



# HV-LSC-ex<sup>2</sup>: velocity field interpolation using extended least-squares collocation

Rebekka Steffen<sup>1</sup> · Juliette Legrand<sup>2</sup> · Jonas Ågren<sup>1,3</sup> · Holger Steffen<sup>1</sup> · Martin Lidberg<sup>1</sup>

Received: 9 June 2021 / Accepted: 9 January 2022  
© The Author(s) 2022

## Abstract

Least-squares collocation (LSC) is a widely used method applied in physical geodesy to separate observations into a signal and noise part but has received only little attention when interpolating velocity fields. The advantage of the LSC is the possibility to filter and interpolate as well as extrapolate the observations. Here, we will present several extensions to the traditional LSC technique, which allows the combined interpolation of both horizontal velocity components (horizontal velocity (HV)-LSC), the separation of velocity observations on different tectonic plates, and the removal of stationarity by moving variance (the latter as HV-LSC-ex(tended)<sup>2</sup>). Furthermore, the covariance analysis, which is required to find necessary input parameters for the LSC, is extended by finding a suitable variance and correlation length using both horizontal velocity components at the same time. The traditional LSC and all extensions are tested on a synthetic dataset to find the signal at known as well as newly defined points, with stations separated on four different plates with distinct plate velocities. The methodologies are evaluated by calculation of a misfit to the input data, and implementation of a leave-one-out cross-validation and a Jackknife resampling. The largest improvement in terms of reduced misfit and stability of the interpolation can be obtained when plate boundaries are considered. In addition, any small-scale changes can be filtered out using the moving-variance approach and a smoother velocity field is obtained. In comparison with interpolation using the Kriging method, the fit is better using the new HV-LSC-ex<sup>2</sup> technique.

**Keywords** Least-squares collocation · Velocity field · Covariance analysis · Plate boundary · GNSS data · Correlation analysis · Moving variance

## 1 Introduction

Collection of ground data in an area of interest is limited by the accessibility of the area (e.g. due to rough topography, harsh weather conditions, political situation, and/or high

costs). Thus, obtained datasets represent only point information. However, for many applications it is desired to have data values in areas where observations are not available and/or cannot be obtained, for example when a continuous surface of a variable is required. The data of this variable have to be estimated at all wished (regular and/or irregular distributed) points of the continuous surface, which is done with a spatial interpolation. Additionally, if values are required outside the cloud of point information, extrapolation is needed. Thus, spatial interpolation and extrapolation allow estimation of the value of a field variable at any point in space or at nodes of a regular grid using scattered point measurements of this variable.

Several interpolation techniques are available (e.g. bilinear interpolation, nearest-neighbourhood, spline interpolation, inverse-distance weighted, radial basis function, Kriging) to estimate a continuous surface from a set of observations. Most of these techniques have in common that the interpolated value is only directly based on the surround-

---

✉ Rebekka Steffen  
rebekka.steffen@lm.se

Juliette Legrand  
juliette.legrand@oma.be

Jonas Ågren  
jonas.agren@lm.se

Holger Steffen  
holger.steffen@lm.se

Martin Lidberg  
martin.lidberg@lm.se

<sup>1</sup> Lantmäteriet, Gävle, Sweden

<sup>2</sup> Royal Observatory of Belgium, Uccle, Belgium

<sup>3</sup> University of Gävle, Gävle, Sweden

ing values and/or the smoothness described by mathematical formulas (named deterministic interpolation techniques). In contrast, Kriging (e.g. Cressie 2007), for example, is a stochastic interpolation method, where the statistical information between the points is used in the interpolation (e.g. Herzfeld 1992; Chilès and Desassis 2018). In physical geodesy, least-squares collocation (LSC) has become a well-known and widely used stochastic interpolation technique, which is similar to Kriging (e.g. Dermanis 1984; Reguzzoni et al. 2005; Fuhrmann 2016). LSC is usually applied for interpolation and combination of different gravity field and height parameters, for example when estimating geoid undulations (e.g. Moritz 1980; Forsberg and Tscherning 1981). Beforehand, a covariance analysis of the entire dataset has to be done and the results are used as input for the interpolation. Thus, the LSC method considers the spatial dependency structure of the data. In addition, the LSC technique applies a filter in a first step to obtain a consistent dataset at observation points and has the advantage of allowing several parameters to be combined in the interpolation and filtering. Nevertheless, the usage of LSC is still limited in the area of GNSS (Global Navigation Satellite System) velocity field interpolations and only a few studies have used this technique (e.g. Japan (Fujii and Xia 1993; El-Fiky et al. 1997), China (Chen et al. 2016), northern Europe (Vestøl 2006; Vestøl et al. 2019), southern Europe (Kahle et al. 1995, 2000; Straub et al. 1997; Caporali et al. 2003), Switzerland (Egli et al. 2007)). In contrast, Kriging has been used more widely (e.g. Kierulf et al. 2013; Bogusz et al. 2014; Ching and Chen 2015; Fuhrmann et al. 2015; Béjar-Pizarro et al. 2016; Robin et al. 2020) and comparison to other interpolation techniques has shown its advantages (e.g. Bogusz et al. 2014). Other studies have extended interpolation techniques to incorporate geophysical information (e.g. Shen et al. 2015; Sandwell and Wessel 2016), but these methods rely on background models, which are also introducing uncertainties and/or errors. Furthermore, when the horizontal velocity field was interpolated using LSC or Kriging, correlation between the horizontal velocity components was ignored in all previous studies.

The aim of this paper is the presentation of the new LSC methodology modified for the special purpose to filter and interpolate horizontal velocities including correlation between the velocity field components (horizontal velocity (HV)-LSC). This correlation has not been considered in velocity field interpolations before. Our method is a pure mathematical model that does not contain geophysical information and is thus suitable to compare with geophysical observations as well as models. In addition, we present a new method for overcoming the problem of non-stationarity, which arises due to the application of a homogeneous and isotropic covariance function. Observations are usually not evenly distributed and large variations in the observational magnitude may lead to a covariance function, which does

not fit all data equally well. Thus, a method has to be defined to account for this inequality. We also present our chosen method to interpolate a velocity field with stations distributed over several tectonic plates that have different plate motion velocities.

Section 2 will present the standard LSC technique and covariance analysis. This is followed by the description of the HV-LSC for horizontal velocity data and corresponding covariance analysis (Sect. 3). Section 4 presents then the method of including plate-boundary constraints and Sect. 5 the methodology of moving variance to overcome non-stationarity of LSC. Finally, we apply the presented methods on a synthetic dataset (Sect. 6). We also compare the interpolated velocity field to the result using Kriging and discuss advantages and disadvantages of LSC. We finish with a conclusion.

## 2 Least-squares collocation

The method of the standard LSC is a stochastic interpolation technique, and their general framework was presented by Moritz (1980), which is used in the following to summarize the technique. LSC is based on the principle that observations  $\mathbf{l}$  can be separated into a signal  $\mathbf{s}$  and a noise  $\mathbf{n}$  (bold symbols refer to the vectorized variable):

$$\mathbf{l} = \mathbf{s} + \mathbf{n}, \quad (1)$$

where the observations, signal and noise are vectors. The observations  $\mathbf{l} = l(P_i)$  are given at  $N_{l(otal)}$  points. In addition, an observational uncertainty  $\mathbf{m} = m(P_i)$  at the same  $N_l$  locations is provided. It is assumed that  $\mathbf{l}$  and  $\mathbf{s}$  are centred random variables; thus, their statistical expectations are zero ( $\mathbb{E}[\mathbf{l}] = 0$ ,  $\mathbb{E}[\mathbf{s}] = 0$ ). One can define variance–covariance matrices of the signal and the noise:

$$\mathbf{C}_{ss} = \mathbb{E}[\mathbf{s}\mathbf{s}^T] \quad \text{and} \quad \mathbf{C}_{nn} = \mathbb{E}[\mathbf{n}\mathbf{n}^T],$$

respectively. The upper index  $T$  indicates matrix transposition. A cross-covariance matrix between the signal and the noise does not exist as the signal and the noise are uncorrelated ( $\mathbf{C}_{sn} = \mathbb{E}[\mathbf{s}\mathbf{n}^T] = 0$ ). The signal can then be obtained by solving the least-squares minimum condition:

$$\mathbf{s}^T \mathbf{C}_{ss}^{-1} \mathbf{s} + \mathbf{n}^T \mathbf{C}_{nn}^{-1} \mathbf{n} = \text{minimum}. \quad (2)$$

The solution to Eq. (1) can be found with the constraint of Eq. (2) using Lagrange multiplier:

$$\mathbf{s} = \mathbf{C}_{ss} (\mathbf{C}_{ss} + \mathbf{C}_{nn})^{-1} \mathbf{l}. \quad (3)$$

The signal  $\mathbf{s}_p$  can also be calculated at unknown points using:

$$\mathbf{s}_p = \mathbf{C}_{ps} (\mathbf{C}_{ss} + \mathbf{C}_{nn})^{-1} \mathbf{l}. \quad (4)$$

These equations are now depending on the observational vector  $\mathbf{l}$  as well as the covariance matrices  $\mathbf{C}_{ss}$  and  $\mathbf{C}_{nn}$  and the cross-covariance matrix  $\mathbf{C}_{ps}$ , which have to be defined in the following.

The covariance matrix of the signal is:

$$\mathbf{C}_{ss} = \begin{bmatrix} K_{P_1, P_1} & K_{P_1, P_2} & \cdots & K_{P_1, P_{N_t}} \\ K_{P_2, P_1} & K_{P_2, P_2} & \cdots & K_{P_2, P_{N_t}} \\ \vdots & \vdots & \ddots & \vdots \\ K_{P_{N_t}, P_1} & K_{P_{N_t}, P_2} & \cdots & K_{P_{N_t}, P_{N_t}} \end{bmatrix} \quad (5)$$

with  $K_{P_i, P_j}$  as the covariance functions, and  $P_i$  and  $P_j$  being the locations of the  $N_t$  observations. The covariance matrix of the observational noise can be determined using:

$$\mathbf{C}_{nn} = \begin{bmatrix} m_{P_1, P_1}^2 & 0 & \cdots & 0 \\ 0 & m_{P_2, P_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & m_{P_{N_t}, P_{N_t}}^2 \end{bmatrix} \quad (6)$$

with the observational uncertainty  $\mathbf{m}$  given at each observation point  $P_i$ . The off-diagonal parts of  $\mathbf{C}_{nn}$  are zero as no correlation between the observational uncertainty at different locations is assumed. Similar to the variance–covariance matrix of the signal  $\mathbf{C}_{ss}$  at the observation points, the signal cross-covariance matrix between the signal  $\mathbf{s}$  at the observation points  $P_i$  and the signal  $\mathbf{s}_p$  at  $M_t$  new points  $Q_i$  can be written as:

$$\mathbf{C}_{ps} = \begin{bmatrix} K_{Q_1, P_1} & K_{Q_1, P_2} & \cdots & K_{Q_1, P_{N_t}} \\ K_{Q_2, P_1} & K_{Q_2, P_2} & \cdots & K_{Q_2, P_{N_t}} \\ \vdots & \vdots & \ddots & \vdots \\ K_{Q_{M_t}, P_1} & K_{Q_{M_t}, P_2} & \cdots & K_{Q_{M_t}, P_{N_t}} \end{bmatrix}, \quad (7)$$

with  $K_{Q_i, P_j}$  as the covariance functions at  $P_j$  and  $Q_i$ . Similarly, the covariance matrix at the new points  $\mathbf{C}_{pp}$  can be written as:

$$\mathbf{C}_{pp} = \begin{bmatrix} K_{Q_1, Q_1} & K_{Q_1, Q_2} & \cdots & K_{Q_1, Q_{M_t}} \\ K_{Q_2, Q_1} & K_{Q_2, Q_2} & \cdots & K_{Q_2, Q_{M_t}} \\ \vdots & \vdots & \ddots & \vdots \\ K_{Q_{M_t}, Q_1} & K_{Q_{M_t}, Q_2} & \cdots & K_{Q_{M_t}, Q_{M_t}} \end{bmatrix}, \quad (8)$$

with  $K_{Q_i, Q_j}$  being the covariance functions at the locations  $Q_i$  and  $Q_j$  of the  $M_t$  new points. All covariance matrices and cross-covariance matrices have to be positive definite.

The advantage of the LSC techniques is the estimation of a standard error, which is based on the covariance matrices

$\mathbf{C}_{ss}$ ,  $\mathbf{C}_{nn}$  and  $\mathbf{C}_{pp}$ , and the cross-covariance matrix  $\mathbf{C}_{ps}$ . The error covariance matrices of the estimated signals  $\mathbf{s}$  and  $\mathbf{s}_p$  are:

$$\mathbf{E}_s = \mathbf{C}_{ss} - \mathbf{C}_{ss} (\mathbf{C}_{ss} + \mathbf{C}_{nn})^{-1} \mathbf{C}_{ss}, \quad (9)$$

$$\mathbf{E}_{s_p} = \mathbf{C}_{pp} - \mathbf{C}_{ps} (\mathbf{C}_{ss} + \mathbf{C}_{nn})^{-1} \mathbf{C}_{ps}^T, \quad (10)$$

respectively. The standard error is then the square root of the diagonal terms of the error covariance matrices. The estimation of the signal and their associated standard errors depends only on the covariance function  $K(d)$  and the observations  $\mathbf{l}$  and associated uncertainties  $\mathbf{m}$ . In the next step, the covariance function  $K(d)$  has to be defined, which is obtained from a covariance analysis.

## 2.1 Covariance analysis

The covariance analysis is used to find parameters of the covariance function, which is required to estimate the entries of the covariance as well as cross-covariance matrices to find the signal and their standard errors. In a first step, empirical variance and covariances are estimated using observations  $\mathbf{l}$ :

$$K(0) = \frac{1}{N_t} \sum_{i=1}^{N_t} l_i^2, \quad (11)$$

$$K(d_p) = \frac{1}{N_p - 1} \sum_{i \neq j}^{N_p} l_i l_j, \quad (12)$$

with  $N_t$  as the total number of observations,  $N_p$  as the number of point pairs with mutual distance within specified distance intervals, and  $d_p$  as the distance between two points  $P_i$  and  $P_j$ . The observations must be reduced by the mean value beforehand. While the first equation (Eq. (11)) is the variance of the observations only, the second equation (Eq. (12)) requires the classification of observation points within specific distance groups (e.g. Goudarzi et al. 2015). Each distance group is defined by an interval of  $(2p-3)\delta$  and  $(2p-1)\delta$ , where  $\delta$  is a user-specified value, and the first interval ( $p = 1$ ) is defined as  $0 \leq d_p \leq \delta$  (a smaller interval is obtained).  $P$  intervals are created ( $p = 1, 2, \dots, P$ ), and the distance between each point pair is distributed to the specific interval. Thus, the empirical covariance  $K(d_p)$  is calculated for each distance group based on observations of station pairs that belong to the specified interval.  $\delta$  depends on the size of the study area and various values should be tested. The uncertainty of the empirical covariances is obtained by calculating a standard deviation for each distance group. The number of station pairs  $N_p$  per distance group is an important parameter to obtain a statistically meaningful empirical covariance estimate. An empirical covariance obtained from a small number

of station pairs  $N_p$  might stand out significantly from others and needs to be neglected before a covariance function is obtained. The obtained empirical covariances are then plotted together with the variance from Eq. (11) over the distance.

Some stations have observational values that are much larger or smaller than the other observations. Those are defined as non-standard values, more precisely when the value is greater than three times the value of the standard deviation of the entire observational dataset (three-sigma rule; e.g. Pukelsheim 1994):

$$|l_j| > 3\sqrt{\frac{\sum_{i=1}^{N_t} l_i^2}{N_t}} \quad (13)$$

Non-standard values are values, which are affected by a local change that is not correlated with the movement of neighbouring values or isolated values located in a significant deformation zone. In practice, only very few values can be considered out of range and rejected for estimating the covariance function.

