, as well as ordinary kriging (OKR) and universal kriging (UKR) known from geology and other geosciences. The objective of this work is to assess advantages and drawbacks of fundamental differences in modeling between these methods in imperfect data conditions. These differences primarily refer to the treatment of the reference field, commonly called ‘mean value’ or ‘trend’ in geostatistical language. The trend in LSC is determined globally before the interpolation, whereas OKR and UKR detrend the observations during the modeling process. The approach to detrending leads to the evident differences between LSC, OKR and UKR, especially in severe conditions such as far from the optimal data distribution. The theoretical comparisons of LSC, OKR and UKR often miss the numerical proof, while numerical prediction examples do not apply cross-validation of the estimates, which is proven to be a reliable measure of the prediction precision and a validation of empirical covariances. Our study completes the investigations with precise parametrization of all these methods by leave-one-out validation. It finds the key importance of the detrending schemes and shows the advantage of LSC prior global detrending scheme in unfavorable conditions of sparse data, data gaps and outlier occurrence. The test case is the modeling of vertical total electron content (VTEC) derived from GNSS station data. This kind of data is a challenge for precise covariance modeling due to weak signal at higher frequencies and existing outliers. The computation of daily set of VTEC maps using the three techniques reveals the weakness of UKR solutions with a local detrending type in imperfect data conditions.

**Keywords** Ionosphere · Total electron content · Least-squares collocation · Simple kriging · Ordinary kriging · Universal kriging

## 1 Introduction

An important aspect in TEC estimation at a regional or global scale is the appropriate interpolation strategy. GNSS dual-frequency observations provide TEC point measurements at the so-called ionosphere pierce points (IPPs), whose density substantially differs across the globe, e.g., due to the large gaps over the oceans. Therefore, it is expected that any interpolation method will render some larger differences in

the modeling over the regions of sparse data. Global TEC models are often approximated by relatively low degrees of the spherical orthogonal functions, e.g., spherical harmonics (Alizadeh et al. 2011; Liu et al. 2018) or other empirical orthogonal functions (Mao et al. 2008). These functions are very sensitive to data gaps, and additionally, the regions with denser data suffer from the high-frequency cutoff of the signal in case of lower-order functions. A better representation of the resolution, in comparison with spherical harmonics, may be provided by spherical splines (Schmidt et al. 2011; Erdogan et al. 2017). However, one should keep in mind that every sum of orthogonal functions applies signal cutoff at the spatial resolution equivalent to the maximum degree of the applied base functions. Together with the development of GNSS remote sensing, a need has arisen for the employment of precise modeling techniques in the interpolation

✉ Wojciech Jarmołowski  
wojciech.jarmolowski@uwm.edu.pl

<sup>1</sup> Faculty of Geoen지니어ing, University of Warmia and Mazury in Olsztyn, ul. Oczapowskiego 2, 10-719 Olsztyn, Poland

<sup>2</sup> School of Geodesy and Geomatics, Wuhan University, Wuhan, China

of atmospheric parameters. Particularly, the local applications of GNSS positioning rely on the stochastic parametric methods of the interpolation like least-squares collocation (LSC) (Odijk 2002; Krypiak-Gregorczyk et al. 2017) or the ordinary kriging (OKR) (Odijk 2002; Wielgosz et al. 2003). Sayin et al. (2008) present a comparison of OKR and universal kriging (UKR) and assess these techniques together, in a study of local TEC phenomena. Other local studies of TEC and other parameters related to the ionosphere with the use of kriging can be found in Stanisławska et al. (1996, 2002). The atmospheric effects like ionospheric or tropospheric delays are nowadays commonly modeled globally, using a large number of GNSS stations and other additional data. In order to enhance the spatial resolution and accuracy of local and global TEC models, Orús et al. (2005, 2007) conducted research on the application of the OKR to global ionosphere mapping. It has been shown there that kriging based on the assumption of an error decorrelation function with distance and direction was the optimal interpolation method for GNSS data. This approach can provide improvements of up to 10% or more, in comparison with other popular techniques. The evidence of the stochastic approach improvement is thus demonstrated globally; however, many factors contribute to compromising this advantage like sparse and irregular global GNSS data, noise of these data or noise of the validation data such as altimetry-derived TEC (Li et al. 2019).

The majority of the stochastic parametric techniques apply the least-squares (LS) rule in linear models. The advantages of the parametric techniques of interpolation and extrapolation result from accurate signal covariance models estimated from real data. These modeling techniques can allocate data noise in noise covariance matrix, separating it from the signal. Different kinds of these techniques may have different detrending schemes, correlation approximations and parametrization methods. This study compares two groups of methods, which substantially differ with respect to the detrending rule. These two methods are compared using the same covariance model and the same parametrization rule, which preserves detrending rule as the only difference. This approach enables the comparison of the influence of prior detrending and synchronous detrending on the modeling results. Two general names can be found among the parametric spatial domain techniques based on the LS rule: LSC and kriging. The former one originates from geodesy and gravity field modeling and can be generalized from the interpolation/extrapolation alone, to the transformation between physically dependent quantities, e.g., physical quantities and their gradients (Moritz 1980, Tscherning and Rapp 1974, Sansó et al. 1999). The latter term has its roots in geology and refers to different forms of interpolation, which have different detrending schemes and names (Olea 1999; Wackernagel 2003; Diggle and Ribeiro 2007; Lichtenstern 2013). LSC and

kriging correspond to each other, in particular when LSC has only an interpolation form (Hofmann-Wellenhof and Moritz 2005, p. 361) and kriging has a form of simple kriging (SKR) (Ligas and Kulczycki 2010). Three selected techniques are assessed in this work: LSC as interpolation that is comparable to SKR, OKR and UKR. Three orders of trends are applied in LSC, which forms the group of three approaches using prior detrending. OKR and two orders of trend applied in UKR compose the second group of methods, which detrend the data synchronously with the interpolation process.

There are some examples in the literature of the common assessment and comparison of LSC and kriging, or different kinds of kriging. Many of the existing works aiming at the evaluation of similarities or differences of LSC/SKR and OKR or UKR are theoretical comparisons. Numerical studies, in turn, often have no detailed parametrization of the covariance matrices, regarding especially the noise variance. Too little attention is also paid to the fact that detrending in kriging has a local character, when the range of interpolation data is limited, which is common in local applications. A synthesis of kriging methods is given by Blais (1982) and includes OKR and UKR theory, as well as an extension to nonstationary problems and nonlinear modeling. Dermanis (1984) mathematically compares LSC and kriging, employing OKR and UKR. His conclusions point to the differences between LSC and kriging resulting from the difference in the detrending approach related to the scaled mean in kriging and differences in linearity conditions. Another theoretical comparison with a numerical example is given by Reguzoni et al. (2005). That study compares a generalized form of kriging with LSC and discusses known application areas of both methods, and also existing limitations of their use in geodesy and other geosciences. The examples of LSC, OKR and additional linear methods used for the interpolation of atmospheric effects for real-time kinematic GNSS positioning can be found in Al-Shaery et al. (2011). The authors assess horizontal and vertical positioning errors after the applications of LSC and OKR in spatial interpolation of the vertical reference station (VRS) data in GNSS real-time kinematic (RTK) positioning. They find some advantage of LSC interpolation in the vertical accuracy, but the errors are not further analyzed. Many geophysical problems suffer from a deficiency of lower-order models of the analyzed phenomena or from inadequate knowledge of the mean. The problem of unbiasedness appears, therefore, as one of the most investigated issues, and it makes a difference between the LSC/SKR and OKR, which together with UKR applies constraints for unbiasedness. Thus, kriging with constrained averages like OKR and UKR is often more willingly applied (Daya and Bejari 2015; Malvić and Balić 2009; Wackernagel 2003). Malvić and Balić (2009) investigate Lagrange multiplier significance in OKR and indicate advantages of OKR related to the Lagrange multiplier application. They also find

some rules for the Lagrange multiplier choice in OKR. The presented study, therefore, completes the above-mentioned investigations with a precise parametrization of LSC and kriging prior to their common assessment. Additionally, the same three selected orders of the polynomial trend are applied in our investigations, respectively, in both groups of methods. We decided to start the comparisons with the lowest trend orders, as they are most often used. The third detrend order, which is the second-order polynomial, as a curved shape, is expected to behave in a different way in global detrending used in LSC and local detrending applied in UKR.

The parametrization, i.e., the estimation of the optimal parameters of LSC and kriging, is a crucial stage prior to the modeling. The effects of the parameter change on the modeling results are widely discussed in the literature (Brooker 1986; Arabelos and Tscherning 1998; O'Dowd 1991; Sadiq et al. 2010). The oldest and well-known method of the parameter selection is the fitting of analytical model into its corresponding empirical representation (Posa 1989; Barzaghi et al. 2003; Samui and Sitharam 2011). This method is quite effective for the signal parameters; however, supplementary investigations have been undertaken in order to determine noise parameters more accurately (Marchenko et al. 2003; Peng and Wu 2014; Jarmołowski 2016, 2017). The estimates of signal and noise covariance parameters are very often achieved by different cross-validation techniques, e.g., hold-out (HO) validation (Arlot and Celisse 2010; Kohavi 1995) or leave-one-out (LOO) validation (Kohavi 1995; Behnabian et al. 2018). Another solution that can be applied in the parametrization procedures is maximum likelihood estimation (MLE) of parameters for kriging (Pardo-Igúzquiza et al. 2009; Todini 2001; Zimmermann 2010) or for LSC (Jarmołowski 2015, 2017).

This study assesses LSC and kriging (OKR and UKR) after implementing the same  $C^4$ -Wendland covariance models, parametrized in exactly the same way by LOO in case of all six approaches. Therefore, the only remaining difference tested in the study is prior detrending in LSC in comparison with synchronous detrending in OKR and UKR. The data conditions are far from the ideal distribution, both from the spatial and from the statistical point of view. By preserving data gaps and outliers, we expect to reveal weak points of one or the other method. Particularly, this study can show strong and weak points of the detrending schemes, as the most evident differences between the corresponding methods are their rules of trend removal. The differences between the models created under rigorous parametrization conditions are analyzed with the use of cross-validation techniques, such as LOO, and comparison with independent data set. Hence, this work aims at revealing the advantages and drawbacks of global and local detrending schemes using unfavorable to severe data conditions.

## 2 Algorithms and applied covariance

As it was mentioned in Sect. 1, in general LSC and kriging terms may refer to different algorithms, which cannot be compared, and these terms can also refer to individual applications. A generalized LSC can employ different data types in one process of interpolation combined with the integration. For instance, gravity anomalies together with deflections of the vertical can be used in geoid modeling with the help of appropriate Stokes kernels (Moritz 1980). Kriging, on the contrary, is usually modified to various forms that, for example, handle higher-order trends existing in the data or work with anisotropic data (Wackernagel 2003). Therefore, we compare collocation and kriging in their most common and most comparable forms referring exactly to the same problem—the interpolation of spatially correlated data with expected values approximated by lower-order trends and covariance expressed by an isotropic covariance model. This kind of interpolation model stems from general LS estimation, typically applied in LS adjustment (Teunissen 2003, p. 5):

$$\mathbf{Z} = \mathbf{F}\boldsymbol{\beta} + \mathbf{e} \quad (1)$$

where  $\mathbf{Z}$  is the observation vector,  $\mathbf{F}$  is the design matrix, which will be explained in the next paragraphs,  $\boldsymbol{\beta}$  is a vector including parameters and  $\mathbf{e}$  is the vector of errors also called noise vector. The stochastic fundamentals of this type of model are explained in details in Koch (1999 p. 153–161). The noise vector often stores uncorrelated random values, and since  $\mathbf{F}\boldsymbol{\beta}$  is deterministic, in order to interpolate correlated quantities, general LS must be extended to LSC (Moritz 1980, p. 111):

$$\mathbf{Z} = \mathbf{F}\boldsymbol{\beta} + \mathbf{s} + \mathbf{n} \quad (2)$$

where  $\mathbf{s}$  is the vector of correlated signal interpolated by LSC, OKR or UKR, and  $\mathbf{n}$  is the vector of uncorrelated noise. Equation (2) is the representative for LSC, OKR and UKR. However, the deterministic part  $\mathbf{F}\boldsymbol{\beta}$  will be separated from the observations in different ways, which is explained in Sects. 2.1–2.3. One should keep in mind that SKR is especially close or even equivalent to LSC, and, therefore, its results are assumed to be equal to those of LSC (Dermanis 1984). LSC works with detrended data, whereas OKR and UKR determine mean or higher-order trend by the constraints applied to the modeling process. LSC with a zero-order trend, which is simply the mean, corresponds to OKR, which uses a scaled mean in theory. For instance, the scaled mean is applied locally in practical formulas. The LSC based on the residuals after first- and second-order polynomial trends removal corresponds to UKR, which also uses the same higher-order polynomials. Nevertheless, there is a slightly

larger difference in the detrending between LSC and UKR, than in the former case. The trend in LSC is applied globally, as based on the whole dataset, whereas UKR applies the trend locally for the selected data subset defined by the maximum sampling distance. Also, the approximation of the signal variance is different for different estimation methods. The UKR method, in contrast to LSC, applies the detrending based on the observations in the covariance matrix used in the single prediction, and therefore it handles the existing non-stationarity in the analyzed stochastic field to a certain level. Hence, the aim of this study is the confrontation of LSC with OKR and UKR in order to empirically inspect the advantages and drawbacks of detrending prior to the modeling or when the trend removal is combined with modeling process in point prediction.

