Research Article

Next Generation Network Real-Time Kinematic Interpolation Segment to Improve the User Accuracy

Mohammed Ouassou,[1,2] Anna B. O. Jensen,[3] Jon G. O. Gjevestad,[2,4] and Oddgeir Kristiansen[1,2]

[1]*Norwegian Mapping Authority, Geodetic Institute, 3511 Hønefoss, Norway*
[2]*Department of Mathematical Sciences and Technology, NMBU, 1432 Akershus, Norway*
[3]*KTH Royal Institute of Technology, 100 44 Stockholm, Sweden*
[4]*Department of Electronics and Telecommunications, NTNU, 7491 Trondheim, Norway*

Correspondence should be addressed to Mohammed Ouassou; mohammed.ouassou@statkart.no

This paper demonstrates that automatic selection of the right interpolation/smoothing method in a GNSS-based network real-time kinematic (NRTK) interpolation segment can improve the accuracy of the rover position estimates and also the processing time in the NRTK processing center. The methods discussed and investigated are inverse distance weighting (IDW); bilinear and bicubic spline interpolation; kriging interpolation; thin-plate splines; and numerical approximation methods for spatial processes. The methods are implemented and tested using GNSS data from reference stations in the Norwegian network RTK service called CPOS. Data sets with an average baseline between reference stations of 60–70 km were selected. 12 prediction locations were used to analyze the performance of the interpolation methods by computing and comparing different measures of the goodness of fit such as the root mean square error (RMSE), mean square error, and mean absolute error, and also the computation time was compared. Results of the tests show that ordinary kriging with the Matérn covariance function clearly provides the best results. The thin-plate spline provides the second best results of the methods selected and with the test data used.

1. Introduction

The use of GNSS and network real-time kinematic positioning to achieve GNSS positions with accuracy at the cm-level is increasing rapidly these years. This is partly due to the development and modernization of the GNSS systems themselves (GPS, GLONASS, Galileo, and Beidou), but it is also caused by a general quest for better position accuracy in many user communities.

High-accuracy GNSS positioning is based on the carrier phase being observable. Using the notation from [1], the basic observation equation that summarizes the relation between observations and error sources is given as follows:

$$\Phi = \lambda^{-1}\left(r - I + T\right) + f\left(\delta t_u - \delta t^s\right) + N + \varepsilon_\phi, \quad (1)$$

where Φ is the phase observation in cycles, λ is the wavelength in meters/cycle, r is the geometric distance between the receiver and satellite in meters, I is the ionospheric signal delay in meters, T is the tropospheric signal delay in meters, f is the frequency in Hertz, δt_u and δt^s are the clock errors of, respectively, the receiver and the satellite, N is the initial number of cycles at the first observation epoch (the *ambiguity*), and ε_ϕ is a noise term given in cycles that mainly accounts for multipath (reflected signals) and receiver noise.

When using the NRTK technique, a network of reference stations is used to estimate the errors in the positioning process, that is, the effects of the ionosphere and troposphere as well as inaccuracies in the satellite position as provided with the broadcast ephemerids from the satellites.

The accuracy of NRTK positioning systems depends on the ability to identify and mitigate the error sources in
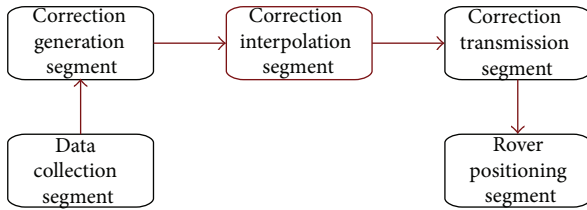
FIGURE 1: Network real-time kinematic segments.

the system as well as the residual biases. The biases include residual effects from the space segment, signal propagation, environment effects, and receiver noise in the reference network. The mitigation process can be carried out by modeling, estimation, or combinations of observables.

The NRTK processing chain can be summarized as follows: the first step is to collect raw measurements from the network of reference stations, solve for the ambiguities within the reference network, and generate error estimates. The next step is to apply the interpolation/smoothing scheme to generate the RTK corrections for the user location. The RTK corrections are then transmitted to users who can then perform real-time positioning with accuracy at the cm-level.

Figure 1 shows all segments involved in the NRTK processing chain. The figure illustrates the so-called virtual reference station (VRS) concept, which was developed by Landau et al. [2]. Other NRTK standards such as for instance the master auxiliary concept (MAC) also exist [3], but we limit the discussion in this paper to the VRS concept.

As the GNSS systems and users become more numerous, the amount of data that needs processing increases as well, which poses some interesting challenges for the NRTK system developers and service providers. This paper focuses on processing large data sets and high quality interpolators/smoothers that can be used to aid the data processing. Let us consider how the RTK processing is carried out. First the user sends his/her position to the control center, and then the network engine chooses a suitable subnetwork which is used to generate corrections, and these corrections are then transmitted back to the user. The first challenge to this model is the number of users, since each user has to be processed independently, and the number of users has increased dramatically in recent years. The solution to this is to construct new models and algorithms. These should be able to process data from large geographical areas, as well as computing the necessary corrections and quality indicators ready for use, so that any RTK user that connects will be served immediately.

In other branches of science and engineering, new analysis tools that satisfy these requirements have already been developed: neural networks, machine learning, classification and regression trees, hierarchical models, and so forth. In this paper, some existing interpolation/smoothing methods are applied to real datasets, and the strengths and weaknesses of each method are identified. The results are then used to combine these methods and construct models that describe the observed variations in the data as well as possible.

Interpolation methods can be divided into two categories: local methods and global methods. The local methods only use a subset of the data for interpolation, which implies that the required processing time is reduced. Conversely, the global techniques use all the data available to generate predictions. In this paper, both these approaches are considered. Referring to Figure 1, the main focus of this paper is directed at the correction interpolation segment and more specifically at the automatic selection of the right interpolation algorithm based on appropriate tests, such that the rover position estimation will be improved.

The rest of the paper is organized as follows: Section 2 gives a full description of the test data using the Norwegian GNSS network data, known as CPOS, and introduces the variational problem in general. Section 3 covers local interpolation algorithms, specifically the inverse-distance weighted and bilinear/bicubic methods by Akima. Section 4 and the following sections deal with global interpolation methods. First, thin-plate splines and the Bayesian model behind the smoothing are reviewed in this section. Section 5 introduces numerical approximation schemes for Gaussian random fields. Section 6 covers spatial interpolation algorithms and specifically the ordinary kriging method. Section 7: the performance parameters are defined in this section. Section 8: the results from Sections 3–6 are generalized in this section. Section 9 is the conclusion and discussion and covers applications of the results developed in Sections 3–6.

2. Test Data and Variational Reconstruction

2.1. Test Data. The main success of network real-time kinematic positioning has been the reduction of correlated errors in the network (e.g., ionospheric, tropospheric, and satellite position errors). This type of errors is collectively referred to as *distance-dependent errors* and can be subdivided into the *dispersive errors* which depend on frequency and the *nondispersive errors* which do not.

The size of the network varies with time, as the individual reference stations and satellites may not deliver data for a while, and observations are typically correlated to each other. Modeling the spatial and temporal variations of such a process is too complex to capture the covariance structure of the data, so often we end up imposing stationarity. In this paper, we apply techniques for handling spatial processes in order to capture the covariance structure in the data, such that high quality synthetic data can be provided. The main clue is to employ the right tool from epoch to epoch, based on some appropriate criteria.