The empirical covariances can now be used to find a covariance function  $K(d)$  and their associated parameters. A function is fitted to the empirical covariances, which needs to fulfil the following conditions (Peter 2000) to obtain a positive-definite matrix:

$$\begin{aligned} \text{(i): } & K = K(d) \Rightarrow K \text{ is only a function of the distance } d, \\ \text{(ii): } & \lim_{d \rightarrow \infty} K(d) = 0, \\ \text{(iii): } & \left. \frac{\partial K}{\partial d} \right|_{d=0} = \text{finite}. \end{aligned} \quad (14)$$

Various covariance functions  $K(d)$  are used in the literature, while the Hirvonen function (Hirvonen 1962):

$$K(d) = C_0 \frac{d_0^2}{d_0^2 + d^2}, \quad (15)$$

and the functions of the first- and second-order Gauss–Markov process (e.g. Moritz 1978):

$$K(d) = C_0 \cdot e^{-\frac{d}{d_0}} \quad (16)$$

and

$$K(d) = C_0 \cdot e^{-\frac{d^2}{d_0^2}}, \quad (17)$$

respectively, are commonly applied, with  $d$  being the distance. The correlation length  $d_0$  and the variance  $C_0$  at  $d = 0$  are the unknown parameters, which have to be estimated using the covariance analysis. The covariance function is then only a function of the distance  $d$  between the points  $P_i$  and

$P_j$ . The covariance function  $K(d)$  together with correlation length  $d_0$  and variance  $C_0$  are then used to find entries for covariance and cross-covariance matrices.

The covariance analysis to find a covariance function is similar to the covariance modelling within Kriging to find a semivariogram (e.g. Cressie 2007). Similarly, a correlation length and a covariance value  $C_0$  (named sill) are identified using various functions, which are called models within the Kriging process. Typically, the functions of the first- and second-order Gauss–Markov process are also used in the covariance modelling (e.g. Bogusz et al. 2014; Fuhrmann 2016; Robin et al. 2020).

The described covariance analysis can be done for any observational dataset; however, the covariance analysis cannot be done jointly for two or more components (e.g. horizontal velocity components). In this case, each component has to be analysed separately. This excludes the possible correlation between the components. Thus, a new iterative scheme is presented in Sect. 3.2, which allows the estimation of empirical covariances for both horizontal velocity components measured by GNSS stations.

### 3 Least-squares collocation of velocity data

At a point  $P_i$ , the horizontal velocity field can be described by a velocity and observational uncertainty in the east direction ( $V_E(P_i)$  and  $m_E(P_i)$ , respectively) and by a velocity and observational uncertainty in the north direction ( $V_N(P_i)$  and  $m_N(P_i)$ , respectively). The observations can be divided into a signal ( $W_E(P_i)$  and  $W_N(P_i)$ ) and a noise ( $\eta_E(P_i)$  and  $\eta_N(P_i)$ ) following Eq. (1):

$$\begin{aligned} V_E(P_i) &= W_E(P_i) + \eta_E(P_i) \\ V_N(P_i) &= W_N(P_i) + \eta_N(P_i). \end{aligned} \quad (18)$$

As the observations and their uncertainties are obtained at several observation points ( $i = 1, 2, \dots, g$ ), we can also write the observations, uncertainties, signal, and noise as vectors:

$$\begin{aligned} \mathbf{V}(P) &= [V_E(P_1), V_N(P_1), \dots, V_E(P_g), V_N(P_g)]^T \\ \mathbf{m}(P) &= [m_E(P_1), m_N(P_1), \dots, m_E(P_g), m_N(P_g)]^T \\ \mathbf{W}(P) &= [W_E(P_1), W_N(P_1), \dots, W_E(P_g), W_N(P_g)]^T \\ \boldsymbol{\eta}(P) &= [\eta_E(P_1), \eta_N(P_1), \dots, \eta_E(P_g), \eta_N(P_g)]^T. \end{aligned} \quad (19)$$

From this follows:

$$\mathbf{V}(P) = \mathbf{W}(P) + \boldsymbol{\eta}(P). \quad (20)$$

We can estimate the signal  $\mathbf{W}$  at known points  $P_i$  (see Eq. (3); Legrand 2007):

$$\mathbf{W}(P) = \mathbf{C}_{W(P)W(P)} [\mathbf{C}_{W(P)W(P)} + \mathbf{C}_{\eta(P)\eta(P)}]^{-1} \mathbf{V}(P), \quad (21)$$

and unknown points  $Q_i$  (see Eq. (4); Legrand 2007):

$$\mathbf{W}(Q) = \mathbf{C}_{W(Q)W(P)} [\mathbf{C}_{W(P)W(P)} + \mathbf{C}_{\eta(P)\eta(P)}]^{-1} \mathbf{V}(P), \quad (22)$$

using equations from standard LSC. The error covariance matrices of the obtained signal follow from Eq. (9) (Legrand 2007):

$$\mathbf{E}_{W(P)} = \mathbf{C}_{W(P)W(P)} - \mathbf{C}_{W(P)W(P)} [\mathbf{C}_{W(P)W(P)} + \mathbf{C}_{\eta(P)\eta(P)}]^{-1} \mathbf{C}_{W(P)W(P)}, \quad (23)$$

and from Eq. (10) (Legrand 2007):

$$\mathbf{E}_{W(Q)} = \mathbf{C}_{W(Q)W(Q)} - \mathbf{C}_{W(Q)W(P)} [\mathbf{C}_{W(P)W(P)} + \mathbf{C}_{\eta(P)\eta(P)}]^{-1} \mathbf{C}_{W(Q)W(P)}^T, \quad (24)$$

for the known and unknown points, respectively. The equations above contain the covariance matrices  $\mathbf{C}_{W(P)W(P)}$ ,  $\mathbf{C}_{W(Q)W(Q)}$  and  $\mathbf{C}_{\eta(P)\eta(P)}$  as well as the cross-covariance matrix  $\mathbf{C}_{W(Q)W(P)}$ , which need to be defined in the following.

The covariance matrix of the noise can be expressed as expected value of the noise at known points  $P$ :

$$\mathbf{C}_{\eta(P)\eta(P)} = \mathbb{E} [\boldsymbol{\eta}(P) \boldsymbol{\eta}(P)^T]. \quad (25)$$

The covariance matrix of the noise is estimated using the observational uncertainties  $\mathbf{m}(P)$ , and as these uncertainties are provided for both horizontal velocity components separately, Eq. (25) has to become:

$$\begin{aligned} \mathbf{C}_{\eta(P_i)\eta(P_j)} &= \mathbb{E} [\boldsymbol{\eta}(P_i) \boldsymbol{\eta}(P_j)^T] \\ &= \begin{bmatrix} \mathbb{E} [\eta_E(P_i) \eta_E(P_j)] & \mathbb{E} [\eta_E(P_i) \eta_N(P_j)] \\ \mathbb{E} [\eta_N(P_i) \eta_E(P_j)] & \mathbb{E} [\eta_N(P_i) \eta_N(P_j)] \end{bmatrix}, \end{aligned} \quad (26)$$

with  $\eta_E(P_{i,j})$  and  $\eta_N(P_{i,j})$  as the noise of the east-west and north-south component, respectively. A correlation of the observational uncertainty between different velocity components is assumed to be non-existent, which leaves diagonal elements in the covariance matrix:

$$\mathbf{C}_{\eta(P_i)\eta(P_j)} = \begin{bmatrix} \mathbb{E} [\eta_E(P_i) \eta_E(P_j)] & 0 \\ 0 & \mathbb{E} [\eta_N(P_i) \eta_N(P_j)] \end{bmatrix}. \quad (27)$$

Similarly, the covariance matrix of the signal can be obtained for the known points  $(P_{i,j})$ :

$$\begin{aligned} \mathbf{C}_{W(P_i)W(P_j)} &= \mathbb{E} [W(P_i) W(P_j)^T] \\ &= \begin{bmatrix} \mathbb{E} [W_E(P_i) W_E(P_j)] & \mathbb{E} [W_E(P_i) W_N(P_j)] \\ \mathbb{E} [W_N(P_i) W_E(P_j)] & \mathbb{E} [W_N(P_i) W_N(P_j)] \end{bmatrix}, \end{aligned} \quad (28)$$

the cross-covariance matrix of the signal for the known  $(P_{i,j})$  and unknown  $(Q_{i,j})$  points:

$$\begin{aligned} \mathbf{C}_{W(Q_i)W(P_j)} &= \mathbb{E} [W(Q_i) W(P_j)^T] \\ &= \begin{bmatrix} \mathbb{E} [W_E(Q_i) W_E(P_j)] & \mathbb{E} [W_E(Q_i) W_N(P_j)] \\ \mathbb{E} [W_N(Q_i) W_E(P_j)] & \mathbb{E} [W_N(Q_i) W_N(P_j)] \end{bmatrix}, \end{aligned} \quad (29)$$

and the covariance matrix of the signal at the unknown points  $(Q_{i,j})$  only:

$$\begin{aligned} \mathbf{C}_{W(Q_i)W(Q_j)} &= \mathbb{E} [W(Q_i) W(Q_j)^T] \\ &= \begin{bmatrix} \mathbb{E} [W_E(Q_i) W_E(Q_j)] & \mathbb{E} [W_E(Q_i) W_N(Q_j)] \\ \mathbb{E} [W_N(Q_i) W_E(Q_j)] & \mathbb{E} [W_N(Q_i) W_N(Q_j)] \end{bmatrix}. \end{aligned} \quad (30)$$

However, the horizontal velocity components are depending on each other and correlate due to the movement of the plates on a sphere. Thus, the non-diagonal term cannot be excluded here.

The expected value  $\mathbb{E}$  can be described by covariance functions  $K_{kl}(k, l = E, N$  with  $E$  being the east-west and  $N$  north-south components), which depend on the spherical distance between the points considered as well as the variance of the velocities:

$$\begin{aligned} \mathbb{E} [W_k(I) W_l(J)] &= K_{kl}(I, J) \\ &= K_{kl}(d_{ij}) \quad (k, l = E, N; I, J = P_i, P_j) \end{aligned}$$

with  $d_{ij}$  as the spherical distance between the points  $P_{i,j}$  and  $Q_{i,j}$ . Thus, Eqs. (28–30) can also be expressed as:

$$\mathbf{C}_{W(P_i)W(P_j)} = \begin{bmatrix} K_{EE}(P_i, P_j) & K_{EN}(P_i, P_j) \\ K_{NE}(P_i, P_j) & K_{NN}(P_i, P_j) \end{bmatrix}, \quad (31)$$

$$\mathbf{C}_{W(Q_i)W(P_j)} = \begin{bmatrix} K_{EE}(Q_i, P_j) & K_{EN}(Q_i, P_j) \\ K_{NE}(Q_i, P_j) & K_{NN}(Q_i, P_j) \end{bmatrix}, \quad (32)$$

$$\mathbf{C}_{W(Q_i)W(Q_j)} = \begin{bmatrix} K_{EE}(Q_i, Q_j) & K_{EN}(Q_i, Q_j) \\ K_{NE}(Q_i, Q_j) & K_{NN}(Q_i, Q_j) \end{bmatrix}. \quad (33)$$



It is therefore necessary to define four different covariance functions:  $K_{EE}$ ,  $K_{EN}$ ,  $K_{NE}$ , and  $K_{NN}$  (Legrand 2007). Suppose a rigid movement of the whole plate, the interpolation of the resulting velocity field has to correspond to a rigid movement as well. To achieve this, we need to consider the geometry of the movement and express the covariance matrices  $\mathbf{C}_{W(I)W(J)}$  (with  $I$  and  $J$  as elements of  $P$  and  $Q$ ) as covariance matrices of the angular velocity field, which contains the rigid movement of the plate. It is therefore necessary to express the horizontal velocities as a function of the angular velocities. Thus, the interpolation is applied over the angular velocity field rather than over the horizontal velocity field itself. This allows: (1) to preserve the characteristics of the spherical motion of the velocity field, and (2) to define a covariance function which is independent from the coordinate system and thus remains valid over large areas, i.e. areas larger than  $10\text{--}15^\circ$  for which the curvature of the Earth has an effect (Haines and Holt 1993). In this case, the correlation of the horizontal velocity field components is based on the movement of the plates on the sphere. This is considered when the interpolation is done on a sphere by using the angular velocity. The interpolation of the angular velocity field is also used by Kreemer et al. (2000, 2003), where the angular velocity field is interpolated over a spherical Earth using bi-cubic Bessel splines. This new method (in the following named HV (horizontal velocity)-LSC), which considers the spatial correlation between the velocities, is based on a LSC method on the sphere. The vertical velocity component is assumed to be uncorrelated with the horizontal velocity components.

### 3.1 Covariance matrices of the horizontal velocity field

A point  $P_i$  can be described by its Cartesian coordinates  $xyz_i = [x_i, y_i, z_i]^T$ , its geographical coordinates (longitude  $\lambda_i$  and latitude  $\phi_i$ ), its velocities in the Cartesian coordinate system  $V_{xyz_i} = [V_{x_i}, V_{y_i}, V_{z_i}]^T$ , its local horizontal velocities  $V_{EN_i} = [V_{E_i}, V_{N_i}]^T$ , and its angular velocity  $\omega_{xyz_i} = [\omega_{x_i}, \omega_{y_i}, \omega_{z_i}]^T$ . The height of point  $P_i$  is assumed to be the same for all observation points and is not considered further. The Cartesian coordinates on a unit sphere are written as:

$$xyz_i = \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} \cos \lambda_i \cos \phi_i \\ \sin \lambda_i \cos \phi_i \\ \sin \phi_i \end{bmatrix}. \quad (34)$$

The local horizontal velocity field can be obtained from the velocity expressed in global coordinates by using a transformation matrix  $R_{xyz_i \rightarrow EN_i}$  for the horizontal components only:

$$V_{EN_i} = R_{xyz_i \rightarrow EN_i} V_{xyz_i}, \quad (35)$$

which can be written as:

$$R_{xyz_i \rightarrow EN_i} = \begin{bmatrix} -\sin \lambda_i & \cos \lambda_i & 0 \\ -\sin \phi_i \cos \lambda_i & -\sin \phi_i \sin \lambda_i & \cos \phi_i \end{bmatrix}. \quad (36)$$

The Cartesian velocity field  $V_{xyz_i}$  can be additionally expressed in terms of the angular velocity field  $\mathbf{V} = \boldsymbol{\omega} \times \mathbf{xyz}_i$ :

$$\begin{aligned} V_{xyz_i} &= \begin{bmatrix} \omega_{x_i} \\ \omega_{y_i} \\ \omega_{z_i} \end{bmatrix} \times \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} \\ &= \begin{bmatrix} 0 & z_i & -y_i \\ -z_i & 0 & x_i \\ y_i & -x_i & 0 \end{bmatrix} \begin{bmatrix} \omega_{x_i} \\ \omega_{y_i} \\ \omega_{z_i} \end{bmatrix} \\ &= \begin{bmatrix} 0 & \sin \phi_i & -\sin \lambda_i \cos \phi_i \\ -\sin \phi_i & 0 & \cos \lambda_i \cos \phi_i \\ \sin \lambda_i \cos \phi_i & -\cos \lambda_i \cos \phi_i & 0 \end{bmatrix} \begin{bmatrix} \omega_{x_i} \\ \omega_{y_i} \\ \omega_{z_i} \end{bmatrix}, \end{aligned} \quad (37)$$

which depends now on the angular velocity and the geographical location of the point  $P_i$ . This can be combined with Eq. (35) to obtain the horizontal velocity field:

$$\begin{aligned} V_{EN_i} &= R_{xyz_i \rightarrow EN_i} \omega_{xyz_i} \times xyz_i \\ &= A_{P_i} \omega_i, \end{aligned} \quad (38)$$

with the matrix  $A_{P_i}$  defined as:

$$\begin{aligned} A_{P_i} &= \underbrace{\begin{bmatrix} -\sin \lambda_i & \cos \lambda_i & 0 \\ -\sin \phi_i \cos \lambda_i & -\sin \phi_i \sin \lambda_i & \cos \phi_i \end{bmatrix}}_{\text{from eq. (36)}} \\ &\quad \underbrace{\begin{bmatrix} 0 & \sin \phi_i & -\sin \lambda_i \cos \phi_i \\ -\sin \phi_i & 0 & \cos \lambda_i \cos \phi_i \\ \sin \lambda_i \cos \phi_i & -\cos \lambda_i \cos \phi_i & 0 \end{bmatrix}}_{\text{from eq. (37)}} \\ &= \begin{bmatrix} -\sin \phi_i \cos \lambda_i & -\sin \phi_i \sin \lambda_i & \cos \phi_i \\ \sin \lambda_i & -\cos \lambda_i & 0 \end{bmatrix}. \end{aligned} \quad (39)$$

