Journal of Geodesy (2019) 93:1251–1269
https://doi.org/10.1007/s00190-019-01240-2

ORIGINAL ARTICLE

Thin plate spline interpolation

Wolfgang Keller[1] · **Andrzej Borkowski**[2]

Received: 25 April 2018 / Accepted: 13 February 2019 / Published online: 27 February 2019

Abstract
The thin plate spline (TPS) is an interpolation approach that has been developed to investigate a frequently occurring problem in geosciences: the modelling of scattered data. In this paper, we carry over the concept of the thin plate spline from the plane to the sphere. To develop the spherical TPS, we utilize the idea of an elastic shell that is attributed with the bending energy and the external energy. The bending energy describes the shape of the membrane, while the external energy reflects deviations between the shell and the data to be modelled. Minimizing both energy terms leads to the variational problem with the solution in the form of the Euler–Lagrange equations. We provide the solution of the variational problem for two cases: (1) total energy minimization over the whole sphere and (2) total energy minimization over a closed region of the sphere. In case (1) we found a closed analytical solution in the form of collocation in a reproducing kernel Hilbert space. The local case (2) solution is based on a discretization of the corresponding Euler–Lagrange equation using the spherical Laplace operator. The performance of the introduced spherical TPS is demonstrated on two real world data sets. It is shown quantitative that the thin plate approach is significantly more effective than Gaussian filter in terms of the GRACE data de-striping. We also show that the TPS can be used effectively for the modelling of the vertical total electron content. It allows the reduction of the computational effort in comparison with well-established planar TPS approximation. Moreover, the harmonicity property of the TPS can be utilized to solve various issues related to Earth gravity modelling.

Keywords Thin plate spline · Approximation · Spline interpolation · Reproducing kernal Hilbert spaces · Biharmonic equation · Finite differences approximation · vTEC interpolation

1 Introduction

The interpolation and smoothing of scattered data is a frequently occurring problem in Geodesy and other Geosciences. Usually, nothing is known about the behaviour of the data in the gaps between the measurements and some hypotheses have to be used to construct a proper interpolation and smoothing algorithm. One frequently used hypothesis is that the solution is smooth. Depending on what is understood by smoothness, different algorithms can be designed. Usually, in geodesy the concept of smoothness is associated with a certain asymptotic behaviour of the spectrum of the interpolation or, in stochastic interpretation, with a behaviour of the its degree variances. Only in a few cases these concepts have a clear geometrical interpretation.

In this paper the concept of smoothness is borrowed from mechanics: imagine the measured data to be elevations above a plane or a sphere and imagine an membrane going through these elevations but is left free otherwise. Then this shell is certainly a smooth interpolation of the given data. The geometric shape of this shell is characterized by the minimum of the total bending energy. The concept of minimal bending energy is well known in the plane and leads directly to the well-established technique of thin plate spline (TPS) interpolation.
In this paper, the idea of thin plate spline interpolation will be carried over from the plane to the sphere. It will be discussed for two cases

1. The minimum will be computed over the whole surface of the sphere (global approach) and
2. the minimum will be computed only over a simply connected region of the sphere (local approach).

✉ Wolfgang Keller
wolfgang.keller@gis.uni-stuttgart.de

Andrzej Borkowski
andrzej.borkowski@upwr.edu.pl

[1] Geodetic Institute, University Stuttgart, Stuttgart, Germany

[2] Institute of Geodesy and Geoinformatics, Wroclaw University of Environmental and Life Sciences, Wrocław, Poland

For each version, plane approximation, global and local spherical approximation examples coming from GRACE data smoothing and vTEC interpolation are studied.

2 Related work

Many approximation and interpolation methods have been developed to solve specific problems of scattered data modelling. These methods were subject of a large number of studies (e.g. Franke and Nielson 1980; Franke 1982) and textbooks (e.g. Hoschek and Lasser 1992). Besides the well-known polynomial and spline interpolation, also radial basis function-based methods, such as kriging and linear prediction interpolation, are used to solve different geodetic interpolations problems.

Methods of radial basis functions represent a set of interpolants having the form,

$$z(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i R(\|\mathbf{x} - \mathbf{y_i}\|), \tag{1}$$

with the parameters α_i and a radial basis function R, that depends on the Euclidean distance between data points $\mathbf{y}$ and an interpolation point $\mathbf{x}$. This class of the interpolants also includes the thin plate splines considered in this paper.

Furthermore, this paper is focused on interpolation issues on the sphere. To apply a planar interpolation technique to the spherical problem, the data has to be projected on the plane or planar approximation has to be utilized locally.

Spherical polynomials (e.g. Sloan and Womersley 2002; Wang and Sloan 2017) are the most commonly used interpolation technique on the unit sphere. However, spherical harmonics have several disadvantages. Oscillating properties and convergence problems can be mentioned in this context. To avoid these limitations, spherical splines have been introduced. Freeden (1990) provides a comprehensive survey to this interpolation tool.

The idea of spline interpolation on the sphere is not a new one. Maybe the earliest contribution is by Wahba (1981). And a lot of contributions came from the Kaiserslautern GeoMathematics group (Freeden 1981, 1982, 1984) or (Freeden and Hermann 1986). All these contributions have in common that they are a minimal norm interpolation in a reproducing kernel Hilbert space with a kernel of the type

$$K(\boldsymbol{\xi}, \boldsymbol{\eta}) = \sum_{n=0}^{\infty} \frac{2n+1}{a_n^2} P_n(\boldsymbol{\xi} \cdot \boldsymbol{\eta}), \quad |\boldsymbol{\xi}| = |\boldsymbol{\eta}| = 1,$$

with P_n denoting the Legendre polynomials. In most cases, the choice of the spherical symbols a_n is driven by the exclusion or inclusion of certain parts of the spherical harmonics spectrum. For these choices, a geometrical interpretation is not obvious. In this paper, it will be shown that the mechanical principle of minimal bending energy in a natural way leads to the choice

$$a_n = n(n+1).$$

In the geodetic community, the spline interpolation on the sphere is known under the name of collocation in reproducing kernel Hilbert spaces. There are numerous contributions to this topic. For instance (Moritz 1987; Tscherning 1978; Forsberg and Tscherning 1981; Tscherning 2001) or (Keller 1998). In all these applications, the spherical symbols are chosen as the degree variances of a statistical auto-covariance model. The determination of these degree variances is based on the very strong assumptions of isotropy and ergodicity. Despite to the fact, that in real cases these assumptions will hardly be fulfilled, the method proved to be very successful.

According to Hubbert et al. (2015), spherical radial basis functions represent a technique that is rapidly emerging and very promising for solving interpolation problems on the surface of a sphere. In Hubbert et al. (2015), the theoretical background is provided and practical details for implementation of radial basis functions to solve spherical real world problems are given.

But already before 2015 radial basis functions were frequently used for the modelling of the gravity field. For instance in the publications (Klees et al. 2008b; Freeden and Michel 1999; Freeden and Schreiner 2005; Freeden et al. 1998) or (Schmidt et al. 2007).

In this paper, we generalize the thin plate spline, which is a variant of radial basis function interpolation to spherical thin plate spline. This problem has also been studied by Hubbert and Morton (2004). The authors propose a strategy of the planar thin plate spline for the sphere using stereographic projection and a Sobolev space on the sphere.

All the previously mentioned contributions are based on a global minimization. For regional applications a regional minimization principle is more appropriate. By developing the Euler equation for the spherical thin plate spline interpolation such a regional minimization is developed. It leads to a boundary value problem for the biharmonic equation on the sphere.

3 Planar thin plate spline approximation

To describe mathematical properties of splines and to make them more plausible, more "natural", mechanical analogies are used frequently. Here, we consider an elastic flat thin plate that underlay stress which distributes internal tension forces due to external forces (Balek and Mizera 2013). The stress causes deformation. We consider elastic deformations

only, that occur if the stress does not exceed a critical value. It means, the deformations are reversible. According to Hook's law the stress is proportional to deformation. Provided that the deformation is small, the relation between deformation and stress is linear. The physical model discussed here is a thin plate, that is represented by a function $f(x, y)$. Applying a bending energy to the plate, the upper layers of the plate are stretched and the lower ones are compressed, both according to the Hook's law. Neglecting a squeezing in the perpendicular direction to stretching (Poisson ratio equal to zero) the bending of the plate can be interpreted as univariate stretching or squeezing in the principal axis direction. The mean squared dilatation is than proportional to the square of the curvature, that can be approximated, in two-dimensional case, by the Hessian H of $f(x, y)$. The deformation energy can be obtain as the integral of $trace(H^2)$ (Balek and Mizera 2013). Therefore, if the shape of a planar thin plate is given by the function $z = f(x, y)$, the integral

$$E_{\text{int}} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}\right)^2\right] \mathrm{d}x\mathrm{d}y \tag{2}$$

represents its bending energy.

In the locations (x_i, y_i), $i = 1, \ldots, n$ the values $f(x_i, y_i)$ have to be as close as possible to the measured data z_i. This means the external energy

$$E_{ext} = \sum_{i=1}^{n}(z_i - f(x_i, y_i))^2 \tag{3}$$

has to be as small as possible. Therefore, we have to solve the following mixed target minimization problem

$$E_{\text{tot}} := E_{\text{ext}} + \alpha E_{\text{int}} \to \min \tag{4}$$

with the tuning parameter α.

To solve the above variational problem, the first variation of (4) has to vanish, $\delta E_{\text{tot}} = 0$. It yields the following associated biharmonic equation, the Euler–Lagrange equation (Eberly 2018):

$$\sum_{i=1}^{n}(z_i - f(x_i, y_i)\delta(x - x_i, y - y_i)) + \alpha\left(\frac{\partial^4 f}{\partial x^4} + 2\frac{\partial^4 f}{\partial x^2 \partial y^2} + \frac{\partial^4 f}{\partial y^4}\right) = 0, \tag{5}$$

where $\delta(x - x_i, y - y_i)$ is the Dirac delta function. The fundamental solution of this equation is the thin plate spline (Duchon 1976; Terzopoulos 1986; Eberly 2018). The Euler–Lagrange differential equation can be solved using a Green function, here in the form $f(r) = r^2 \ln(r)$. Evaluating the solution at data points $\mathbf{z}$, it can be written in the matrix form:

$$\mathbf{z} = (\mathbf{A} + \alpha\mathbf{I})\boldsymbol{\lambda} + \mathbf{N}\mathbf{d}, \tag{6}$$

where $\mathbf{A}$ consists of Green functions, $\mathbf{I}$ is the $n \times n$ identity matrix and $N_i = [1, x_i, y_i], \quad i = 1, 2, \ldots, n$ are the rows of $\mathbf{N}$. The bivariate polynomials form the null space for the internal energy. In order to generate a unique solution, the solution has to be in the orthogonal complement of the null space. The functions $\{1, x, y\}$ form a basis of the null space. The condition of being in the orthogonal complement of the null space is formulated by (7).

$$\mathbf{N}^\top\boldsymbol{\lambda} = \mathbf{0}. \tag{7}$$

The combined Eq. (6) is a linear system with two unknown vectors: $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_n]^\top$ and $\mathbf{d} = [d_{00}, d_{10}, d_{01}]^\top$.

The system of equations (6) and (7) allows to compute the unknown parameter vectors separately. Alternatively, (6) and (7) can be formulated as the following linear equation system (e.g. Borkowski and Keller 2005)

$$\begin{bmatrix} \alpha & a_{12} & a_{13} & \ldots & a_{1n} & 1 & x_1 & y_1 \\ a_{21} & \alpha & a_{23} & \ldots & a_{2n} & 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & \alpha & 1 & x_n & y_n \\ 1 & 1 & 1 & \ldots & 1 & 0 & 0 & 0 \\ x_1 & x_2 & x_3 & \ldots & x_n & 0 & 0 & 0 \\ y_1 & y_2 & y_3 & \ldots & y_n & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ d_{00} \\ d_{10} \\ d_{01} \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{8}$$

with $r_i^2 = (x - x_i)^2 + (y - y_i)^2$ and $a_{ij} = r_{ij}^2 \ln r_{ij}$; $i, j = 1, 2, \ldots, n$. The effect of the additionally orthogonality condition is the modification of the thin plate spline from

$$f(x, y) = \sum_{i=1}^{n}\lambda_i r_i^2 \ln r_i$$

to

$$f(x, y) = \sum_{i=1}^{n}\lambda_i r_i^2 \ln r_i + d_{00} + d_{10}x + d_{01}y.$$

Having measured the data z_i, the Eq. (8) allows the determination of the TPS parameters: $\lambda_i, d_{00}, d_{10}, d_{01}$. Of course, it is possible to determine the parameters λ_i and d_{ij} separately when applying block matrix elimination. For $\alpha = 0$ TPS is the interpolation spline, otherwise (8) is an approximation function controlled by the smoothing parameter α.