We prefer to work with real data, and since the real network error estimates were not made available, we decided to analyze the ionospheric path delays for CPOS RTK network, given by absolute TEC values. If the ionosphere data is replaced with the full network corrections, the same algorithms should still function very well. Ionospheric path delay is considered the single largest source of inaccuracy for positioning and navigation, so the quality of the NRTK corrections is strongly affected in the case of moderate to high ionosphere activity. To test the algorithms against each other, a large ionospheric data set from the Norwegian CPOS network is investigated. The data is generated by a first-order geometry-free approach (Section 8.1).
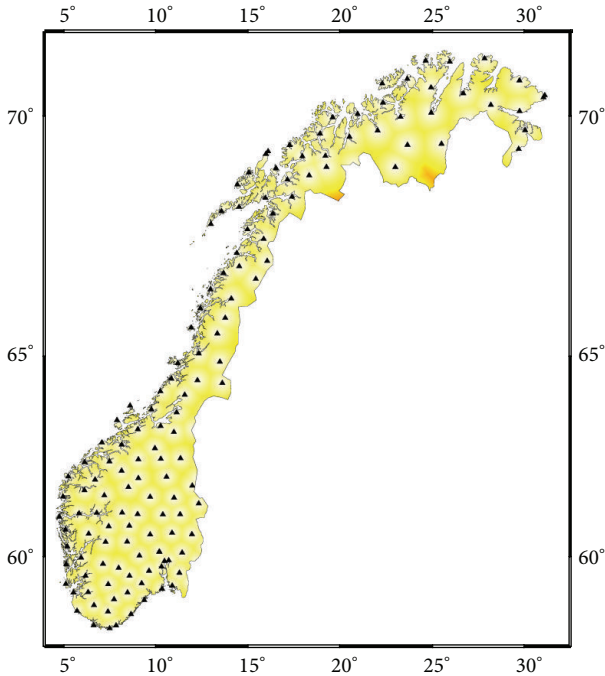
FIGURE 2: Norwegian CPOS RTK network. Filled triangles mark station locations. Height difference between sites up to 1000 meters.

At the time of writing, the CPOS RTK network contains approximately 180 stations on the Norwegian mainland (for a listing of stations of the CPOS RTK network, see the appendix.). The algorithms were tested with different station configurations (50, 75, and 110 stations), equipped with geodetic dual frequency GNSS receivers, which track both GPS and GLONASS satellites. In this investigation, however, only the GPS portion of the data was used. The distribution of the CPOS RTK network reference stations is given in Figure 2.

2.2. Variational Reconstruction

2.2.1. Problem Formulation. Let us assume that the observed noisy measurements $\{z_i\}$ at the locations $\{s_i : s_i \in \mathbb{R}^2\}$ is a random function, with the mean $m(s)$ and variance σ^2. Our goal is then to predict the value at other locations $\{s_k\}$ where we have no observations, under the assumption that the predicted values should resemble its neighbors. To achieve this, we can either interpolate or construct a smooth function $g(s)$ that represents the variation in the data and is robust against outliers.

The data that will be modeled is a pure spatiotemporal process, namely, the absolute total electron count (TEC). Assuming weak stationarity of the process under study, the mean and variance are not functions of the spatial location s. The model used to describe the variation in data in this paper, however, is assumed to have the form

$$z(s) = m(s) + \varepsilon(s). \tag{2}$$

The mean function $m(s)$, often referred to as the *trend* or *deterministic part*, determines the large-scale variation in the data. The function $\varepsilon(s)$ is called the *random part* and determines the small-scale variation. This process model will be assumed in the subsequent discussion of all the different interpolation/smoothing techniques presented in this paper.

Some data analysts prefer the *Cressie decomposition* [4, ch. 3] of the observed random field $Z(s)$, which takes the form

$$Z(s) = m(s) + T(s) + \varepsilon(s), \tag{3}$$

where $s \in \mathbb{R}^2$ is the spatial location; $Z(s)$ is the observation; $m(s) = X\beta$ is the trend (the mean component of the model); $T(s)$ is a stationary Gaussian process with variance σ^2 (partial sill), and a correlation function parameterized in its simplest form by ϕ (the range parameter); and finally ε is an error term, with a variance parameter τ^2 (nugget variance).

2.2.2. Model Parameter Estimations. Once the model is defined, the next step is to estimate the model parameters. In general, this is done numerically by minimizing the negative log-likelihood function. The most used optimization methods are, respectively, the conjugate gradient method, the quasi-Newtonian method, and the Nedler-Mead method. The details of these methods will not be treated in this paper, but the interested reader is referred to references [5, 6].

The algorithm may not converge to correct parameter values when called with the default options. The user should therefore try different initial values, and if the parameters have different orders of magnitude, a scaling of the parameters may be necessary. If such problems arise, some possible workarounds include

(i) rescaling data values by dividing by a constant,

(ii) rescaling coordinates by subtracting values and/or dividing by constants,

(iii) bootstrapping to accelerate the convergence. This method is used in our implementation of the kriging algorithm in Section 6.

2.2.3. Model Validation. In the field of statistical analysis, an appropriate way of analyzing data is to divide it into three distinct subsets. The training dataset is used to construct the model, the validation data is used to check the model, and the last data set is used to challenge the model. The main purpose is to determine whether or not our model is an accurate representation of the real world data. This process is called the *model validation assessment*. The most famous methods are the family of cross-validation, generalized maximum likelihood (GML) methods, Akaike information criterion (AIC), Bayesian information criterion (BIC), and so forth.

In our implementation, the generalized cross-validation is used to determine the optimal smoothing parameter λ (see Section 4). The computation AIC and BIC are computed in Section 6, when maximum likelihood estimation is used instead of weighted least squares in the kriging algorithm. The GML methods will be used in future work.

3. IDW and Akima Interpolation

3.1. IDW Interpolation. The inverse-distance weighted (IDW) scheme is an exact interpolator. It honors the data by assigning weights to all neighboring points according to their distance from the prediction location s_0. Locations that are closer to s_0 receive higher weights, and locations that are far from s_0 are given lower weights; this mechanism is administered by the parameter r in the IDW predictor formula. The user can freely choose the number of observations used to perform the interpolation. This is done by defining a radius R around the prediction location s_0.

The IDW predictor scheme is defined as

$$\widehat{z}\left(s_0\right) = \frac{\sum_{i=1}^{N} z\left(s_i\right) d_{0i}^{-r}}{\sum_{i=1}^{N} d_{0i}^{-r}}. \qquad (4)$$

Here, s_0 is the prediction location, N is the number of observations, $z(s_i)$ are the neighboring observations, r is the weight decision parameter, and d is the distance (either spherical or Euclidean).

The IDW method is originally due to Shepard [7], which described a global method. All derived IDW methods are either generalizations or variations of this method. The basic Shepard's method can be expressed as

$$F\left(x, y\right) = \frac{\sum_{k=1}^{N} w_k\left(x, y\right) f_k}{\sum_{k=1}^{N} w_k\left(x, y\right)}, \qquad (5)$$

where typically the weight w_k is the inverse Euclidean distance $w_k = 1/d_k^2 = \{(x - x_k)^2 + (y - y_k)^2\}^{-1/2}$. We will however define a disk with center (x_k, y_k) and a radius R and set the weight w_k to zero outside of this disk. A natural scheme suggested by many authors, like, for example, Renka and Brown [8], is given by the expression

$$w_k\left(x, y\right) = \left\{ \frac{\left(R_w - d_k\right)_+}{R_w d_k} \right\}^2, \qquad (6)$$

where

$$\left(R_w - d_k\right)_+ = \begin{cases} R_w - d_k & \text{if } d_k < R_w \\ 0 & \text{if } d_k \geq R_w. \end{cases} \qquad (7)$$

Impose the constraints such that

(i) the sum of all weights w_k inside the disk R_w should be normalized to unity, that is, $\sum w_k = 1$,

(ii) the predictor is a linear combination of the observations.

If the variance of the predictor is then controlled such that it is at a minimum, the IDW behaves almost like the local kriging interpolator (Section 6); however the covariance structure is not preserved.

For the implementation, the package *gstat* from Edzer Pebesma is used to carry out IDW (see Table 3 for more information).