The signal  $W(P_i)$  can now be expressed at the observation points  $P_i$ :

$$W(P_i) = A_{P_i} \omega(P_i). \quad (40)$$

Similarly, the signal  $W(Q_i)$  is:

$$W(Q_i) = A_{Q_i} \omega(Q_i) \quad (41)$$

at the new points  $Q_i$ . These two equations can be used to find covariance matrices for the HV-LSC. The covariance matrix  $\mathbf{C}_{W(P)W(P)}$  can be written as follows:

$$\begin{aligned}
C_{W(P_i) W(P_j)} &= \mathbb{E} \left[ W(P_i) W(P_j)^T \right] \\
&= \mathbb{E} \left[ A_{P_i} \omega(P_i) (A_{P_j} \omega(P_j))^T \right] \\
&= \mathbb{E} \left[ A_{P_i} \omega(P_i) \omega(P_j)^T A_{P_j}^T \right] \\
&= A_{P_i} \mathbb{E} \left[ \omega(P_i) \omega(P_j)^T \right] A_{P_j}^T. \quad (42)
\end{aligned}$$

The term in the middle ( $\mathbb{E} [\omega(P_i) \omega(P_j)^T]$ ) is the covariance matrix of the angular velocity field  $C_{\omega(P_i) \omega(P_j)}$ :

$$\begin{aligned}
C_{W(P_i) W(P_j)} &= \begin{bmatrix} -\sin \phi_{P_i} \cos \lambda_{P_i} & -\sin \phi_{P_i} \sin \lambda_{P_i} & \cos \phi_{P_i} \\ \sin \lambda_{P_i} & -\cos \lambda_{P_i} & 0 \end{bmatrix} \begin{bmatrix} -\sin \phi_{P_j} \cos \lambda_{P_j} & \sin \lambda_{P_j} \\ -\sin \phi_{P_j} \sin \lambda_{P_j} & -\cos \lambda_{P_j} \\ \cos \phi_{P_j} & 0 \end{bmatrix} K(P_i, P_j) \\
&= \begin{bmatrix} \sin \phi_{P_i} \sin \phi_{P_j} \cos(\lambda_{P_i} - \lambda_{P_j}) + \cos \phi_{P_i} \cos \phi_{P_j} \sin \phi_{P_i} \sin(\lambda_{P_i} - \lambda_{P_j}) \\ \sin \phi_{P_j} \sin(\lambda_{P_j} - \lambda_{P_i}) \\ \cos(\lambda_{P_i} - \lambda_{P_j}) \end{bmatrix} K(P_i, P_j). \quad (47)
\end{aligned}$$

$$\begin{aligned}
C_{\omega(P_i) \omega(P_j)} &= \begin{bmatrix} \mathbb{E}[\omega_x(P_i) \omega_x(P_j)] & \mathbb{E}[\omega_x(P_i) \omega_y(P_j)] & \mathbb{E}[\omega_x(P_i) \omega_z(P_j)] \\ \mathbb{E}[\omega_y(P_i) \omega_x(P_j)] & \mathbb{E}[\omega_y(P_i) \omega_y(P_j)] & \mathbb{E}[\omega_y(P_i) \omega_z(P_j)] \\ \mathbb{E}[\omega_z(P_i) \omega_x(P_j)] & \mathbb{E}[\omega_z(P_i) \omega_y(P_j)] & \mathbb{E}[\omega_z(P_i) \omega_z(P_j)] \end{bmatrix}. \quad (43)
\end{aligned}$$

In the following, we assume that there is no correlation between the components  $\omega_x(P_i)$ ,  $\omega_y(P_i)$ , and  $\omega_z(P_i)$  of the vector  $\omega$ . This is a rather strong assumption that is needed to obtain a solution for Eq. (42). However, as the neglect of the correlation between the north-south and east-west components in standard LSC is an even stronger assumption, our HV-LSC approach still represents a step forward compared to standard LSC. In addition, we assume that there is no correlation between two different components of the vectors  $\omega(P_i)$  and  $\omega(P_j)$  for  $i \neq j$ , and due to revolution symmetry follows  $\mathbb{E}[\omega_x(P_i) \omega_x(P_j)] = \mathbb{E}[\omega_y(P_i) \omega_y(P_j)] = \mathbb{E}[\omega_z(P_i) \omega_z(P_j)]$ . We can therefore write  $C_{\omega(P_i) \omega(P_j)}$  for  $i = j$  and  $i \neq j$  with  $K(P_i, P_j) = \mathbb{E}[\omega_x(P_i) \omega_x(P_j)] = \mathbb{E}[\omega_y(P_i) \omega_y(P_j)] = \mathbb{E}[\omega_z(P_i) \omega_z(P_j)]$  as:

$$\begin{aligned}
C_{\omega(P_i) \omega(P_j)} &= \begin{bmatrix} K(P_i, P_j) & 0 & 0 \\ 0 & K(P_i, P_j) & 0 \\ 0 & 0 & K(P_i, P_j) \end{bmatrix} \\
&= K(P_i, P_j) I_3. \quad (44)
\end{aligned}$$

Similarly follows for the covariance matrix  $C_{\omega(Q_i) \omega(Q_j)}$  and the cross-covariance matrix  $C_{\omega(Q_i) \omega(P_j)}$ :

$$\begin{aligned}
C_{\omega(Q_i) \omega(P_j)} &= K(Q_i, P_j) I_3 \\
C_{\omega(Q_i) \omega(Q_j)} &= K(Q_i, Q_j) I_3. \quad (45)
\end{aligned}$$

The covariance matrices are now depending only on one covariance function  $K(I, J)$  with  $I, J = P_i, Q_j$ . This can be now substituted into Eq. (42):

$$\begin{aligned}
C_{W(P_i) W(P_j)} &= A_{P_i} \mathbb{E} \left[ \omega(P_i) \omega(P_j)^T \right] A_{P_j}^T \\
&= A_{P_i} C_{\omega(Q_i) \omega(P_j)} A_{P_j}^T \\
&= A_{P_i} K(P_i, P_j) I_3 A_{P_j}^T \\
&= A_{P_i} A_{P_j}^T K(P_i, P_j), \quad (46)
\end{aligned}$$

which becomes:

by applying Eq. (39). In addition, we know that the covariance matrix  $C_{W(P_i) W(P_j)}$  is also a function of the four covariance functions  $K_{EE}$ ,  $K_{NN}$ ,  $K_{EN}$ , and  $K_{NE}$  (see Eq. (31)):

$$C_{W(P_i) W(P_j)} = \begin{bmatrix} K_{EE}(P_i, P_j) & K_{EN}(P_i, P_j) \\ K_{NE}(P_i, P_j) & K_{NN}(P_i, P_j) \end{bmatrix}. \quad (48)$$

Thus, we can define the functions  $f_{EE}$ ,  $f_{EN}$ ,  $f_{NE}$ , and  $f_{NN}$  so that:

$$\begin{aligned}
C_{W(P_i) W(P_j)} &= \mathbb{E} \left[ W(P_i) W(P_j)^T \right] \\
&= f_{kl}(P_i, P_j) K(P_i, P_j); k, l = E, N, \quad (49)
\end{aligned}$$

with:

$$\begin{aligned}
f_{EE}(P_i, P_j) &= \sin \phi_{P_i} \sin \phi_{P_j} \cos(\lambda_{P_i} - \lambda_{P_j}) \\
&\quad + \cos \phi_{P_i} \cos \phi_{P_j} \quad (50)
\end{aligned}$$

$$f_{EN}(P_i, P_j) = \sin \phi_{P_i} \sin(\lambda_{P_i} - \lambda_{P_j}) \quad (51)$$

$$f_{NE}(P_i, P_j) = \sin \phi_{P_j} \sin(\lambda_{P_j} - \lambda_{P_i}) \quad (52)$$

$$f_{NN}(P_i, P_j) = \cos(\lambda_{P_i} - \lambda_{P_j}), \quad (53)$$

obtained from the matrix components in Eq. (47). The four different covariance functions are now replaced by one covariance function and an additional term, which depends on the geographical coordinates and varies for each component combination. Similar equations can be obtained for the cross-covariance and covariance matrices  $C_{W(Q_i) W(P_j)}$  and  $C_{W(Q_i) W(Q_j)}$ , respectively:

$$C_{W(Q_i) W(P_j)} = \mathbb{E} \left[ W(Q_i) W(P_j)^T \right]$$

$$= f_{kl}(Q_i, P_j) K(Q_i, P_j); \quad k, l = E, N, \quad (54)$$

$$\begin{aligned} C_{W(Q_i) W(Q_j)} &= \mathbb{E} [W(Q_i) W(Q_j)^T] \\ &= f_{kl}(Q_i, Q_j) K(Q_i, Q_j); \quad k, l = E, N. \end{aligned} \quad (55)$$

We can now find the (cross-)covariance matrices in order to estimate the signal at observational points as well as at new points, and in both cases the signal is estimated for both horizontal velocity field components at the same time and the correlation between the components is even taken into account. In a next step, the covariance function  $K(P_i, P_j)$  has to be defined.

### 3.2 Covariance analysis of the horizontal velocity field

Two points  $P_i$  and  $P_j$  are separated by a spherical distance  $d_{ij}$ , and  $\hat{K}(d_{ij})$  is an estimator of the covariance function  $K(P_i, P_j)$  for this distance. The point  $P_i$  is characterized by a velocity  $V_k(P_i)$ , which can be divided into a signal  $W_k(P_i)$  and noise  $\eta_k(P_i)$ :

$$V_k(P_i) = W_k(P_i) + \eta_k(P_i); \quad k = E(ast), N(orth). \quad (56)$$

The expectation of the signal at the two points is then:

$$\begin{aligned} \mathbb{E} [W_k(P_i) W_l(P_j)] &= f_{kl}(P_i, P_j) K(P_i, P_j); \\ k, l &= E, N, \end{aligned} \quad (57)$$

with the functions  $f_{kl}(P_i, P_j)$  as in Eqs. (50–53).

For the sake of simplicity, we introduce the following notations:  $s$  for the component  $k$  at the point  $P_i$  and  $t$  for the component  $l$  at the point  $P_j$ . Thus, we have the following equations:

$$\begin{aligned} V_s &= V_k(P_i); \quad W_s = W_k(P_i); \\ \eta_s &= \eta_k(P_i) \Rightarrow V_s = W_s + \eta_s \\ V_t &= V_l(P_j); \quad W_t = W_l(P_j); \\ \eta_t &= \eta_l(P_j) \Rightarrow V_t = W_t + \eta_t. \end{aligned} \quad (58)$$

Similarly,  $f_{kl}(P_i, P_j)$  is equivalent to  $f_{st}$ . The expectation of the observation can then be written as:

$$\begin{aligned} \mathbb{E} [V_s V_t] &= \mathbb{E} [(W_s + \eta_s) (W_t + \eta_t)] \\ &= \mathbb{E} [W_s W_t + \eta_t W_s + \eta_s W_t + \eta_s \eta_t] \\ &= \mathbb{E} [W_s W_t] + \mathbb{E} [\eta_t W_s] + \mathbb{E} [\eta_s W_t] + \mathbb{E} [\eta_s \eta_t], \end{aligned} \quad (59)$$

which becomes:

$$\mathbb{E} [V_s V_t] = f_{st} K(st) + C_{st} \quad (60)$$

with  $\mathbb{E} [W_s W_t] = f_{st} K(st)$  (based on the notations introduced before) and  $C_{st}$  as the covariance between the noises  $\eta_s$  and  $\eta_t$  ( $C_{st} = \mathbb{E} [\eta_s \eta_t]$ ). The two terms in the middle are zero ( $\mathbb{E} [\eta_t W_s] = 0$  and  $\mathbb{E} [\eta_s W_t] = 0$ ) as signal and noise are independent of each other (e.g. Moritz 1980). We can now obtain an expression for the covariance function:

$$K(st) = \frac{\mathbb{E} [V_s V_t] - C_{st}}{f_{st}} = \mathbb{E} \left[ \frac{V_s V_t - C_{st}}{f_{st}} \right], \quad (61)$$

which can be used in the following to determine the estimator of the covariance function.

The estimator of the covariance function  $K$  at a distance  $d_p$  is of the form:

$$\hat{K}(d_p) = \sum_{d_{st} \in d_p} \alpha_{st} \frac{V_s V_t - C_{st}}{f_{st}}, \quad \sum_{d_{st} \in d_p} \alpha_{st} = 1 \quad (62)$$

with  $\alpha_{st}$  as the weighting coefficient and summation over all distance pairs  $d_{st}$  within the distance group defined by the distance range  $d_p$  (see Sect. 2.1). The empirical covariances  $\hat{K}(d_p)$  have to be defined for each distance group with  $N_p$  as the number of points per groups (see Sect. 2.1). Equation (62) can be simplified by setting:

$$\beta_{st} = \frac{V_s V_t - C_{st}}{f_{st}}, \quad (63)$$

and we obtain:

$$\hat{K}(d_p) = \sum_{d_{st} \in d_p} \alpha_{st} \beta_{st}, \quad \sum_{d_{st} \in d_p} \alpha_{st} = 1. \quad (64)$$

The empirical covariances are now a function of the parameters  $\alpha_{st}$ , which need to be defined in the following, and  $\beta_{st}$ .  $\beta_{st}$  can be determined by applying Eq. (63) and using the observations, the covariances of the observational uncertainties as well as the functions  $f_{st}$  (Eqs. (50–53)). Thus, the  $\beta$  value is calculated for each point pair  $(P_i, P_j)$  four times due to the horizontal velocity field components ( $\beta_{EE}$ ,  $\beta_{EN}$ ,  $\beta_{NE}$ , and  $\beta_{NN}$ ). All  $\beta$  values that belong to a group of points with similar distance  $d_{st}$  are combined in a vector  $\boldsymbol{\beta} = [\beta_{st}]_{d_{st} \in d_p} = [\beta_{11}, \dots, \beta_{st}, \dots, \beta_{N_p N_p}]^T$ .

Similarly, we can define the vector  $\boldsymbol{\alpha} = [\alpha_{st}]_{d_{st} \in d_p} = [\alpha_{11}, \dots, \alpha_{st}, \dots, \alpha_{N_p N_p}]^T$  as well as the matrix  $\boldsymbol{\Sigma}_\beta = [Cov(\beta_{st}, \beta_{uv})]_{(d_{st} \in d_p, d_{uv} \in d_p)}$  as the covariance associated to the vector  $\boldsymbol{\beta}$ , with  $u$  and  $v$  defined in the same way as  $s$  and  $t$  in Eq. (58). The vector  $\boldsymbol{\alpha}$  then becomes:

$$\boldsymbol{\alpha} = \frac{\boldsymbol{\Sigma}_\beta^{-1} \mathbb{1}}{\mathbb{1}^T (\boldsymbol{\Sigma}_\beta^{-1}) \mathbb{1}} \quad (65)$$



which minimizes  $\text{Var}(\widehat{K}(d_p))$  by imposing  $\sum_{d_{st} \in d_p} \alpha_{st} = 1$  (see supplementary material, Eqs. (S3–S10)).  $\mathbf{1}$  is a vector of ones.

The covariance of the vector  $\beta$  is by definition:

$$\Sigma_\beta = [\text{Cov}(\beta_{st}, \beta_{uv})]_{(d_{st} \in d_p, d_{uv} \in d_p)}, \quad (66)$$

with:

$$\text{Cov}(\beta_{st}, \beta_{uv}) = \mathbb{E}[(\beta_{st} - \mathbb{E}[\beta_{st}])(\beta_{uv} - \mathbb{E}[\beta_{uv}])]. \quad (67)$$

This can be simplified to (see supplementary material, Eqs. (S11–S16)):

$$\begin{aligned} \text{Cov}(\beta_{st}, \beta_{uv}) &= \frac{1}{f_{st} f_{uv}} (f_{su} f_{tv} K(su) K(tv) + f_{sv} f_{tu} K(sv) K(tu) \\ &\quad + f_{tv} K(tv) C_{su} + f_{sv} K(sv) C_{tu} + f_{tu} K(tu) C_{sv} \\ &\quad + f_{su} K(su) C_{tv} + C_{su} C_{tv} + C_{sv} C_{tu}), \end{aligned} \quad (68)$$

with the functions  $f_{st}$  and  $f_{uv}$  defined as in Eqs. (50–53) and the covariances of the observational uncertainties as defined in Eq. (6). The function  $K$  is the covariance function and has to be defined before. A-priori estimates of the covariance  $C_0$  and correlation length  $d_0$  have to be used. Thus, the result of  $\text{Cov}(\beta_{st}, \beta_{uv})$  will depend on the covariance function and their parameters chosen in the beginning.