Finally, the TPS-value in any point can be calculated according to:

$$f(x, y) = \sum_{i=1}^{n} \lambda_i r_i^2 \ln r_i + d_{00} + d_{10}x + d_{01}y. \tag{9}$$

4 Bending energy on the sphere

In this section, the concept of bending energy will be carried over from the plane to the unit sphere. If $\mathcal{F}\{f\}$ denotes the Fourier transform of the function f, the Sobolev space $H^2(\mathbb{R}^2)$ is defined as the set of all functions f with

$$(1 + |\boldsymbol{\omega}|^2)\mathcal{F}\{f\}(\boldsymbol{\omega}) \in L^2(\mathbb{R}^2). \tag{10}$$

In $H^2(\mathbb{R}^2)$ the integration by parts simplifies to

$$\int_{\mathbb{R}^2} \frac{\partial f}{\partial x_i} \cdot g \, \mathrm{d}\mathbf{x} = -\int_{\mathbb{R}^2} f \cdot \frac{\partial g}{\partial x_i} \mathrm{d}\mathbf{x}, \quad i = 1, 2. \tag{11}$$

If (11) is applied to twice to the following expression

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial^2 f}{\partial x \partial y} \cdot \frac{\partial^2 f}{\partial x \partial y} \mathrm{d}x\mathrm{d}y,$$

once with respect to x and once with respect to y, we can conclude

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 \mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial^2 f}{\partial x \partial y} \cdot \frac{\partial^2 f}{\partial x \partial y} \mathrm{d}x\mathrm{d}y \\
&= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial f}{\partial y} \cdot \frac{\partial^3 f}{\partial x^2 \partial y} \mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} \mathrm{d}x\mathrm{d}y.
\end{aligned} \tag{12}$$

After this preparations the expression for the bending energy in the plane can be rewritten

$$\begin{aligned}
E_{int} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\frac{\partial^2 f}{\partial x^2}\right)^2 \\
&\quad + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}\right)^2 \mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} + \left(\frac{\partial^2 f}{\partial y^2}\right)^2 \mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\frac{\partial^2 d}{\partial x^2} + \frac{\partial^2 d}{\partial y^2}\right)^2 \mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (\Delta f)^2 \mathrm{d}x\mathrm{d}y.
\end{aligned} \tag{13}$$

This makes it easy to carry over the concept of bending energy from the plane to the sphere: the planar Laplace operator

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

has to be replaced by the Laplace–Beltrami operator

$$\Delta_S = \frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial}{\partial\vartheta}\right) + \frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\lambda^2}. \tag{14}$$

As a consequence, we arrive at the following thin plate spline principle on the sphere

$$\min\left\{\int_S (\Delta_S f)^2 \mathrm{d}S \mid f(\vartheta_i, \lambda_i) = z_i, \quad i = 1, \ldots, n\right\} \tag{15}$$

for data z_i measured at the locations (ϑ_i, λ_i).

5 Thin plate spline interpolation on the sphere

5.1 Reproducing kernel Sobolev spaces on the sphere

Let us denote by $C_0^\infty(S)$ the set of all infinite often differentiable functions φ on the sphere with vanishing mean value

$$\int_S \varphi \mathrm{d}S = 0. \tag{16}$$

Obviously,

$$\langle f, g\rangle := \int_S \Delta_S f \cdot \Delta_S g \mathrm{d}S \tag{17}$$

is a scalar product in $C_0^\infty(S)$. For the norm, derived from this scalar product holds

$$\|f\|^2 = \int_S (\Delta_S f)^2 \mathrm{d}S. \tag{18}$$

The completion of $C_0^\infty(S)$ in the norm (18) is a Sobolev space, which will be denoted by $H_0^{2,2}(S)$. If we denote by

$Y_{l,m}$ the fully normalized surface spherical harmonics, then the functions

$$Z_{l,m} = \frac{1}{l(l+1)} Y_{l,m} \tag{19}$$

form an orthonormal set on $H_0^{2,2}(S)$. Since $H_0^{2,2}(S)$ is a subset of $L^2(S)$ and because the $Y_{l,m}$ form a complete orthonormal system in $L^2(S)$ the functions $Z_{l,m}$ are complete in $H_0^{2,2}(S)$. Hence, $H_0^{2,2}(S)$ is separable and has a reproducing kernel

$$\begin{aligned} K(\vartheta_1, \lambda_1; \vartheta_2, \lambda_2) &= \sum_{l=1}^{\infty} \sum_{m=-l}^{l} Z_{l,m}(\vartheta_1, \lambda_1) \overline{Z_{l,m}(\vartheta_2, \lambda_2)} \\ &= \sum_{l=1}^{\infty} \frac{1}{l^2(l+1)^2} \sum_{m=-l}^{l} Y_{l,m}(\vartheta_1, \lambda_1) \overline{Y_{l,m}(\vartheta_2, \lambda_2)} \\ &= \sum_{l=1}^{\infty} \frac{2l+1}{l^2(l+1)^2} P_l(\cos\psi) \end{aligned} \tag{20}$$

with ψ denoting the spherical angle between the two arguments of the kernel

$$\cos\psi = \cos\vartheta_1 \cos\vartheta_2 + \sin\vartheta_1 \sin\vartheta_2 \cos(\lambda_1 - \lambda_2). \tag{21}$$

5.2 Thin plate spline interpolation in reproducing kernel spaces

In reproducing kernel Hilbert spaces, there is a closed solution for the minimization problem (15). In order to derive this closed solution, some subsets of $H_0^{2,2}(S)$ have to be introduced. First of all, we denote the set of all interpolating functions as

$$\begin{aligned} H_{0,z}^{2,2}(S) &:= \{u \in H^{2,2}(S) \mid u(\boldsymbol{\xi}_i) = z_i,\ i = 1, \ldots, n\} \\ \boldsymbol{\xi}_i &= \begin{bmatrix} \sin(\vartheta_i)\cos(\lambda_i) \\ \sin(\vartheta_i)\sin(\lambda_i) \\ \cos(\vartheta_i) \end{bmatrix}. \end{aligned} \tag{22}$$

The set of all functions with zero values in the interpolation nodes will be denoted by

$$H_{0,0}^{2,2}(S) := \{u \in H^{2,2}(S) \mid u(\boldsymbol{\xi}_i) = 0,\ i = 1, \ldots, n\}. \tag{23}$$

The first observation is that the linear span of the kernel functions at the interpolation nodes is the orthogonal complement of $H_{0,0}^{2,2}$: If we denote the linear span by

$$V = span\{K(\boldsymbol{\xi}_1, \bullet), \ldots, K(\boldsymbol{\xi}_n, \bullet)\} \tag{24}$$

than for every $u \in H_{0,0}^{2,2}$ and for every $v \in V$ holds

$$\begin{aligned} \langle u, v \rangle &= \left\langle u, \sum_{i=1}^{n} \alpha_i K(\boldsymbol{\xi}_i, \bullet) \right\rangle \\ &= \sum_{i=1}^{n} \alpha_i \langle u, K(\boldsymbol{\xi}_i, \bullet) \rangle \\ &= \sum_{i=1}^{n} \alpha_i u(\boldsymbol{\xi}_i) \\ &= 0 \end{aligned}$$

Both V and $H_{0,0}^{2,2}(S)$ are subspaces of $H_0^{2,2}(S)$. While V is m-dimensional, $H_{0,0}^{2,2}(S)$ is of infinite dimension.

The thin plate spline interpolator is that element of $H_z^{2,2}(S)$ with the smallest norm, i.e. that element with the smallest distance from the zero element 0 of $H_0^{2,2}(S)$. From the geometry of the problem (see Fig. 1), it is clear that the orthogonal projection of the minimal norm interpolator to the sub-space $H_{0,0}^{2,2}(S)$ is the zero element 0. Since V is orthogonal to $H_{0,0}^{2,2}(S)$, the minimal norm interpolator is the intersection of the sets V and $H_{0,z}^{2,2}(S)$. Because the minimal norm interpolator f is an element of V it has the representation

$$f = \sum_{i=1}^{n} \alpha_i K(\boldsymbol{\xi}_i, \bullet)$$

and because it is also an element of the interpolation space $H_{0,z}^{2,2}(S)$ it has to fulfil the conditions

$$z_j = f(\boldsymbol{\xi}_j) = \sum_{i=1}^{n} \alpha_i K(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j), \quad j = 1, \ldots, n. \tag{25}$$

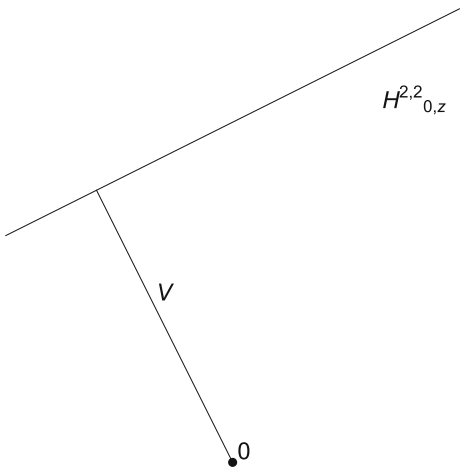

Fig. 1 Sketch of the minimal norm interpolation problem

This is a system of linear equations for the weights α_i with the solution

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} K(\xi_1, \xi_1) & \dots & K(\xi_1, \xi_n) \\ & \ddots & \\ K(\xi_n, \xi_1) & \dots & K(\xi_n, \xi_n) \end{bmatrix}^{-1} \cdot \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}. \tag{26}$$

Putting everything together, we find the interpolating spherical thin plate spline by

$$\begin{aligned} f &= [K(\xi_1, \bullet), \dots, K(\xi_1, \bullet)]\boldsymbol{\alpha} \\ &= [K(\xi_1, \bullet), \dots, K(\xi_1, \bullet)] \\ &\quad \begin{bmatrix} K(\xi_1, \xi_1) & \dots & K(\xi_1, \xi_n) \\ & \ddots & \\ K(\xi_n, \xi_1) & \dots & K(\xi_n, \xi_n) \end{bmatrix}^{-1} \cdot \mathbf{z}. \end{aligned} \tag{27}$$

This is the well-known collocation in reproducing kernel Hilbert spaces.

All in all, the spherical thin plate spline is a linear combination of kernel functions

$$f = \sum_{i=1}^{m} \alpha_i K(\xi_i, \bullet)$$

The kernel functions themselves are series of Legendre polynomials

$$K(\xi_i, \bullet) = \sum_{n=0}^{\infty} \frac{2n+1}{(n(n+1))^2} P_n(\xi_i \cdot \bullet),$$

which means that the spherical thin plate spline is also a series of Legendre polynomials

$$f(\bullet) = \sum_{n=0}^{\infty} \frac{2n+1}{(n(n+1))^2} \left(\sum_{i=1}^{m} \alpha_i P_n(\xi_i \cdot \bullet) \right).$$

Unfortunately, there is no closed expression for the series expansion of the kernel. For numerical computations, the kernel has to be truncated. For local approximations, as discussed in this paper, a truncation $n = 40$ was sufficiently accurate. For spherical distances smaller than $4\,^\circ$, series truncated at $n = 40$ is practically identical to the series truncated at $n = 200$.

So far, it was always supposed that the data $\mathbf{z}$ are samples from a function, belonging to $H_0^{2,2}(S)$. If this was the case, the condition

$$\sum_{i=1}^{m} z_i = 0$$

would automatically be fulfilled. If for practical applications the condition is violated and a rudimentary remove-restore technique has to be applied: the mean value of the data has to be subtracted prior to interpolation and added to the interpolation function once it is computed.