3.2. Akima Algorithms. Bilinear or bicubic spline interpolation is applied using different versions of algorithms by Akima [9, 10]. Given a set of data points in a plane, our aim is to fit a smooth curve that passes through the given points considered as reference points. The method is local and is based on a piecewise function composed of a set of polynomials and applicable up to the third degree on each interval. The method produces remarkable results with minimum processing time. For a detailed mathematical formulation, please refer to references [9, 10].

3.2.1. Basics of Spline Interpolation/Smoothing. An understanding of the basic elementary building blocks of a 1D spline facilitates the understanding of 2D and 3D splines, for instance, the TPS (Section 4).

Given real numbers $\{t_1, t_2, \dots, t_n\} \in [a, b]$, a function g defined on interval $[a, b]$ is a cubic spline if and only if the following conditions are satisfied:

(i) the function g is a cubic polynomial on each interval $(a, t_1), (t_1, t_2), \dots, (t_n, b)$;

(ii) the function g and its first and second derivatives are continuous at each of the points t_i.

Condition (ii) implies that the cubic polynomials from condition (i) fit together on each t_i, where the t_i are called *knots*. Together these two conditions imply that $g(x)$ is a function with continuous first and second derivatives on the whole interval $[a, b]$.

For some given real constants a_i, b_i, c_i, d_i, the cubic spline function g can be expressed as

$$g\left(t\right) = a_i + b_i\left(t - t_i\right) + c_i\left(t - t_i\right)^2 + d_i\left(t - t_i\right)^3, \qquad (8)$$

where the index $i = 0, 1, 2, \dots, n$. The end-point knots correspond to the boundaries of the function domain; that is, $t_0 = a$ and $t_{n+1} = b$.

Finding a smoothing spline is not an easy task. Reinsch (1967) proposed an algorithm and showed that the solution of the minimum principle is actually cubic splines. The basic idea is to construct a nonsingular system of linear equations of the second derivative λ_i of $\widehat{g}$. The resulting equations are computationally efficient because of their banded structure. For an excellent exposition of the material, see also [11].

3.2.2. Output Result from Akima. Figure 3 shows the output from Akima with bilinear interpolation.

4. Thin Plate Spline Method

4.1. Mathematical Preliminaries. In this section, our main interest is not to construct a function $m(s_i)$ that exactly interpolates the data z_i at n distinct points but to find an attractive way to smooth noisy data. The method of thin-plate splines (TPS) will be used for this purpose.

Duchon [12] was the first to build the theoretical foundation for the TPS method. The name TPS comes from the physical situation of bending a thin surface, where the method minimizes the bending energy of a thin plate fixed
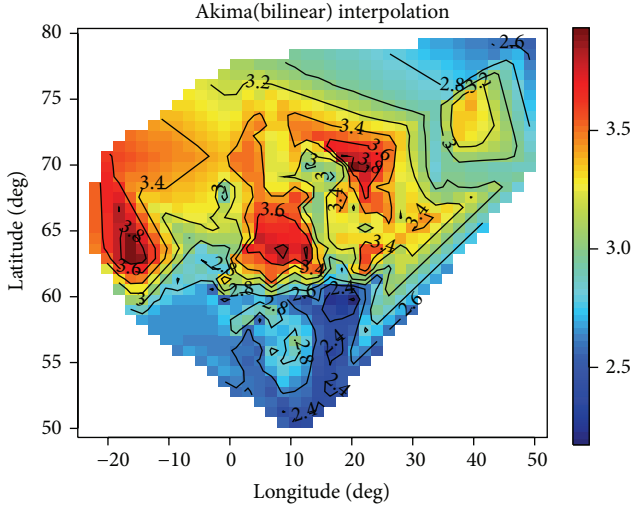
FIGURE 3: L1-VTEC variation surface generated by AKIMA (bilinear) interpolation. Color scale is expressed in meters. High ionospheric activity (year: 2013, doy: 152), 110 reference stations.

at the data sites. For our application, the TPS method is used to minimize the cost function:

$$J_m(f) = \int_{\mathbb{R}^d} \left\| D^m f(s) \right\|^2 ds, \tag{9}$$

where D^m is a vector of partial differentiation operators of order m. In the two-dimensional case, that is, when $d = 2$, $m = 2$, and $f = f(x, y)$, the TPS penalty function can be written as

$$J_2(f) = \iint_{\mathbb{R}^2} \left\{ \left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right\} dx\, dy. \tag{10}$$

Let $\Delta^4(x, y)$ denote the differential operator in the integrand of (10). The thin-plate spline $f(x, y)$, which is the solution to the variational problem of minimizing the penalty $J_2(f)$, can then be found by solving the biharmonic equation $\Delta^4 f(x, y) = 0$.

The goal is to find the function f_λ in Sobolev space [13, p. 250] that minimizes the following expression:

$$S(f) = n^{-1} \sum_{i=1}^{n} (z(s_i) - f(s_i))^2 + \lambda J_2(f), \tag{11}$$

where n is the total number of observations, J_2 is a smoothness penalty (the cost function), and λ is the smoothing parameter. The smoothing parameter is a trade-off factor between the rate of change of the residual error and local variation. Optimal minimization of $S(f)$ results in a good compromise between smoothness and goodness of fit.

Once the curve approximation of the data has been constructed, generating values at any location, where no observations are available, is accomplished by simply indexing the variables x and y and fetching the corresponding value. This is a major advantage of smoothing methods over

interpolation methods; no extra interpolations are required after the curve has been constructed for a given epoch.

Validation is carried out by the GCV (Section 2.2.3). Let the cross-validation function $\hat{g}_{n,\lambda}^{[k]}$, with $k = 1, 2, \ldots, n$, be defined as

$$\hat{g}_{n,\lambda}^{[k]} = \arg\min_{\hat{g} \in W_2^{(m)}} \sum_{i=1}^{n} (z_i - \hat{g}(x_i))^2 + \lambda \int_0^1 \left\{ \hat{g}^{(m)}(x_i) \right\}^2 dx, \tag{12}$$

where $\lambda > 0$. The OCV (ordinary cross-validation) and OCV MSE (ordinary cross-validation mean square error) $V_0(\lambda)$, respectively, are defined as

$$\begin{aligned} \text{OCV}(\lambda) &= \arg\min_{\lambda \in \mathbb{R}^+} V_0(\lambda), \\ V_0(\lambda) &= \frac{1}{n} \sum_{i=1}^{n} \left\{ \hat{g}_{n,\lambda}^{[k]}(x_i) - z_i \right\}^2. \end{aligned} \tag{13}$$

The determination of the GCV (general cross-validation) goes as follows. First, the expression for V_0 has to be rewritten. There exists an $n \times n$ matrix $A(\lambda)$, the smoothing/influence/sensitivity matrix with the property. Consider

$$\begin{bmatrix} \hat{g}_{n,\lambda}(x_1) \\ \vdots \\ \hat{g}_{n,\lambda}(x_n) \end{bmatrix} = A(\lambda) Z, \tag{14}$$

such that $V_0(\lambda)$ can be written as

$$V_0(\lambda) = n^{-1} \sum_{j=1}^{n} \frac{\left(a_{kj} z_j - z_k \right)^2}{\left(1 - a_{kk} \right)^2}, \tag{15}$$

where $k, j \in \{1, 2, \ldots, n\}$ and a_{kj} is element $\{k, j\}$ of $A(\lambda)$.

Definition 1 (generalized cross-validation (GCV)). Let $A(\lambda)$ be the smoothing matrix defined in (14); then the GCV function is given by the expression

$$\text{GCV}(\lambda) = \frac{n^{-1} \left\| (I - A(\lambda)) z \right\|^2}{\left\{ n^{-1} \operatorname{tr}(I - A(\lambda)) \right\}^2}. \tag{16}$$

4.2. Estimation of the Smoothing Parameter λ. The smoothing parameter λ plays a central role in the TPS method. By adjusting the value of λ, one can get the desired level of smoothness at the cost of accuracy at the data sites. When we set this parameter to zero, the problem is reduced to an interpolation with no smoothing. On the other hand, when the smoothing parameter tends to infinity, the method yields a plane which is least-square fitted to the data. The smoothness penalty method can be chosen by any criteria, but the most popular criterion is GCV (generalized cross-validation), also known as the "left-out one" method. The GCV criterion selects the smoothing parameter λ that minimizes the GCV function, equation (16), that is $\hat{\lambda} = \arg\min_{\lambda \in \mathbb{R}^+} \text{GCV}(\lambda)$.