To summarize, we have a velocity field and their associated uncertainties. The spherical distances between stations are estimated for all pairs of points that can be formed. The pairs are then classified by distance and assigned to a distance group. For each distance group, we compute the vector  $\beta$ :

$$\beta = \left[ \beta_{kl}(P_i, P_j) = \frac{V_k(P_i) V_l(P_j) - C_{\eta_k(P_i) \eta_l(P_j)}}{f_{kl}(P_i, P_j)} \right]_{\substack{d_{P_i P_j} \in d_p \\ k, l = E, N}}, \quad (69)$$

and its covariance matrix  $\Sigma_\beta$  (see Eq. (68)). The empirical covariance at distance  $d_p$  is:

$$\widehat{K}(d_p) = \alpha^T \beta, \quad (70)$$

with  $\alpha$  as in Eq. (65). The variance of the empirical covariance has the form:

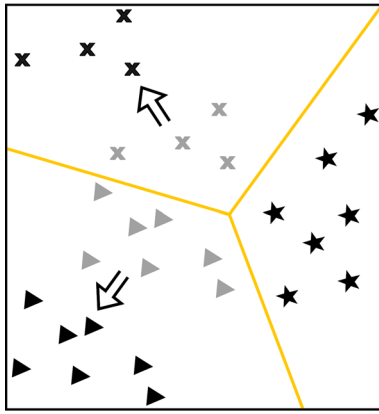
$$S_{\widehat{K}(d_p)}^2 = \frac{1}{\mathbf{1}^T (\Sigma_\beta^{-1})^T \mathbf{1}}. \quad (71)$$

We can now compute the empirical covariance  $\widehat{K}(d_p)$  at distance  $d_p$  and its associated standard deviation  $S_{\widehat{K}(d_p)}$ . The values are then plotted over the distance, and we can obtain the parameters (covariance  $C_0$  and correlation length  $d_0$ ) of the covariance function as described in Sect. 2.1 and done for

a standard covariance analysis. Again, outliers are removed based on the three-sigma rule (see Eq. (13)). As  $\widehat{K}(d_p)$  is a function of the covariance matrix of  $\beta$ , which in turn is a function of the chosen covariance function and their parameters, it is therefore necessary to calculate  $\widehat{K}(d_p)$  by iteration, which is not needed for a standard covariance analysis. However, the new method allows the usage of both horizontal velocity field components and their correlations, as well as the estimation of a realistic standard deviation of the empirical covariances.

## 4 Plate-boundary constraints

The Earth consists of several moving tectonic plates, which lead to diverse GNSS velocities for nearby locations on different tectonic plates, especially if they move in opposing directions. During an interpolation the velocities are usually smoothed and the difference in the observations due to a plate boundary is not visible in the interpolated velocity field. In order to avoid a smoothing of the velocities, it is necessary to define the plate boundaries and perform the interpolation separately. This can be done by increasing the distance between the stations (Fig. 1). Thus the distance of, for example, two stations that are only a few kilometres apart but on different tectonic plates, will be increased to several hundred kilometres so that this distance exceeds the correlation length of the stations on these plates. After that, their velocities will not interfere with each other anymore. To avoid a shift of stations on the same tectonic plate, all stations on one tectonic plate are moved together keeping their original distances to each other (Fig. 1). The locations of the stations that are located on the tectonic plate in the centre (so the one that is surrounded by several other tectonic plates with GNSS stations) are not moved. Stations on different tectonic plates are then moved away from this central tectonic plate (Fig. 1) by a specified distance. To avoid the mixing of stations that had originally a large distance and being on different tectonic plates, the station movement is done iteratively until all stations on different tectonic plates have a large distance to each other. This allows the interpolation of all GNSS stations at the same time, but station velocities on different tectonic plates do not interfere with each other. Thus, the interpolated velocity field can show a strong gradient along a plate boundary. Known tectonic plate boundaries (e.g. Bird 2003) as well as other pre-defined geological boundaries or fault zones can be used. The increase in distance between the stations allows the possibility to include further developments where boundaries are identified by strong gradients (Egli et al. 2007) rather than by pre-defining their locations as it is done now. It is thus also suitable for local studies where a plate boundary or fault is not known but might exist. However, this approach is not valid for a global interpolation as all stations are located on

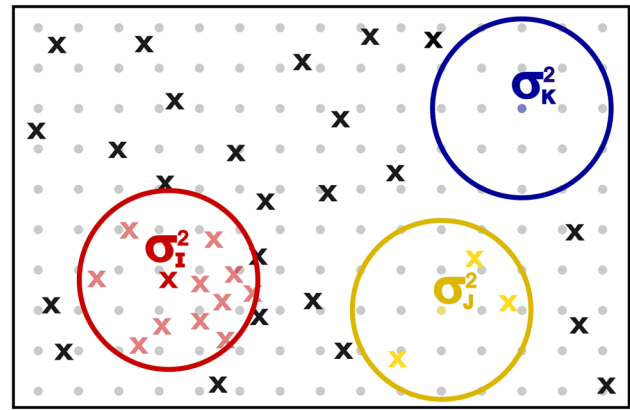


**Fig. 1** Sketch of the implementation of plate-boundary constraints. The map has three plates with stations shown as crosses on the upper plate, triangles on the lower plate and stars on the right plate. The grey symbols show the original location of the stations, while the black symbols are the new locations using plate-boundary constraints. The right plate is kept as central plate and stations on the upper and left plate are moved. The arrows indicate the movement direction of the stations. The plate boundaries are shown as solid yellow lines

plates, which in turn are surrounded by other plates. Thus the iterative scheme will not resolve. The chosen value that is used to increase the distance between adjoining tectonic plates must be larger than the correlation length  $d_0$ . As the covariance function decreases remarkably after the correlation length, interpolated values will approach zero. Thus, the interpolated velocities on one tectonic plate are really free from velocities on different tectonic plates. Therefore, the chosen distance should be determined after the covariance analysis by finding a value where the covariance is approaching zero. While the signal variance and the cross-covariances can vary for different plates, an additional approach needs to include these changes, which is described in the next section.

## 5 Moving variance

The estimation of the parameters of the covariance function provides general information about the observed velocity field. However, when station density and the statistical behaviour of the signal vary significantly in different areas, then the general variance and correlation length are not representative for the entire region. This means that the assumption that the signal covariance function is homogeneous and isotropic is not fulfilled; thus, the estimated covariance depends on the magnitude of the signal and areas with small or opposite velocities are not well reflected in the covariance analysis, if one area has larger velocities. This is known as non-stationarity and several approaches to overcome this problem have been published before (e.g. Tscherning 1999; Darbeheshti and Featherstone 2010).



**Fig. 2** Estimation of the sample variance based on a random set of observations. The crosses show the location of the observations and the dots the location of the new points. The light-red and red crosses are used for the estimation of the sample variance  $\sigma_I^2$  at the observational point  $I$  (marked by the red cross). The circle surrounding point  $I$  is shown by the dark-red circle. The sample variance  $\sigma_J^2$  at the grid point  $J$  (yellow dot) is obtained by the observations (yellow crosses) within the yellow circle. The dark blue circle marks the example where no sample variance can be obtained as no observations (crosses) are within the dark-blue circle. Thus, the sample variance  $\sigma_K^2$  would be equal to  $\rho_0$

Here, we introduce and refine the approach of moving variance that was suggested by Vestøl et al. (2019). Within this approach, the sample variance of an area surrounding the station is estimated by including the station itself and all stations that lie within a circle specified by a radius (Fig. 2):

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n l_i^2, \quad (72)$$

with the observations  $l_i$  within the circle. The moving variance is used together with results from a correlation analysis (see below) instead of a covariance analysis. The observation vector  $\mathbf{I}$  has to be reduced by its mean value beforehand as described in Sect. 2.1.

The correlation at 0 km ( $\rho_0$ ) is equal 1.0 as the correlation with the observations itself is a perfect positive correlation. Thus, the variance–covariance and covariance matrices have a maximum value of 1.0. In order to obtain a similar maximum value after applying moving variance, the observations and uncertainties have to be normed by the maximum value of the absolute observations. Thus the observations have now a maximum or minimum of exactly  $\pm 1.0$ . The sample variance has to be calculated for each station by keeping the same radius of the circle to define the region for the sample-variance calculation. In addition, it is required to define a minimum value ( $n_{\min}$ ) for the number of stations that have to be located within the circle. If less than  $n_{\min}$  observations are within the circle, a fill value has to be defined, which allows the sample-variance calculation despite not having

enough observations (Fig. 2, yellow circle if  $n_{\min}$  is larger than three). A user-defined value can be used. Otherwise, the fill value is estimated based on the mean value of the stations that are located within the circle. The fill value is then equal to the average of mean value and variance of the entire dataset ( $\rho_0$ ; see Eq. (75) below). The latter is the correlation factor at 0 km and thus equal to 1.0. Thus, the fill value is larger than the mean value based on the stations located within the circle and thus the sample variance within this circle is increased and approaches the variance of the entire dataset ( $\rho_0 = 1.0$ ). For the calculation of the sample variance at the unknown points, the circle might not include any observation (Fig. 2, blue circle). In this case, the variance of the entire dataset ( $\rho_0$ ) is used as sample variance. Thus, for all observation and new points a sample variance can be defined. The sample variance for each point is then used to find the sample standard deviation ( $S = \sqrt{\sigma^2}$ ), which is included in the estimation of the covariance matrix. A comparison to the observations can be used to find a suitable radius and minimum number of stations that have to be located within this circle ( $n_{\min}$ ). Various values should be tested to find the best fit.

From before we know that the elements of the covariance matrix  $\mathbf{C}$  are obtained from the covariance function  $K(d)$  (see Eq. (5)):

$$\mathbf{C} = \begin{bmatrix} K_{P_1, P_1} & K_{P_1, P_2} & \cdots & K_{P_1, P_{N_t}} \\ K_{P_2, P_1} & K_{P_2, P_2} & \cdots & K_{P_2, P_{N_t}} \\ \vdots & \vdots & \ddots & \vdots \\ K_{P_{N_t}, P_1} & K_{P_{N_t}, P_2} & \cdots & K_{P_{N_t}, P_{N_t}} \end{bmatrix}, \quad (73)$$

with  $P_i, P_j$  as the locations of the observation or new points ( $i, j = 1, \dots, N_t$ ). By applying the moving variance, the covariance matrix becomes:

$$\mathbf{C} = \begin{bmatrix} S_{P_1} S_{P_1} \cdot r_{P_1, P_1} & S_{P_1} S_{P_2} \cdot r_{P_1, P_2} & \cdots & S_{P_1} S_{P_{N_t}} \cdot r_{P_1, P_{N_t}} \\ S_{P_2} S_{P_1} \cdot r_{P_2, P_1} & S_{P_2} S_{P_2} \cdot r_{P_2, P_2} & \cdots & S_{P_2} S_{P_{N_t}} \cdot r_{P_2, P_{N_t}} \\ \vdots & \vdots & \ddots & \vdots \\ S_{P_{N_t}} S_{P_1} \cdot r_{P_{N_t}, P_1} & S_{P_{N_t}} S_{P_2} \cdot r_{P_{N_t}, P_2} & \cdots & S_{P_{N_t}} S_{P_{N_t}} \cdot r_{P_{N_t}, P_{N_t}} \end{bmatrix}, \quad (74)$$

with  $S_{P_i, P_j}$  ( $i, j = 1, \dots, N_t$ ) as the sample standard deviation for each point and  $r(d)$  as the correlation function at the location of the observations or new points  $P_i, P_j$  (instead of the covariance function  $K(d)$ ). Thus, while the original covariance matrix depends only on the distance between the stations and the chosen covariance function and their parameters, the moving-variance covariance matrix depends on the observations again and a correlation function. The correlation function  $r(d)$  is more homogeneous and isotropic than the original covariance function (Vestøl et al. 2019), and the same function for the entire dataset is used. However, parameters for the correlation function have to be defined, which is explained below. The new correlation function

( $r(d)$ ) together with the sample standard deviations ( $S_{P_i, P_j}$ ) is applied for all signal covariance and cross-covariance matrices ( $\mathbf{C}_{ss}$ ,  $\mathbf{C}_{ps}$  and  $\mathbf{C}_{pp}$ ).

## 5.1 Correlation analysis

The application of a moving-variance covariance matrix also involves the determination of a correlation function, which can be obtained by dividing the covariances by the standard deviation (Vestøl 2006):

$$r(0) = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{l_i^2}{S_{l_i}^2} = \rho_0$$

$$r(d_p) = \frac{1}{N_p - 1} \sum_{i \neq j}^{N_p} \frac{(l_i - \bar{l}_i)(l_j - \bar{l}_j)}{S_{l_i} S_{l_j}}, \quad (75)$$

with  $\bar{l}_{i,j}$  and  $S_{l_i, l_j}$  being the mean and standard deviation of the observations within the specific distance group. The observations are reduced by the mean value before the correlation analysis is carried out. The correlations are estimated for each distance group as for the covariances and at the end plotted in a correlation diagram.

The combined covariance analysis using the approach described in Sect. 3.2 has to be adjusted as well to obtain a correlation function instead. Here, Eq. (63) is modified by removing the mean value of the observations within specific distance groups (as in Eq. (75)). Thus, Eq. (63) becomes:

$$\beta_r = \frac{(V_k(P_i) - \bar{V}_k(P_i))(V_l(P_j) - \bar{V}_l(P_j)) - C_{\eta_k(P_i)\eta_l(P_j)}}{f_{kl}(P_i, P_j)} \bigg|_{\substack{d_{P_i, P_j} \in d_p \\ k, l = E, N}} \quad (76)$$

which is then used to find an estimator for the empirical correlation  $\hat{r}(d_p)$  of the correlation function at distance  $d_p$ . Equation (70) is also modified by taking into account the standard deviations of the observations within each distance group:

$$\hat{r}(d_p) = \frac{\alpha^T \beta_r}{S_{V_k(P_i)} S_{V_l(P_j)}} \bigg|_{\substack{d_{P_i, P_j} \in d_p \\ k, l = E, N}} \quad (77)$$

The determination of the standard deviation of the empirical correlations is not affected in the correlation analysis as the calculation of the variance of the estimator ( $S_{\hat{K}(d_p)}^2$ ; Eq. (71)) is only dependent on  $\Sigma_\beta$  (Eq. (66)). However,  $\Sigma_\beta$  is a function of the a-priori chosen correlation function and their parameters, the latitude and longitude-dependent functions  $f_{st}$  and  $f_{uv}$  defined as in Eqs. (50–53), and the covariances of the observational uncertainties. Thus, the standard deviation determined in the correlation analysis is determined in the

same way as in the covariance analysis. The new parameters are then used together with the correlation function to estimate the moving-variance covariance and cross-covariance matrices and carry out a (HV-)LSC to find the signal at the observation points and new points.

## 6 Synthetic dataset

The presented methods above are partly new and need to be verified. A simple synthetic dataset is created where 130 stations are distributed over four separate tectonic plates (Fig. 3). Rotational background velocities in the east-west (EW) and north-south (NS) direction are assigned for each plate (see Fig. S1). We assume angular velocities for the four tectonic plates and keep one plate fixed (plate 1) so that all velocities on the three remaining plates (plate 2, plate 3, and plate 4) are with respect to the fixed plate. We also add a random noise to the assigned background velocities by using the Python function *numpy.random* (Harris et al. 2020). The random noise represents the variations in the amplitude of the velocity that are usually observed for GNSS stations, e.g. due to local effects and/or tectonics. In addition, an observational uncertainty is created at each point, which is randomly distributed between predefined values using the Python function *numpy.random* as well. The application of a real dataset (obtained from GNSS solutions) is the purpose of a forthcoming paper.