For practical applications, the fact is important that for the spherical thin plate spline a harmonic continuation can easily be found

$$f(r, \bullet) = \sum_{n=0}^{\infty} \frac{2n+1}{(n(n+1))^2} r^{-n-1} \left(\sum_{i=1}^{m} \alpha_i P_n(\xi_i \cdot \bullet) \right).$$

So far, only the exact interpolation has been discussed. If an additional smoothing has to be carried out, the Eq. (26) changes into

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \left[\begin{bmatrix} K(\xi_1, \xi_1) & \dots & K(\xi_1, \xi_n) \\ & \ddots & \\ K(\xi_n, \xi_1) & \dots & K(\xi_n, \xi_n) \end{bmatrix} + \gamma \mathbf{I} \right]^{-1} \cdot \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}. \tag{28}$$

The bigger the smoothing parameter γ, the smaller the weights $\boldsymbol{\alpha}$ and consequently the smoother the solution f.

5.3 Spherical thin plate interpolation over simply connected regions of the sphere

So far the internal energy minimization was always carried out over the complete surface of the sphere. Even in the case that the data is given only in a small subregion of the sphere. For the case of local data coverage, with the data given only at the boundary of the region, it would be more adequate, to perform the minimization only over the region of interest.

In this special case, the reproducing kernel property cannot longer be used and instead of the direct solution, the solution of the corresponding Euler equations has to be found. The Euler equation will include second order derivatives. If the derivatives are understood in the generalized sense, we could stay in $H_0^{2,2}(S.)$. But in order not to complicate the derivations, we switch to $C^2(S_0)$, which has no practical consequences. This means that from now on we work in $C^2(S_0)$ instead of $H_0^{2,2}(S)$.

For this reason, the first variation of the minimization functional

$$I(u) := \int_{S_0} (\Delta_S u)^2 \, \mathrm{d}S \tag{29}$$

has to be found. Using the definition of the first variation we get

$$\delta I(u,h) := \lim_{\epsilon\to 0} \frac{I(u+\epsilon h) - I(u)}{\epsilon} \tag{30}$$

$$= \lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_{S_0} (\Delta_S u)^2 + 2\epsilon \Delta_S u \cdot \Delta_S h + \epsilon^2 (\Delta_S h)^2 - (\Delta_S u)^2 \mathrm{d}S \tag{31}$$

$$= 2\int_{S_0} \Delta_S u \cdot \Delta_S h \mathrm{d}S + \lim_{\epsilon\to 0} \epsilon \int_{S_0} (\Delta_S h)^2 \mathrm{d}S \tag{32}$$

$$= 2\int_{S_0} \Delta_S u \cdot \Delta_S h \mathrm{d}S. \tag{33}$$

The Euler equations for the minimization problem ask for a function u such, that the first variation disappears for every function h:

$$0 = \delta I(u,h) = \int_{S_0} \Delta_S u \cdot \Delta_S h \mathrm{d}S, \quad \forall h. \tag{34}$$

If we insert the spherical harmonics expansions for the functions u', h', which coincide with u, h on S_0 and are zero elsewhere,

$$u' = \sum_{l,m} u_{l,m} Y_{l,m}, \quad h' = \sum_{p,q} h_{p,q} Y_{p,q}, \tag{35}$$

we obtain

$$\begin{aligned} 0 &= \int_{S_0} \Delta_S u \cdot \Delta_S h \mathrm{d}S \\ &= \int_S \sum_{l,m} l(l+1) u_{l,m} Y_{l,m} \cdot \sum_{p,q} p(p+1) h_{p,q} Y_{p,q} \mathrm{d}S \\ &= \sum_{l,m} l^2 (l+1)^2 u_{l,m} h_{l,m} \\ &= \int_S \sum_{l,m} l^2 (l+1)^2 u_{l,m} Y_{l,m} \cdot \sum_{p,q} h_{p,q} Y_{p,q} \mathrm{d}S \\ &= \int_S \Delta_S^2 u' \cdot h' \mathrm{d}S \\ &= \int_{S_0} \Delta_S^2 u \cdot h \mathrm{d}S. \end{aligned} \tag{36}$$

Since the equation has to hold for every function h, this is only possible for

$$\Delta_S^2 u(x) = 0, \quad x \in S_0. \tag{37}$$

Equation (37) has to be supplemented by the boundary conditions

$$u|_{\partial S_0} = f, \quad \frac{\partial u}{\partial n}|_{\partial S_0} = g. \tag{38}$$

This means the regional thin plate spline approximation is the solution of a biharmonic boundary value problem on a subset $S_0 \subset S$.

Since the normal derivative of the unknown function u is unknown, a homogeneous normal derivative boundary condition is used. This facilitates the increase of smoothness of the solution.

5.4 Discretization of biharmonic operator

For an arbitrary simply connected subset S_0 of the surface S of the sphere an analytic solution of (37), (38) is impossible. A numerical approximate solution has to be found. For the discretization, we write (37), (38) as a cascaded problem for the Laplace–Beltramo operator.

$$\Delta_S v = 0 \tag{39}$$

$$\Delta_S u = v \tag{40}$$

$$u|_{\partial S_0} = f \tag{41}$$

$$\frac{\partial u}{\partial n}|_{\partial S_0} = 0. \tag{42}$$

For the discretization of the spherical Laplace operator, a equiangular grid in S_0 is constructed

$$G = \left\{ \boldsymbol{\xi}_{i,j} = \begin{bmatrix} \sin(ih_\vartheta)\cos(jh_\lambda) \\ \sin(ih_\vartheta)\sin(jh_\lambda) \\ \cos(ih_\vartheta) \end{bmatrix} \right. \left. \mid i = 0,\ldots,N-1,\ j = 0,\ldots,M-1 \right\} \bigcap S_0 \tag{43}$$

with

$$h_\vartheta = \frac{\pi}{N}, \quad h_\lambda = \frac{2\pi}{M}. \tag{44}$$

As the next step, the Laplace operator of v in a grid point $\boldsymbol{\xi}_{i,j}$ is to be approximated by the weighted mean of the values of v in the neighbouring grid points

$$\Delta_S v(\boldsymbol{\xi}_{i,j}) \approx \sum_{l=-1}^{1} \sum_{k=-1}^{1} a_{k,l} v(\boldsymbol{\xi}_{i+k,j+l}). \tag{45}$$

Then the values $v(\boldsymbol{\xi}_{i+k,j+l})$ are replaced by their Taylor expansions around $\boldsymbol{\xi}_{i,j}$

$$v(\boldsymbol{\xi}_{i+k,j+l}) = v(\boldsymbol{\xi}_{i,j}) + kh_\vartheta \frac{\partial v}{\partial \vartheta}(\boldsymbol{\xi}_{i,j}) + lh_\lambda \frac{\partial v}{\partial \vartheta}(\boldsymbol{\xi}_{i,j}) + \cdots \tag{46}$$

This leads to the following equation

$$\begin{aligned}
&\Delta_S v(\boldsymbol{\xi}_{i,j}) \\
&\quad = \left[\frac{1}{\sin(\vartheta)}\frac{\partial}{\partial\vartheta}\left(\sin(\vartheta)\frac{\partial v}{\partial\vartheta}\right) + \frac{1}{\sin^2(\vartheta)}\frac{\partial^2 v}{\partial\lambda^2}\right](\boldsymbol{\xi}_{i,j}) \qquad (47)
\end{aligned}$$

$$\begin{aligned}
&= v(\boldsymbol{\xi}_{i,j}) \sum_{k=-1}^{1}\sum_{l=-1}^{1} a_{k,l} \\
&\quad + h_\lambda \frac{\partial v}{\partial\lambda}(\boldsymbol{\xi}_{i,j}) \\
&\qquad \left(-a_{-1,-1} + a_{-1,1} - a_{0,-1} + a_{0,1} - a_{1,-1} + a_{1,1}\right) \\
&\quad + \frac{h_\lambda^2}{2}\frac{\partial^2 v}{\partial\lambda^2}(\boldsymbol{\xi}_{i,j}) \\
&\qquad \left(a_{-1,-1} + a_{-1,1} + a_{0,-1} + a_{0,1} + a_{1,-1} + a_{1,1}\right) \\
&\quad + h_\vartheta \frac{\partial v}{\partial\vartheta}(\boldsymbol{\xi}_{i,j}) \\
&\qquad \left(-a_{-1,-1} - a_{-1,0} - a_{-1,1} + a_{1,-1} + a_{1,0} + a_{1,1}\right) \\
&\quad + \frac{h_\vartheta^2}{2}\frac{\partial^2 v}{\partial\vartheta^2}(\boldsymbol{\xi}_{i,j}) \\
&\qquad \left(a_{-1,-1} + a_{-1,0} + a_{-1,1} + a_{1,-1} + a_{1,0} + a_{1,1}\right) \\
&\quad + \text{mixed terms} \qquad (48)
\end{aligned}$$

A comparison of the partial derivatives of v on the left and on the right side of (47) yields the following linear equations for the weights $a_{k,l}$

$$\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
-h_\lambda & 0 & h_\lambda & -h_\lambda & 0 & h_\lambda & -h_\lambda & 0 & h_\lambda \\
\frac{h_\lambda^2}{2} & 0 & \frac{h_\lambda^2}{2} & \frac{h_\lambda^2}{2} & 0 & \frac{h_\lambda^2}{2} & \frac{h_\lambda^2}{2} & 0 & \frac{h_\lambda^2}{2} \\
-h_\vartheta & -h_\vartheta & -h_\vartheta & 0 & 0 & 0 & h_\vartheta & h_\vartheta & h_\vartheta \\
h_\vartheta h_\lambda & 0 & -h_\vartheta h_\lambda & 0 & 0 & 0 & -h_\vartheta h_\lambda & 0 & h_\vartheta h_\lambda \\
-\frac{h_\vartheta h_\lambda^2}{2} & 0 & -\frac{h_\vartheta h_\lambda^2}{2} & 0 & 0 & 0 & \frac{h_\vartheta h_\lambda^2}{2} & 0 & \frac{h_\vartheta h_\lambda^2}{2} \\
\frac{h_\vartheta^2}{2} & \frac{h_\vartheta^2}{2} & \frac{h_\vartheta^2}{2} & 0 & 0 & 0 & \frac{h_\vartheta^2}{2} & \frac{h_\vartheta^2}{2} & \frac{h_\vartheta^2}{2} \\
-\frac{h_\vartheta^2 h_\lambda}{2} & 0 & \frac{h_\vartheta^2 h_\lambda}{2} & 0 & 0 & 0 & -\frac{h_\vartheta^2 h_\lambda}{2} & 0 & \frac{h_\vartheta^2 h_\lambda}{2} \\
\frac{h_\vartheta^2 h_\lambda^2}{4} & 0 & \frac{h_\vartheta^2 h_\lambda^2}{4} & 0 & 0 & 0 & \frac{h_\vartheta^2 h_\lambda^2}{4} & 0 & \frac{h_\vartheta^2 h_\lambda^2}{4}
\end{bmatrix}
\cdot
\begin{bmatrix} a_{-1,-1} \\ a_{-1,0} \\ a_{-1,1} \\ a_{0,-1} \\ a_{0,0} \\ a_{0,1} \\ a_{1,-1} \\ a_{1,0} \\ a_{1,1} \end{bmatrix}
=
\begin{bmatrix} 0 \\ 0 \\ \sin^{-2}\vartheta \\ \cot\vartheta \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad (49)$$

with the solution

$$\begin{bmatrix} a_{-1,-1} & a_{-1,0} & a_{-1,1} \\ a_{0,-1} & a_{0,0} & a_{0,1} \\ a_{1,-1} & a_{1,0} & a_{1,1} \end{bmatrix}
= \begin{bmatrix} 0 & \frac{2-h_\vartheta\cot(\vartheta)}{2h_\vartheta^2} & 0 \\ \frac{\csc^2(\vartheta)}{h_\lambda^2} & -\frac{2\csc^2(\vartheta)}{h_\lambda^2} - \frac{2}{h_\vartheta^2} & \frac{\csc^2(\vartheta)}{h_\lambda^2} \\ 0 & \frac{h_\vartheta\cot(\vartheta)+2}{2h_\vartheta^2} & 0 \end{bmatrix}. \qquad (50)$$

This is the typical structure of a 5-point difference operator for the approximation of the Laplace operator. In contrast to the planar case on the sphere, the weights depend on the co-latitude ϑ.