The GCV function $V(\lambda)$ is the predicted mean square error and can be viewed as a weighted version of the $OCV(\lambda) = V_0(\lambda)$:

$$V(\lambda) = \frac{1}{n} \sum_{i=1}^{n} \left\{ \hat{g}_{n,\lambda}^{[i]}(x_i) - z_i \right\}^2 W_k(\lambda),$$

$$W_k(\lambda) = \left\{ \frac{I - a_{kk}(\lambda)}{n^{-1} \operatorname{tr}(I - A(\lambda))} \right\}^2. \tag{17}$$

In geodesy, it is often interesting to estimate the accuracy f_λ. Two loss functions are considered: the mean square prediction error $T(\lambda)$, and the stricter Sobolev error is defined as $W(\lambda) = \|f_\lambda - f\|_W^2$,

$$W(\lambda) = \int_a^b \left\{ f_\lambda - f \right\}^2 dG + \int_a^b \left\{ f_\lambda^{(m)} - f^{(m)} \right\}^2 dG. \tag{18}$$

The performance of an estimator is often well characterized by the risk function, defined as the expectation value of the loss function:

$$R_T(\lambda) = \mathbb{E}(T(\lambda)), \qquad R_W(\lambda) = \mathbb{E}(W(\lambda)). \tag{19}$$

In this analysis, the GCV is used to estimate the smoothing parameter λ. Figure 12 shows the smoothed surface generated by the TPS with GCV.

For implementation, the CRAN package *rgcvpack* is used to implement the TPS algorithm (see Table 3 for more information).

5. Numerical Approximation Methods

Numerical approximation techniques will assist us in processing huge data sets with convergence. The main idea is based on the pioneering work of Besag [14].

Let us assume that our observations at different locations follow a multivariate Gaussian distribution with mean μ_z and variance-covariance Σ_z. Then the continuously Gaussian fields have the distribution

$$f(z) = \{2\pi\}^{-n/2} |\Sigma_z|^{-1} \exp\left\{ -\frac{1}{2}(z - \mu_z)^T \Sigma_z^{-1}(z - \mu_z) \right\}. \tag{20}$$

Approximating the continuous Gaussian random field by the discrete Gauss-Markov random field is accomplished by introducing the Markov property. This is done as follows: we say that two locations s_i and s_j are conditionally independent if and only if

$$x_i \perp x_j \mid x_{-(i,j)}. \tag{21}$$

This property is very important when constructing the precision matrix Q of the GMRF. That is, if we know what happens nearby, we can ignore everything that lies further away. Consider

$$x_i \perp x_j \mid x_{-(i,j)} \implies Q_{i,j} = 0. \tag{22}$$

That is, element (i, j) of Q is zero if the process at location i is conditionally independent of a process at j given the
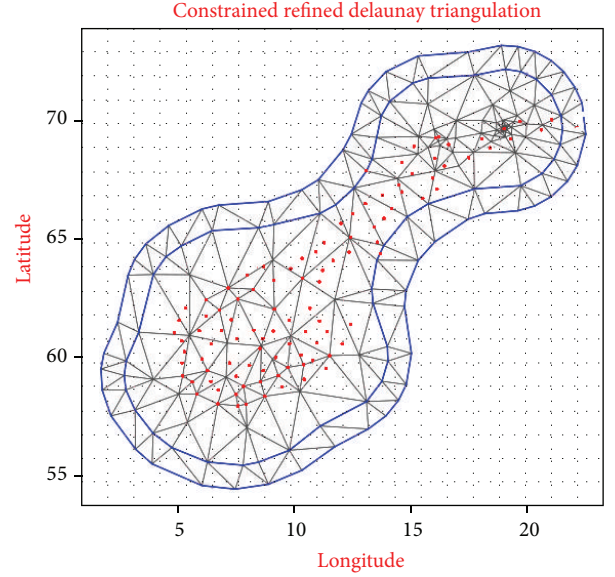

FIGURE 4: Mesh construction on top of which the GMRF is built. Red circles represent the configured reference stations. The mesh in this figure is used to construct the neighborhoods which are key to specifying GMRF conditional independence structure.

process at all locations except $\{i, j\}$. Figure 4 illustrates the concept of the GMRF.

The sparse precision matrix makes the GMRF computationally effective, but it is difficult to construct reasonable precision matrices. As a conclusion, the GMRF is a Gaussian field with a sparse precision matrix $Q = \Sigma^{-1}$. For an excellent description of the theory and applications of GMRF, the reader is referred to, for example, Rue and Held [15].

The integrated nested Laplace approximation (INLA) method developed by Håvard Rue is used to implement the GMRF (see Table 3 for more information).

6. Kriging Interpolator

The kriging interpolator is a linear spatial interpolation algorithm and is primarily used in geostatistics. In recent years, however, the interpolator has been applied in many new areas, such as geophysics and climate data analysis.

Given the observations $\{Z(s_i)\}_{i=1}^N$, we want to predict the value of $Z(s_0)$ where no observations have been made. Our goal is to find an estimator $\widehat{Z}_0 = \widehat{Z}(s_0) = \sum_{i=1}^N w_i Z(s_i)$ such that the following requirements are met.

(i) Unbiasedness. This means that $\mathbb{E}(Z(s_0)) = \mathbb{E}(Z_0)$ and is accomplished if $\sum_{i=1}^N w_i = 1$ and the mean is stationary.

(ii) Minimum Prediction Variance. We make some assumptions about the mean value of the random field $Z(s)$. If the mean is unknown but constant across the entire region of interest, we have *ordinary kriging*. Otherwise, the method is known as *simple kriging*.

Any estimator that meets the conditions of unbiasedness and minimum prediction variance is said to be a BLUP

(best linear unbiasedness predictor). Let us examine the components of the MSPE (mean square prediction error). Consider

$$
\begin{aligned}
\sigma_{S_0}^2 &= \text{var}\left(Z_0 - \widehat{Z}_0\right) \\
&= \text{var}\left(Z_0\right) + \text{var}\left(\widehat{Z}_0\right) - 2\,\text{cov}\left(Z_0, \widehat{Z}_0\right) \\
&= \sigma^2 + \text{var}\left(\sum_{i=1}^{N} w_i Z_i\right) - 2\,\text{cov}\left(Z_0, \sum_{i=1}^{N} w_i Z_i\right) \\
&= \sigma^2 + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \,\text{cov}\left(Z_i, Z_j\right) - 2\sum_{i=1}^{N} w_i \,\text{cov}\left(Z_i, Z_0\right) \\
&= \sigma^2 + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j C_{ij} - 2\sum_{i=1}^{N} w_i C_{i0}.
\end{aligned}
\tag{23}
$$

We want to minimize $\text{var}[Z(s_0) - \widehat{Z}(s_0)]$ subject to the constraint $\sum_{i=1}^{N} w_i = 1$.

The procedure is well defined by the method of Lagrange multipliers. Form the Lagrangian L,

$$
L = \text{var}\left(Z\left(s_0\right) - \widehat{Z}\left(s_0\right)\right)^2 - 2\lambda \sum_{i=1}^{N} \left(w_i - 1\right).
\tag{24}
$$

We then take the partial derivatives of L with respect to the weights w_i and to λ, set the equations to be equal to zero, and solve them; we get

$$
\begin{bmatrix} w_1 \\ \vdots \\ w_N \\ \lambda \end{bmatrix} = \begin{bmatrix} C_{11} & \cdots & C_{1N} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C_{1N} & \cdots & C_{NN} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} C_{10} \\ \vdots \\ C_{N0} \\ 1 \end{bmatrix}.
\tag{25}
$$

Equation (25), which is the kriging equation, is used to compute the weights. The computation of weights is based on the covariances among locations in the sample (region of interest) and the covariances between sample locations and the location to be predicted. To be specific.