### 2.1 Least-squares collocation (LSC)/simple kriging (SKR)

It is assumed in LSC that the mean or trend of data  $Z(x)$  is known, i.e.,

$$\mu = \mathbb{E}(Z(x)), \tag{3}$$

where  $\mathbb{E}$  denotes the expectation operator. Furthermore,  $Z(x)$  is assumed to be second-order stationary and having known covariance function

$$C(h) = \mathbb{E}(Z(x)Z(x+h)) - \mu^2. \tag{4}$$

The empirical covariance function calculation rule is explained in Hofmann-Wellenhof and Moritz (2005, p. 347). In case of LSC, the unbiasedness condition is assumed to be provided automatically by the trend removal, i.e.,

$$\mathbb{E}(\tilde{Z}_{LSC}(x_p) - Z(x_p)) = \mu + \sum_{i=1}^n \omega_i^{LSC} \mathbb{E}(Z(x_i) - \mu) - \mathbb{E}(Z(x_p)) = \mu - \mu = 0, \tag{5}$$

and therefore any additional constraints on weights  $\omega_i^{LSC}$  are not necessary.  $Z(x_i)$  denotes the arbitrary data value,  $Z(x_p)$  is data in some selected position and  $\tilde{Z}_{LSC}(x_p)$  is the LSC estimate at this selected position. The term  $\mu$  is expanded later (Eq. 8–10) to the polynomial case, as the detrending in this study is based on zero-, first- and second-order polynomials removed prior to LSC prediction. However, one should keep in mind that the trend in LSC can be also derived from the other sources or functions of higher orders, e.g., spherical harmonic expansion, popular in the case of gravity and geoid modeling (Sadiq et al. 2010; Jarmołowski 2019). The assumption of the minimum prediction variance (Hofmann-Wellenhof and Moritz 2005) provides the condition for LSC

weights  $\omega_j^{LSC}$  and enables the creation of the system of equations, which in matrix form reads.

$$\mathbf{C}\boldsymbol{\omega}_{LSC} = \mathbf{c}_p, \tag{6}$$

where  $\boldsymbol{\omega}_{LSC}$  denotes the vector of LSC weights. Finally, with the use of matrices, the prediction is

$$\begin{aligned} \tilde{Z}_{LSC}(x_p) &= \mu + \mathbf{c}_p^T \mathbf{C}^{-1} (\mathbf{Z} - \mu \mathbf{I}) \\ &= \mathbf{c}_p^T \mathbf{C}^{-1} \mathbf{Z} + \mu (1 - \mathbf{c}_p^T \mathbf{C}^{-1} \mathbf{I}) \end{aligned} \tag{7}$$

where  $\mathbf{c}_p$  preserves the covariances prediction-data,  $\mathbf{C}$  is the data covariance matrix and  $\mathbf{Z}$  is the data vector of  $n$  points. The variable  $\mu$  denotes the mean as the simplest form of the trend function removed in the numerical study. In order to remove a higher-order trend, we must determine a set of deterministic functions of the coordinates  $f_0, f_1 \dots f_L$ , as follows:

$$\mathbf{F} = \begin{bmatrix} 1 & f_1(x_1) & \dots & f_L(x_1) \\ 1 & f_1(x_2) & \dots & f_L(x_2) \\ \vdots & \vdots & \dots & \vdots \\ 1 & f_1(x_n) & \dots & f_L(x_n) \end{bmatrix} \tag{8}$$

$\mathbf{F}$  is in fact previously introduced design matrix (Eqs. 1–2). Two successive orders of the trend are applied in the study as first- and second-order polynomials. They are removed before LSC interpolation and implemented intrinsically in the UKR method. The second-order polynomial trend applied locally reads:

$$\mathbf{F} = \begin{bmatrix} 1 & \varphi_1 & \lambda_1 & \varphi_1^2 & \lambda_1^2 & \varphi_1 \lambda_1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \varphi_n & \lambda_n & \varphi_n^2 & \lambda_n^2 & \varphi_n \lambda_n \end{bmatrix}. \tag{9}$$

Note that the first three columns correspond to first-order trend. In LSC trend matrix,  $\mathbf{F}$  can be applied to compose the so-called projection matrix  $\mathbf{P}$  (Pardo-Igúzquiza et al. 2009) or orthogonal projection operator (Koch 1999, p. 65; Teunissen 2003, p. 8).

$$\mathbf{P} = \mathbf{I}_n - \mathbf{F} [\mathbf{F}^T \mathbf{F}]^{-1} \mathbf{F}^T \tag{10}$$

which multiplied by the data vector  $\mathbf{Z}$  replaces  $\mathbf{Z} - \mu \mathbf{I}$  and removes a higher-order trend. This removal is principally related with Eq. (2), and in practice is equal to LS adjustment of the polynomial surface of arbitrary order, determined by the base functions such as those in Eq. (9). The relation of  $\mathbf{P}$  with general LS solution can be found in Koch (1999, pp. 153–154), where two solutions of the parameter vector  $\boldsymbol{\beta}$  (Eq. 2) can also be found. The simplified, unweighted projection operator based on unweighted LS solution is applied

here. This is because the covariance matrix is not directly available before detrending process, and because a very similar unbiasedness of Eq. (10) with low-order trends was previously proven in practice (Jarmołowski and Bakuła 2014; Pardo-Igúzquiza et al. 2009). The variance of the prediction, i.e., a posteriori error at selected point  $p$ , can be calculated from

$$\sigma'_{LSC}{}^2 = C_0 + \omega_{LSC}^T C \omega_{LSC} - 2\omega_{LSC}^T c_p = C_0 - c_p^T C^{-1} c_p \tag{11}$$

with  $C_0$  referring to signal covariance at distance equal to zero. The above a posteriori error (Eq. 11), in case of noisy data, is strongly related to a priori noise that together with the signal covariance contributes to the covariance matrix  $C$ . Therefore, a realistic estimate of a priori noise must be introduced to the noise covariance matrix. In the LSC theory, the total covariance matrix  $C$  is composed of the signal covariance matrix  $C_s$  and the noise covariance matrix  $C_n$  (Moritz 1980, p. 102):

$$C = C_s + C_n \tag{12}$$

where  $C_s$  is based on the selected signal covariance model, and  $C_n$ , under the assumption of non-correlated noise, is a diagonal matrix based on data a priori noise standard deviation ( $\sigma_{LSC}$ ) in the following way:

$$C_n = \sigma_{LSC}^2 \cdot I_n. \tag{13}$$

Therefore finally, introducing a priori noise variance of the data gives an extended form of the a posteriori error (Moritz 1980, p. 105):

$$\sigma'_{LSC}{}^2 = C_0 + \sigma_{LSC}^2 - c_p^T C^{-1} c_p. \tag{14}$$

Different covariance models and corresponding variograms can be applied in the modeling process. Isotropic covariance models are the most popular and easy to apply. Some theoretical considerations about the positive definiteness and other properties of the selected models for the spheres can be found in Huang et al. (2011), Gneiting (2013) or in Guinness and Fuentes (2016). This numerical study applies  $C^4$ -Wendland covariance function in LSC and  $C^4$ -Wendland variogram in OKR and UKR. The  $C^4$ -Wendland covariance function applied in LSC is

$$C^{C4W}(\psi) = C_0 \cdot \left( 1 + \tau \frac{\psi}{\alpha_{LSC}} + \frac{\tau^2 - 1}{3} \frac{\psi^2}{\alpha_{LSC}^2} \right) \left( 1 - \frac{\psi}{\alpha_{LSC}} \right)_+^\tau \tag{15}$$

where  $\psi$  is the spherical distance, and  $\alpha_{LSC}$  is named correlation length in LSC theory (Hofmann-Wellenhof and Moritz

2005). The shape parameter  $\tau$  has to be  $\geq 6$  in order to guarantee positive-definiteness on the sphere (Gneiting 2013). Therefore, this parameter is equal 6.5 in the study.

### 2.2 Ordinary kriging (OKR)

The theory of OKR method starts without the mean or trend removal before the covariance estimation (Olea 1999; Wackernagel 2003; Lichtenstern 2013). Instead, the mean is eliminated by the subtraction in the estimation of the semi-variogram:

$$\gamma(h) = \frac{1}{2} \mathbb{E} \left( (Z(x+h) - Z(x))^2 \right). \tag{16}$$

The calculation of empirical semivariogram is given by Wackernagel (2003, p. 47). The unbiasedness of OKR must be forced by the equality

$$\sum_{i=1}^n \omega_i^{OKR} = 1. \tag{17}$$

Therefore, the bias in OKR also converges to zero, as follows:

$$\begin{aligned} &\mathbb{E} \left( \tilde{Z}_{OKR}(x_p) - Z(x_p) \right) \\ &= \mathbb{E} \left( \sum_{i=1}^n \omega_i^{OKR} Z(x_i) - Z(x_p) \sum_{i=1}^n \omega_i^{OKR} \right) \\ &= \sum_{i=1}^n \omega_i^{OKR} \mathbb{E} (Z(x_i) - Z(x_p)) = 0. \end{aligned} \tag{18}$$

The system of equations similar to Eq. (6), which leads to the solution of  $\omega_i$  values, is extended to

$$\Gamma \omega_{OKR} + \lambda_{OKR} I = \gamma_p \tag{19}$$

$$\omega_{OKR}^T I = 1 \tag{20}$$

where  $\lambda_{OKR}$  is the so-called Lagrange multiplier. The solution must be additionally supported by the condition in Eq. (17).

The block matrix of variogram between the data  $\bar{\Gamma}$  and data prediction vector of variogram  $\bar{\gamma}_p$  are therefore composed as follows:

$$\bar{\Gamma} = \begin{bmatrix} & & & 1 \\ & \Gamma & & \\ & & & \\ & & & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \text{ and } \bar{\gamma}_p = \begin{bmatrix} \gamma_p \\ \vdots \\ 1 \end{bmatrix}, \text{ as well } \bar{Z} = \begin{bmatrix} Z \\ \vdots \\ 0 \end{bmatrix}. \tag{21}$$



Consequently, the final OKR prediction reads (Olea 1999, p. 48)

$$\tilde{Z}_{OKR}(x_p) = \bar{\mathbf{Z}}^T \bar{\mathbf{\Gamma}}^{-1} \bar{\mathbf{y}}_p. \tag{22}$$

Contrary to LSC, a posteriori error estimate at the selected point  $p$  can be derived from a slightly different equation (Olea 1999, p. 48):

$$\begin{aligned} \sigma'_{OKR}{}^2 &= -\omega_{OKR}^T \mathbf{\Gamma} \omega_{OKR} + 2\omega_{OKR}^T \mathbf{y}_p \\ &= \omega_{OKR}^T (2\mathbf{y}_p - \mathbf{\Gamma} \omega_{OKR}) = \bar{\mathbf{y}}_p^T \bar{\mathbf{\Gamma}}^{-1} \bar{\mathbf{y}}_p. \end{aligned} \tag{23}$$

It was mentioned in Sect. 2 that covariance models are equivalent for all predictions in the study. The  $C^4$ -Wendland variogram selected for the analysis of signal model parameters in OKR and UKR reads:

$$\gamma^{C4W}(\psi) = C_0 \cdot \left( 1 - \left( 1 + \tau \frac{\psi}{\alpha_K} + \frac{\tau^2 - 1}{3} \frac{\psi^2}{\alpha_K^2} \right) \left( 1 - \frac{\psi}{\alpha_K} \right)_+^\tau \right), \tag{24}$$

and all the parameters have the same meaning as in Eq. (15); however,  $\alpha_K$  is often called range in relation to kriging methods. The shape parameter  $\tau$  in the variograms applied in the study is also equal to 6.5, as well as in the covariance function case. The noise is introduced in OKR equations by the nugget parameter ( $\sigma_K$ ) and applied in the following way:

$$\gamma(x_i, x_j) = \gamma^{C4W} + \sigma_K^2 \text{ if } i \neq j \tag{25}$$

with  $\sigma_K$  referring to natural surveying conditions, where the signal is not the only quantity that contributes to data, as there is no measurement device that can provide errorless observations. The signal variance  $C_0$  is often called a partial sill in the kriging theory.