The discretization of the biharmonic operator is obtained, if in the discrete Laplace operator

$$\Delta_S v(\boldsymbol{\xi}_{i,j}) \approx \sum_{l=-1}^{1}\sum_{k=-1}^{1} a_{k,l} v(\boldsymbol{\xi}_{i+k,j+l}) \qquad (51)$$

each value $v(\boldsymbol{\xi}_{p,q})$ is replaced by the corresponding 5-point differences star

$$v(\boldsymbol{\xi}_{p,q}) = \sum_{l=-1}^{1}\sum_{k=-1}^{1} a_{k,l} u(\boldsymbol{\xi}_{p+k,q+l}). \qquad (52)$$

This leads to the 13-point difference star for the biharmonic operator

$$\Delta_S^2 u(\boldsymbol{\xi}_{i,j}) = \sum_{k=-2}^{2}\sum_{l=-2}^{2} b_{k,l} u(\boldsymbol{\xi}_{i+k,j+l}) \qquad (53)$$

with the coefficients $b_{k,l}$ given in the "Appendix".

5.5 Treatment of the boundary ∂S_0

Inside the region S_0, the solution of the biharmonic equation is unknown. Only in the grid points $\boldsymbol{\xi}_{ij}$ an approximate solution can be computed by a finite-differences approximation of the biharmonic operator. There is a difference between those grid points where the complete stencil is inside the region S_0 and those where a part of the stencil is outside the region S_0. Both situations are discussed in the following. If the approximate solution of the biharmonic equation in the grid-point $\boldsymbol{\xi}_{l,k}$ is denoted by $u_{l,k} \approx u(\boldsymbol{\xi}_{l,k})$, then each inner grid-point $\boldsymbol{\xi}_{i,j}$ generates a linear equation for the unknown values $u_{l,k}$:

$$\sum_{l=-2}^{2}\sum_{k=-2}^{2} b_{l,k} u_{i+l,j+k} = 0. \qquad (54)$$

The situation is more complicated for grid points $\boldsymbol{\xi}_{i,j}$, which are close to the boundary ∂S_0. Because some of the neigh-

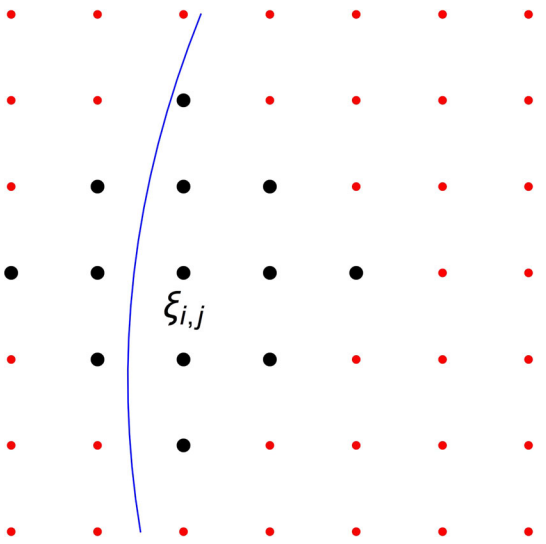

Fig. 2 Biharmonic difference star for a point $\xi_{i,j}$ close to the boundary. Blue line represents the boundary and black dots are the points influenced by the difference star

bouring points $\xi_{i+l,j+k}$ might lie outside S_0 (see Fig. 2). Obviously, for some points left to the blue boundary values of the solution are needed, which are a priori not available, since the given data are the boundary values. Therefore, the reproducing kernel thin plate spline interpolation is used to predict the missing values from the values given on the boundary.

In the end, for each grid point in the interior, one linear equation can be derived. The resulting system is sparse and numerically stable and provides a solution, which is not strongly affected by measurement errors and represents a solution where the minimization of the internal energy is not carried out over the whole sphere, but only over the subset S_0 of the sphere.

6 Applications

6.1 Between ground-tracks interpolation of GRACE solutions

The discussion of spherical harmonics interpolation from scattered data has a long tradition in geodesy. Basically, two problems have to be addressed:

- The aliasing effect due to the discrete sampling of a non band-limited signal and
- the effect of irregular sampling.

In Sansó (1990) it is shown that if the number of sampling points tends to infinity, for a regular sampling the aliasing error tends to zero. But for an irregular distribution of data points a significant bias in the estimated spherical harmonics coefficients remains. Due to that bias, the data in the sampling points is reproduced, while in the gaps between the sampling points large errors are generated.

This problem got a new attention, when gravity field solutions from the 61/4 repeat orbit of GRACE were computed. Wagner et al. (2006) showed that due to the sparse sampling during the repeat mode a resolution only up to degree and order $L = 30$ is possible. It was presumed that the loss of orthogonality of Legendre functions and sine/cosine function for this irregular grid is responsible for the degradation. This assumption was tested in Weigelt et al. (2009), with the result that the number of sampling points in a 1° stripe does not significantly change between repeat and non- repeat orbits. This confirms the result of Sansó (1990) that the irregular sampling contributes to the degradation.

If GRACE is in a repeat orbit the distance between adjacent ground tracks can be rather large (see Fig. 3). Implicitly, the usual spherical harmonics solution does an interpolation in East–West direction by trigonometric polynomials. Because the polynomial interpolation tends to an overshooting with increasing distance from the data points, the quality of this interpolation is rather poor and contributes to the undesired striping in the GRACE solution (Klees et al. 2008a; Weigelt et al. 2009; Zheng et al. 2012). Of course, the irregular data sampling is not the only source of the striping. Much more important are inaccurate background models or calibration errors.

The only place where the GRACE data is given is along the ground tracks. Using the energy-balance approach, the potential differences at orbital altitude are converted into geoid heights, using the reproducing kernel collocation (as a modification of the reproducing kernel thin plate interpolation), as already discussed in Sect. 5.5. That means that the geoid variations are only known along the ground tracks. In between the ground tracks nothing is known, it is only reasonably to assume that the solution is smooth. Under this assumption, a thin plate spline interpolation seems to be an appropriate tool.

For the test of the thin plate spline interpolation, the spherical harmonics coefficients $c_{l,m}^{(i)}, s_{l,m}^{(i)}, \quad l = 0, \ldots, 90, m = 0, \ldots, 90, i = 1, \ldots, 12$ from the GFZ GRACE solutions for the year 2004 were selected. By averaging an annual mean of the coefficients $\overline{c}_{l,m}, \overline{s}_{l,m}, \quad l = 0, \ldots, 90, m = 0, \ldots, 90$ was computed. Followed by the derivation of the monthly change for the month September 2004: $\delta c_{l,m} = c_{l,m}^{(9)} - \overline{c}_{l,m}, \delta s_{l,m} = s_{l,m}^{(9)} - \overline{c}_{l,m}, \quad l, m = 0, \ldots, 90.$

Using the coefficient $c_{2,0}^{(9)}$ only, a one month orbit was computed with a temporal spacing of 10 s. In the ground-track points $(\vartheta, \lambda_j), \quad j = 1, \ldots, 259{,}200$ of this orbit, the changes in geoidal heights for September 2004 were computed, using the coefficients $\delta c_{l,m}, \delta s_{l,m}, \quad l, m = 0, \ldots, 90.$

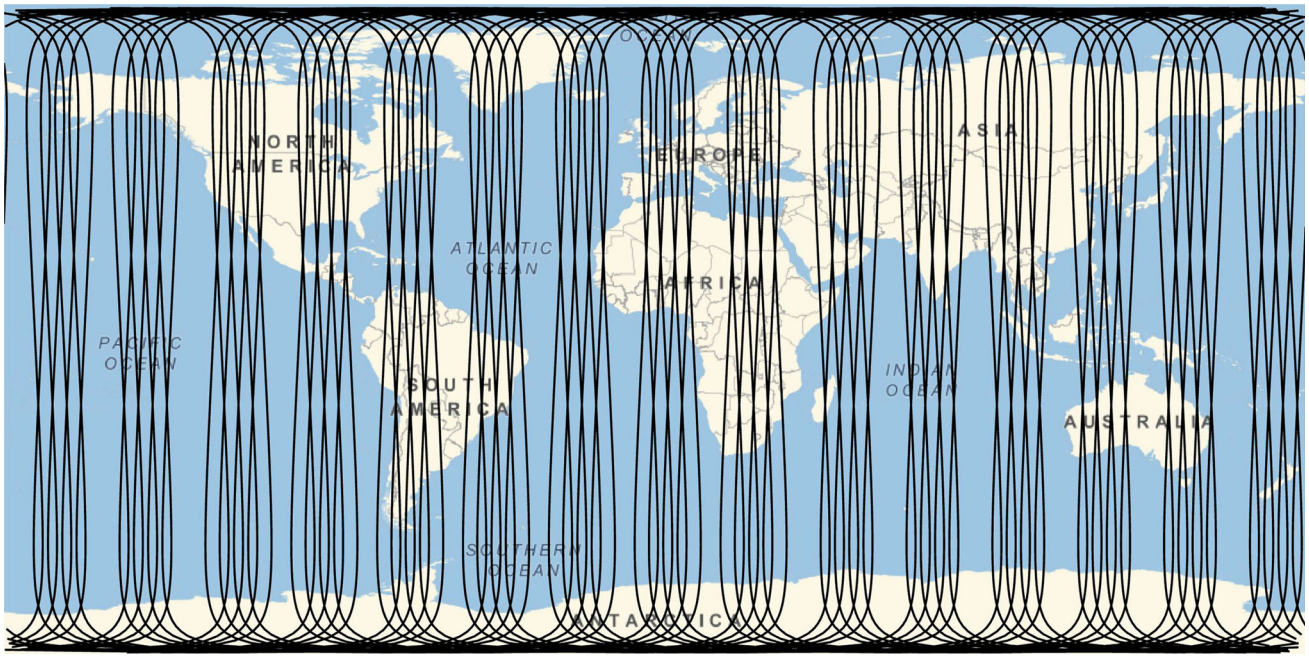

Fig. 3 GRACE ground track for the 61/4 repeat orbit in September 2004

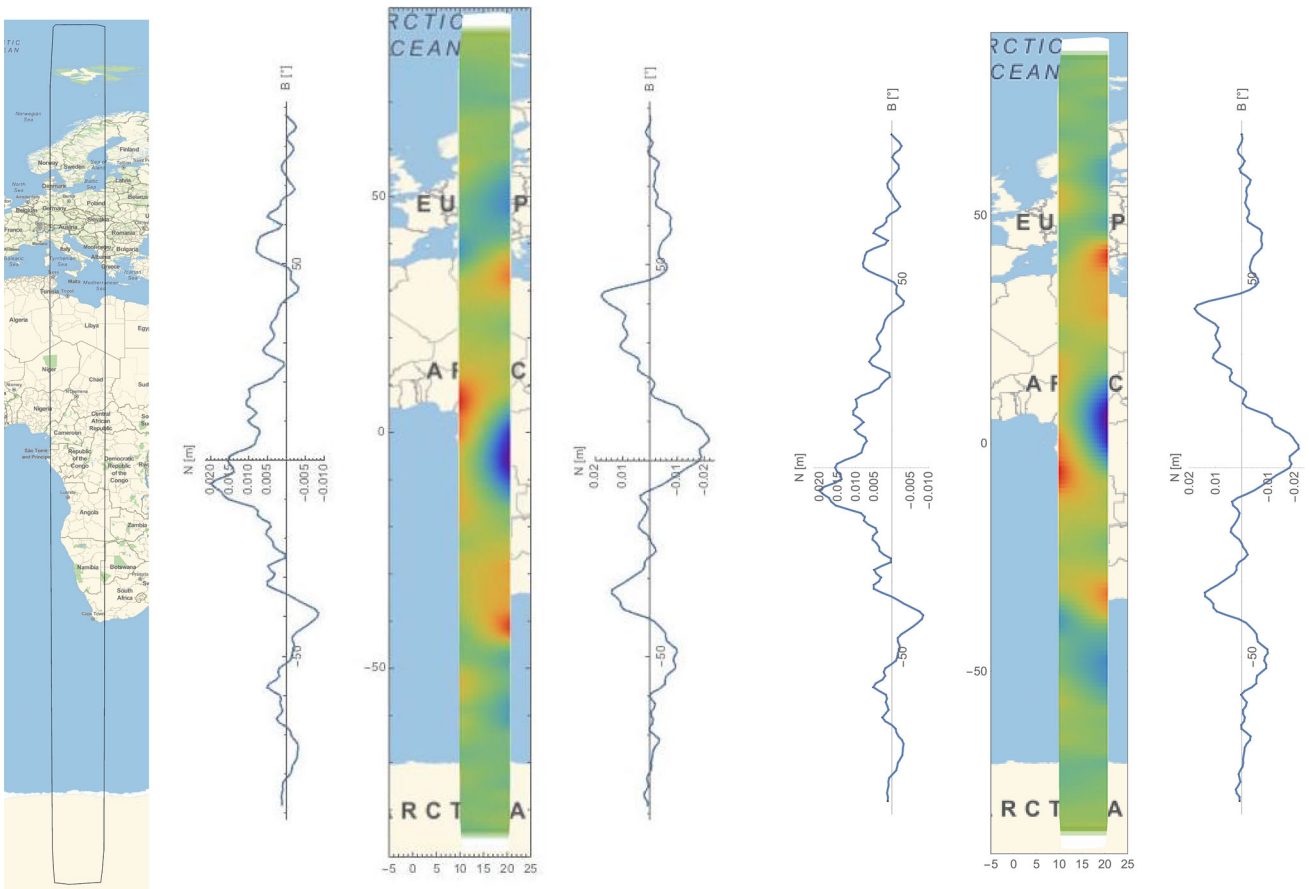

Fig. 4 Interpolation area (left), spherical thin plate spline solution (middle) and biharmonic solution (right)