(1) Covariances among the locations in the sample:

$$
C_{ij} = C\left(s_i, s_j\right) = \text{Cov}\left(Z_i, Z_j\right) \quad \forall i, j.
\tag{26}
$$

The covariance matrix of the N sample values read

$$
\mathbf{C} = \begin{bmatrix} C_{11} & \cdots & C_{1N} \\ \vdots & \ddots & \vdots \\ C_{1N} & \cdots & C_{NN} \end{bmatrix}.
\tag{27}
$$

(2) Covariances between the sample locations and the prediction point:

$$
C_{i0} = C\left(s_i, s_0\right) = \text{Cov}\left(Z\left(s_i\right), Z\left(s_0\right)\right) \quad \forall i.
\tag{28}
$$

The vector of covariances between the sample locations and the prediction point read

$$
\mathbf{c} = \begin{bmatrix} C_{10} & C_{20} & \cdots & C_{N0} \end{bmatrix}^{\mathrm{T}}.
\tag{29}
$$

Equation (25) becomes

$$
\begin{bmatrix} \mathbf{w} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{C} & \mathbf{1} \\ \mathbf{1}^{\mathrm{T}} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{c} \\ 1 \end{bmatrix},
\tag{30}
$$

where $\mathbf{w}$ is $1 \times N$ vector of weights and $\mathbf{1} = [1 \cdots 1]$ is a vector of the same dimensions.

6.1. Directional Effects. Another form of nonstationarity lies in the covariance structure. One specific way to relax the stationarity assumption is to allow directional effects. For instance, the correlation decay rate at increasing distances may be allowed to depend on the relative orientation between pairs of locations. The simplest form of directional effects in the covariance structure is called *geometrical anisotropy*. This arises when a stationary covariance structure is transformed by a differential stretching and rotation of the coordinate axes. Hence, geometrical anisotropy is defined by two additional parameters. Algebraically, a model with geometrical anisotropy in spatial coordinates $x = (x_1, x_2)$ can be converted to a stationary model in coordinates $y = (y_1, y_2)$ by the transformation

$$
\left(y_1 \;\; y_2\right) = \left(x_1 \;\; x_2\right)\begin{pmatrix} \cos\left(\psi_A\right) & -\sin\left(\psi_A\right) \\ \sin\left(\psi_A\right) & \cos\left(\psi_A\right) \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \psi_R^{-1} \end{pmatrix}.
\tag{31}
$$

ψ_A is called the *anisotropy angle* and $\psi_R > 1$ the *anisotropy ratio*. The direction with the slowest correlation decay is called the *principal axis*.

6.2. Choice of Covariance Function. The spatial correlation between measurements at different locations is described by the semivariogram functions:

$$
\begin{aligned}
\gamma\left(h\right) &= \frac{1}{2}\text{var}\left(Z\left(s_i\right) - Z\left(s_j\right)\right) \\
&= \frac{1}{2}\left\{\text{var}\left(Z\left(s_i\right)\right) + \text{var}\left(Z\left(s_j\right)\right)\right. \\
&\quad \left. -2\,\text{cov}\left(Z\left(s_i\right), Z\left(s_j\right)\right)\right\} \\
&= C\left(0\right) + C\left(h\right),
\end{aligned}
\tag{32}
$$

where $C(0)$ is the variance and $C(h)$ is the covariance. The variogram and the covariance contain the same information and can be used interchangeably.

In this study, the spatial correlation function $C(h)$ is defined by the Matérn function and is given by

$$
C\left(h\right) = \frac{\sigma^2}{2^{\nu-1}\Gamma\left(\nu\right)}\left(\kappa\|h\|\right)^{\nu} K_{\nu}\left(\kappa\|h\|\right).
\tag{33}
$$

$h = \|s_i - s_j\| \in \mathbb{R}^{+}$ is the Euclidean spatial distance between locations. K_{ν} is the modified Bessel function of
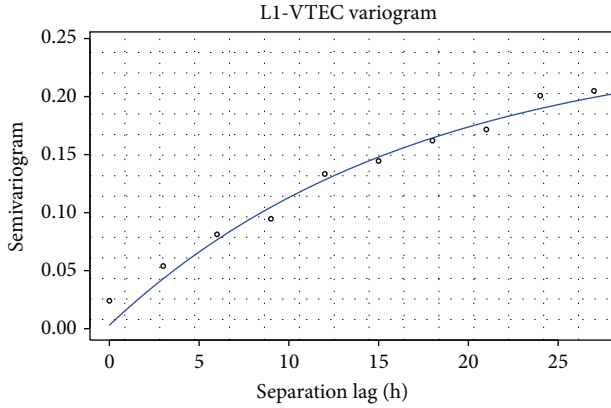
FIGURE 5: An example of the Matérn covariance function, using data with moderate ionosphere activity. Empirical variogram (small circles) and Matérn model fitted by maximum likelihood (solid line). The parameter values used were: smoothness κ = 0.5, partial sill σ^2 = 0.185, and range parameter ϕ = 4.05.

the second kind [16], the order $v > 0$ measures the degree of the smoothness of the process, and κ is the scaling parameter related to the distance of decorrelation (dependency becomes almost 0).

$C(h)$ is obtained from spectral densities [17, p. 31] of the form

$$f(\omega) = \phi\left(\alpha^2 + \omega^2\right)^{-(v+1/2)} \quad v > 0, \ \phi > 0, \ \alpha > 0. \quad (34)$$

Figure 5 shows the empirical semivariogram $\gamma(h)$ with the Matérn covariance function, which fits the L1-VTEC data well. It also works in a wide range of circumstances; including low, moderate, and high ionospheric activities, tested with several different reference station configurations, more specifically 75, 100, and 115 stations.

6.3. Computation of the Inverse Matrix. The kriging equation (25) requires the inverse of the covariance matrix to be computed, and this is detrimental to the performance of the algorithm for large data sets. The operation may occasionally even fail to invert the matrix. Numerical methods with optimization algorithms will help us avoid this, for instance, factorization methods, ill-conditioned test, and other suitable methods.

7. Performance Parameters

In order to carry out the performance analysis of each individual algorithm, an averaging weighted reference signal X was constructed. It is defined as a linear combination of values generated by algorithms with different weights, that is, $X = \sum_{i=1}^{5} w_i \text{Alg}_i$, under the normalization constraint $\sum_{i=1}^{5} w_i = 1$. Five algorithms are involved to construct the reference signal X.

The weights are chosen according to algorithm performance measured in terms of minimum value and stability of variance, functionality, correctness, and processing time. Figure 6 shows the variance of two algorithms. We see
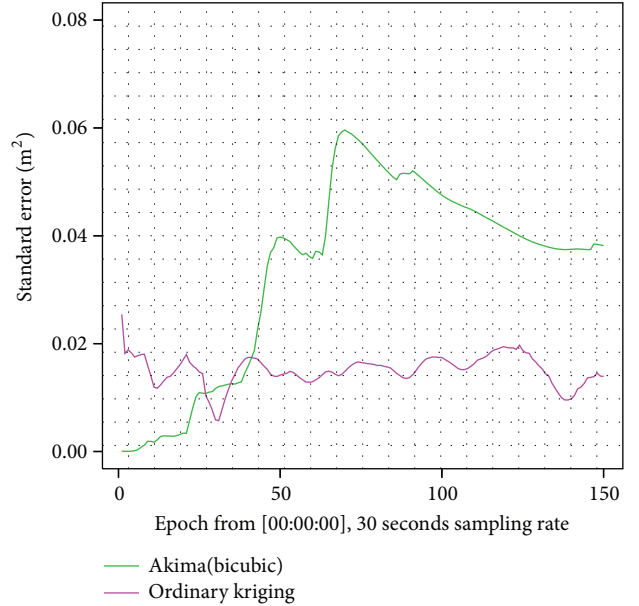
FIGURE 6: The ordinary kriging algorithm (BLUP, pink curve) has a stable and minimum variance compared to the AKIMA algorithm (green curve). Normal ionospheric activity (year 2013, doy: 153), 110 reference stations.