### 2.3 Universal kriging (UKR)

In UKR modeling, the signal can be decomposed into a deterministic trend function  $\mu(x)$  of the order  $l$ , and a residual function  $Y(x_i)$ , such that

$$Z(x_i) = \mu(x_i) + Y(x_i) = \sum_{l=0}^L a_l f_l(x_i) + Y(x_i). \tag{26}$$

The unbiasedness condition is satisfied from

$$\begin{aligned} \mathbb{E}(\tilde{Z}_{UKR}(x_p) - Z(x_p)) &= \sum_{i=1}^n \omega_i^{UKR} (\mathbb{E}(\mu(x_i)) + \mathbb{E}(Y(x_i))) \\ &\quad - (\mathbb{E}(\mu(x_p)) + \mathbb{E}(Y(x_p))) \end{aligned}$$

$$\begin{aligned} &= \sum_{i=1}^n \omega_i^{UKR} \mu(x_i) - \mu(x_p) = 0 \\ &\Leftrightarrow \sum_{l=0}^L a_l \left( \sum_{i=1}^n \omega_i^{UKR} f_l(x_i) - f_l(x_p) \right) = 0 \end{aligned} \tag{27}$$

where  $a_l$  are the coefficients of the deterministic trend. If we denote the vector of this trend by  $\mathbf{a}$ , and the matrix of deterministic functions by  $\mathbf{F}$ , we can apply the first- or second-order trend based on Eq. (9). The matrix  $\mathbf{F}$  is implemented intrinsically in UKR for the local correlated data sample selected for the prediction in point  $p$ .

Using the matrix notation, we obtain the following system:

$$\mathbf{\Gamma}_Y \omega_{UKR} + \mathbf{F} \lambda_{UKR} = \mathbf{y}_{Yp} \tag{28}$$

$$\mathbf{F}^T \omega_{UKR} = \mathbf{f}_p,$$

where  $\lambda_{UKR}$  is Lagrange parameter vector. The block matrix form of these equations, similarly to OKR, gives the solution of the UKR prediction (Olea 1999, p. 105, Wackernagel 2003, pp. 68–69):

$$\begin{bmatrix} \omega_{UKR} \\ \lambda_{UKR} \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma}_Y & \mathbf{F} \\ \mathbf{F}^T & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{y}_{Yp} \\ \mathbf{f}_p \end{bmatrix} = \bar{\mathbf{\Gamma}}^{-1} \bar{\mathbf{y}}_{Yp} \tag{29}$$

where matrix  $\mathbf{\Gamma}_Y$  is based on  $C^4$ -Wendland variogram and matrix  $\mathbf{F}$  is expressed by Eq. (9).  $\mathbf{y}_{Yp}$  and  $\mathbf{f}_p$  are vectors calculated from the variogram and deterministic functions for the point of interest  $p$ , respectively. Summarizing, the prediction can be written in a similar way to OKR (Olea 1999, p. 103):

$$\tilde{Z}_{UKR}(x_p) = \bar{\mathbf{Z}}^T \bar{\mathbf{\Gamma}}_Y^{-1} \bar{\mathbf{y}}_{Yp}. \tag{30}$$

The variance of the prediction error will be then (Olea 1999, p. 103):

$$\sigma'_{UKR}{}^2 = -\omega_{UKR}^T \mathbf{\Gamma}_Y \omega_{UKR} + 2\omega_{UKR}^T \mathbf{y}_{Yp} = \bar{\mathbf{y}}_{Yp}^T \bar{\mathbf{\Gamma}}_Y^{-1} \bar{\mathbf{y}}_{Yp}. \tag{31}$$

The  $\sigma_K$  parameter is implemented in the variogram in an analogous way as in the case of OKR in the previous section.

### 3 LOO validation

LOO validation is based on the prediction in the position of data point, with exclusion of this point value from the dataset being used for the prediction (Kohavi 1995). The differences between  $n$  data values and the predictions made in the same

positions are used as a measure of reliable prediction precision. The difference in case of LSC is calculated between the residual  $Z'_P = Z_P - \mu \mathbf{I}$  in point  $P$  and its estimate excluding this single compared value, i.e.,

$$\text{LOO}_P(C_0, CL, \sigma | \mathbf{Z}_{n \times 1}^r) = \left\{ \begin{array}{c} Z'_P - \tilde{Z}'_P \\ \tilde{Z}'_P = \mathbf{c}_P^T \cdot (\mathbf{C}_{s(n-1) \times (n-1)} + \mathbf{C}_{n(n-1) \times (n-1)})^{-1} \cdot \mathbf{Z}_{(n-1) \times 1}^r \\ Z'_P \notin \mathbf{Z}_{(n-1) \times 1}^r \end{array} \right\}, \quad (32)$$

where  $Z'_P$  is the point residual value and  $\tilde{Z}'_P$  is the LSC estimate in the same location. The root mean square of all  $n$  differences (RMSLOO) calculated by Eq. (32) is a measure of the quality of the estimated model. RMSLOO is affected by the noise present in the dataset, as it comes from the comparison with noisy data. Therefore, if the prediction is optimal in least-squares sense, RMSLOO can be assessed as an empirical measure of the noise. In the proposed study, LOO differences are calculated for some range of parameters  $CL$  and  $\sigma$ . Equation (32) describes the validation for LSC, whereas OKR and UKR have analogous rule for comparison of  $Z_P$  and its estimates.

#### 4 VTEC test data

The study uses point VTEC values calculated from GNSS observations from the Crustal Movement Observation Network of China (CMONOC). The CMONOC network includes around 1000 permanent stations distributed almost homogeneously over most of China's territory. This number of stations provides dense local coverage of ionosphere piercing points (IPPs) (Chen et al. 2017), which is quite a challenging dataset for testing of local data modeling by stochastic techniques.

The estimation of VTEC at IPPs is derived using the precise point positioning (PPP) approach. The PPP method and the related VTEC estimation procedure are based on the same algorithm as PPP implementation in previous works of the authors, investigating the calculation of ionospheric maps (Jarmołowski et al. 2019; Ren et al. 2016). This article, in turn, is focused on the extended analysis of the parametric spatial stochastic modeling of the VTEC signal resulting from the PPP method. The noise coming from the PPP VTEC estimation is a useful property of the data, and it is thoroughly investigated and estimated by LOO validation in the individual modeling methods. Additionally, the datasets are cleaned of large outliers and intentionally sparsed out in relation to the original set of IPPs with the use of  $1^\circ$  grid for selection of the closest data. The selected datasets for the successive data epochs are distributed approximately homogeneously in the horizontal direction. This is made intentionally for the

purpose of the study on the stochastic methods, in order to help the parameter estimation, especially if we assume the stationarity of the process and average uncorrelated value

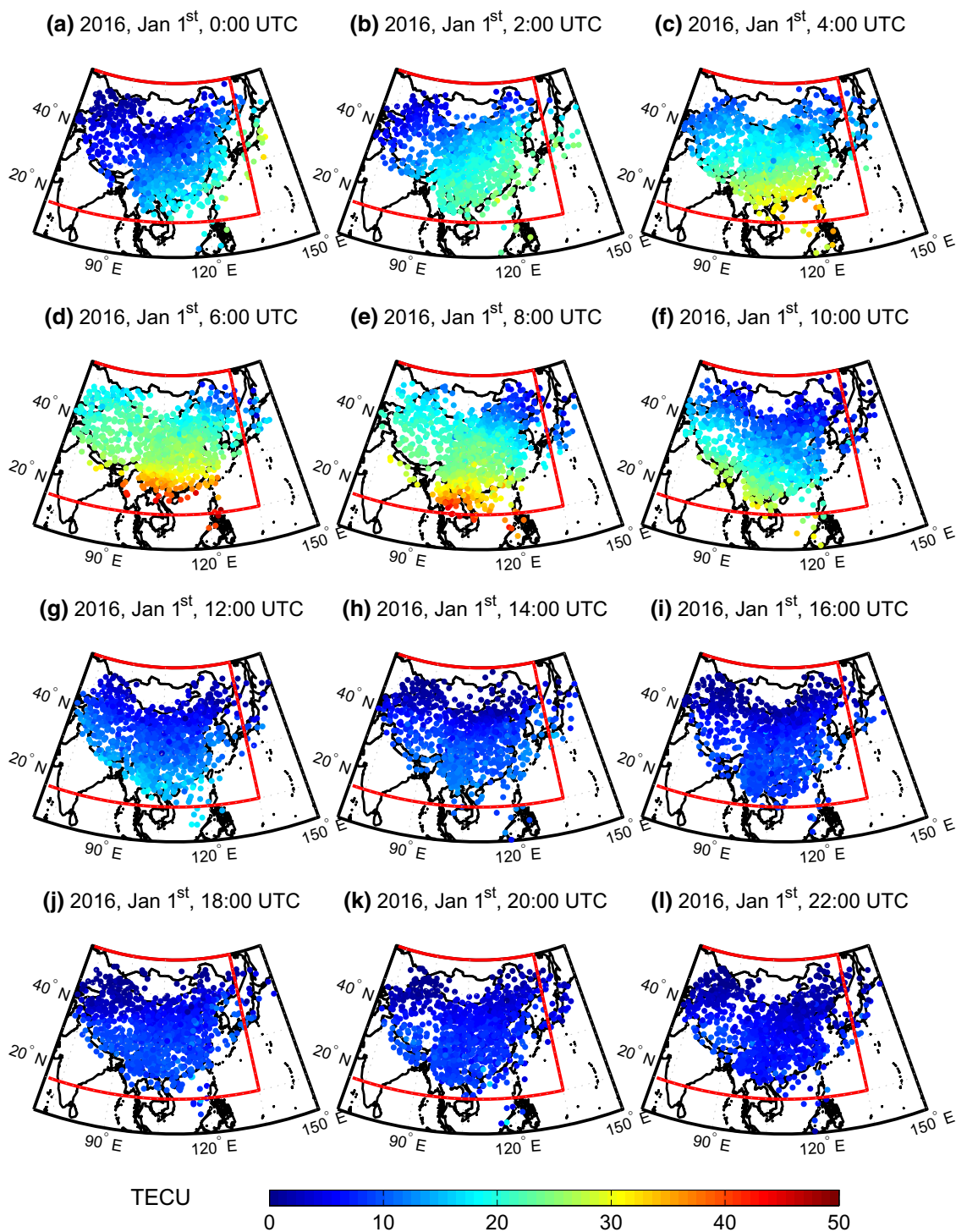
of the noise. The estimation of the parameters by LOO is difficult, as we deal with low residual signal variance with respect to the noise in case of VTEC observations. Most of the signal referring to equatorial TEC anomaly is preserved at lower frequencies and removed with the polynomial deterministic trend. Figure 1 shows all epochs of the data used in the numerical study, separated by 2 h. This relatively large interval was selected to shorten the time of estimation; however, it is also known that many of global ionospheric maps (GIMs) have the same interval (Roma-Dollase et al. 2018). Figure 1 presents the IPPs after the removal of the outliers and sparsing with  $1^\circ$  grid, used in LOO estimation of the parameters. The red line in Fig. 1 separates the rectangular area of the compact, enclosed region of approximately homogeneous spatial distribution, and data at the margins, where the estimation can suffer from gaps and sparsity. These two sets of unequal size will be investigated separately in further numerical tests.

Figure 2a presents the removal of the outliers, which is done by a specific application of LOO validation (Tscherning 1991). The LSC estimation with some roughly assumed, overestimated a priori noise variance has been made at data points. Then, the data values furthest from the predictions were removed, if these differences extended some assumed threshold. This threshold was quite rigorous in the study, and it was set to 3 TECU, as the amount of data was large enough to allow us to freely discard the outliers. Figure 2b reveals more clearly the effect of data sparsing, which is mostly noticeable in the center of the area, as IPPs are always concentrated mostly in the center due to the GNSS satellites tracks alignment.

#### 5 Comparison of covariance parameters and accuracy indicators between LSC, OKR and UKR

The numerical analysis comprises a few associated items:

LOO estimation of  $\sigma$  and  $\alpha$  parameters in LSC with zero-, first- and second-order polynomial removal, and in OKR and UKR after first- and second-order polynomial removal



**Fig. 1** Datasets in sample GPS epochs, red rectangle separates approximately homogeneous data from sparse data at the margins of the area

(Figs. 3, 5, and 6). A posteriori errors  $\sigma'$  are presented jointly in Fig. 6.