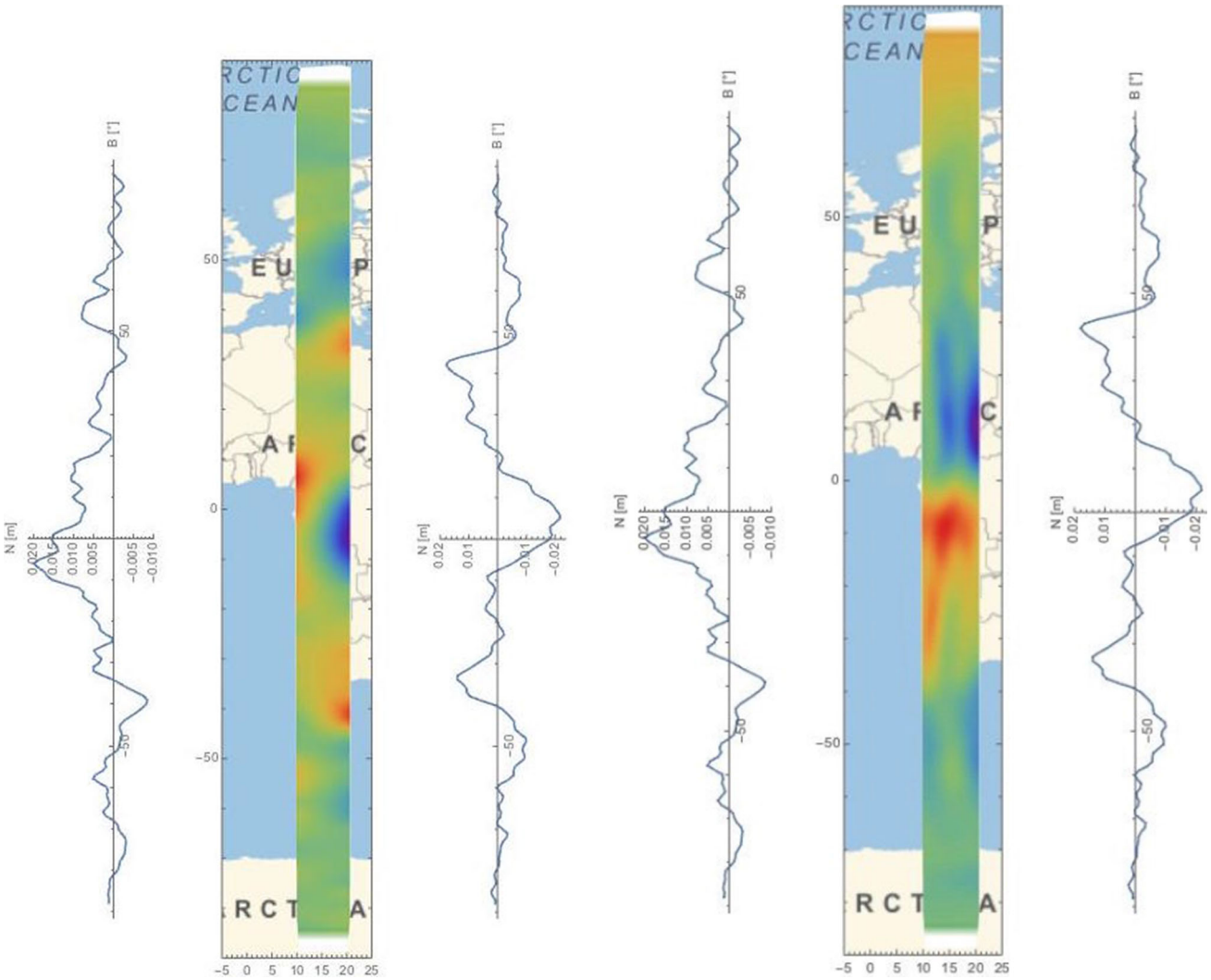

Fig. 5 Biharmonic solution (left) and Gaussian smoothed global solution (right)

Two adjacent ground tracks were considered the boundaries of a region S_0 of the surface of the Earth and a thin plate spline interpolation from the boundary, where the data is given, to the interior was carried out and the result is compared to the global Gaussian smoothed standard solution and to the global application of the Kusche filter (Kusche 2007).

As a test area we use the ground-track gap over western Africa and from the geoid variations along the boundaries we interpolate into the interior of the region. Both by spherical thin plate spline interpolation and by the solution of the biharmonic equation. It is clearly visible (see Fig. 4) that both solutions follow closely the values on the East- and on the West boundary. Visually, the biharmonic solution is slightly smoother than the thin plate spline solution.

Both solutions are significantly better than a 500 km Gaussian smoothed global solution, which is than restricted to the area. From Fig. 5, it is visible that the Gaussian smoothed global solution is much stripier and that it does not pick up the data on the boundaries very well. In Fig. 6, the biharmonic solution and the Kusche-smoothed solution are compared. The biharmonic and the Kusche-smoothed solution show a similar behaviour. That confirms the applicability of the biharmonic interpolation.

The fact that the biharmonic solution outperforms the Gaussian smoothing does not depend on the location of the area S_0 on the sphere. Also for the ground-track gap over Japan and Australia a similar behaviour can be observed (see Fig. 7).

Besides the purely visual assessment of de-striping performance of the biharmonic solution and the Gaussian smoothed solution, also quantitative performance measures can be given. Here, the wavelet decompositions of the two solution can be used.

A direction sensitive wavelet transform as for instance developed in Antoine et al. (2002) or Demanet and Vandergheynst (2003) is essentially the two-dimensional continuous wavelet transform on $\mathbb{R}^2$

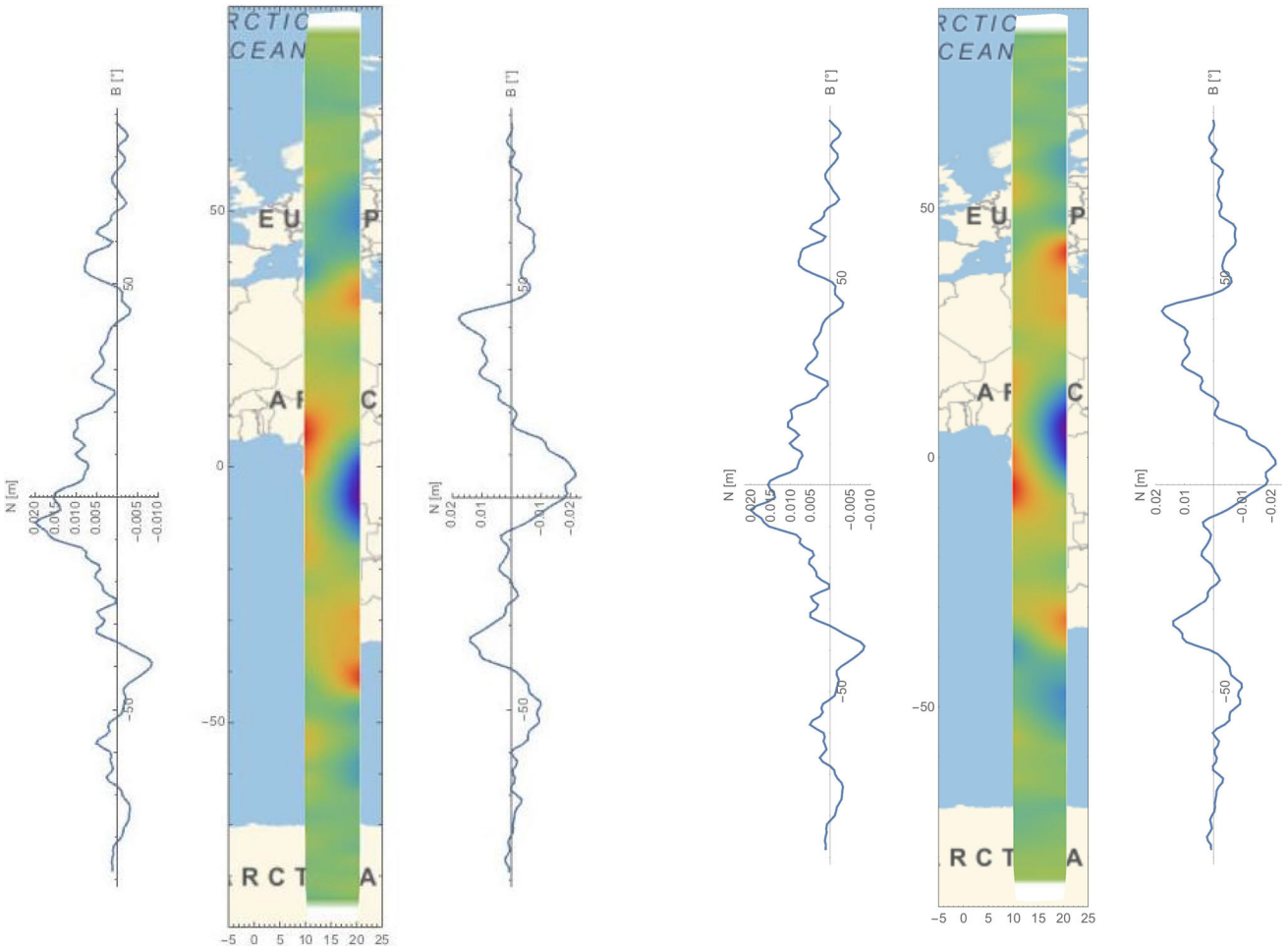

Fig. 6 Biharmonic solution (left) and Kusche-smoothed global solution (right)

$$\mathcal{W}\{f\}(a,\vartheta,\mathbf{b}) = \frac{1}{\sqrt{c_\psi}} \int_{\mathbb{R}^2} \frac{1}{a} \psi\left(\mathbf{R}(\vartheta)(\mathbf{x}-\mathbf{b})\right) \cdot f(\mathbf{x}) \mathrm{d}\mathbf{x}, \tag{55}$$

projected to the sphere by inverse stereographic projection. This wavelet transform provides information not only about the location $\mathbf{b}$ and the scale a of dominant signal features but also about their orientation ϑ. Because the horizontal extension of the signal is much smaller than its vertical extension the azimuthal resolution of the directional wavelet analysis is very limited. For a coarse assessment of the anisotropy it is sufficient to consider the ratio of the energy contained in the vertical features and the energy contained in the horizontal features. Therefore, the much simpler two-dimensional discrete wavelet transform was applied.