TABLE 1: Definition of algorithm weights.

Algorithm type	Weights	Comments
OK	w_1 = .25	Produces trustworthy results with MCF
TPS	w_2 = .25	A real competitor of kriging, always delivers, even when the covariance structure is poor
GMRF	w_3 = .25	Trustworthy results, like OK and TPS
Akima	w_4 = .15	Handles small variations perfectly. The weight is reduced compared to OK, GMRF, and TPS
IDW	w_5 = .10	No covariance structure is preserved. The assigned weight is reduced compared to OK, GMRF, TPS, and Akima

that ordinary kriging has a minimum and stable variance; therefore its weight is higher than for the Akima bicubic spline. Table 1 summarizes the weight assignment for the algorithms.

7.1. Quality of Service Parameters Definitions. For each one of the quality of service (QoS) parameters whose values are negotiable, the worst case performance must be specified. In some cases, the minimum or the maximum values are preferable, in other cases the averaged value.

The criteria chosen for performance evaluation in this paper are based on comparing the reference signal X to the

TABLE 2: Quality of service (QoS) parameter definitions.

Parameters	Explanations
MAE	Mean absolute error
MSE	Mean square error
RMSE	Root mean square error
NSE	Nash-Sutcliffe efficiency
KGE	Kling-Gupta efficiency
R	Pearson correlation coefficient
β	Changes in locations (mean)
α	Changes in scale (variance)
CT	Computation time is set to 5 seconds. In NRTK processing the values most used are:
	ionosphere: 10 seconds
	geometrical: 15–20 seconds

output Y_k from algorithm k. Analysis is based on statistical monitoring and detecting the changes in spatial location, scale, and level. The full list is given in Table 2.

The required length of time series before we can carry out the goodness of fit is a critical parameter. With the data sets used for testing, values in the range of 100–200 epochs were acceptable.

All algorithms compete about the QoS, the one with highest score is selected as the winner, and the corrections from this algorithm are used. 12 locations (can be regarded as VRS) are chosen inside the CPOS RTK network for testing, and one location is chosen randomly for each run to compute the QoS. The mathematical definitions of the QoS parameters are given in Table 2.

(i) Mean Absolute Error (MAE). MAE measures the average absolute error and is defined below. Ideally, this value should be as small as possible. Consider

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |X_i - Y_i|. \tag{35}$$

(ii) Mean Square Error (MSE). This measures the average squared error and is defined below. This value should also be as close to zero as possible. Consider

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (X_i - Y_i)^2. \tag{36}$$

(iii) Root Mean Square Error (RMSE). RMSE between reference signals X_i and Y_i gives the standard deviation of the algorithm prediction error and minimum value is preferable. Consider

$$\text{RMSE} = \left(\frac{1}{N} \sum_{i=1}^{N} (X_i - Y_i)^2 \right)^{1/2}. \tag{37}$$

(iv) Nash-Sutcliffe Efficiency (NSE) [18]. NSE determines the relative magnitude of the noise variance compared to the

observed data variance. Unity means a perfect match, and zero means that algorithm predictions are as accurate as the mean of the observed information, while negative values imply that the observed mean is better than the predicted one. Consider

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N} (X_i - Y_i)^2}{\sum_{i=1}^{N} (X_i - \overline{X_i})^2}. \tag{38}$$

(v) Kling-Gupta Efficiency (KGE) [19]. KGE was developed by Gupta et al. as a goodness of fit that decomposes the NSE to facilitate the analysis of correlation, bias, and variability. Consider

$$\text{KGE} = 1 - \left\{ (R-1)^2 (\alpha-1)^2 (\beta-1)^2 \right\}^{1/2}. \tag{39}$$

Three components are involved in computation of this index.

(a) R is the Pearson product moment correlation coefficient, which ideally should tend to unity. This quantity is defined by the expression

$$R = \frac{N \left(\sum X_i Y_i \right) - \left(\sum X_i \right) \left(\sum Y_i \right)}{\left\{ \left[N \sum (X_i)^2 - \left(\sum X_i \right)^2 \right] \left[N \sum (Y_i)^2 - \left(\sum Y_i \right)^2 \right] \right\}^{1/2}}. \tag{40}$$

(b) β represents the change in locations. This index is defined as the ratio between distribution locations (means), and the ideal value is unity. Consider

$$\beta = \frac{\mu_O}{\mu_C}. \tag{41}$$

(c) Variability ratio (VR) represents changes in scale (variances). This index is defined as the ratio between distribution standard deviations, and the ideal value is again unity. Consider

$$\text{VR} = \frac{\sigma_C / \mu_C}{\sigma_O / \mu_O}. \tag{42}$$

(vi) Computation Time (CT). Algorithms with high quality data and minimum CT are preferable.

(vii) Coefficient of Determination (R2). $0 \leqslant R2 \leqslant 1$ and gives the portion of the variance of one variable that is predictable from the other variable.

(viii) Spearman Correlation Coefficient (ρ). $-1 \leqslant \rho \leqslant 1$ is a nonparametric test used to measure the degree of associations between two variables.

8. Implementation and Analyses

Packages used in the implementation are downloaded from the Comprehensive R Archive Network (CRAN). Table 3 gives a full description of each package.

Table 3: Implementation packages.

Algorithm	Package name	Download	Comments
Kriging	geoR	CRAN	Prediction by OK (i) AIC computation (ii) BIC computation (iii) MLE estimation (iv) WLS estimation (v) Anisotropy
GMRF	INLA	CRAN	Prediction by GMRF
TPS	rgcvpack	CRAN	Prediction by TPS (i) GCV to estimate λ
AKIMA	akima	CRAN	Bilinear interpolation Bicubic interpolation
IDW	gstat	CRAN	IDW interpolation
Normality	Moments	CRAN	Jarque-Bera normal test
Smoothing	KernSmooth	CRAN	Wand Kernel smoothing

8.1. Data Preprocessing. The data used in this investigation is ionospheric data obtained from the Norwegian CPOS RTK network. The TEC values are generated using the GFA (geometry-free approach) algorithm. The algorithm takes two steps to process.

(a) From IGS [20], the global ionospheric model (GIM) is available. The ionospheric path delay and differential code biases (DCBs) for each satellite are retrieved from the IONEX (IONosphere map EXchange) format. This information is used to estimate the hardware biases of the reference stations, by using the code observable.

(b) From the previous step, we then use the biases in phase measurement and compute the ionospheric delay.

The procedure is described in more detail in [21].

8.2. Interpretation of Results. The test results show that ordinary kriging with the Matérn covariance function is the most appropriate choice under normal circumstances and produces a smooth solution with acceptable accuracy. The Matérn covariance function is well-behaved even for nonstationary fields and is governed by three parameters: location, scale, and shape.

Stein [17, p. 12] recommended the use of the Matérn model due to its flexibility (ability to model the smoothness of physical processes and possibility to handle nonstationarity).

The processing time of this algorithm increases as the number of observations increase. Another approach is to exclude observations that are far away from the interpolation points and use only a subset of the data for the interpolation. This approach is called *local kriging*.

In order to increase the convergence of OK, we have incorporated the bootstrapping algorithm on historical data to get a very good guess for initial values. Figure 9 illustrates the concept.

One major obstacle of this algorithm is the computation of the inverse matrix in the kriging equation (25). Using the numerical approximation of the inverse matrix, the computation time will improve considerably, as mentioned previously in Section 6.3.