The first plate has 76 stations and is fixed (background velocity is 0 mm/a, Fig. S1). However, velocities vary between  $-1$  and  $1$  mm/a in both directions due to the added noise. The observational uncertainties of the velocities are assigned to be between  $0.2$  and  $0.5$  mm/a for the EW and NS component. The second plate (19 stations) moves with about  $2$  mm/a to the north and  $0.5$  mm/a to the west in the northern part of the plate, while a eastward direction is obtained in the southern part of plate 2 (Fig. S1). The velocities are with respect to plate 1. The same noise range as for plate 1 is used for stations on plate 2 (Fig. 3b), but the observational uncertainty is chosen to be a bit larger on the second plate, reaching values of up to  $0.6$  mm/a for both horizontal velocity components. The third plate, which is the smallest tectonic plate and surrounded by the other three plates (Fig. 3), moves with up to  $4$  mm/a to the west and south with respect to plate 1. The 13 stations on this plate have the same additional noise as the other two plates. The observational uncertainty of both simulated velocity components is between  $0.2$  and  $0.8$  mm/a. Velocities on the fourth plate (22 stations) are slightly smaller and the plate moves with up to  $3$  mm/a to the east and up to  $1$  mm/a to the north with respect to plate 1 as well (Fig. 3b). The north component becomes larger on the east side of plate 4 (Fig. S1). The observational uncertainty of the velocities is between  $0.2$  and  $0.5$  mm/a for both horizontal components.

Looking at the simulated velocity field (Fig. 3b, Fig. S1), the plate boundaries are of different type. Boundaries between plates 1 and 2, plate 3 and 4, and plate 4 and 2 are convergent plate boundaries, while between plates 1 and 3, as well as 2 and 3 divergent plate boundaries exist. Plate 4 and 1 are separated by a transform zone. The distance between stations varies between  $11$  km and almost  $7000$  km, while having a mean distance of around  $3100$  km. However, stations are located in three areas only (Fig. 3a), where mean station distance is much smaller ( $800$ – $1100$  km) and also maximum distance varies between  $2000$  and  $3100$  km.

The synthetic velocity field is analysed to find the covariance parameters (covariance and correlation length) using a standard covariance analysis (see Sect. 2.1) and the combined covariance analysis of the horizontal velocity field (see Sect. 3.2). The parameters are then used to find the velocity field at the observation points as well as grid points using LSC. The grid has a resolution of  $0.25^\circ$  in the longitude and latitude. The covariance analysis and the LSC are based on the condition that the observations have an expectation of  $0$  mm/a. Thus, the data have to be reduced by their mean values, which are  $-0.025$  mm/a for the EW component and  $-0.01$  mm/a for the NS component. The LSC is then used in the remove-compute-restore principle (e.g. Vestøl et al. 2019) where the mean value is removed from the synthetic dataset before the signal is obtained and added afterwards to the obtained signal (restored). In a first step, a standard LSC is done using the velocity components separately as well as the results from the standard covariance analysis. This is followed by using the HV-LSC technique where both horizontal components are obtained at the same time. In addition, the HV-LSC techniques are applied together with plate-boundary constraints (see Sect. 4) as well as with the moving-variance technique (HV-LSC-ex<sup>2</sup>, see Sect. 5). For the latter, correlation parameters are estimated in a correlation analysis. The final interpolated velocity field is also compared to results using Kriging.

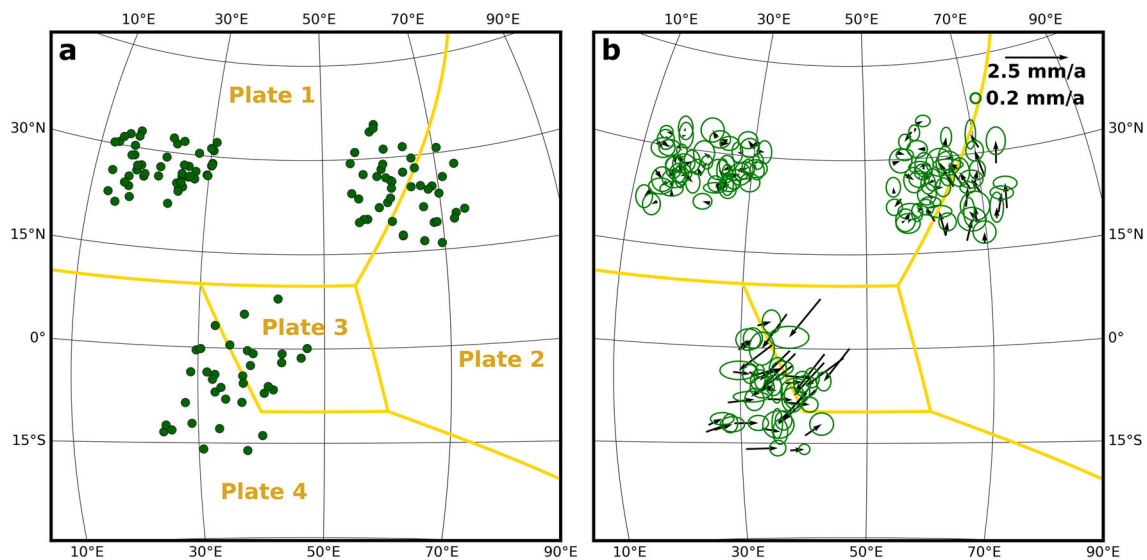
### 6.1 Covariance analysis

The empirical covariances are determined using Eqs. (11) and (12). The covariance function is then obtained by determining the unknown parameters:  $C_0$  and  $d_0$ . The fit between function and empirical covariances is defined by the Pearson's correlation coefficient (*PCC*; Freedman et al. 2007) and the misfit  $\phi$ , which is obtained using:

$$\phi = \frac{\sqrt{\frac{1}{N_\delta} \sum (K_{\text{curve}}(d_p) - K_{\text{emp}}(d_p))^2}}{C_0}, \quad (78)$$

with  $N_\delta$  being the number of distance groups,  $K_{\text{curve}}(d_p)$  the value of the covariance function at  $d_p$ , and  $K_{\text{emp}}(d_p)$  the





**Fig. 3** Synthetic dataset created for a scenario with four tectonic plates. The stations (green dots) and the plate numbers are shown in (a). The horizontal velocity field is shown as a vector field in (b). The green

ellipses around the vectors represent the generated uncertainties. The scale for the vectors and uncertainty ellipses is shown in the top-right corner of (b). The plate boundaries are shown as solid yellow lines

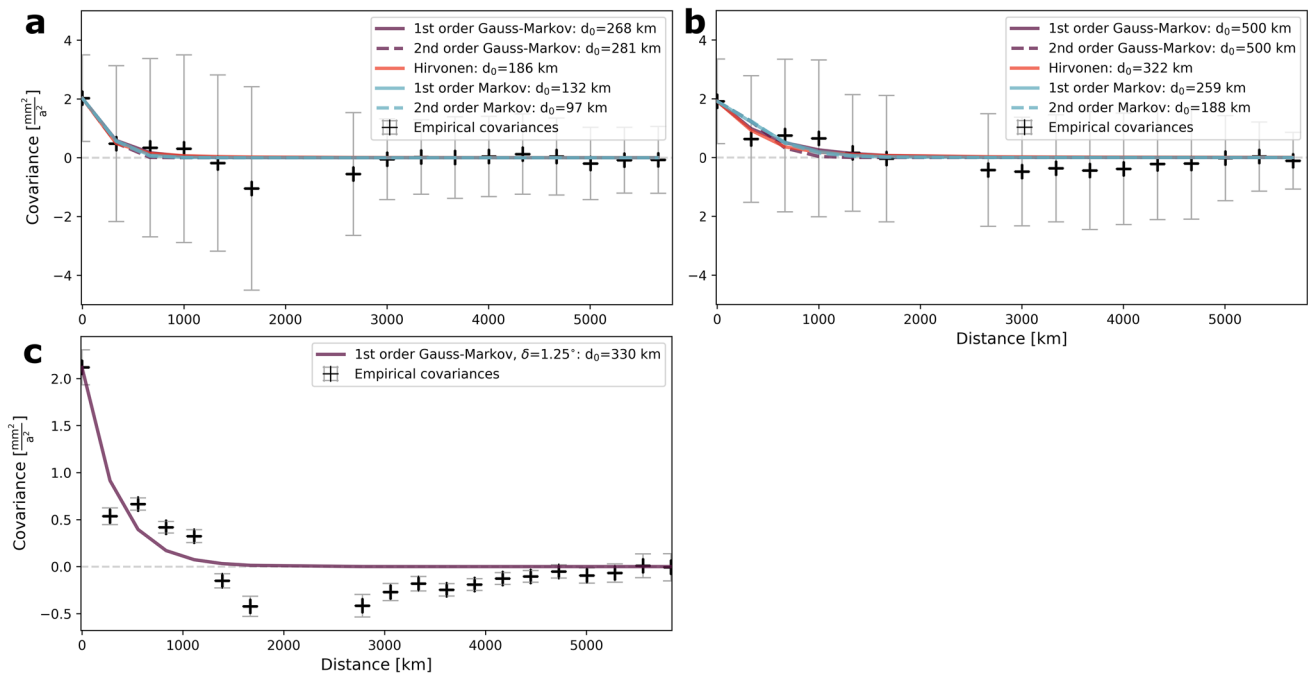
empirical covariance at  $d_p$ . The division by the variance  $C_0$  allows the comparison of the different  $\phi$  values for the various velocity components. The misfit is also estimated for the first three points (denoted as  $\phi_3$ ) only to be able to differentiate between similar fits. The fit to the first three points is the most important as the covariance and correlation are the largest for points in the close surrounding of an observation point. The correlation length is depending on the correlation between signals that are in close distance rather than by stations that are furthest away, and thus, the covariances closest to the y-axis are mainly used for computing the correlation length. The best-fitting parameters are then chosen based on all three values: *PCC* (should be close to 1),  $\phi$  (should be close to 0),  $\phi_3$  (should be close to 0).

The variances  $C_0$  of the synthetic data are  $2.025 \text{ mm}^2/\text{a}^2$  and  $1.913 \text{ mm}^2/\text{a}^2$  for the EW and NS velocity components, respectively. Using a  $\delta$  of  $1.5^\circ$  (Fig. 4), correlation lengths  $d_0$  of  $268 \pm 102 \text{ km}$  (for the EW component) and  $500 \pm 130 \text{ km}$  (for the NS component) are obtained from fitting a first-order Gauss–Markov function (see Eq. (16)). The *PCC*s are 0.87 and 0.92 for the EW and NS components, respectively (Table 1). A different  $\delta$  of  $0.75^\circ$  (Fig. S2) results in a change of the correlation lengths to  $250 \pm 45 \text{ km}$  (EW) and  $250 \pm 81 \text{ km}$  (NS). The fit for the EW component is better using a smaller  $\delta$  value (*PCC* increased to 0.90 and  $\phi$  decreased to 0.108), but the fit for the NS component is worse. In total, several  $\delta$  values as well as five different type of functions have been tested, whose results are shown in Fig. 5 as well as in the supplementary material (Fig. S2). The  $\delta$  parameter is chosen to vary between  $0.5^\circ$  and  $5.0^\circ$  based on the size of the study area. A

smaller  $\delta$  value would result in too small distance groups with not enough points per group to obtain a realistic estimate of the empirical covariance. In contrast, a larger  $\delta$  of more than  $5.0^\circ$  would lead to too few empirical covariance estimates and the fitting of a covariance function is not possible. Distance group with not enough points are not plotted and not considered in the covariance analysis. The obtained parameters and their misfits  $\phi$ ,  $\phi_3$  and *PCC*s are summarized in Fig. 5, Table 1, and the supplementary Excel spreadsheet. The usage of a  $\delta$  of  $1.5^\circ$  and a first-order Gauss–Markov function are providing the best fit based on *PCC*,  $\phi$ , and  $\phi_3$  considering both horizontal components (Fig. 4a, b; Table 1). In addition, a  $\delta$  of  $1.5^\circ$  provides several values within the first 1000 km to be able to define the covariance function. The obtained correlation lengths are smaller than the calculated mean distances (800–1100 km), but larger than minimum distances (11–82 km). Based on the correlation lengths, covariances are not  $0 \text{ mm}^2/\text{a}^2$  at the mean distances for the three areas. Thus, the correlation lengths cover the various distances between the stations well.

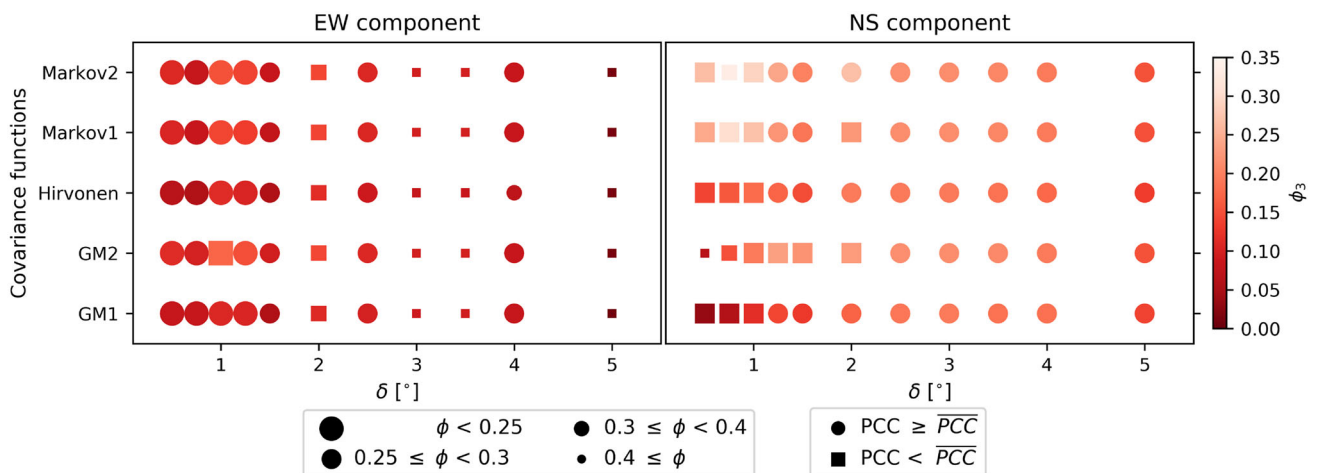
The results obtained from the standard covariance analysis of the horizontal components (function of the first-order Gauss–Markov process) can now be used for the first iteration of the new covariance analysis where both components and their correlation are taken into account. Various  $\delta$  values are tested again. The combined covariance analysis gives a correlation length of  $330 \pm 67 \text{ km}$  and a covariance of  $2.120 \text{ mm}^2/\text{a}^2$  using a  $\delta$  of  $1.25^\circ$  (Fig. 4c) after one iteration (*PCC* = 0.93,  $\phi$  = 0.107, and  $\phi_3$  = 0.127). These two values are now used in the following for the HV-LSC. The





**Fig. 4** Covariance analysis of the synthetic dataset. The empirical covariance is shown as black cross with the grey bar as the standard deviation. The solid and dashed lines show the fitted covariance functions with the parameters written in the top-right corner. (a) and (b) are the covariance functions of the EW and NS velocity components, respectively. (c) shows the covariance function obtained from the combined covariance analysis using both horizontal components at the same time as well as the correlations between them. A first-order Gauss–Markov

function is used in (c). The solid and dashed purple lines are the functions of the first- and second-order Gauss–Markov process, respectively, the Hirvonen function is shown as solid orange line, and the functions of the first- and second-order Markov model are visualized as solid and dashed light-blue lines for covariance functions in (a) and (b). Distance groups with less than 130 station pairs are not considered in the covariance analysis (and therefore not plotted here)



**Fig. 5** Results of the covariance analysis for various  $\delta$  values and covariance functions (GM1 - first-order Gauss–Markov, GM2 - second-order Gauss–Markov, Markov1 - first-order Markov, Markov2 - second-order Markov2). The colourscale of the circles and squares represents the value of  $\phi_3$  (dark red colours stand for a better fit). The size of the circles and squares is related to  $\phi$ , with larger circles/squares for better fits of empirical covariances to the covariance function. The  $PCC$  value

is visualized by circles and squares: when the  $PCC$  between empirical covariances and covariance function is above the average of all  $PCC$  values obtained in the covariance analysis a circle is used. In contrast, a square is used when the  $PCC$  is below the average. The best fit of empirical covariances to covariance function is found, when a large, dark-red circle is obtained

**Table 1** Results of the covariance analysis using a  $\delta$  of  $1.5^\circ$  for the horizontal velocity field as well as the first- and second-order Gauss–Markov function (Eqs. (16) and (17)), Hirvonen function (Eq. (15)), and the functions of first- and second-order Markov model

Covariance function	EW				NS			
	$d_0$ [km]	<i>PCC</i>	$\phi$	$\phi_3$	$d_0$ [km]	<i>PCC</i>	$\phi$	$\phi_3$
First-order Gauss–Markov	<b>268</b> $\pm$ 102	0.87	0.158	0.057	<b>500</b> $\pm$ 130	0.92	0.152	0.130
Second-order Gauss–Markov	281 $\pm$ 67	0.86	0.162	0.095	500 $\pm$ 111	0.85	0.182	0.221
Hirvonen	186 $\pm$ 78	0.87	0.160	0.055	322 $\pm$ 86	0.89	0.163	0.147
First-order Markov	132 $\pm$ 40	0.86	0.160	0.079	259 $\pm$ 57	0.89	0.165	0.188
Second-order Markov	97 $\pm$ 27	0.86	0.160	0.085	188 $\pm$ 41	0.88	0.171	0.202

$d_0$  is the correlation length given in km, *PCC* is the Pearson correlation coefficient, and  $\phi$  and  $\phi_3$  are the misfit values (see Eq. (78)). The final values for the horizontal velocity field are marked in bold. The equations for the functions of the first- and second-order Markov model are given in the supplementary material (Eqs. (S1) and (S2))

same values are obtained within one or two iterations, if the starting correlation length is varied. A slightly different correlation length can be obtained when the Hirvonen function is used and/or a different  $\delta$  (see Fig. S3).