If we denote by $\mathbf{c}^{(0)}$ either the biharmonic or the Gaussian smoothed solution, its wavelet decomposition yields

$$\mathbf{c}^{(j+1)} = H_R H_C \mathbf{c}^{(j)}, \quad \mathbf{d}_1^{(j+1)} = G_R H_C \mathbf{c}^{(j)}, \tag{56}$$

$$\mathbf{d}_2^{(j+1)} = H_R G_C \mathbf{c}^{(j)}, \quad \mathbf{d}_3^{(j+1)} = G_R G_C \mathbf{c}^{(j)}. \tag{57}$$

Here, H, G are the smoothing and the differencing operators of Mallat's algorithm, respectively (e.g. Keller 2004). The indices R, C indicate the application of these operators to the rows and to the columns. Due to the small number of data points in the rows, the Haar wavelet transform was applied, with the smoothing and differencing operators according to

$$H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \dots & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \dots & 0 & 0 \\ & & & \ddots & & \\ \frac{1}{\sqrt{2}} & 0 & \dots & & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{58}$$

and

$$G = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & \dots & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & \dots & 0 & 0 \\ & & & \ddots & & \\ -\frac{1}{\sqrt{2}} & 0 & \dots & & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{59}$$

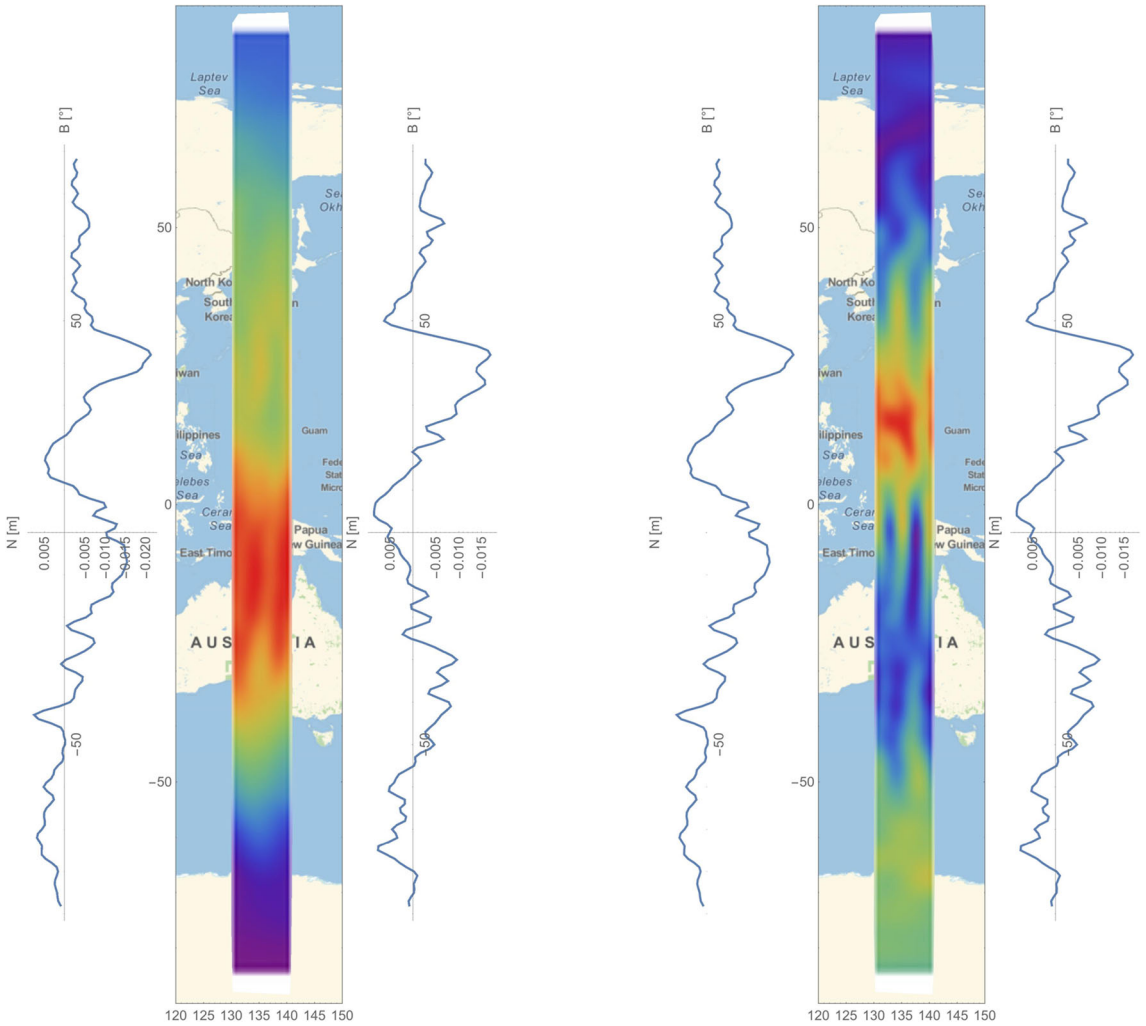

Fig. 7 Biharmonic solution (left) and Gaussian smoothed global solution (right)

On the scale j, the East–West and the North–South features are pronounced in $\mathbf{d}_2^{(j)}$ and in $\mathbf{d}_1^{(j)}$, respectively. Hence

$$a_j = \frac{\mathrm{Var}(\mathbf{d}_2^{(j)})}{\mathrm{Var}(\mathbf{d}_1^{(j)})} \tag{60}$$

measures the anisotropy on scale j. An a_j value close to one means isotropy on scale j, bigger values indicate vertical, smaller values indicate horizontal features. In Fig. 8, the isotropy measures both for the biharmonic solution and for the Gaussian smoothed solution are displayed. For both solutions, the strongest anisotropy is on scale 4, which corresponds to a scale size of about 16°. But the anisotropy for the Gaussian smoothed solution is four times bigger than for the biharmonic solution, which indicates that the biharmonic solution provides a much more effective de-striping.

6.2 Modelling of the total electron content

Ionospheric delay modelling is essential for precise Global Navigation Satellite System (GNSS) positioning. In order to calculate very accurate ionospheric corrections, reliable information about the parameters of the ionosphere such as total electron content (TEC) is needed. There exist various global, regional and local TEC models, mostly characterized by low spatial and temporal resolution. Recently, a high-accuracy regional model has been introduced by Krypiak-Gregorczyk et al. (2017). This model is based exclusively on precise un-differenced multi-GNSS carrier phase data and planar TPS interpolation. It is assumed that the geometry-free linear combination $L4$ of dual-frequency carrier phase measurements consists of a slant ionospheric delay ΔI and carrier phase bias B (Leick et al. 2015),

$$L4_r^s = L1_r^s - L2_r^s = B_r^s - u\Delta I_r^s \tag{61}$$

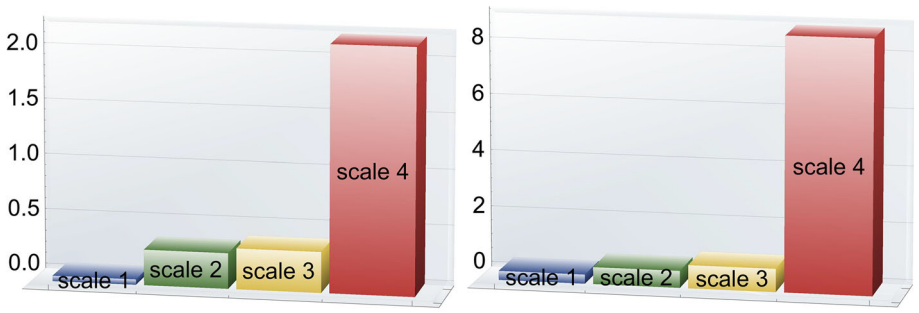

Fig. 8 Anisotropy measures for the biharmonic/Kusche-smoothed solution(left) and Gaussian smoothed global solution (right)

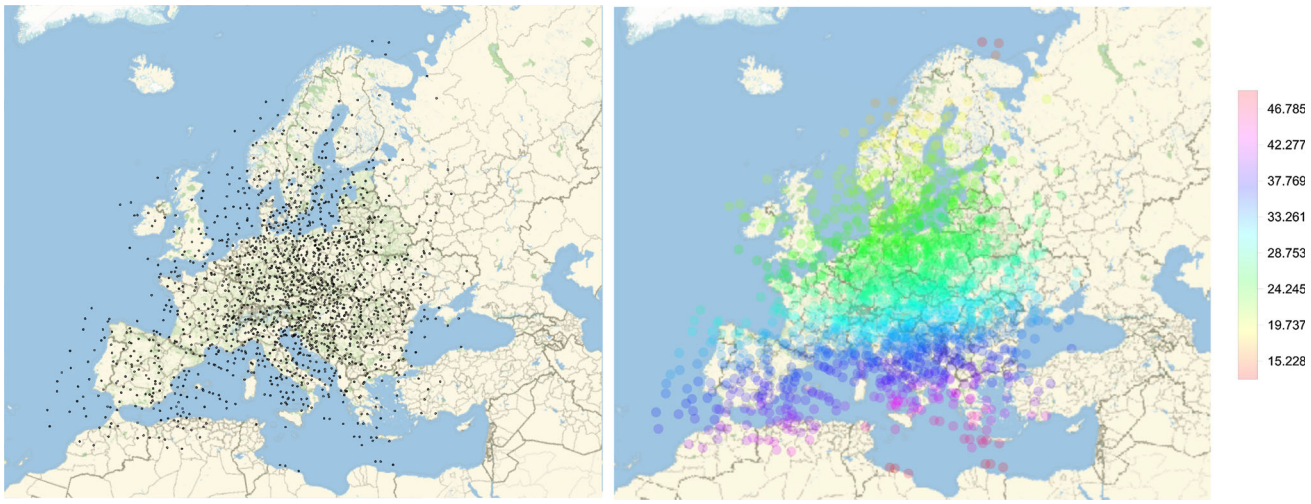

Fig. 9 Location of ionosphere pierce points (left) and colour-coded vTEC values in TECU there (right)

with the constant factor $u = 1/f_1^2 - 1/f_2^2 = -0.6477$ that maps $L4$ slant ionospheric delay onto $L1$ delay. $L1$, f_1 and $L2$, f_2 are phase measurements and signal frequencies, respectively. The indices s and r stand for satellites and receivers, respectively. The bias B comprises of hardware delay and carrier phase ambiguities and is constant in time. In order to estimate this parameter, the ionosphere is parametrized in intervals 10–20 min by means of a functional model. In this study, we use two-dimensional polynomial of second degree. Further details related to ionosphere parameterization can be found in Krypiak-Gregorczyk and Wielgosz (2018). Therefore, for each continuous data arc, the unknown epoch depended coefficients of the applied ionosphere model and time-constant carrier phase bias values have to be estimated. It is performed using the least squares adjustment of observations from all available GNSS stations. After figuring out of bias values, the slant ionosphere delays can be calculated using $L4$ observations according to

$$\Delta I_r^s = \frac{B_r^s - L4_r^s}{u}. \tag{62}$$

In the next step, the ΔI values are converted to slant TEC values. Utilizing the single layer model mapping function vertical TEC (vTEC) values can be figured out. Finally, we receive a set of vertical TEC at the ionosphere pierce points (IPP) $\{\phi, \lambda, vTEC\}$. An example of such scattered data is visualized in Fig. 9, which indicates the distribution of IPP and colour-coded the values of vTEC observed in these points. To achieve these results, observations from over 200 European stations of permanent ground GNSS network were used.

The last important step of the presented approach is the spatial vTEC interpolation on a grid. This is done in two independent ways:

1. using planar spherical thin plate splines and
2. using spherical thin plate splines.

For the testing purpose of the spherical TPS, we use vTEC data for DoY 75/2015, 11:40 GPST (GPS system time). Further information related to vTEC processing, including accuracy assessment and comparison with other modelling approaches are given in Krypiak-Gregorczyk and Wielgosz (2018).