The WLSE (weighted least square estimation) algorithm is preferable to maximum likelihood or restricted maximum likelihood and works in most cases, regardless of the distribution of the observations. If the observations have a Gaussian distribution, the WLS and ML/REML yield the same results.

We are often faced with a nonstationary process where we are interested in estimating the spatial covariance for the entire random field. Guttorp and Sampson [22] proposed a two-step approach for solving this problem, a nonparametric algorithm to estimate the spatial covariance structure for the entire random field without assuming stationarity. The interested reader is referred to [23, pp. 93–95].

When the covariance structure preserves sparsity, numerical approximation methods are preferable to all other methods, as they require less memory and computation time.

TPS algorithm is preferred when performing smoothing rather than interpolating data.

8.3. Algorithms Delay Comparison. In this subsection, the delays caused by different algorithms are investigated. The test results are shown in Figure 7. The number of observations varied between 700 and 3000 ionospheric piercing points (IPPs). The data used for the plot are the first 255 epochs with a resolution of 30 seconds from high ionospheric activities, from the year 2013 and day of year 152.

The delay caused by local methods is shown on the right of Figure 7 and is much lower compared to global methods.

The GMRF has the highest delay over the OK and the TPS. The only challenge of TPS is to select a good smoothing parameter, λ. The modified cross-validation, the generalized cross-validation, and robust GCV all work well.

The IDW methods are local interpolation techniques and use only a subset of the data set to perform the interpolation. The benefit of these methods is the reduced computation time.

8.4. QoS Results. Statisticians are usually more interested in smoothing data than interpolating it. When the data is noisy, the TPS smoothing scheme works best. One major advantage of this approach is that once the curve that represents the variation in the data is constructed, we can retrieve the value at any other location without reinterpolating the data set.

Figure 8 shows the result for an arbitrary prediction point with coordinates (lon, lat) = (5.0, 57.0). The reference signal X is compared to the predicted values generated by the ordinary kriging algorithm with MCF. The computed quality of service parameters (QoS) are presented below the plot.

The results are summarized in Table 4 where the QoS parameters are provided for each of the interpolation algorithms tested. An arbitrary epoch has been picked for the test. High scores are highlighted in bold font. The result shows that the ordinary kriging has the best performance. The TPS comes in second place and is the only real competitor to the ordinary kriging for this case. As kriging has the best performance, the corrections from this algorithm will be used to generate synthetic observations of the user in the field. This
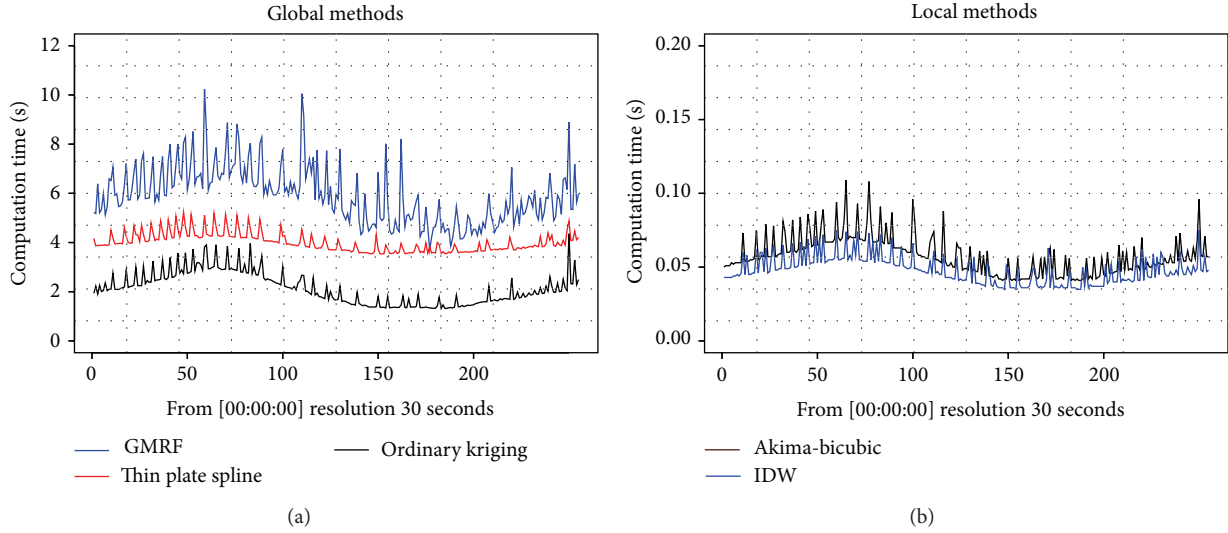
(a)

(b)

FIGURE 7: The left panel shows the CT from global methods, the GMRF (blue curve), TPS (red), and OK (brown). The right panel shows the computation time (CT) caused by the local interpolation methods, the AKIMA (brown) and IDW (blue).

GoF's:
ME = −0.01
MAE = 0.03
RMSE = 0.04
NSE = 0.92
R2 = 0.92
R = 0.96
KGE = 0.93
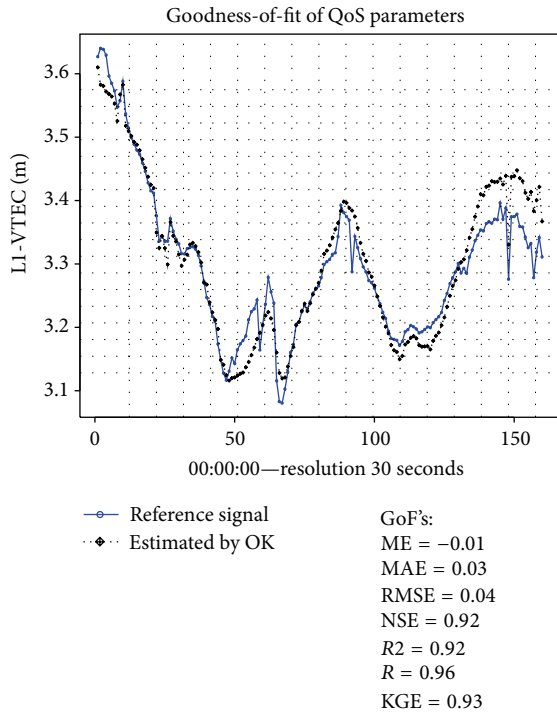
FIGURE 8: QoS computation. Time series of predicted values by ordinary kriging (black curve) and the reference signal (blue curve). Goodness of fit parameters are given below the plot.

comparison to determine the best interpolation algorithm is performed for each epoch.

8.5. Practical Considerations

8.5.1. Ordinary Kriging. The library geoR from CRAN (Comprehensive R Archive Network) is used to implement the spatial interpolation. In order to produce correct results, the
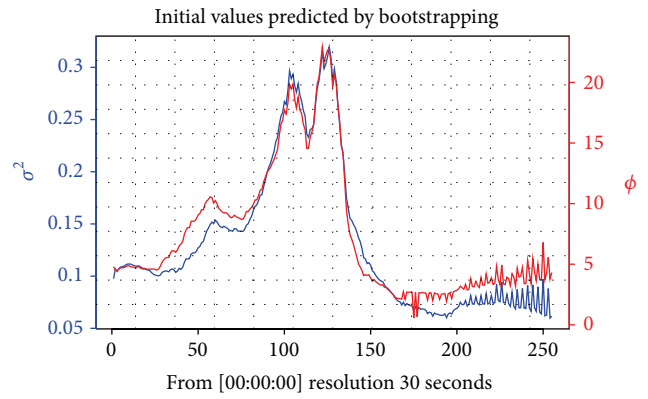
FIGURE 9: Bootstrapping to accelerate the convergence of OK. Partial sill (σ^2, blue curve) and range parameter (ϕ, orange curve).

TABLE 4: Computed QoS by different algorithms.