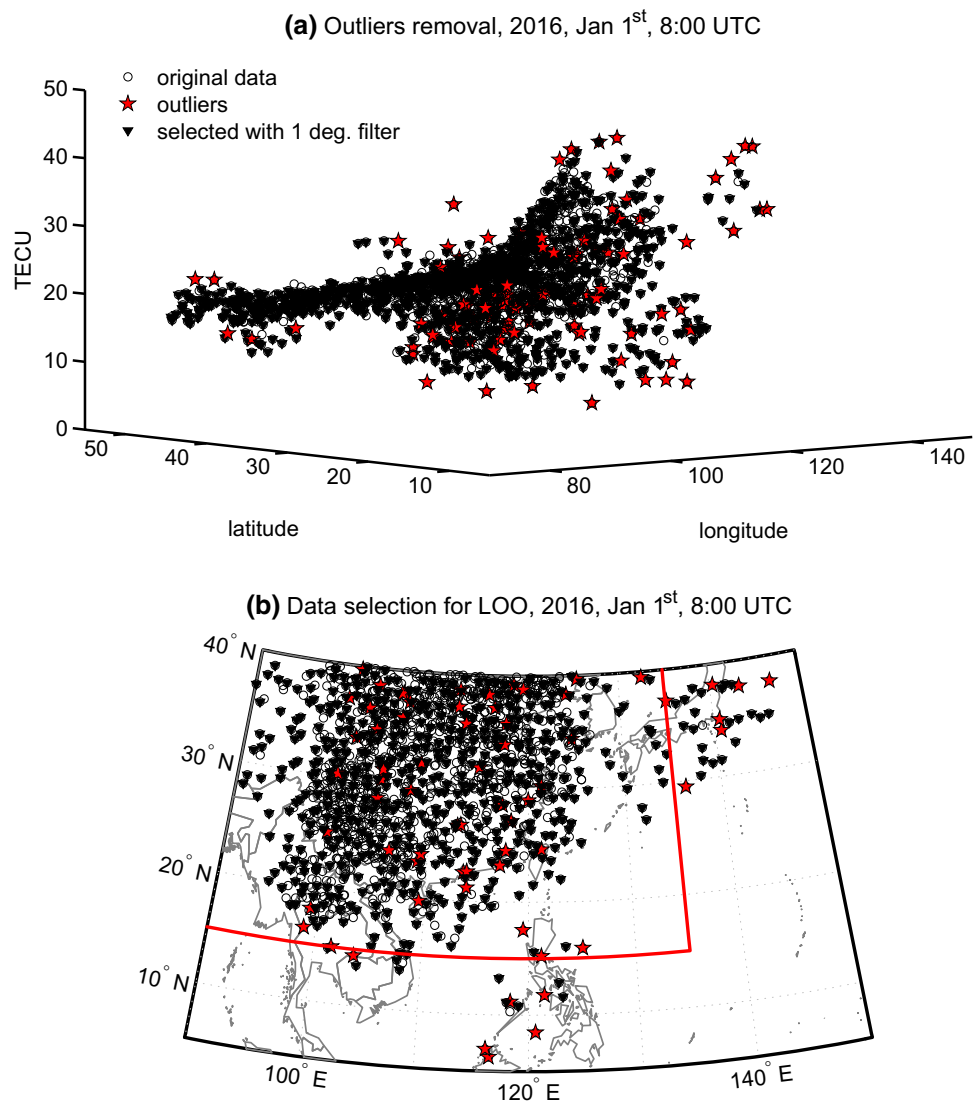
Empirical covariance and semi-variance calculation for the same six approaches, and comparative  $\alpha$  estimation by analytical model fitting (Fig. 4).

Comparison of minimum empirical accuracy of the prediction from LOO process, for all six approaches in terms of RMSLOO (Figs. 7, 8, 9).

Comparison of gridded models predicted with optimal estimated parameters, including various data conditions, i.e.,



**Fig. 2** Example data epoch describing sparsing and outliers removal: **a** outlier spatial distribution is visible and **b** sparsing effect is better visible in a horizontal view



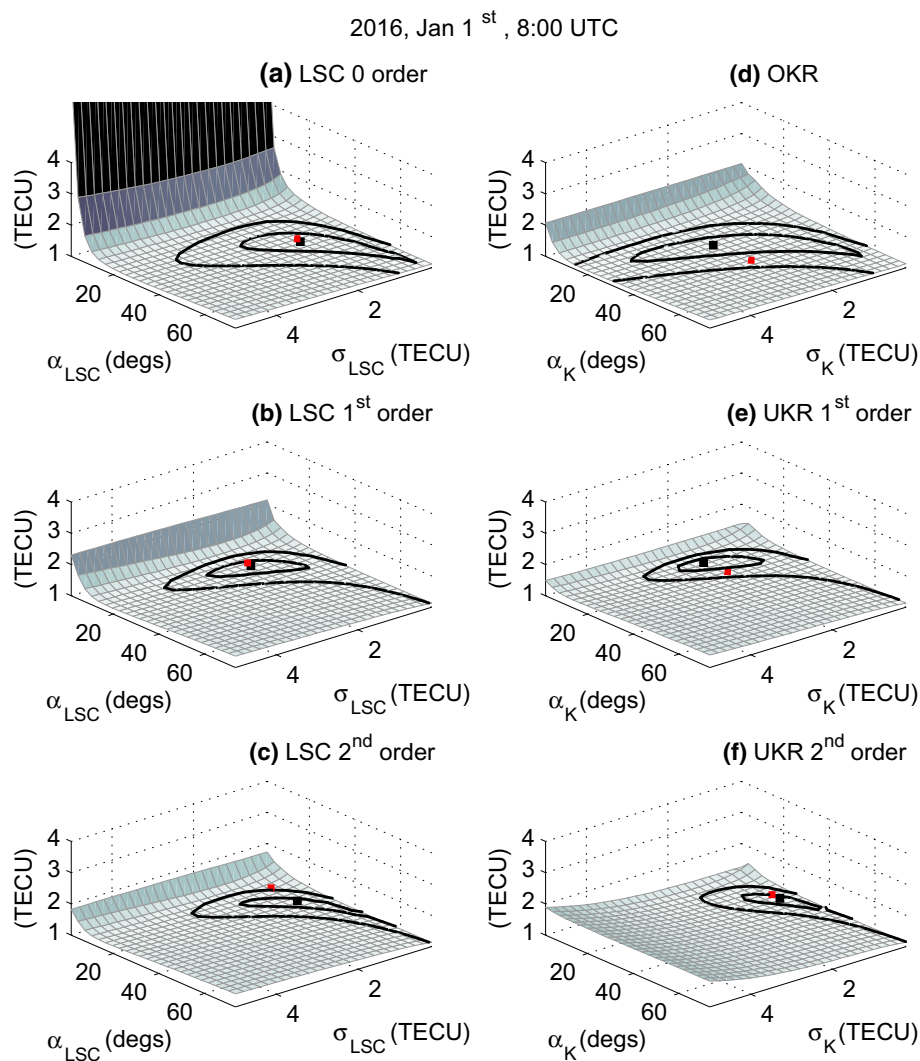
sparse data regions, extrapolated regions, outlier occurrences and different solar conditions (Fig. 10, 11, 12).

This study applies precise LOO parametrization before the comparison of the two detrending approaches, different in LSC, OKR and UKR (Fig. 3). The estimation of  $\alpha_{LSC}$  and  $\alpha_K$  parameters by the covariance function fitting has been also compared to that coming from LOO (Fig. 4). The optimal parameters in the modeling are crucial if we want to be sure that we analyze and compare the best possible solutions of LSC to those from OKR and UKR. The RMSLOO parameter, as coming from the best solution, is also a kind of relative measure of the obtained accuracy. It is also used to find the most accurate and robust way of the trend removal (Figs. 7, 8, 9). Additionally, the unused data from the dataset remaining after data selection are applied in additional comparisons and provide another measures of accuracy differences (Figs. 10, 11, 12). Therefore, the process of the validation can appear

as combined together with the parametrization step, as it uses similar techniques based on cross-validation.

The parametrizations of LSC, OKR and UKR often assume homogeneous uncorrelated noise, and therefore a single value of  $\sigma_{LSC}$  or  $\sigma_K$  is determined for a single stationary process. The theory of LSC uses a priori noise as a name of this parameter, while it is named nugget in kriging techniques. Generally, they are corresponding parameters. However, technically their implementation in the covariance matrices differs, and therefore they are denoted by different symbols in this work:  $\sigma_{LSC}$  in LSC and  $\sigma_K$  in both kriging methods, as OKR and UKR are based on the same variogram and noise model (Eqs. 24 and 25). The second pair of parameters corresponding one to another is correlation length in LSC ( $\alpha_{LSC}$ ) and range in kriging techniques ( $\alpha_K$ ). There is a third covariance parameter present in both Eqs. (15) and (24), denoted as  $C_0$ . This parameter stands for the signal variance, and it is approximated by the residual data variance in this

**Fig. 3** Example epoch (January 1, 2016, 8:00 AM) of  $\alpha_{LSC}/\alpha_K$  and  $\sigma_{LSC}/\sigma_K$  parameter estimation by LOO for: **a** LSC after mean removal, **b** LSC after first-order trend removal, **c** LSC after second-order trend removal, **d** OKR, **e** UKR with first-order trend, **f** UKR with second-order trend. The contours with interval of 0.02 TECU indicate minima of RMSLOO. The black dot indicates minimum RMS in LOO and corresponding  $\alpha_{LSC}/\alpha_K$  and  $\sigma_{LSC}/\sigma_K$  estimates. The red dot indicates alternative  $\alpha_{LSC}/\alpha_K$  estimate from the manual covariance function fitting in Fig. 4

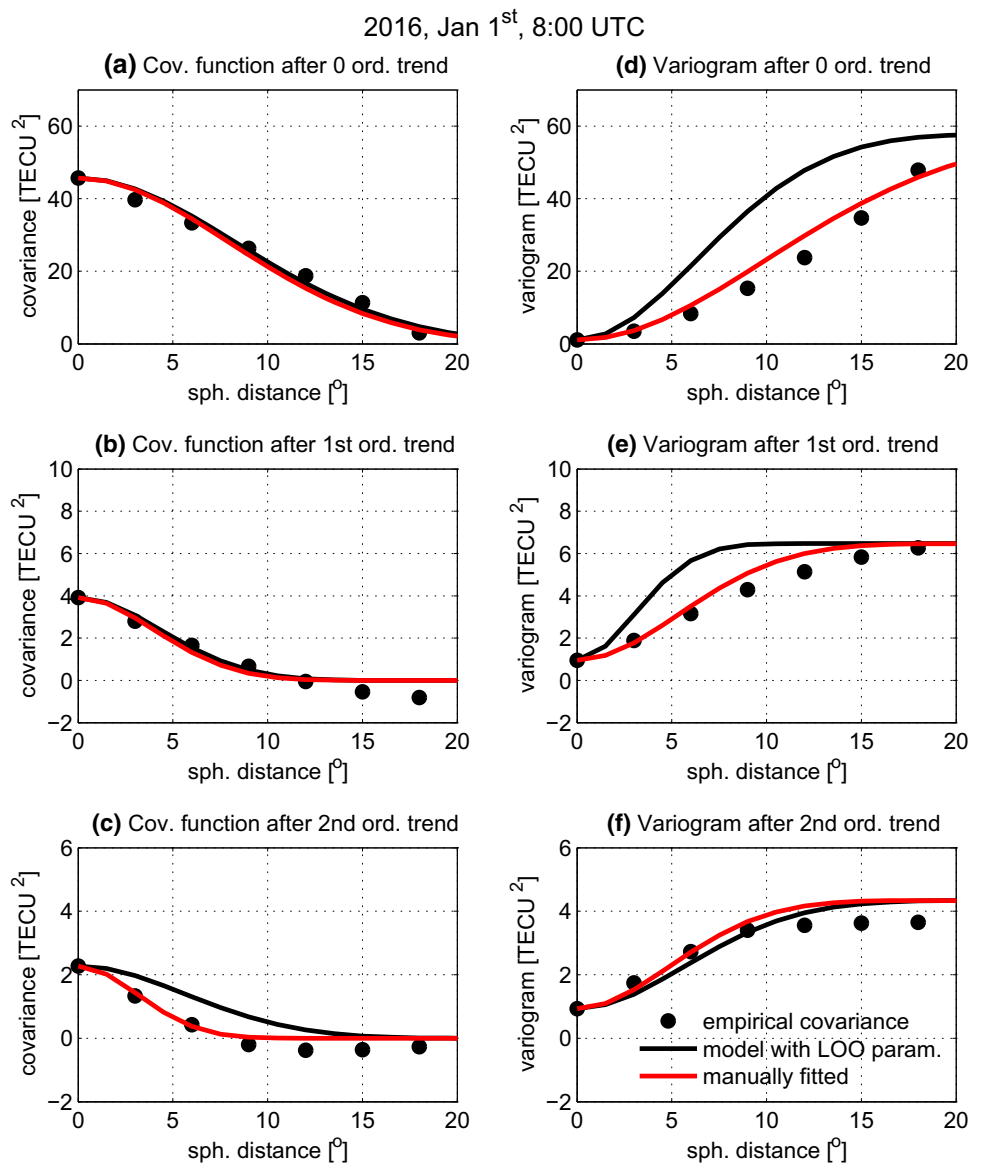


study. This way we limit LOO parametrization to two parameters and make it more reliable in practical computation. This simplification introduces some uncertainty of  $\sigma_{LSC}/\sigma_K$  parameters. However, its insignificance on the modeling results is proven later in this section. The parametrization of  $C_0$  together with  $\sigma_{LSC}$  or  $\sigma_K$  is difficult by LOO, as these parameters are dependent in the applied kind of covariance models (Jarmołoski and Bakula 2014).