## 6.2 Correlation analysis

We also have to find parameters for the correlation function when using the moving-variance approach. We test the same  $\delta$  values and functions as in the covariance analysis, but instead calculate the empirical correlation (Eq. (75)). The correlation at  $d=0$  km ( $\rho_0$ ) and the correlation length  $d_{0,r}$  are determined by finding the best fit of a correlation function to the empirical correlations by calculating the *PCC*,  $\phi$ , and  $\phi_3$  values. The best fits are again obtained using a function of the first-order Gauss–Markov process for various  $\delta$ s (Fig. S4, supplementary Excel spreadsheet). This information is used to obtain the combined correlation length, which is found to be  $369 \pm 139$  km using a  $\delta$  of  $0.75^\circ$  (Fig. 6). The correlation  $\rho_0$  is 1.0. Empirical correlations decrease to zero for small distances, but are not zero for larger distances. Thus, the correlation function does not represent the behaviour of the data well, when stations are more than 3000 km away from each other. However, at this distance it is expected that velocities on Earth vary due to existence of moving plates.

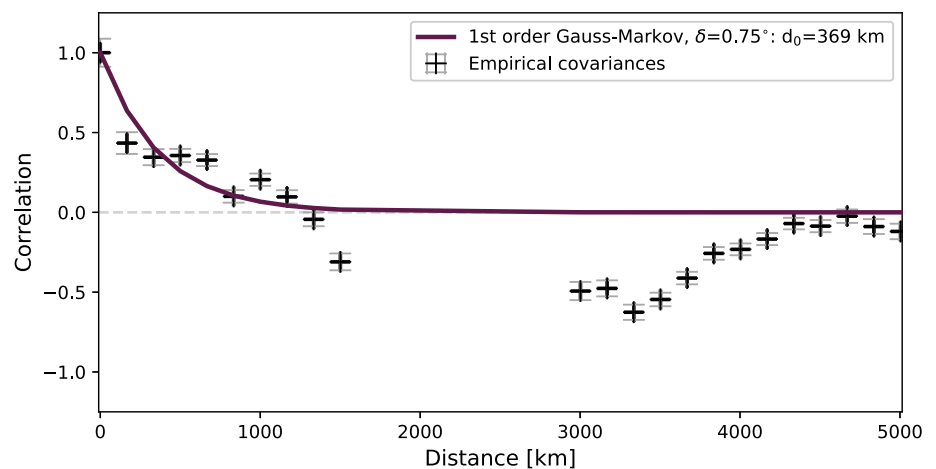
## 6.3 Least-squares collocation at known points

LSC of the horizontal velocity field is done for the four methods presented here: standard LSC (Fig. 7a–c), HV-LSC (Fig. 7d–f), HV-LSC with plate-boundary constraints applied (Fig. 7g–i), and HV-LSC with plate-boundary constraints and moving variance applied (termed HV-LSC-ex<sup>2</sup>; Fig. 7j–l). In all cases a filtering is only applied where the observations (**I**, in this case the synthetic data) are split up into a signal part and a noise part (using Eqs. (3) and (21)). Additionally, a standard error of the obtained signal is estimated (using Eqs. (4) and (23)).

Differences of collocated to synthetic dataset are then computed to analyse the performance of each method. As a synthetic dataset is used, which is based on a background velocity and an additional random noise, the differences between the estimated signal (collocated data) to the background velocity (Fig. S1(A)) can be obtained (Fig. 7). The analysis allows us to identify the best method that is able to find the background velocity. This comparison is only possible for a synthetic dataset and is not working for real observations as the “true” signal is not known. Thus, we also compare the collocated velocity field with the synthetic data including noise. In addition, we apply a leave-one-out cross-validation (LOOCV) to check the stability of the collocated velocity field. Within LOOCV a station is removed and the velocity at this station is interpolated using the remaining station velocities by applying the specific LSC technique. The obtained velocities from LOOCV are then used to find the difference to the collocated velocity where all data points were used (Fig. 7c, f, i, l), which allows us to analyse the effect of the applied LSC technique. However, the LOOCV is not able to judge the method, which needs to be defined by analysing the fit to the synthetic data (with and without noise).

The synthetic horizontal velocity field has minimum and maximum velocities of  $-4.392$  mm/a (EW) and  $-4.101$  mm/a (NS) as well as  $3.363$  mm/a (EW) and  $2.880$  mm/a (NS) including noise (Table 2), respectively. A LSC of the velocities with correlation lengths of 268 km (EW component) and 500 km (NS component) and without taking the correlation between the velocities into account provides values, which are similar to the synthetic velocity field (Fig. 7a, Table 2). Differences are larger than 1 mm/a for both components comparing the collocated velocity field with the background velocity field without noise (Fig. 7b, Table 3). The LOOCV shows differences of up to  $\pm 3$  mm/a for the standard LSC (Fig. 7c, Table 3). These large differences are found for stations close to plate boundaries where the station velocity

**Fig. 6** Combined correlation analysis using the first-order Gauss–Markov covariance function for a  $\delta$  of  $0.75^\circ$ . Distance groups with less than 130 station pairs are not considered in the covariance analysis (and therefore not plotted here)



itself cannot be interpolated well enough, if one station is removed.

In the next step, the velocity field is obtained by applying the HV-LSC technique and using a correlation length of 330 km (Fig. 7d–f). The differences to the background velocity field are smaller than for the standard LSC (Fig. 7e, f, Table 3), thus showing the improvement when using the HV-LSC method. LOOCV for HV-LSC shows a mixed behaviour with both improvement and deterioration in both components (see Table 3) compared to applying standard LSC. However, LOOCV is only applied to see, if the chosen LSC is able to replicate their results when using only a subset of the stations, and it cannot be used to decide on the best method.

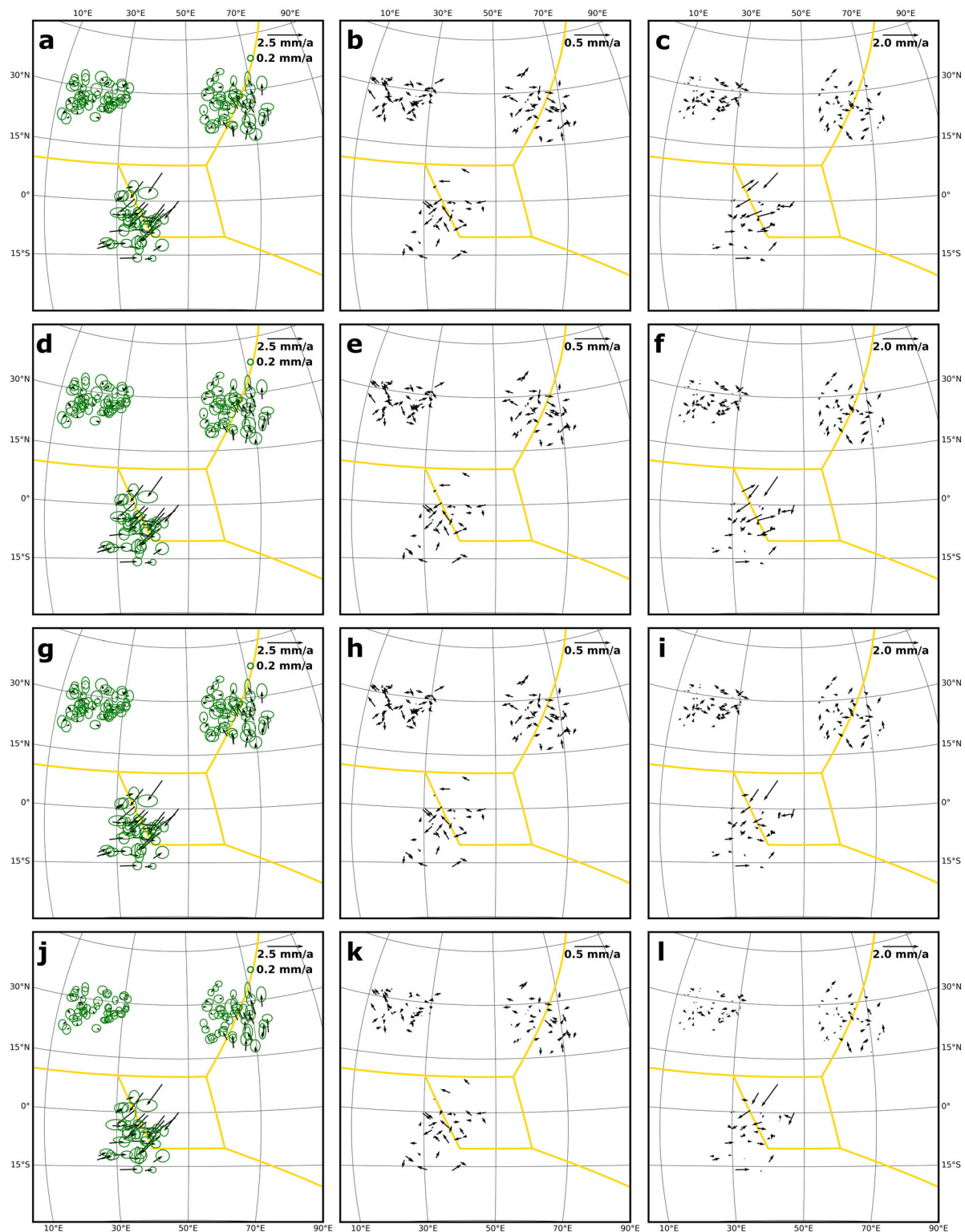
The additional application of plate boundary constraints provides a better fitting collocated velocity field along plate boundaries (Fig. 7g–i, Table 3). The stations on plate 3 are fixed and stations on plate 1, 2, and 4 are virtually moved to the north, east, and south, respectively, to simulate the independency of velocities on various plates. The differences in the velocities for stations on plate 1 and 2 are greatly reduced. Results from the LOOCV show that the collocation is more stable and the differences to the collocated velocity field are reduced by 1 mm/a (Fig. 7i, Table 3). However, the velocities are still noisy with large variations in the surrounding of a station (see plate 1, Fig. 7h). Thus, we apply the moving-variance approach on top of the HV-LSC with plate-boundary constraints: HV-LSC-ex<sup>2</sup>.

Based on the results of the correlation analysis, a correlation length of 369 km is used when the moving-variance approach is applied. In addition, various radii (250–1050 km) and number of minimum stations (2, 3, 5, 7 and 10) have been tested. The fit of the velocity field to the background velocity at known points is used to decide on the best-fit radius and the number of minimum stations that have to lie within the circle. The overall best fit is obtained for a radius of 850 km using at least 7 stations in the circle (see supplementary Excel spreadsheet). The covariance  $\rho_0$  is only 10% at a radius of

850 km. Thus, such a boundary condition could be used in future studies, where the “true” signal is not known.

The new filtered horizontal velocity field using HV-LSC together with plate-boundary constraints and moving-variance approach (HV-LSC-ex<sup>2</sup>) shows consistent velocities on all plates, pointing into similar directions (Fig. 7j). The fit to the synthetic dataset is improved and LOOCV indicates that the collocated values are stable. The advantage of the moving-variance approach becomes especially visible for plate 1 (Fig. 7k) where the differences to the background velocity are much smaller than those for the other LSC techniques. Differences of the collocated velocity field to the results of a LOOCV are small for plate 1 (Fig. 7l). The minimum standard errors of the collocated velocity field are also a bit smaller than the standard errors obtained from the other LSCs, but their maxima are larger and more similar to the uncertainties of the synthetic dataset. The result using HV-LSC-ex<sup>2</sup> produced a smoother result by filtering out the noise.

Smoothing in any LSC largely depends to a rather large extent on the signal to noise ratio. However, smoothing from the LSC is not always wished and Kotsakis (2007) provided a methodology to reduce the smoothing. But, in some cases, the more general behaviour of the observations is tried to be resolved. The moving-variance approach will smooth the data more in some areas, but less in other areas, depending on how large the moving variance is at specific locations. Velocity field solutions can also include local signals due to unstable stations (e.g. due to the installation on the roof of a building) and those have to be smoothed to obtain a velocity model that fits the overall movement of the area. This is especially important considering that velocity field solutions and/or models are used to estimate the strain rates. These would be increased due to small-scale changes, which are not due to a true change in strain rate. Thus, a smoothed velocity field can be an advantage.