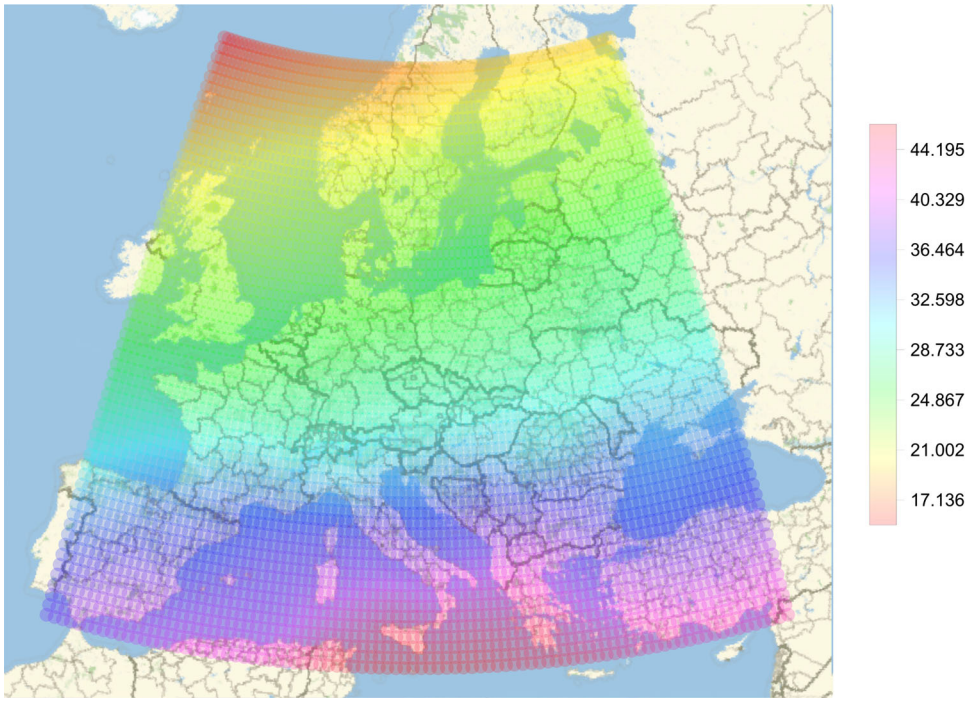

Fig. 10 Planar TPS gridded vTEC values in TECU

6.2.1 Gridding by planar thin plate splines (TPS)

Since the data is given on a sphere and the planar TPS requires planar data the conversion of the one into the other is done in the following way:

1. Since TPS (9) is a planar function, vTEC data scattered on the sphere related to IPP, has to be mapped onto the plane. The Universal Transverse Mercator (UTM) projection is used for this purpose.
2. Then, the parameters of TPS are determined according to (8).
3. Using UTM, the spherical grid ϕ, λ is mapped on to the plane.

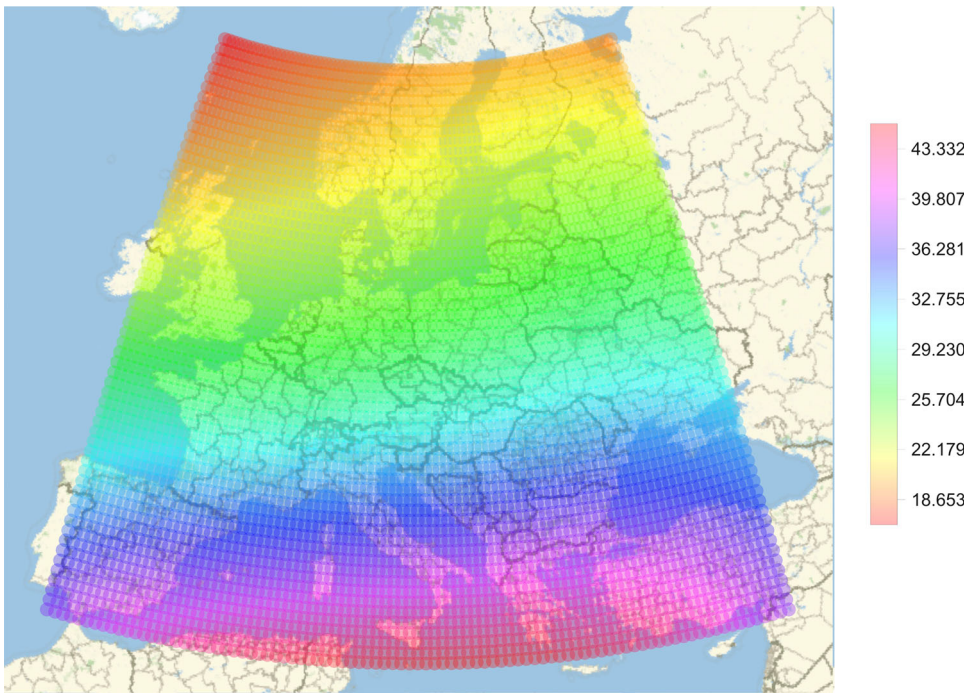

Fig. 11 Spherical TPS gridded vTEC values in TECU

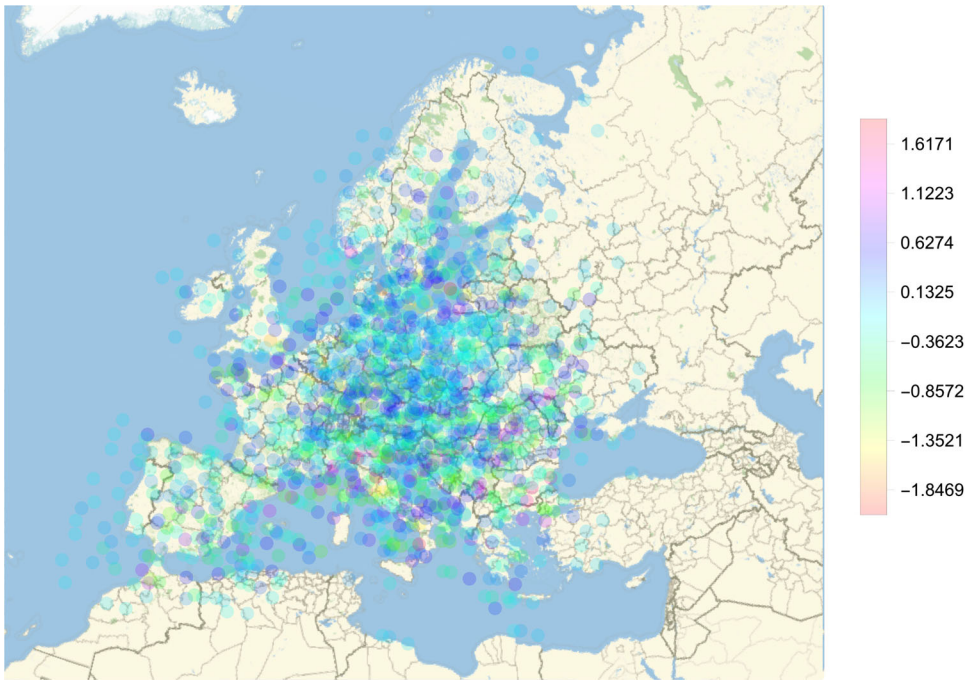

Fig. 12 Residuals of the spherical TPS in the data points in TECU

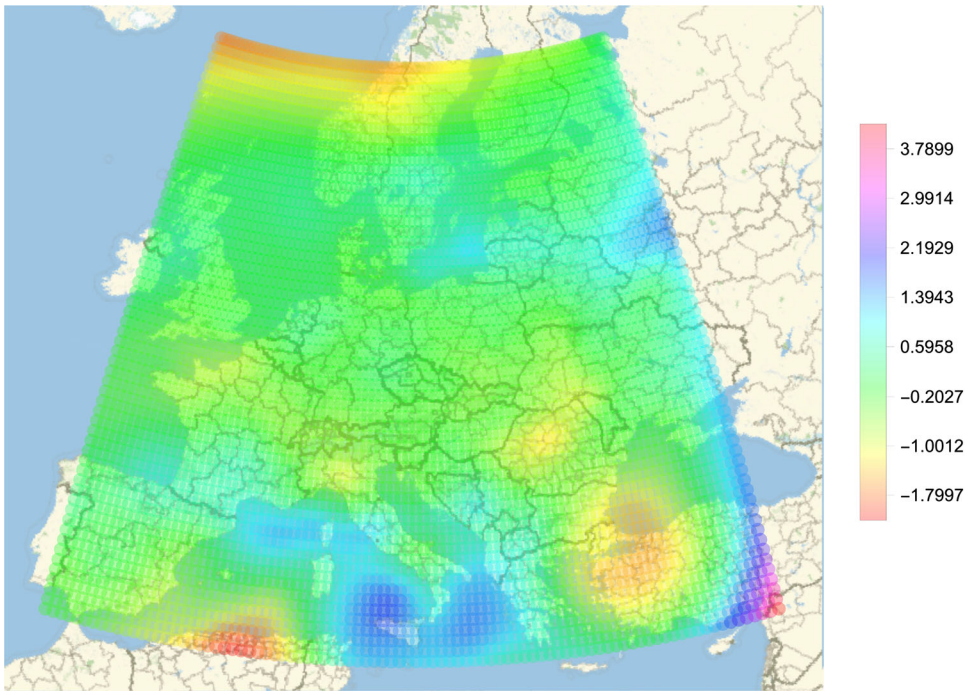

Fig. 13 Difference between the planar and the spherical solution in TECU

4. vTEC values are computed for the grid.
5. vTEC values calculated by TPS are merged with the spherical grid.

The received grid of vTEC values for our demonstration data is shown in Fig. 10. An extended validation of the outlined here approach is given in Krypiak-Gregorczyk et al. (2017) and Wielgosz et al. (2017). The comparison of the TEC-modelling results with five other, most popular, well-established models has justified the introduced approach.

Because the thin plate spline is the planar interpolation model using a projection (UTM) is needed, what is connected with additional computations. This additional computational effort related to mapping the original data (steps 1, 3, 5) can be reduced, if we use the proposed spherical thin plate spline.

6.2.2 Gridding by spherical thin plate splines

Using the same data and the method described in Sect. 5, the gridded data is shown in Fig. 11. Comparing Figs. 10 and 11, one can conclude that there is no visible difference between the outcomes of the planar and the spherical gridding. An indication of the quality of the gridding is the residuals of the interpolated vTEC values in the data points. These residuals are shown in Fig. 12. The residuals behave irregularly and their RMS is about 1% of the signal, which indicates a fairly good interpolation. The difference between the planar and the spherical TPS is shown in Fig. 13. Except at the boundaries, both solutions are practically identical. There is a very small offset of about 0.028 TECU between the two solutions which is way below the data accuracy. The RMS values 7.46 TECU for the planar and 7.30 TECU for the spherical solution are practically identical. Therefore, there seems to be no need to convert spherical data into planar data and then to apply the planar TPS interpolation. A direct application of the spherical TPS interpolation generates practically the same results.

7 Conclusions

In this contribution, we have proposed a new approach of smooth data interpolation on a sphere. This approach is a spherical generalization of the planar thin plate spline interpolation. In contrary to many other spherical spline models, that are solutions of a minimal norm interpolation in a reproducing kernel Hilbert space we have utilized a concept from the mechanics. According to this concept, measurement data are approximated by an shell that has both to be smooth and has to be as close as possible to the data. These properties are described by the internal and the external energy of the shell. Both energies express deviations between the data and the spline function and smoothing properties of the spline, respectively. Minimizing of the internal and the external energies leads to the variational problem that has the solution in the form of the Euler–Lagrange equation. The variational problem has been solved for two cases:

- as a global solution, minimizing the total energy for the whole sphere and
- as a local solution, minimizing the total energy over a simply connected region of the sphere.

For the global case, we found the close solution of the variational problem in the form of collocation in reproducing kernel Hilbert space. For the local case, the direct analytical solution of the variational problem is not possible. Therefore, the local approach is based on a discretization of the corresponding Euler–Lagrange equation using the spherical Laplace operator.

The performance of the local spherical thin plate spline is demonstrated on GRACE data. The between ground-track interpolation of this data is compared with the standard Gaussian smoothing method. Utilizing wavelets-based measure, it is shown quantitative that the thin plate approach is significantly more effective than Gaussian filter in terms of de-striping of the GRACE data. Moreover, it turned out that the global thin plate spline is also the effective interpolator of the GRACE data, even if it produces slightly smoother solution in comparison with the local thin plate approach.

Additionally, we deployed the global spherical thin plate spline to model the total electron content over Europe, despite the fact that the data is given only regional. But because the data is given in the whole region and not only at its boundary, the local thin plate spline interpolation is not applicable. A spherical thin plate spline is an effective tool in this context that allows for gridding the vertical TEC data and to reduce computational afford in comparison with well-established planar TPS approximation. Nevertheless, it turned out that the spherical approach has worse extrapolation properties what demonstrates by the occurrence of small deviations on the border of the investigating area when missing data points.