This study is based on LOO method of parametrization, which is composed of interpolation in the positions of data points with the use of 100 nearest points that are located at most  $20^\circ$  from the investigation point, excluding the data in the place of interest. The 60 points are the closest points, and the remaining 40 points are a sparsed subset selected from more distant points located up to  $20^\circ$  from the investigated point. This way we can assure the influence of the distant signal and limit the calculation time. The spherical distance of  $20^\circ$  was empirically selected as the smallest that enables the calculation of the covariance parameters in

case of zero-order trend. This is because it approximately corresponds to the distance at which the covariance model for zero-order detrended data is noticeably larger than zero (Fig. 4a). Subsequently, the estimate is compared to the data value and renders a single LOO value. Then, the RMS of all LOO differences (RMSLOO) in data points is calculated for the selected parameters, and the procedure is repeated for the selected range of parameters in order to find their best set with the minimum value of RMSLOO (Behnabian et al. 2018; Krypiak-Gregorczyk et al. 2017; Jarmołoski 2016). The estimation of  $C_0$  in LSC is always based on the whole set of residual data, so the covariance function (Eq. 15) is parametrized globally for the whole residual dataset. In the same way, we estimate  $C_0$  and variogram (Eq. 24) in OKR, as its estimation in OKR is based on the residuals detrended by the mean value only. This does not change the residuals variance regardless of whether the mean is estimated from the whole dataset or its subset. The residuals of the whole dataset are more representative from the statistical point of view, as

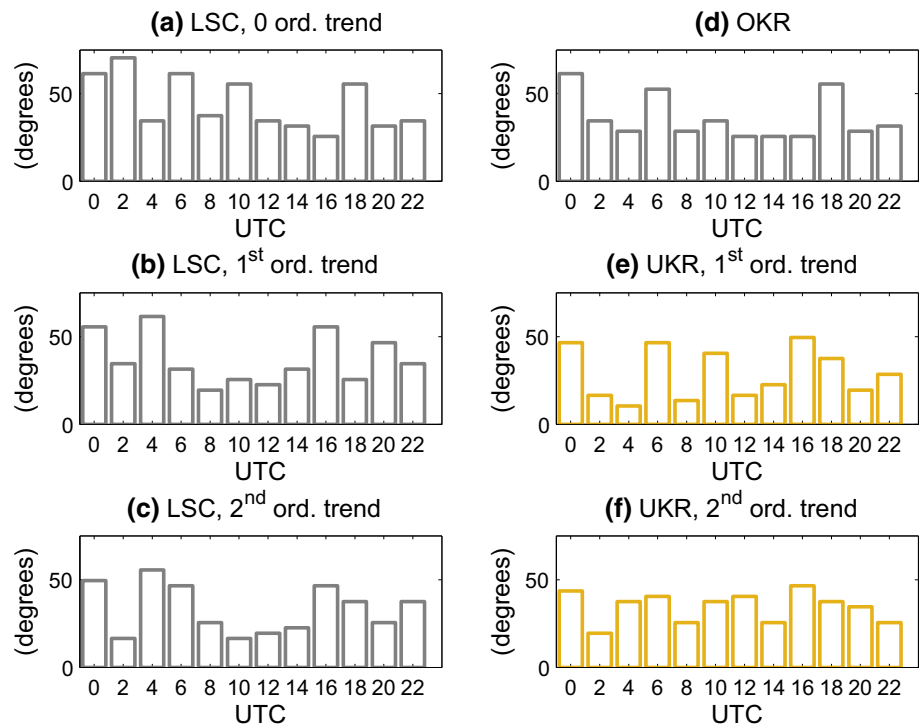
**Fig. 4** Covariance functions and variograms for the six modeling schemes. The dots describe empirical covariance/variogram estimates, the black line is the  $C^4$ -Wendland model with parameters from LOO, and the red line is manual fit of the same model



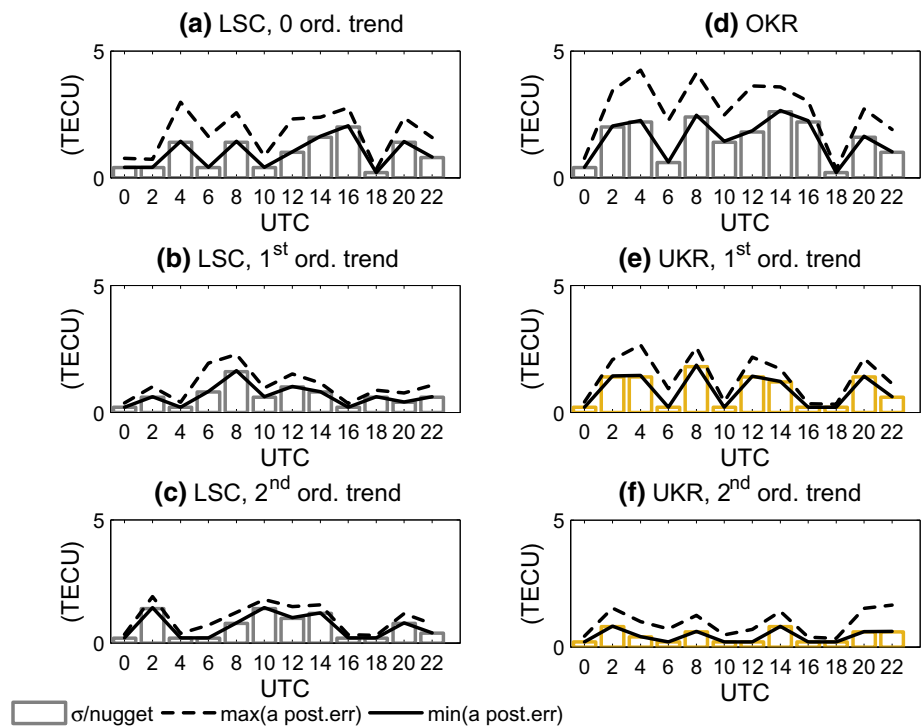
more Gaussian-distributed data can be applied in variogram parametrization. The  $C_0$  parameter in UKR cannot be calculated for the whole dataset, as detrending is applied locally in the variogram matrix (Eq. 29), which is created using local points, and the remaining data cannot be included in the detrending. Of course, the point prediction could involve all the data in UKR; however, such inclusion would make the prediction equivalent to LSC. This study, however, refers to the local prediction, where the residuals are correlated within some spatial distance only, and the limitation of the sampling is crucial for practical implementation in the engineering and science applications. The detailed illustrations of the parametrizations for the selected time epoch are shown in Fig. 3, and the estimated parameters for the whole day at 2-h intervals are presented in Figs. 5 and 6.

The minima of the RMSLOO surfaces drawn for the selected parameters ranges, i.e.,  $\alpha_{LSC} = \alpha_K \in \{0.5:1:24.5\}^\circ$  and  $\sigma_{LSC} = \sigma_K \in \{0.1:0.2:5\}$  TECU, indicate the optimal parameters on the axes of the parameters, depending on the method (Fig. 3) and also on the epoch (Figs. 5 and 6). However, the minima of RMSLOO calculated with the use of these optimal parameters differ significantly only between the epochs, but not between the methods (Fig. 7). This strongly indicates a minor difference between the methods, and the differences in specific cases will be observed later, in relation to Figs. 8, 9, 10 and 11. It can be noted in Fig. 1 that the removal of the mean cannot produce Gaussian residuals, nor even close to Gaussian from local VTEC data, and the parametrization and modeling are biased. An insufficient detrending with the use of mean in LSC and OKR produces the largest RMSLOO at two edges of the parameter space,

**Fig. 5**  $\alpha_{LSC}/\alpha_K$  parameter estimated for, January 1, 2016, at 2-h intervals. The methods are in the same order as in Fig. 4



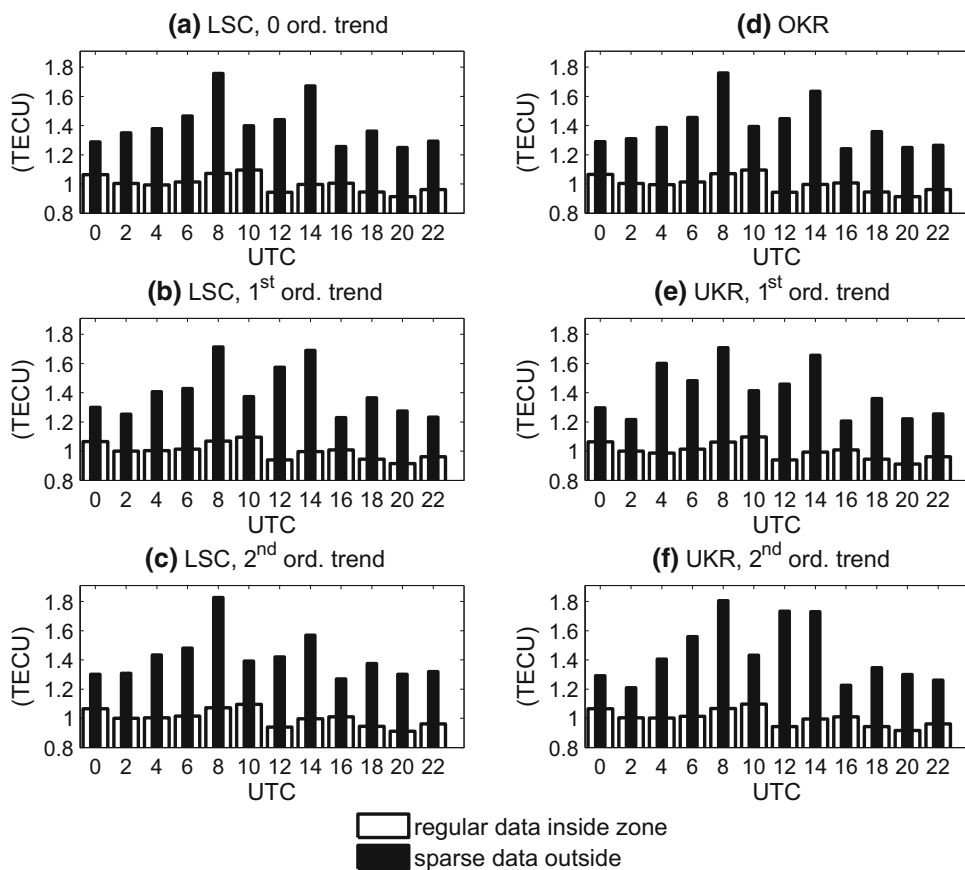
**Fig. 6**  $\sigma_{LSC}/\sigma_K$  parameter (bars) and minima (solid line) and maxima (dashed line) of  $\sigma'_{LSC}/\sigma'_{OKR}/\sigma'_{UKR}$  estimated at points for January 1, 2016, at 2-h intervals