**Fig. 7** Results of LSC for the horizontal velocity field at known points. Left column (a, d, g, j) shows the obtained velocity field using various LSC techniques, middle column (b, e, h, k) their differences to the background velocities (Fig. S1(A)), and right column (c, f, i, l) the differences of collocated to LOOCVs velocities using various LSC techniques. (a–c) Standard LSC separately for each horizontal velocity field component, (d–f) HV-LSC, (g–i) HV-LSC together with plate-boundary

constraints, (j–l) HV-LSC together with plate-boundary constraints and moving variance (HV-LSC-ex<sup>2</sup>). The length of the velocity arrow is different for each column: 2.5 mm/a for the LSC results, 0.5 mm/a for the differences to the background velocities, and 2.0 mm/a for the differences between the results of the LSCs to the LOOCVs. The solid yellow lines mark the plate boundaries

**Table 2** Statistics of the synthetic dataset and LSC of the horizontal velocity field at the known points

	EW				NS			
	Min	Max	Mean	StdDev	Min	Max	Mean	StdDev
<b>Data</b>								
Synthetic data	− 4.392	3.363	− 0.025	1.473	− 4.101	2.880	− 0.010	1.438
Synthetic data w/o noise	− 3.805	2.380	− 0.042	1.439	− 3.566	2.134	0.027	1.330
LSC	− 3.760	3.233	− 0.005	1.371	− 3.840	2.710	− 0.001	1.365
HV-LSC	− 3.774	3.244	− 0.005	1.373	− 3.822	2.741	− 0.001	1.369
HV-LSC with plate boundaries	− 3.926	3.244	− 0.010	1.395	− 3.824	2.745	− 0.001	1.375
HV-LSC-ex <sup>2</sup>	− 4.301	3.320	− 0.015	1.396	− 4.029	2.812	− 0.010	1.379
<b>Standard error</b>								
Synthetic data	0.201	0.777	0.366	0.111	0.202	0.649	0.382	0.096
LSC	0.195	0.681	0.338	0.092	0.193	0.575	0.336	0.076
HV-LSC	0.194	0.684	0.335	0.091	0.198	0.589	0.349	0.081
HV-LSC with plate boundaries	0.194	0.684	0.335	0.091	0.198	0.590	0.350	0.081
HV-LSC-ex <sup>2</sup>	0.173	0.716	0.296	0.100	0.165	0.635	0.299	0.097

All values are given in mm/a

**Table 3** Statistics of the difference between the background velocities (synthetic dataset w/o noise) and the results of LSCs at known points as well as the differences between the results of the LSCs to the LOOCVs

	Synthetic dataset w/o noise – LSC				LSC – LOOCV			
	Min	Max	Mean	StdDev	Min	Max	Mean	StdDev
<b>EW</b>								
LSC	− 1.098	0.885	− 0.037	0.552	− 2.670	2.823	− 0.002	0.875
HV-LSC	− 1.072	0.880	− 0.037	0.545	− 2.545	2.829	− 0.003	0.837
HV-LSC with plate boundaries	− 1.037	0.880	− 0.033	0.541	− 2.126	2.287	− 0.014	0.736
HV-LSC-ex <sup>2</sup>	− 0.940	0.868	− 0.027	0.440	− 2.493	2.233	− 0.017	0.636
<b>NS</b>								
LSC	− 0.874	1.053	0.028	0.454	− 2.480	1.437	− 0.018	0.600
HV-LSC	− 0.905	0.985	0.028	0.475	− 3.127	1.327	− 0.028	0.687
HV-LSC with plate boundaries	− 0.903	0.934	0.027	0.472	− 3.127	1.300	− 0.034	0.640
HV-LSC-ex <sup>2</sup>	− 0.770	0.878	0.037	0.380	− 3.497	1.300	− 0.028	0.565

All values are given in mm/a

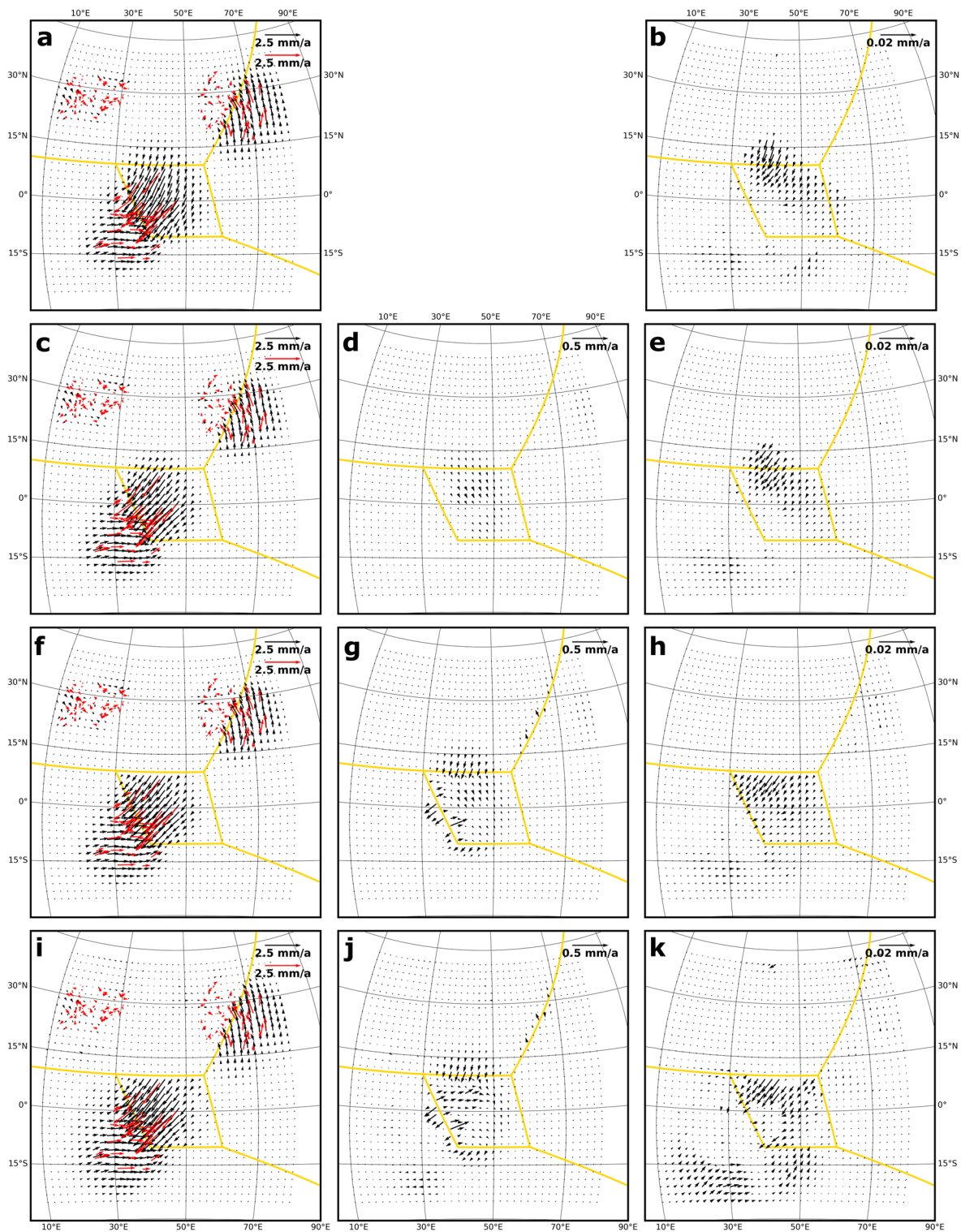
## 6.4 Least-squares collocation at grid points

The main purpose of LSC is the interpolation and extrapolation of existing data to arbitrary chosen points (e.g. a grid). Here, we create a grid spanning a longitude range of  $-10^\circ$  to  $80^\circ$  and a latitude range of  $-25^\circ$  to  $41^\circ$  with a resolution of  $0.25^\circ$  (Fig. 8, however, shows only every eighth grid point to increase visibility). The differences of HV-LSCs to standard LSC are calculated (Fig. 8d, g, j). In addition, we run a Jackknife resampling (e.g. Tukey 1958) for a grid with a resolution of  $2.0^\circ$ . Here, for every grid point the specific LSC is done by leaving one synthetic data point out. Thus, for every grid point a LOOCV is done based on the entire synthetic dataset. The final result is then obtained by calculating a mean for each grid point, which is compared to the

LSC result. Furthermore, the minimum and maximum values from the Jackknife resampling can be compared to the minimum and maximum values of the synthetic data (Table S1). These values together can be used to verify the interpolation method. We use a grid resolution of  $2.0^\circ$  only for Jackknife resampling as a resampling for a grid with a resolution of  $0.25^\circ$  involves applying LSC for almost 15 Million points. Such a calculation would run for  $\sim 134$  days using 32 CPUs. Thus, we decrease the grid resolution to allow a faster running time of the Jackknife resampling (less than 5 h).

The interpolated horizontal velocity field using a standard LSC (Table 4) shows large velocities over areas where no stations are located (Fig. 8a). In particular velocities in the southern part of plate 1 are largely affected by the velocities on plate 3. The same can also be observed for interpolated





**Fig. 8** Results of LSC at newly defined grid points for the horizontal velocity field. Left column (**a**, **c**, **f**, **i**) shows the obtained velocity field, middle column (**d**, **g**, **j**) presents the difference of the interpolated velocity field to the velocity field using a standard LSC (**a**), and right column (**b**, **e**, **h**, **k**) shows differences to the mean of Jackknife resampling. The arrows are plotted at every eighth grid point for the left and middle columns, and every grid point for the right column (a grid of 2.0° is obtained). Red arrows in (**a**), (**c**), (**f**) and (**i**) are the synthetic data. (**a**–**b**) standard LSC separately for each horizontal velocity field component. (**c**–**e**) shows HV-LSC, where (**d**) is (**a**) minus (**c**), but (**e**)

is the difference between (**c**) and the mean velocity field from Jackknife resampling using HV-LSC. (**f**–**h**) shows HV-LSC together with plate-boundary constraints, where (**g**) is (**a**) minus (**f**), but (**h**) shows the difference of the mean velocity field from Jackknife resampling using HV-LSC with plate-boundary constraints to (**f**). (**i**–**k**) shows HV-LSC-ex<sup>2</sup> (moving-variance radius of 850 km), where (**j**) is (**a**) minus (**i**), and (**k**) is again the difference between the mean velocity field from Jackknife resampling using HV-LSC-ex<sup>2</sup> to (**i**). The length of the velocity arrow varies throughout the entire figure (see upper left in each sub-figure). The solid yellow lines mark the plate boundaries

**Table 4** Statistics of LSC of the horizontal velocity field at grid points

	EW				NS			
	Min	Max	Mean	StdDev	Min	Max	Mean	StdDev
<b>Data</b>								
Synthetic data (grid points w/o noise)	− 3.861	2.667	0.397	1.633	− 3.748	2.300	0.291	1.517
LSC	− 3.683	3.152	− 0.028	0.594	− 3.825	2.690	− 0.066	0.820
HV-LSC	− 3.714	3.179	− 0.030	0.665	− 3.773	2.707	− 0.042	0.686
HV-LSC with plate boundaries	− 3.856	3.178	− 0.014	0.700	− 3.771	2.712	− 0.023	0.694
HV-LSC-ex <sup>2</sup>	− 4.231	3.267	0.007	0.783	− 4.273	2.781	0.016	0.789
<b>Standard error</b>								
Synthetic data	0.201	0.777	0.366	0.111	0.202	0.649	0.382	0.096
LSC	0.285	1.423	1.334	0.166	0.263	1.383	1.216	0.228
HV-LSC	0.274	1.456	1.340	0.195	0.289	1.456	1.340	0.194
HV-LSC with plate boundaries	0.274	1.456	1.343	0.193	0.289	1.456	1.344	0.192
HV-LSC-ex <sup>2</sup>	0.064	1.456	0.932	0.528	0.060	1.456	0.923	0.534

All values are given in mm/a

**Table 5** Statistics of the difference between LSC techniques at grid points as well as to the mean values of Jackknife resampling

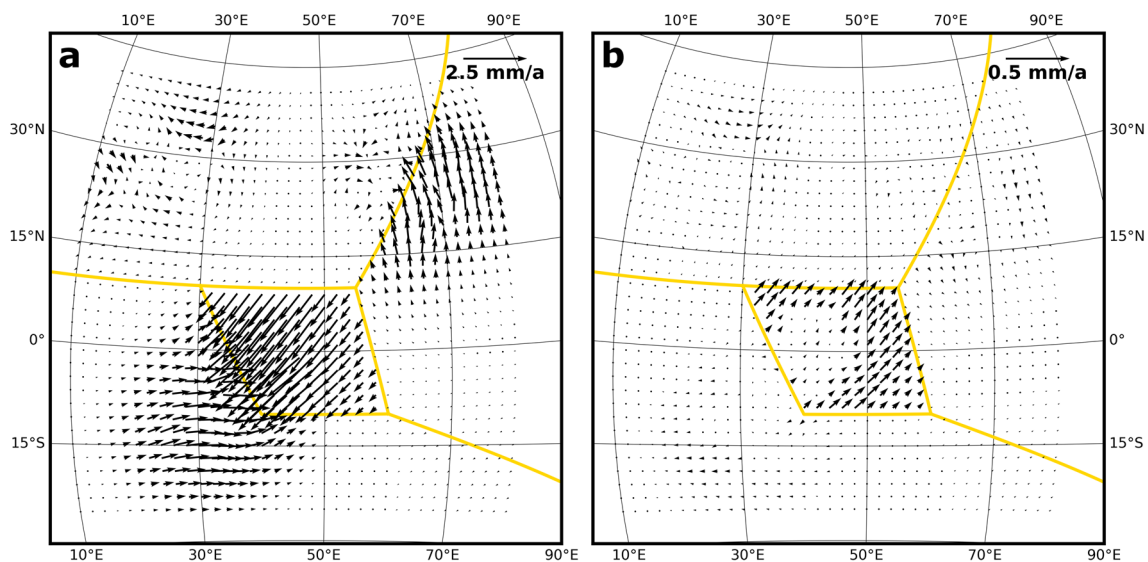
	Difference to standard LSC				LSC – Jackknife resampling			
	Min	Max	Mean	StdDev	Min	Max	Mean	StdDev
<b>EW</b>								
LSC	—	—	—	—	− 0.144	0.091	0.000	0.022
HV-LSC	− 0.295	0.455	0.003	0.093	− 0.149	0.090	0.000	0.025
HV-LSC with plate boundaries	− 3.326	4.524	− 0.014	0.213	− 0.161	0.090	0.002	0.027
HV-LSC-ex <sup>2</sup>	− 3.340	4.897	− 0.034	0.306	− 0.197	0.167	0.005	0.037
<b>NS</b>								
LSC	—	—	—	—	− 0.305	0.093	− 0.004	0.035
HV-LSC	− 0.880	0.485	− 0.025	0.186	− 0.221	0.044	− 0.004	0.027
HV-LSC with plate boundaries	− 2.585	1.662	− 0.044	0.275	− 0.229	0.068	− 0.001	0.029
HV-LSC-ex <sup>2</sup>	− 2.543	1.848	− 0.082	0.262	− 0.263	0.137	0.005	0.038

All values are given in mm/a

velocities using HV-LSC where the correlation between the horizontal velocity components is considered (Fig. 8c). The differences of this interpolated velocity field to standard LSC interpolation are up to 1.0 mm/a (Fig. 8d, Table 5). The largest differences can be found for the area with high velocities around the equator. The advantage of interpolating over a sphere becomes visible. Thus, HV-LSC has to be preferred when interpolating horizontal velocities as the correlation of both components is taken into account and the sphericity of the Earth is considered. Differences to Jackknife resampling of HV-LSC are similar as for standard LSC (Table 5) showing the stability of the interpolated values.

When plate-boundary constraints are considered, differences to the interpolated velocity field using standard LSC are increased to more than 4 mm/a (Table 5). Velocities on different plates do not interfere with each other and strong velocity gradients along the boundaries are obtained due to

the different assumed plate velocities and the application of plate-boundary constraints (Fig. 8f). Thus, plate velocities of different plates do not interfere with each other anymore, showing the improvement of this implementation. For example, large velocities in the southern part of plate 1 from standard LSC and HV-LSC are largely decreased when plate boundaries are considered. The velocity field on plate 1 shows a similar south-west direction for the entire plate, with small variations in areas where synthetic data are located. The differences of the interpolated velocity field using HV-LSC together with plate boundaries to the background velocities are decreased by this application. While values between −5 and 4 mm/a have been obtained for the previous LSCs, differences are now between −4 and 3.5 mm/a (Table 5 and S9). The largest improvement can be seen for plates 1, 2, and 4, as velocities on these plates had been affected by the large velocities on plate 3. The Jackknife resampling also shows



**Fig. 9** Comparison of Kriging to the interpolation using HV-LSC-ex<sup>2</sup> (with plate-boundary constraints and moving variance) for the horizontal velocity field. **(a)** Kriging using *PyKriging*, **(b)** final LSC model minus

**(a).** Only every eighth grid point is plotted (a grid of 2.0° is obtained). The solid yellow lines mark the plate boundaries

largest differences along the plate boundaries (Fig. 8h) as the missing of one station had a larger effect. The absolute differences between the velocity magnitudes are also smaller than for the previously applied LSCs. The simple differences between the LSC velocity field and mean velocity field from Jackknife resampling are similar to the previously applied LSC methods. But, it is visible that the velocity model at grid points is largely improved when plate-boundary constraints are applied, although small-scale changes in the horizontal field direction remain, especially on plate 1. Obtained standard errors are larger than uncertainties of the synthetic data (Table 4). While the maximum standard error is determined by the covariance and thus must be larger than the input uncertainty, the minimum standard error depends on the performance of LSC.