Finally, it should be mentioned that the introduced spherical thin plate spline can easily be continued to a function harmonic outside the sphere. Therefore, it can be used to investigate various issues related to the Earth gravity modelling.

Of course, spherical thin plate splines are not the only method to interpolate and smooth data given on the sphere. Another promising technique has its origin in the diffusion equation. Since the Green function of the diffusion equation is the Gaussian hat function, the solution of the diffusion equation convolves the data with this kernel function and therefore smoothes the data. Particularly interesting is, that with the spatial variation of the diffusion parameter also a spatial variation of the intensity of smoothing can be generated. This approach was not followed here, but we refer the reader to Cunderlik et al. (2016) and Cunderlik et al. (2013)

Appendix

$$
\begin{aligned}
b_{-2,-2} &= 0\\
b_{-2,-1} &= 0\\
b_{-2,0} &= \frac{(h_\vartheta \cot(h_\vartheta - \vartheta) + 2)(2 - h_\vartheta \cot(\vartheta))}{4h_\vartheta^4}\\
b_{-2,1} &= 0\\
b_{-2,2} &= 0
\end{aligned}
$$

$$
b_{-1,-2} = 0
$$

$$b_{-1,-1} = \frac{(2 - h_\vartheta \cot(\vartheta))\csc^2(h_\vartheta - \vartheta)}{2h_\vartheta^2 h_\lambda^2} + \frac{(2 - h_\vartheta \cot(\theta))\csc^2(\vartheta)}{2h_\vartheta^2 h_\lambda^2}$$

$$b_{-1,0} = \frac{(2 - h_\vartheta \cot(\vartheta))\left(-\frac{2\csc^2(h_\vartheta - \vartheta)}{h_\lambda^2} - \frac{2}{h_\vartheta^2}\right)}{2h_\vartheta^2} + \frac{(2 - h_\vartheta \cot(\vartheta))\left(-\frac{2\csc^2(\vartheta)}{h_\lambda^2} - \frac{2}{h_\vartheta^2}\right)}{2h_\vartheta^2}$$

$$b_{-1,1} = \frac{(2 - h_\vartheta \cot(\vartheta))\csc^2(h_\vartheta - \vartheta)}{2h_\vartheta^2 h_\lambda^2} + \frac{(2 - h_\vartheta \cot(\vartheta))\csc^2(\vartheta)}{2h_\vartheta^2 h_\lambda^2}$$

$$b_{-1,2} = 0$$

$$b_{0,-2} = \frac{\csc^4(\vartheta)}{h_\lambda^4}$$

$$b_{0,-1} = \frac{2\csc^2(\vartheta)\left(-\frac{2\csc^2(\vartheta)}{h_\lambda^2} - \frac{2}{h_\vartheta^2}\right)}{h_\lambda^2}$$

$$b_{0,0} = \frac{2\csc^4(\vartheta)}{h_\lambda^4} + \left(-\frac{2\csc^2(\vartheta)}{h_\lambda^2} - \frac{2}{h_\vartheta^2}\right)^2 + \frac{(2 - h_\vartheta \cot(h_\vartheta - \vartheta))(2 - h_\vartheta \cot(\vartheta))}{4h_\vartheta^4} + \frac{(h_\vartheta \cot(\vartheta) + 2)(2 - h_\vartheta \cot(h_\vartheta + \vartheta))}{4h_\vartheta^4}$$

$$b_{0,1} = \frac{2\csc^2(\vartheta)\left(-\frac{2\csc^2(\vartheta)}{h_\lambda^2} - \frac{2}{h_\vartheta^2}\right)}{h_\lambda^2}$$

$$b_{0,2} = \frac{\csc^4(\vartheta)}{h_\lambda^4}$$

$$b_{1,-2} = 0$$

$$b_{1,-1} = \frac{(h_\vartheta \cot(\vartheta) + 2)\csc^2(\vartheta)}{2h_\vartheta^2 h_\lambda^2} + \frac{(h_\vartheta \cot(\vartheta) + 2)\csc^2(h_\vartheta + \vartheta)}{2h_\vartheta^2 h_\lambda^2}$$

$$b_{1,0} = \frac{(h_\vartheta \cot(\vartheta) + 2)\left(-\frac{2\csc^2(\vartheta)}{h_\lambda^2} - \frac{2}{h_\vartheta^2}\right)}{2h_\vartheta^2} + \frac{(h_\vartheta \cot(\vartheta) + 2)\left(-\frac{2\csc^2(h_\vartheta + \vartheta)}{h_\lambda^2} - \frac{2}{h_\vartheta^2}\right)}{2h_\vartheta^2}$$

$$b_{1,1} = \frac{(h_\vartheta \cot(\vartheta) + 2)\csc^2(\vartheta)}{2h_\vartheta^2 h_\lambda^2} + \frac{(h_\vartheta \cot(\vartheta) + 2)\csc^2(h_\vartheta + \vartheta)}{2h_\vartheta^2 h_\lambda^2}$$

$$b_{1,2} = 0$$

$$b_{2,-2} = 0$$

$$b_{2,-1} = 0$$

$$b_{2,0} = \frac{(h_\vartheta \cot(\vartheta) + 2)(h_\vartheta \cot(h_\vartheta + \vartheta) + 2)}{4h_\vartheta^4}$$

$$b_{2,1} = 0$$

$$b_{2,2} = 0.$$

References

Antoine J-P, Demanet L, Jaques L, Vandergheynst P (2002) Wavelets on the sphere: implementation and approximation. Appl Comput Harmonic Anal 13:177200

Balek V, Mizera I (2013) Mechanical models in nonparametric regression. From probability to statistics and back: high-dimensional models and processes—a Festschrift in honor of Jon A. Wellner, 5–19, Institute of Mathematical Statistics, Beachwood, Ohio, USA, 2013. https://doi.org/10.1214/12-IMSCOLL902

Borkowski A, Keller W (2005) Global and local methods for tracking the intersection curve between two surfaces. J Geod 79:1–10

Cunderlik R, Mikula K, Tunega M (2013) Nonlinear diffusion filtering of data on the Earth's surface. J Geod 87(2013):143–160

Cunderlik R, Collar M, Mikula K (2016) Filters for geodesy data based on linear and nonlinear diffusion. GEM Int J Geomath 7(2016):239–274

Demanet L, Vandergheyst P (2003) Gabor wavelets on the sphere. In Wavelets: applications in signal and image processing X (proceedings volume). SPIE, 2003. https://doi.org/10.1117/12.506436

Duchon J (1976) Interpolation des fonctions de deux varianles suivant le principle de la flexion des plaques minces. R.A.I.R.O. Analyse Numerique 10:5–12

Eberly D (2018) Thin-plate spline. Geometric tools, Redmond WA 98052. https://www.geometrictools.com/. Accessed 21 Jan 2019

Forsberg R, Tscherning CC (1981) The use of height data in gravity field approximation by collocation. JGR 86:7843–7854

Franke R (1982) Scattered data interpolation test of some methods. Math Comput 38:181–200

Franke R, Nielson G (1980) Smooth interpolation of large sets of scattered data. Int J Numer Methods Eng 15:1691–1704

Freeden W (1981) On spherical spline interpolation and approximation. Math Methods Appl Sci 3:551–575

Freeden W (1982) On spline methods in geodetic approximation problems. Math Methods Appl Sci 4:382–396

Freeden W (1984) Spherical spline interpolation: basic theory and computational aspects. J Comput Appl Math 11:367–375

Freeden W (1990) Spherical spline approximation and its application in physical geodesy. In: Vogel A, Ofoegbu CO, Gorenflo R, Ursin B (eds) Geophysical data inversion methods and applications. Theory and practice of applied geophysics. Vieweg+Teubner Verlag, Wiesbaden

Freeden W, Hermann P (1986) Uniform approximation by spherical spline interpolation. Math Z 193:265–275

Freeden W, Michel V (1999) Constructive approximation and numerical methods in geodetic research today an attempt at a categorization based on an uncertainty principle. J Geod 73:452–465
Freeden W, Schreiner M (2005) Spaceborne gravitational field determination by means of locally supported wavelets. J Geod 79:431–446
Freeden W, Glockner O, Schreiner M (1998) Spherical panel clustering and its numerical aspects. J Geod 72:586–599
Hoschek J, Lasser D (1992) Grundlagen der geometrischen Datenverarbeitung. B.G. Teubner, Stuttgart
Hubbert S, Morton TM (2004) A Duchon framework for the sphere. J Approx Theory 129:28–57
Hubbert S, Le Gia QT, Morton TM (2015) Spherical radial basis functions, theory and applications. Springer, Berlin
Keller W (1998) Collocation in reproducing kernel Hilbertspaces of a multiscale analysis. Phys Chem Earth 23:25–29
Keller W (2004) Wavelets in geodesy and geodynamics. Walter de Gruyter, Berlin, p 279
Klees R, Revtova EA, Gunter BC, Ditmar P, Oudman E, Winsemius HC, Savenije HHG (2008a) The design of an optimal filter for monthly GRACE gravity models. Geophys J Int 175:417–432. https://doi.org/10.1111/j.1365-246X.2008.03922.x
Klees R, Tenzer R, Prutkin I, Wittwer T (2008b) A data-driven approach to local gravity field modelling using spherical radial basis functions. J Geod 82:457–471
Krypiak-Gregorczyk A, Wielgosz P (2018) Carrier phase estimation of geometry-free linear combination of GNSS signals for ionospheric TEC modeling. GPS Solut 22:24. https://doi.org/10.1007/s10291-018-0711-4
Krypiak-Gregorczyk A, Wielgosz P, Borkowski A (2017) Ionosphere model for european region based on multi-GNSS data and TPS interpolation. Remote Sens 2017(9):1221. https://doi.org/10.3390/rs9121221
Kusche J (2007) Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models. J Geod 81:733–749
Leick A, Rapoport L, Tatarnikov D (2015) GPS satellite surveying, 4th edn. Wiley, Hoboken, p 509
Moritz H (1987) Least-squares collocation. Rev Geophys 16:421–430
Sansó F (1990) On the aliasing problem in the spherical harmonic analysis. Bull Géod 64:313–330
Schmidt M, Fengler M, Meyer-Gürr T, Eicker A, Kusche J, Sánches L, Shin-Chan H (2007) Regional gravity modeling in terms of spherical base functions. J Geod 81:17–38
Sloan IH, Womersley RS (2002) Good approximation on the sphere, with application to geodesy and the scattering of sound. J Comput Appl Math 149:227–237
Terzopoulos D (1986) Regularization of inverse visual problems involving discontinuities. IEEE Trans Pattern Anal Mach Intell PAMI–8(4):413–423
Tscherning CC (1978) Collocation and least squares methods as a tool for handling gravity field dependent data obtained through space research techniques. Bull Geod 52:199–212
Tscherning CC (2001) Computation of spherical harmonic coefficients and their error estimates using least-squares collocation. J Geod 75:12–18
Wagner C, McAdoo D, Klokočnik J, Kostelecký J (2006) Degradation of geopotential recovery from short repeat-cycle orbits: application to GRACE monthly fields. J Geod 80:94–103
Wahba G (1981) Spline interpolation and smoothing on the sphere. SIAM J Sci Stat Comput 2:5–16
Wang H, Sloan IH (2017) On filtered polynomial approximation on the sphere. J Fourier Anal Appl 23:863–876. https://doi.org/10.1007/s00041-016-9493-7
Weigelt M, Sideris MG, Sneeuw N (2009) On the influence of the ground track on the gravity field recovery from high-low satellite-to-satellite tracking missions: CHAMP monthly gravity field recovery using the energy balance approach revisited. J Geod 83:1131–1143. https://doi.org/10.1007/s00190-009-0330-5
Wielgosz P, Krypiak-Gregorczyk A, Borkowski A (2017) Regional ionosphere modeling based on multi-GNSS data and TPS interpolation. In Proceedings of the baltic geodetic congress (BGC Geomatics), Gdansk, Poland, 22–25 June 2017, pp 287–291
Zheng W, Hsu H, Zhong M, Yun M (2012) Efficient accuracy improvement of GRACE global gravitational field recovery using a new inter-satellite range interpolation method. J Geodyn 53:1–7. https://doi.org/10.1016/j.jog.2011.07.003