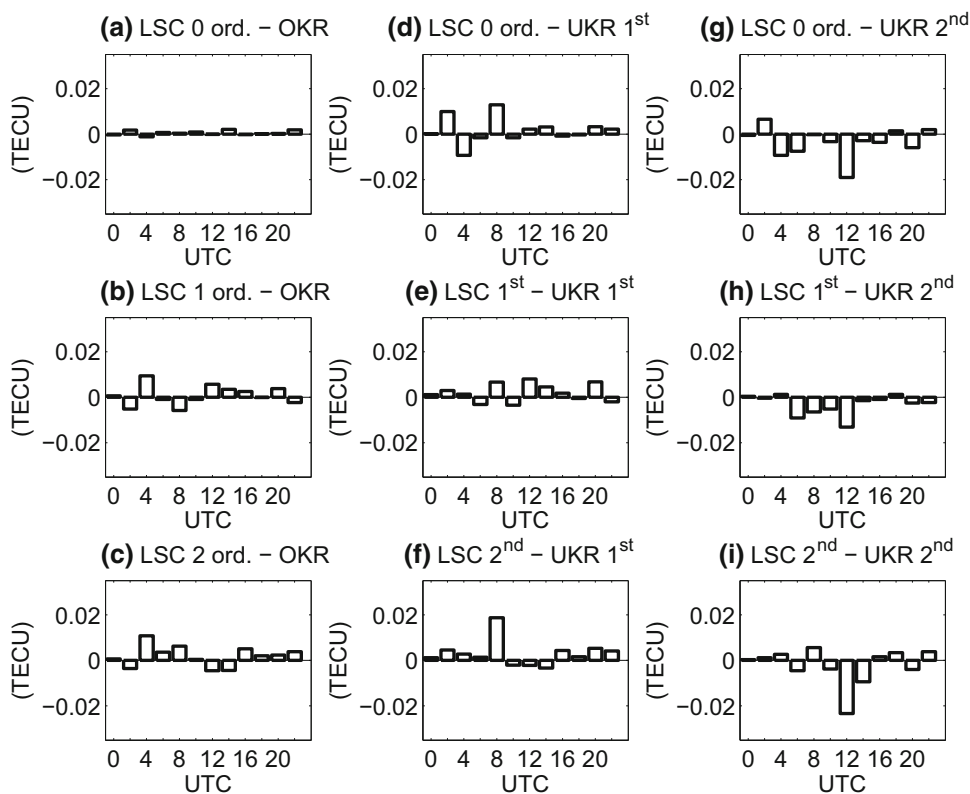
i.e., for small  $\alpha_{LSC}/\alpha_K$  and small  $\sigma_{LSC}/\sigma_K$  (Fig. 3a, d). This automatically requires a larger  $\alpha_{LSC}/\alpha_K$  and  $\sigma_{LSC}/\sigma_K$  with the lowest-order trend (mean) in order to keep the same accuracy as for the higher orders of detrending in LSC or UKR (Fig. 3b, c, e, f). Additionally, as it seems from all

parametrizations in Fig. 3, it is even better to overestimate these parameters, rather than underestimate. Moreover, the surface of RMSLOO in Fig. 3f indicates that the parameters for UKR with the second-order trend must be determined with particular accuracy. This is because an overestimation

**Fig. 7** RMSLOO minima estimated for January 1, 2016, at 2-h intervals. White bars show the values for data inside the red rectangle from Fig. 1, black bars show values at points outside the red zone

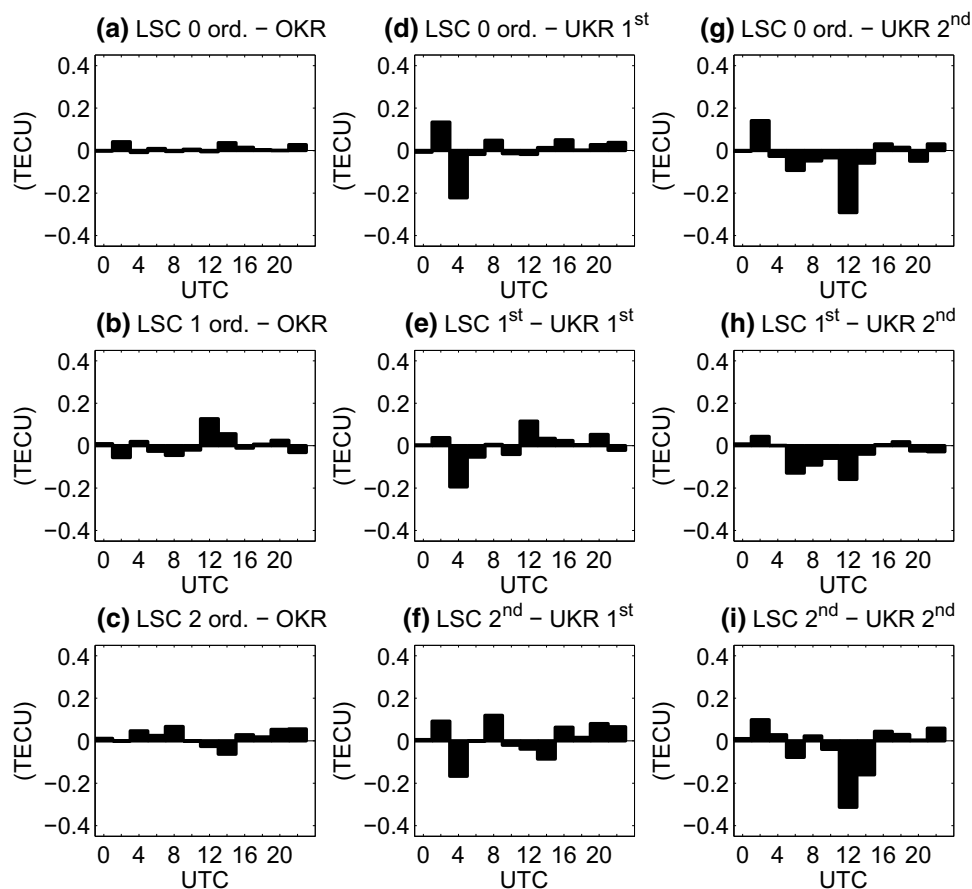


**Fig. 8** Differences of RMSLOO minima for LSC and between different six modeling approaches for the data inside the red zone from Fig. 1





**Fig. 9** Differences of RMSLOO minima for LSC and between different six modeling approaches for the data outside the red zone from Fig. 1



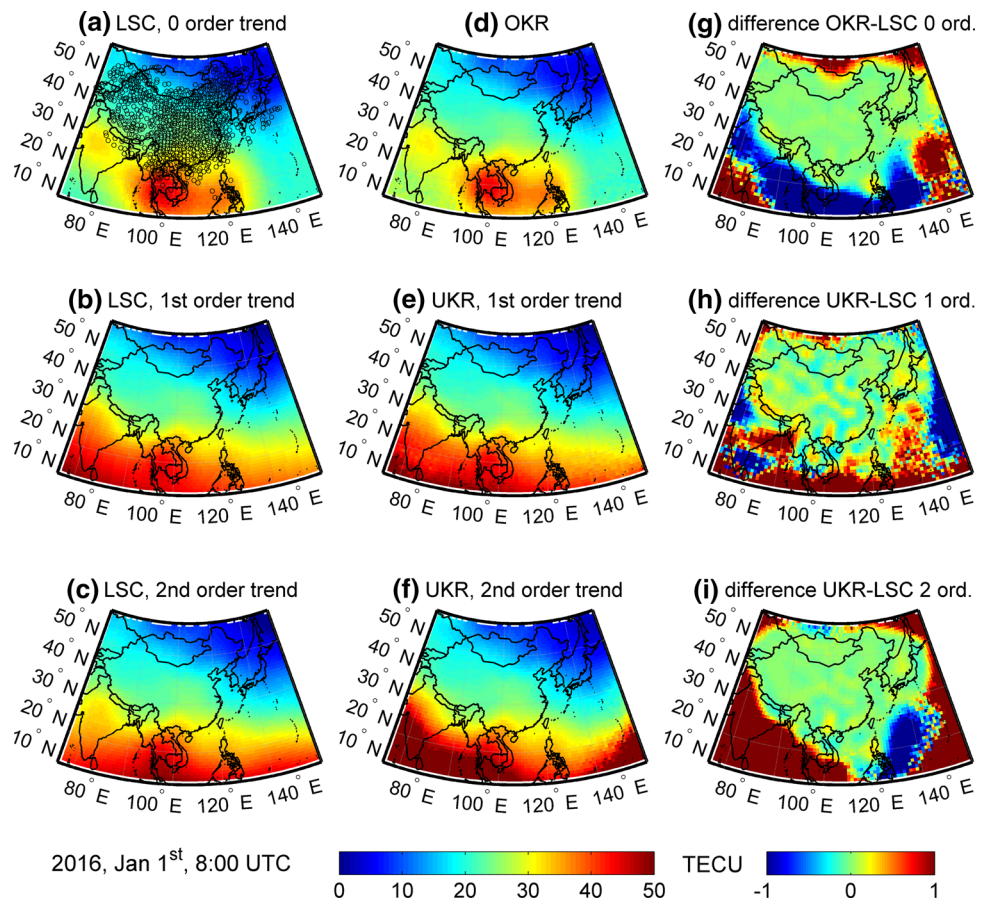
of the noise  $\sigma_K$  increases the RMSLOO and significantly decreases the model accuracy. This happens due to the problems with the local detrending in Eq. (30), when the limited number of data and the use of large noise cause a significantly worse fit of the polynomial trend of the second order.

In addition to the LOO parametrization, the calculation of empirical covariance functions (Hofmann-Wellenhof and Moritz 2005, p. 347) and empirical semivariograms (Wackernagel 2003, p. 47) is done for the same six approaches as in Fig. 3. The sampling interval of the empirical functions was set to  $3^\circ$ , in order to make it larger than average data resolution, which is not everywhere close to  $1^\circ$ . The empirical variograms have been calculated from the residual data, as well as empirical covariance functions in order to keep them consistent. This way the covariance (Eq. 15) and semivariogram (Eq. 24) models based on the parameters from LOO can be compared with the models that fit the empirical values best (Fig. 4). The shape of empirical covariance function strongly depends on the sampling rate, especially for smaller distances, which impedes the assessment of  $\sigma_{LSC}/\sigma_K$  this way. Therefore, we assume the determination of  $\alpha_{LSC}/\alpha_K$  from the empirical covariances/variograms only. Figure 4 shows various coincidences of LOO and the empirical model fit. Worse fit of empirical variograms occurring especially in

Fig. 4d, e can originate from the local detrending schemes in LOO, compared to empirical semivariograms determined from globally detrended data.

The estimated parameters  $\alpha_{LSC}/\alpha_K$  in Fig. 5 and  $\sigma_{LSC}/\sigma_K$  in Fig. 6 are indicated by the minima of the RMSLOO that are placed inside the contours visible on the RMSLOO surfaces drawn for the selected example epoch in Fig. 3. These parameters correspond to the best interpolation processes in the sense of LOO, and only these parameters are further applied in the calculation of the RMSLOO estimates of the modeling error in Fig. 7, their differences in Figs. 8 and 9, as well as in the example predictions in Figs. 10 and 11. The parametrization of LSC and OKR defines global  $\alpha_{LSC}$  (Fig. 5a–c) and  $\sigma_{LSC}$  (Fig. 6a–c) parameters or  $\alpha_K$  (Fig. 5d) and  $\sigma_K$  (Fig. 6d), respectively. This is because  $C_0$  is determined for the whole dataset every time. It is different from UKR, where  $\alpha_K$  and  $\sigma_K$  should differ for different points. So the parameters indicated by the RMSLOO (Fig. 3e, f) in the single analyzed epoch are in fact a kind of average, which gives optimum fit of the prediction to the data when applied to all point predictions. It means that individual point predictions arrive at the same parameters even though they have various  $C_0$  values calculated from the locally detrended residuals (Eq. 29). The  $\alpha_K$  parameters for the whole day are shown in Fig. 5e,

**Fig. 10** Example epoch (January 1, 2016, 8:00 AM) of VTEC interpolated by six methods and differences between respective trend orders in LSC and kriging for data free of outliers



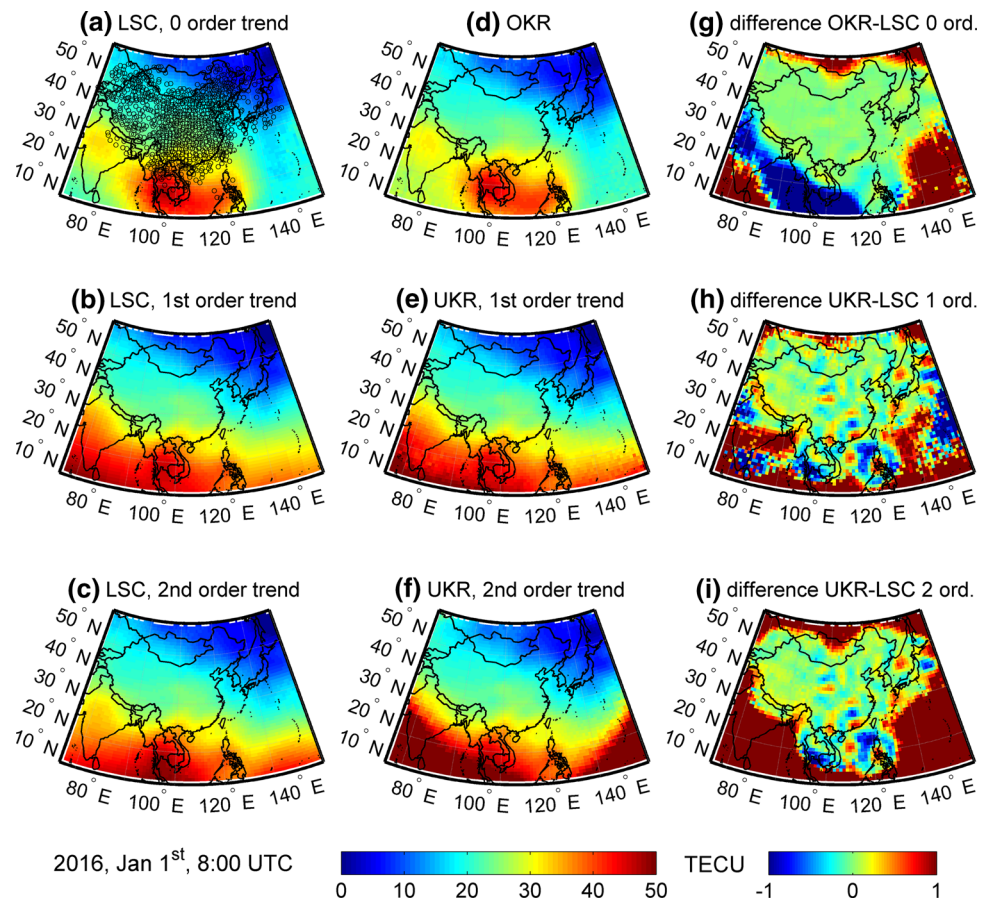
f in orange, to highlight their average character based on the variable local trend and the local  $C_0$ .