The velocity model and its standard error can be improved using the moving-variance method. The velocity model (Fig. 8i) has smaller velocities on plate 1 compared to the previous interpolations, which is consistent with the assumption of 0 mm/a background velocity in both directions (plate velocities were corrected with respect to plate 1 during the creation of the synthetic dataset). Differences in the southern part of plate 4 are obtained, which can be attributed to the moving-variance radius. The results become more similar to the background velocities as interpolated velocities decrease over a larger area to 0 mm/a compared to using (HV-)LSC with and without plate boundaries. In general, the velocity field is improved when applying moving variance as well after using HV-LSC together with plate-boundary con-

straints. The velocity field shows similar velocities for most areas of the plates and reduces variations due to local effects.

## 6.5 Comparison to other interpolation techniques

In addition to LSC other interpolation techniques can be applied, e.g. Kriging, which will be tested and compared in the following. The difference (Fig. 9, Table S2) is then obtained to the final grid using HV-LSC with plate-boundary constraints and moving variance (HV-LSC-ex<sup>2</sup>, Fig. 8i). We use the *PyKriging* package from Python (Murphy et al. 2021) and the changed station coordinates from the introduction of plate-boundary constraints. Thus differences between HV-LSC-ex<sup>2</sup> and Kriging are only due to the application of both horizontal velocity components at the same time, the application of moving variance as well as the choice of the covariogram model used in Kriging (an exponential covariogram model is used here; detailed parameters are provided in the supplement).

The largest difference between Kriging and HV-LSC-ex<sup>2</sup> is obtained for plate 3. Both interpolations use the moved station coordinates to take into account the plate boundaries. The velocities from the Kriging interpolation on plate 3 are larger in areas where no station is located. This fits better to the background velocity grid (Fig. S1(B)). Velocities on plate 1 from Kriging are larger than the velocities obtained using HV-LSC-ex<sup>2</sup>. As all plate velocities have been corrected with respect to plate 1, the background velocity for this plate is 0 mm/a in both directions. A better approximation to this is obtained using HV-LSC-ex<sup>2</sup>. The interpolation using

Kriging also requires the definition of specific parameters (values obtained within a variogram analysis beforehand). A more in-depth analysis could thus lead to slightly different results and a better fit to the final LSC model. However, based on our knowledge, Kriging does not allow the usage of both horizontal velocity field components together. Therefore, the correlation between the components is ignored. The newly presented HV-LSC technique has thus an advantage over Kriging when modelling the horizontal velocity field.

## 7 Conclusion

LSC is a well-known technique to interpolate and combine different components of a dataset. It has been used previously to interpolate velocity fields; however, the advantage of using several components at the same time has, to our knowledge, not been utilized. Here, we presented an extension of the standard LSC technique by allowing the combined interpolation of the horizontal velocity field where the correlation between the components is considered. It also requires the estimation of input parameters for the HV-LSC where both components have been used together in a covariance analysis. Thus, a new approach for finding a combined correlation length is presented. In addition, we apply plate-boundary constraints, where the distance of stations on different tectonic plates can be increased. Thus, the velocities of these stations do not interfere with each other any more in a velocity field interpolation and a strong velocity gradient at the plate boundary can be obtained. Furthermore, the problem of non-stationarity due to a varying density of station networks and velocity magnitudes, is considered by applying moving variance. This new method also requires new input parameters, which have to be found in a correlation analysis. The correlation analysis is also extended to allow the combined usage of both horizontal velocity components. Additionally, a radius has to be defined in which the moving variance is estimated. This radius is defined by finding the fit of the interpolated velocity field to the input data, and a remaining correlation of 10% was shown to perform best. In general, the HV-LSC-ex<sup>2</sup> method has its advantage of obtaining an interpolated velocity field by using only the location of the stations as well as knowledge about existing plate boundaries together with the observed velocities at the GNSS stations. It is thus not depending on geophysical background information, which can induce further errors/uncertainties, and can be applied to interpret the changes in the velocity with respect to geophysical models (e.g. seismic tomography models, density models).

We apply the new techniques to a synthetic dataset that has stations distributed on four plates. Station velocities for both horizontal components are created. The interpolated velocity field shows the best fit to the synthetic data, if HV-LSC-

ex<sup>2</sup>, i.e. HV-LSC together with plate-boundary constraints and moving variance, is used. The largest increase in fit can be obtained for the horizontal velocity field when plate-boundary constraints are applied. Usage of both horizontal velocity field components and integration of the correlation between the components has been shown to perform better than standard LSC, where such a correlation is not considered. The comparison to Kriging also shows that HV-LSC-ex<sup>2</sup> performs better for interpolation and extrapolation and should be preferred when using velocity data. However, the presented method to introduce plate-boundary constraints requires knowledge about the location of the plate boundaries. Thus, only existing geological boundaries can be used and fault zones with strong velocity gradients can be missed. In those cases, it would be necessary to use the adaptive LSC by Egli et al. (2007), which can be easily applied in the presented method as all stations are interpolated together and the distance between them is only increased. In addition, a suitable radius has to be chosen carefully for the moving variance and several values should be tested to find a velocity field that is free of unrealistic velocities. Especially correlation length and moving-variance radius have a large effect on the final velocity model. Thus, suitable values have to be determined advisedly before.

The run time of the HV-LSC implemented in Python2.7 is around 25% longer compared to standard LSC implemented in Python2.7 as well. The usage of moving variance increases the run time by additional 20% using the same implementation with Python2.7 as before. Speed can be increased, if parallelization is used, which is suitable to estimate the signal at the newly defined points. The covariance analysis of both horizontal components also requires large computational power as the matrices can reach dimensions of more than 1 million entries in both directions, if an input dataset with many stations is used. Thus, the new covariance analysis is not always applicable. It would be required to split the dataset and make a covariance analysis of several sub-regions.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00190-022-01601-4>.

**Acknowledgements** We thank the editor in chief Jürgen Kusche, the handling editor Yunzhong Shen, and three anonymous reviewers for their constructive comments and remarks that have improved the manuscript. We would also like to thank Olav Vestøl for discussions about the correlation analysis, and Olivier Jamet and Corné Kreemer for discussions about the implementation of the correlation of the horizontal velocity field components into the LSC. A supplement is added to the paper, which includes further derivations, tables, and figures. An Excel spreadsheet summarizes the results from the synthetic data analysis.

**Author Contributions** RS and ML designed the research topic; RS performed the research; JL and JA contributed to the methodology; RS,



HS, and ML discussed the results. RS and JL wrote the manuscript with input from all authors.

**Data Availability** The synthetic dataset as well as the Python2.7 scripts are made available via GitHub ([https://github.com/rebkkasteffenlm/HV-LSC-ex2\\_Python2p7](https://github.com/rebkkasteffenlm/HV-LSC-ex2_Python2p7)).

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**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

- Béjar-Pizarro M, Guardiola-Albert C, García-Cárdenas RP, Herrera G, Barra A, López Molina A, Tessitore S, Staller A, Ortega-Becerril JA, García-García RP (2016) Interpolation of GPS and geological data using InSAR deformation maps: method and application to land subsidence in the Alto Guadalentín Aquifer (SE Spain). *Remote Sens* 8:965. <https://doi.org/10.3390/rs8110965>
- Bird P (2003) An updated digital model of plate boundaries. *Geochem Geophys Geosyst* 4:1027. <https://doi.org/10.1029/2001GC000252>
- Bogusz J, Klos A, Grzempowski P, Kontny B (2014) Modelling the velocity field in a regular grid in the area of Poland on the basis of the velocities of European Permanent Stations. *Pure Appl Geophys* 171:809–833. <https://doi.org/10.1007/s00024-013-0645-2>
- Caporali A, Martin S, Massironi M (2003) Average strain rate in the Italian crust inferred from a permanent GPS network-II. Strain rate versus seismicity and structural geology. *Geophys J Int* 155:254–268. <https://doi.org/10.1046/j.1365-246X.2003.02035.x>
- Chen G, Zeng A, Ming F, Jing Y (2016) Multi-quadric collocation model of horizontal crustal movement. *Solid Earth* 7:817–825. <https://doi.org/10.5194/se-7-817-2016>
- Chilès J-P, Desassis N (2018) Fifty years of kriging. In: Daya Sagar B, Cheng Q, Agterberg F (eds) *Handbook of mathematical geosciences*. Springer, Cham, pp 589–612. [https://doi.org/10.1007/978-3-319-78999-6\\_29](https://doi.org/10.1007/978-3-319-78999-6_29)
- Ching K-E, Chen K-H (2015) Tectonic effect for establishing a semi-dynamic datum in Southwest Taiwan. *Earth Planet Space* 67:207. <https://doi.org/10.1186/s40623-015-0374-0>
- Cressie NAC (2007) *Statistics for spatial data*. Wiley, New York
- Darbeheshi N, Featherstone WE (2010) A review of non-stationary spatial methods for geodetic least-squares collocation. *J Spat Sci* 55:185–204. <https://doi.org/10.1080/14498596.2010.521971>
- Dermanis A (1984) Kriging and collocation—a comparison. *Manuscr Geod* 9:159–167
- Egli R, Geiger A, Wiget A, Kahle H-G (2007) A modified least-squares collocation method for the determination of crustal deformation: first results in the Swiss Alps. *Geophys J Int* 168:1–12. <https://doi.org/10.1111/j.1365-246X.2006.03138.x>
- El-Fiky GS, Kato T, Fujii Y (1997) Distribution of vertical crustal movement rates in the Tohoku district, Japan, predicted by least-squares collocation. *J Geod* 71:432–442. <https://doi.org/10.1007/s001900050111>
- Forsberg R, Tscherning CC (1981) The use of height data in gravity field approximation by collocation. *J Geophys Res* 86:7843–7854. <https://doi.org/10.1029/JB086iB09p07843>
- Freedman D, Pisani R, Purves R (2007) *Statistics (international student edition)*. Pisani, R. Purves, 4th edn. WW Norton & Company, New York
- Fuhrmann T (2016) *Surface Displacements from fusion of geodetic measurement techniques applied to the Upper Rhine Graben Area*. Ph.D. thesis, Karlsruhe Institute of Technology, Germany
- Fuhrmann T, Caro Cuenca M, Knöpfler A, van Leijen FJ, Mayer M, Westerhaus M, Hanssen RF, Heck B (2015) Estimation of Small Surface Displacements in the Upper Rhine Graben Area from a Combined Analysis of PS-InSAR, Levelling and GNSS Data. *Geophys J Int* 203:614–631. <https://doi.org/10.1093/gji/ggv328>
- Fujii Y, Xia S (1993) Estimation of distribution of the rates of vertical crustal movements in the Tokai District with the aid of least squares prediction. *J Phys Earth* 41:239–256. <https://doi.org/10.4294/jpe1952.41.239>
- Goudarzi MA, Cocard M, Santerre R (2015) GeoStrain: An Open Source Software for Calculating Crustal Strain Rates. *Comput Geosci* 82:1–12. <https://doi.org/10.1016/j.cageo.2015.05.007>
- Haines AJ, Holt WE (1993) A Procedure for Obtaining the Complete Horizontal Motions Within Zones of Distributed Deformation from the Inversion of Strain Rate Data. *J Geophys Res* 98:12057–12082. <https://doi.org/10.1029/93JB00892>
- Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, van Kerkwijk MH, Brett M, Haldane A, Fernández del Río J, Wiebe M, Peterson P, Gérard-Marchant P, Sheppard K, Reddy T, Weckesser W, Abbasi H, Gohlke C, Oliphant TE (2020) Array programming with NumPy. *Nature* 585:357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Herzfeld UC (1992) Least-squares collocation, geophysical inverse theory and geostatistics: a bird's eye view. *Geophys J Int* 111:237–249. <https://doi.org/10.1111/j.1365-246X.1992.tb00573.x>
- Hirvonen RA (1962) On the statistical analysis of gravity anomalies. Ohio State University report 19, 24 pp
- Kahle H-G, Müller MV, Geiger A, Danuser G, Mueller S, Veis G, Billiris H, Paradissis D (1995) The strain field in northwestern Greece and the Ionian Islands: results inferred from GPS measurements. *Tectonophysics* 249:41–52. [https://doi.org/10.1016/0040-1951\(95\)00042-L](https://doi.org/10.1016/0040-1951(95)00042-L)
- Kahle H-G, Cocard M, Peter Y, Geiger A, Reilinger R, Barkaa A, Veis G (2000) GPS-derived strain rate field within the boundary zones of the Eurasian, African, and Arabian Plates. *J Geophys Res* 105:23353–23370. <https://doi.org/10.1029/2000JB900238>
- Kierulf HP, Ouassou M, Simpson MJR, Vestøl O (2013) A continuous velocity field for Norway. *J Geod* 87:337–349. <https://doi.org/10.1007/s00190-012-0603-2>
- Kotsakis C (2007) Least-squares collocation with covariance-matching constraints. *J Geod* 81:661–677. <https://doi.org/10.1007/s00190-007-0133-5>
- Kreemer C, Haines J, Holt WE, Blewitt G, Lavalée D (2000) On the determination of a global strain rate model. *Earth Planet Space* 52:765–770. <https://doi.org/10.1186/BF03352279>
- Kreemer C, Holt WE, Haines JA (2003) An integrated global model of present-day plate motions and plate boundary deformation. *Geo-*



- phys J Int 154:8–34. <https://doi.org/10.1046/j.1365-246X.2003.01917.x>
- Légrand J (2007) Champ de vitesses de l’ITRF, Propriétés cinématiques de la croûte terrestre et Condition de non rotation globale. Ph.D. thesis, Observatoire de Paris, France
- Moritz H (1978) Least-Squares collocation. *Rev Geophys Space* 16:421–430. <https://doi.org/10.1029/RG016i003p00421>
- Moritz H (1980) Advanced physical geodesy, vol 509. Herbert Wichmann Verlag, Karlsruhe, Germany & Abacus Press, Kent, Great Britain
- Murphy B, Müller S, Yurchak R (2021) GeoStat-Framework/PyKrig: v1.6.1 (v1.6.1). Zenodo. <https://doi.org/10.5281/zenodo.5380342>
- Peter YF (2000) Present-day crustal dynamics in the Adriatic-Aegean plate boundary zone inferred from continuous GPS-measurements. Ph.D. thesis, ETH Zurich, Switzerland
- Pukelsheim F (1994) The three sigma rule. *Am Stat* 48:88–91. <https://doi.org/10.2307/2684253>
- Reguzzoni M, Sansó F, Venuti G (2005) The theory of general kriging, with applications to the determination of a local geoid. *Geophys J Int* 162:303–314. <https://doi.org/10.1111/j.1365-246X.2005.02662.x>
- Robin CMI, Craymer M, Ferland R, James TS, Lapelle E, Piraszewski M, Zhao Y (2020) NAD83v70VG: a new national crustal velocity model for Canada. Geomatics Canada, Open File 0062. <https://doi.org/10.4095/327592>
- Sandwell DT, Wessel P (2016) Interpolation of 2-D vector data using constraints from elasticity. *Geophys Res Lett* 43:10703–10709. <https://doi.org/10.1002/2016GL070340>
- Shen Z-K, Wang M, Zeng Y, Wang F (2015) Optimal interpolation of spatially discretized geodetic data. *Bull Seismol Soc Am* 105(4):2117–2127. <https://doi.org/10.1785/0120140247>
- Straub C, Kahle H-G, Schindler C (1997) GPS and geologic estimates of the tectonic activity in the Marmara Sea region, NW Anatolia. *J Geophys Res* 102:27587–27601. <https://doi.org/10.1029/97JB02563>
- Tscherning CC (1999) Construction of anisotropic covariance functions using Riesz-representers. *J Geod* 73:332–336. <https://doi.org/10.1007/s001900050250>
- Tukey J (1958) Bias and confidence in not quite large samples. *Ann Math Stat* 29:614–623. <https://doi.org/10.1214/aoms/1177706647>
- Vestøl O (2006) Determination of postglacial land uplift in Fennoscandia from leveling, tide-gauges and continuous GPS stations using least squares collocation. *J Geod* 80:248–258. <https://doi.org/10.1007/s00190-006-0063-7>
- Vestøl O, Ågren J, Steffen H, Kierulf H, Tarasov L (2019) NKG2016LU: a new land uplift model for Fennoscandia and the Baltic Region. *J Geod* 93:1759–1779. <https://doi.org/10.1007/s00190-019-01280-8>