The  $\sigma_K$  parameters determined by the smallest RMSLOO in UKR are also presented in orange (Fig. 6e, f), as they come from the same LOO validation with  $C_0$  different at every point. Figure 6 shows the determination of  $\sigma_{LSC}/\sigma_K$  for all the investigated methods, through the whole day at 2-h intervals. The drawback of the slightly inaccurate  $C_0$ , as  $C_0$  is determined from the noisy data and consequently affected  $\sigma_{LSC}$  and  $\sigma_K$ , has no practical influence on the modeling precision determined by the RMSLOO, which is explained in Pardo-Igúzquiza et al. 2009 and Jarmołowski and Bakula (2014). They show there that  $C_0$  and  $\sigma_{LSC}$  (or  $\sigma_K$ ) are dependent on each other, and the modeling results depend on the relation between these two parameters. Besides  $\sigma_{LSC}/\sigma_K$ , Fig. 6 also presents the minima and maxima of a posteriori error estimates of the individual prediction methods, denoted  $\sigma'_{LSC}$ ,  $\sigma'_{OKR}$  and  $\sigma'_{UKR}$ . The minima of  $\sigma'_{LSC}$ ,  $\sigma'_{OKR}$  and  $\sigma'_{UKR}$  are very close to  $\sigma_{LSC}$  or  $\sigma_K$  for all methods, because at most of the points a posteriori estimates are close to a priori noise values, and only some individual values at the margins of the area obtain worse  $\sigma'$  estimates. The estimates  $\sigma'_{LSC}$ ,  $\sigma'_{OKR}$  and  $\sigma'_{UKR}$  are strongly dependent on  $\sigma_{LSC}$  or  $\sigma_K$ , and therefore an overestimated  $C_0$ , which leads to an overestimated  $\sigma_{LSC}$

or  $\sigma_K$ . In consequence, this provides overestimated a posteriori errors, which are smaller in practice. The especially large noise indicators (both a priori and a posteriori), in case of the detrending by the mean (Fig. 6a, d), are suspected to be related to the bias that comes from insufficient detrending and the influence of far zone correlation. The removal of the mean appears as not sufficient for the local areas, especially in case of TEC data, as its largest anomalies, extending over tens of degrees, need to be removed in order to obtain approximately Gaussian residuals.

The LOO validation process within the selected ranges of  $\alpha_{LSC}/\alpha_K$  and  $\sigma_{LSC}/\sigma_K$  parameters enables us to find the smallest values of the RMSLOO for each epoch, which identify the optimal covariance parameters. These RMSLOO minima are presented in Fig. 7a–f, where no apparent differences can be found between the methods, despite the differences in the parametrization (Figs. 5 and 6). Figure 7 describes the RMSLOO calculated for the points inside the red rectangle from Fig. 1 (white bars) and outside it (black bars). The differences between the estimates of  $\alpha_{LSC}/\alpha_K$  and  $\sigma_{LSC}/\sigma_K$  are often significant (Figs. 5, 6), whereas the differences of RMSLOO minima between the methods are smaller (Fig. 7). This was theoretically provided and justified in Sansó et al. (1999), who notice that the error estimates are

**Fig. 11** Example epoch (January 1, 2016, 8:00 AM) of VTEC interpolated by the six methods and differences between respective trend orders in LSC and kriging for data with restored outliers



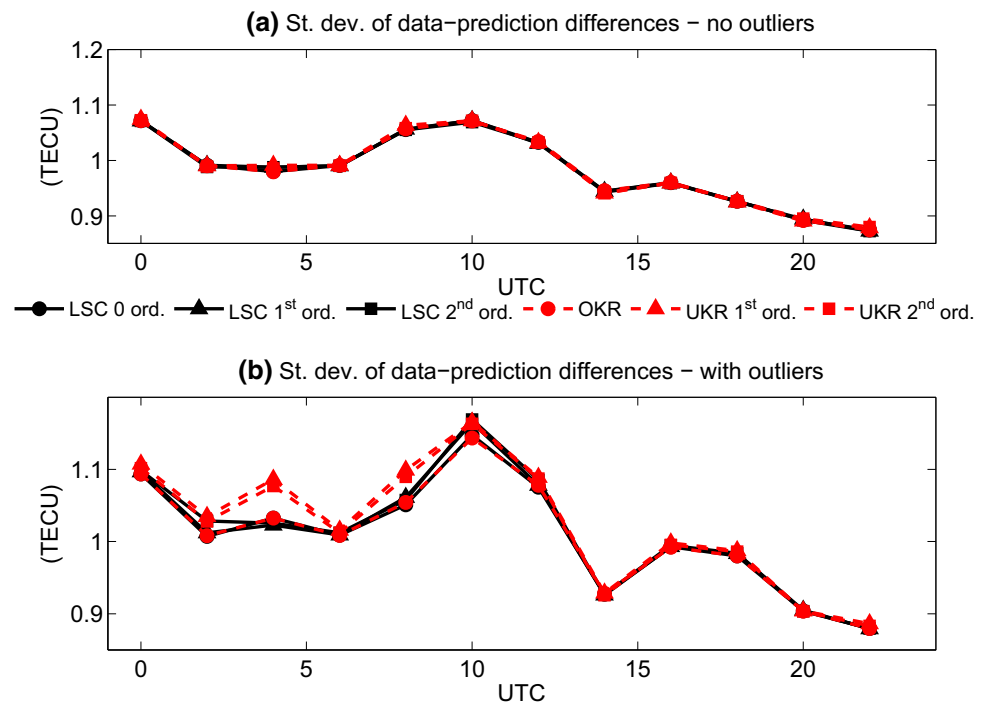
more sensitive to parameters change than estimated values. It is evident at the same time that individual optimal parameters estimated for particular methods noticeably affect the minima of RMSLOO only at data margins (Fig. 7). This means that all the predictions are at a similar level of quality if we apply a precise LOO parametrization for homogeneous dense data interpolation. Referring to the noise estimates, it must be concluded that the RMSLOO is a better estimate of the noise than a priori noise estimates  $\sigma_{LSC}/\sigma_K$ , and a posteriori errors  $\sigma'_{LSC}/\sigma'_{OKR}/\sigma'_{UKR}$ , which are strongly related to a priori values, due to the above-mentioned error sensitivity.

In order to assess possible smaller-scale advantages of some methods, the values of RMSLOO minima from kriging are subtracted from the RMSLOO minima from LSC. Figure 8g–i shows increasing number of negative differences of the RMSLOO minima, which indicates small advantage of LSC with respect to UKR, in terms of a smaller average RMSLOO minimum. This indicates the problems with higher-order detrending inside UKR variogram matrix, for locally selected data, which rapidly increase at the margins of the data, i.e., outside the red zone from Fig. 1 (Fig. 9g–i).

Aside from analyzing the RMSLOO, which is an empirical measure of the accuracy in terms of cross-validation, the VTEC models are created using different six modeling

schemes for large areas extended to sparse marginal data regions and extrapolation regions. These grids allow for the observation of the differences in the worsening conditions, when we interpolate at the data margins, where the data lose their homogeneous distribution. Two sets of models are calculated in 2-h intervals, as well as the covariance parameters above: one set without outliers visible in Fig. 2 and the other including these outliers in the estimation process. The parametrization used in both cases is based on the same LOO validation from Figs. 5 and 6. Figure 10 shows the models calculated without outliers, and Fig. 11 shows the models including outliers in the estimation. The range of data selection for the estimation is extended to 40° in order to enable the covering of the extrapolated areas. Figures 10a and 11a show the data used in the interpolation and extrapolation in order to track where the extrapolated regions are. In fact, these regions obtain most of the signal from the trend restoring process in LSC, whereas in OKR and UKR they are affected by inadequate trend modeling, as they are based on limited data. The differences between LSC and kriging for the respective detrending orders do not exceed 1 TECU inside the interpolation zone, for the models free of outliers (Fig. 10g, i). The extrapolations coincide even better for first-order trend (Fig. 10h) than for the mean removal option (Fig. 10g). The

**Fig. 12** Standard deviations of comparison between values interpolated from models created by the six methods in the positions of free of outliers data points and data: **a** the models are interpolated from data without outliers, and **b** the models are interpolated from data including outliers



extrapolation fails completely in UKR with the second-order trend, as this polynomial obtains improperly exaggerated curvature from the local data, especially at longer distances from the data points.

The differences between the respective models created with the restored outliers reach or exceed 1 TECU in case of UKR (Fig. 11h, i). Of course, every modeling process is preceded by an effort of outliers elimination; however, there can appear a case of inefficient outliers detection/elimination or small data samples, which does not allow for the outliers removal. These can be the cases for which this extension of the study can indicate a safer method of the interpolation. Nevertheless, the differences between the models in Fig. 11h, i do not indicate which one fits the potentially good physical field better.

The only independent source of the information about the ‘true’ physical value of VTEC, aside from the data points free of outliers used for LOO and models, is data remaining after the selection process shown by black circles in Fig. 2. Therefore, the models from Fig. 10a–f and those from Fig. 11a–f are interpolated back using simple bilinear interpolation from  $1^\circ$  grids in the positions of previously unused data free of outliers. Then, the differences between unused data points and interpolated grids from Fig. 10 are compared to differences between the same data and grids from Fig. 11 in all epochs with a 2-h time interval. The results can be viewed in Fig. 12, which provide three interesting observations. The models without the outliers fit with the same accuracy for LSC and kriging across the day (Fig. 12a). The models estimated with outliers fit with similar accuracy for LSC and

kriging only in time where the sun has left the investigated area (Fig. 12b). When the equatorial anomaly passes the China region, a predominant number of LSC solutions are better than UKR solutions based on first- and second-order local polynomial trends (Fig. 12b). The most suspected reason of LSC advantage is global detrending, which, based on a large dataset, becomes more robust to the individual outliers with respect to the local trend fitting.

## 6 Conclusions

This work investigates LSC, OKR and UKR and proves their consistency under the conditions of precise parametrization, homogeneous spatial distribution, lack of outliers and equivalent detrending, which numerically confirms the comments provided by Dermanis (1984). This consistency is observed from the minima of RMSLOO, which are comparable for all six methods tested in this work with regard to dense data without significant outliers and far from data gaps. However, the margins of the data where the gaps start to occur, as well as the places where outliers remain among the correlated data, show a significantly better validation results when modeled with LSC/SKR. Additionally, the daylight hours and equatorial anomaly pass turn out to be challenging for UKR TEC modeling. The unbiasedness constraint applied in OKR and UKR is often hastily considered as a substantial advantage in comparison with LSC and SKR. The actual study shows a worse performance of the local detrending applied in UKR in comparison with the global trend removal applied in LSC/SKR.



The superiority of global detrending in LSC is confirmed numerically for the sparse data regions and also in case of the outlier occurrences.

This study investigates the lowest orders of the trend, as they are most often used in practice, and the aim of this work was to assess LSC, OKR and UKR together. Extended data areas and specific properties of the signal in terms of its variance can, however, require higher-order trends so as to obtain homogeneous statistical distribution of the residuals. Otherwise, due to the imperfect distribution of the residuals, the estimation of the parameters can generate lower accuracy in some cases, as demonstrated in this paper.

Therefore, unbiased detrending has been proven to be crucial in stochastic modeling. The drawback of UKR is such that the polynomials fit worse when based on a limited subset of the data, especially if more noisy or distant data are applied. It is a common practice in gravity field modeling that detrending is performed in the so-called remove–restore method, with the use of lower-order trend from spherical harmonic expansion of geopotential functionals, i.e., using existing global geopotential models (GGM). In TEC modeling, the equatorial anomaly includes the predominant part of the TEC signal and, therefore, in times of the increasing number of GIMs, their applicability in VTEC data detrending can become especially valuable. This is of course associated with the application of LSC/SKR instead of OKR or UKR. The drawback of GIMs lies in their relatively low order of spherical harmonic expansion, which may still introduce some bias in detrending of local TEC observations.

**Acknowledgements** The study is supported by grant no. UMO-2017/27/B/ST10/02219 from the Polish National Center of Science. We thank CMONOC for providing the GNSS data necessary for calculating TEC. The authors wish to thank the Editor-in-Chief and Associate Editor for their regardful assistance and three anonymous reviewers for their hard work on the article review.

**Author contributions** WJ designed the research, processed the data and prepared the paper draft, PW supervised the research and revised the manuscript, XR calculated TEC data from GNSS, and AKG contributed to the data analysis.

**Data availability statement** TEC data are available from the authors upon a reasonable request.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alizadeh MM, Schuh H, Todorova S, Schmidt M (2011) Global ionosphere maps of VTEC from GNSS, satellite altimetry, and formosat-3/COSMIC data. *J Geod* 85(12):975–987. <https://doi.org/10.1007/s00190-011-0449-z>
- Al-Shaery A, Lim S, Rizos C (2011) Investigation of different interpolation methods used in network-RTK for virtual reference station technique. *J Glob Position Syst* 10:136–148
- Arabelos D, Tscherning CC (1998) The use of least squares collocation method in global gravity field modeling. *Phys Chem Earth* 23(1):1–12
- Arlot S, Celisse A (2010) A survey of cross-validation procedures for model selection. *Stat Surv* 4:40–79. <https://doi.org/10.1214/09-SS054>
- Barzaghi R, Borghi A, Ducarme B, Everaerts M (2003) Quasi-geoid BG03 computation In Belgium. *Newton's Bulletin* (1)
- Behnbabian B, Mashhadi Hossaini M, Malekzadeh A (2018) Simultaneous estimation of cross-validation errors in least squares collocation applied for statistical testing and evaluation of the noise variance components. *J Geod* 92:1329. <https://doi.org/10.1007/s00190-018-1122-6>
- Blais JAR (1982) Synthesis of kriging estimation methods. *Manuscripta Geodaetica* 7(4):325–353
- Brooker PI (1986) A parametric study of robustness of kriging variance as a function of range and relative nugget effect for a spherical semivariogram. *Math Geol* 18:477–488
- Chen J, Yang S, Yan W, Wang J, Chen Q, Zhang Y (2017) Recent results of the Chinese CMONOC GNSS network. In: *Proceedings of the ION Pacific PNT conference (ION PNT 2017)*, Honolulu, Hawaii, pp 539–546.
- Daya AA, Bejari H (2015) A comparative study between simple kriging and ordinary kriging for estimating and modeling the Cu concentration in Chehlkureh deposit, SE Iran. *Arab J Geosci* 8:6003–6020
- Dermanis A (1984) Kriging and collocation - a comparison. *Manuscr Geodaet* 9(3):159–167
- Diggle P, Ribeiro PJ Jr (2007) *Model-based geostatistics*. Springer, New York
- Erdogan E, Schmidt M, Seitz F, Durmaz M (2017) Near real-time estimation of ionosphere vertical total electron content from GNSS satellites using B-splines in a Kalman filter. *Ann Geophys* 35:263–277
- Guinness J, Fuentes M (2016) Isotropic covariance functions on spheres: some properties and modeling considerations. *J Multiv Anal* 143:143–152
- Gneiting T (2013) Strictly and non-strictly positive definite functions on spheres. *Bernoulli* 19(4):1327–1349
- Hofmann-Wellenhof B, Moritz H (2005) *Physical geodesy*. Springer, New York
- Huang C, Zhang H, Robeson SM (2011) On the validity of commonly used covariance and variogram functions on the sphere. *Math Geosci* 43:721–733
- Jarmołowski W (2015) Least squares collocation with uncorrelated heterogeneous noise estimated by restricted maximum likelihood. *J Geod* 89:577–589. <https://doi.org/10.1007/s00190-015-0800-x>
- Jarmołowski W (2016) Estimation of gravity noise variance and signal covariance parameters in least squares collocation with considering data resolution. *Ann Geophys* 59(1):S0104
- Jarmołowski W (2017) Fast estimation of covariance parameters in least-squares collocation by fisher scoring with Levenberg–Marquardt optimization. *Surv Geophys* 38(4):701–725
- Jarmołowski W (2019) On the relations between signal spectral range and noise variance in least-squares collocation and simple kriging:



- example of gravity reduced by EGM2008 signal. *B Geofis Teor Appl.* <https://doi.org/10.4430/bgta0265>
- Jarmołowski W, Bakula M (2014) Precise estimation of covariance parameters in least-squares collocation by restricted maximum likelihood. *Stud Geophys Geod* 58:171–189
- Jarmołowski W, Ren X, Wielgosz P, Krypiak-Gregorczyk A (2019) On the advantage of stochastic methods in the modeling of ionospheric total electron content: South-East Asia case study. *Meas Sci Technol* 30(4):044008. <https://doi.org/10.1088/1361-6501/ab0268>
- Koch KR (1999) Parameter estimation and hypothesis testing in linear models, ed 2. Springer, Berlin. <https://doi.org/10.1007/978-3-662-03976-2>
- Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Melish CS (ed) *IJCAI-95: Proceedings of the fourteenth international joint conference on artificial intelligence*, vol 2. Morgan Kaufmann, pp 1137–1143. <https://ijcai.org/Past%20Proceedings/IJCAI-95-VOL2/PDF/016.pdf>
- Krypiak-Gregorczyk A, Wielgosz P, Jarmołowski W (2017) A new TEC interpolation method based on the least squares collocation for high accuracy regional ionospheric maps. *Meas Sci Technol* 28(4):045801. <https://doi.org/10.1088/1361-6501/aa58ae>
- Li W, Huang L (2019) Zhang S and Chai Y (2019) Assessing global ionosphere TEC maps with satellite altimetry and ionospheric radio occultation observations. *Sensors* 19:5489. <https://doi.org/10.3390/s19245489>
- Lichtenstern A (2013) Kriging methods in spatial statistics, Bachelor's Thesis, Department of Mathematics, Technische Universität München
- Ligas M, Kulczycki M (2010) Simple spatial prediction – least squares prediction, simple kriging, and conditional expectation of normal vector. *Geodesy and Cartography* 59(2):69–81
- Liu T, Zhang B, Yuan Y, Li M (2018) Real-Time Precise Point Positioning (RTPPP) with raw observations and its application in real-time regional ionospheric VTEC modeling. *J Geod* 92(11):1267–1283
- Malvić T, Balić D (2009) Linearity and Lagrange linear multiplier in the equations of ordinary kriging. *Nafta* 60(1):31–43
- Mao T, Wan W, Yue X, Sun L, Zhao B, Guo J (2008) An empirical orthogonal function model of total electron content over China. *Radio Sci* 43:RS2009. <https://doi.org/10.1029/2007RS003629>
- 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. In: *Proceedings of the 5th international Antarctic geodesy symposium AGS'03*, Lviv, Ukraine, SCAR Report No. 23. <https://www.scar.org/publications/reports/23/>
- Moritz H (1980) *Advanced physical geodesy*. Herbert Wichmann Verlag, Karlsruhe
- Odiijk D (2002) *Fast precise GPS positioning in the presence of ionospheric delays*. PhD thesis, Delft University of Technology, Delft, Netherlands
- O'Dowd RJ (1991) Conditioning of coefficient matrices of ordinary kriging. *Math Geol* 23(5):721–739
- Olea RA (1999) *Geostatistics for engineers and earth scientists*. Springer, New York, p 303
- Orús R, Hernández-Pajares M, Juan JM, Sanz J (2005) Improvement of global ionospheric VTEC maps by using kriging interpolation technique. *J Atmos Sol Terr Phys* 67:1598–1609
- Orús R, Cander LR, Hernández-Pajares M (2007) Testing regional vertical electron content maps over Europe during 17–21 January 2005 sudden weather event. *Radio Sci* 42:RS3004
- Pardo-Igúzquiza E, Mardia KV, Chica-Olmo M (2009) MLMATERN: a computer program for maximum likelihood inference with the spatial Matérn covariance model. *Comput Geosci* 35:1139–1150
- Peng C-Y, Wu CFJ (2014) On the choice of nugget in kriging modeling for deterministic computer experiments. *J Comput Graph Stat* 23:151–168. <https://doi.org/10.1080/10618600.2012.738961>
- Posa D (1989) Conditioning of the stationary kriging matrices for some well-known covariance models. *Math Geol* 21(7):755–765
- Reguzzoni M, Sansó F, Venuti G (2005) The theory of general kriging, with applications to the determination of a local geoid. *Geophys J Int* 162(4):303–314. <https://doi.org/10.1111/j.1365-246X.2005.02662.x>
- Ren X, Zhang X, Xie W, Zhang K, Yuan Y, Li X (2016) Global Ionospheric modelling using Multi-GNSS: BeiDou, Galileo, GLONASS and GPS. *Sci Rep* 6:33499. <https://doi.org/10.1038/srep33499>
- Roma-Dollase D, Hernández-Pajares M, Krankowski A, Kotulak K, Ghoddousi-Fard R, Yuan Y, Li Z, Zhang H, Shi C, Wang C, Feltens J, Vergados P, Komjathy A, Schaer S, García-Rigo A, Gómez-Cama JM (2018) Consistency of seven different GNSS global ionospheric mapping techniques during one solar cycle. *J Geod* 92(6):691–706. <https://doi.org/10.1007/s00190-017-1088-9>
- Sadiq M, Tscherning CC, Ahmad Z (2010) Regional gravity field model in Pakistan area from the combination of CHAMP, GRACE and ground data Using least squares collocation: A case study. *Adv Space Res* 46:1466–1476
- Samui P, Sitharam TG (2011) Application of geostatistical methods for estimating spatial variability of rock depth. *Engineering* 3:886–894
- Sansó F, Venuti G, Tscherning CC (1999) A theorem of insensitivity of the collocation solution to variations of the metric of the interpolation space. In Schwarz KP (ed) *Geodesy beyond 2000. the challenges of the first decade*, international association of geodesy symposia, vol 121. Springer, Berlin, pp 233–240
- Sayin I, Arikani F, Arikani O (2008) Regional TEC mapping with random field priors and kriging. *Radio Sci* 43:RS5012. <https://doi.org/10.1029/2007RS003786>
- Schmidt M, Dettmering D, Mößner M, Wang Y, Zhang J (2011) Comparison of spherical harmonic and B spline models for the vertical total electron content. *Radio Sci* 46:RS0D11
- Stanisławska I, Juchnikowski G, LjR C (1996) Kriging method for instantaneous mapping at low and equatorial latitudes. *Adv Space Res* 18(6):217–220
- Stanisławska I, Juchnikowski G, LjR C, Ciruolo L, Zbyszynski Z, Świętek A (2002) The kriging method of TEC instantaneous mapping. *Adv Space Res* 29(6):945–948
- Teunissen PJG (2003) *Adjustment theory: an introduction*, 2nd edn. Delft University Press, Delft, Series on Mathematical Geodesy and Positioning
- Todini E (2001) Influence of parameter estimation uncertainty in Kriging. Part 1 – Theoretical development. *Hydrol Earth System Sci* 5:215–232
- Tscherning CC, Rapp RH (1974) Closed covariance expressions for gravity anomalies, geoid undulations and deflections of the vertical implied by anomaly degree variance models, OSU Report No. 208, Department of Geodetic Science and Surveying, The Ohio State University
- Tscherning CC (1991) The use of optimal estimation for gross-error detection in databases of spatially correlated data. *Bulletin d'Information (Bur Gravim int)* 68:79–89
- Wackernagel H (2003) *Multivariate geostatistics: an introduction with applications*, 3rd edn. Springer, Berlin
- Wielgosz P, Grejner-Brzezinska DA, Kashani I (2003) Regional ionosphere mapping with kriging and multiquadric methods. *J Glob Position Syst* 2:48–55
- Zimmermann R (2010) Asymptotic behavior of the likelihood function of covariance matrices of spatial gaussian processes. *J App Math.* Article ID 494070. <https://doi.org/10.1155/2010/494070>