QoS Parameters	Akima bilinear	Akima bicubic	TPS	GMRF	Kriging
ME	−0.050	0.030	**0.010**	−0.030	0.020
MAE	0.090	0.300	**0.070**	0.090	0.080
MSE	**0.020**	0.410	**0.020**	**0.020**	**0.020**
RMSE	0.150	0.640	**0.130**	0.140	0.140
NSE	0.660	0.440	**0.790**	0.480	0.740
R2	0.780	0.660	0.820	**0.840**	0.780
ρ	.920	.930	.940	**.950**	**.950**
KGE	0.925	0.589	0.911	0.955	**0.976**
R	0.958	0.829	0.953	0.959	**0.980**
β	0.997	1.002	0.994	1.005	**1.001**
α	0.937	0.626	1.076	1.016	**1.013**

data analyst must handle many small exceptions that may otherwise result in unexpected output.

Figure 10: The quantile-quantile (q-q) plot for visualizing normality test. Jarque-Bera test failed and maximum likelihood is not used to estimate parameters.

Figure 11: Nonparametric smoothing with an Epanechnikov kernel is used to determine the L1-VTEC distribution. The distribution is not Gaussian. The weighted least square is used in this case to estimate parameters.

(i) Performance Analysis. The variance of the estimated grids are analyzed. If the variance is very small, this ensures stability of the algorithm.

(ii) Parameter Estimations. In order to increase/accelerate the convergence of ordinary kriging, we have incorporated the bootstrapping algorithm on historical data to get very good estimates of the in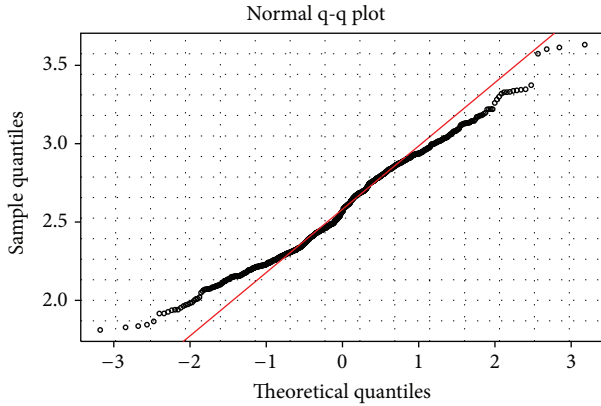itial values. Figure 9 shows the first 255 estimates for the parameters partial sill (σ^2) and range parameter (ϕ), for moderate ionospheric activity, and the network configuration with 75 reference receivers.

(iii) Anisotropy. The optimal estimated kriging weights (negative weights) and variances are very sensitive to anisotropy. Our aim is to ensure that the spatial process does not depend on direction. The geometry anisotropy correction is applied by transforming a set of coordinates according to the geometric anisotropy parameters. The package gstat does not provide automatic fitting of anisotropy parameters, while the *geoR* package transforms/backtransforms a set of coordinates according to the estimated geometric anisotropy parameters.

(iv) Normality Test. The MLE (maximum likelihood estimation) procedure requires that the observations are Gaussian distributed, this assumption is violated in most cases. Therefore the Jarque-Bera test is used as a test of the normality and is based on the third and fourth moments of a distribution, called skewness and kurtosis coefficients; the interested is referred to [24]. If the test fails, the weighted least square method is used to estimate the parameters. Figures 10 and 11 from a configuration with 100 sites and high ionospheric activity confirm that the L1-VTEC distribution is not normally distributed.

Based on the tests and checks mentioned above, the ordinary kriging is assigned a weight of 0.25 when computing the QoS values.

8.5.2. Test Results. Figure 12 shows the smoothed curve generated by TPS software with GCV. Once the curve is determined, we can easily retrieve any value inside the coverage area without extra computation compared to other
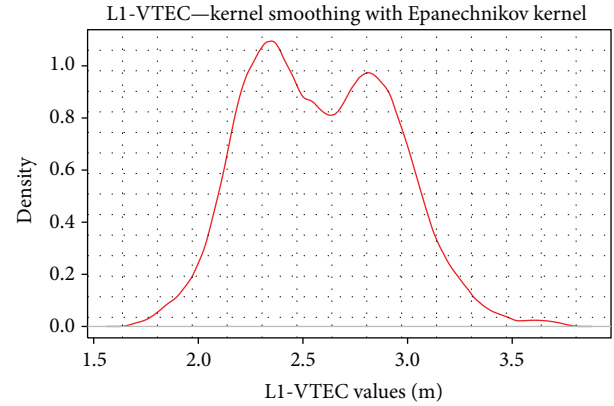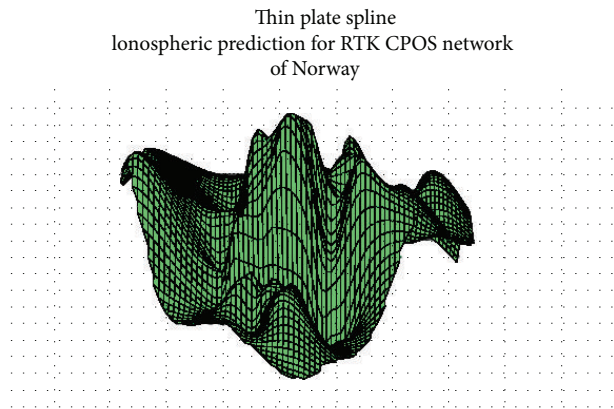
Figure 12: L1-VTEC variation surface generated by the TPS algorithm. The modes are well represented in this case.

interpolation methods. In addition, this describes the variation very well.

Once the smoothing parameter λ is determined by GCV, TPS is the real competitor of the kriging algorithm and the weight assigned to it is 0.25.

9. Conclusion

A significant improvement of the rover position estimation can be achieved by applying the right interpolation/smoothing algorithm at the NRTK interpolation segment. This will reduce the information loss under prediction of the user error level and will provide high quality of virtual reference station data from epoch to epoch.

Five methods have been suggested to generate the rover correction. The study shows that the kriging interpolator, the ordinary kriging with the Matérn covariance function, is the most appropriate choice for weighted spatial linear interpolation, while TPS is a strong competitor of OK when the aim is to smooth, not to interpolate, the data. After performing matrix sparsity tests, the GMRF is computationally effective,

TABLE 5: CPOS network reference station characteristics.

Site name	4chars	ID	Lon	Lat	Receiver	Antenna
Aas	aasc	189	10.78	59.66	NETR9	TRM57971.00
Adamselv	adac	147	26.70	70.41	NETR9	TRM57971.00
Akrahamn	akrc	13	5.19	59.26	NETR9	TPSCR3_GGD
Alesund	ales	135	6.20	62.48	NETR8	TRM55971.00
Alta	altc	70	23.30	69.98	NETR5	TRM55971.00
Alvdal	alvc	51	10.63	62.11	NETR5	TRM55971.00
Andenes	ande	69	16.13	69.33	NETR5	TRM55971.00
Ardal	arda	177	7.70	61.24	NETR9	TRM57971.00
Arendal	arec	27	8.71	58.41	NETR9	TRM55971.00
Arnes	arnc	99	11.48	60.12	NETR5	TRM55971.00
Asane	asac	169	5.34	60.47	NETR9	TRM57971.00
Balsfjord	balc	71	19.23	69.24	NETR5	TRM55971.00
Bardu	barc	64	18.35	68.86	NETR5	TRM55971.00
Batsfjord	batc	132	29.73	70.64	NETR5	TRM57971.00
Bergen	brgs	150	5.27	60.29	NETR8	TRM55971.00
Berkak	berc	23	10.01	62.83	NETR5	TRM55971.00
Bjarkoy	bjac	63	16.57	69.00	NETR5	TRM55971.00
Bjorli	bjoc	127	8.20	62.26	NETR9	TRM57971.00
Bjugn	bjuc	61	9.81	63.77	NETR5	TRM55971.00
Bleikvassli	blec	40	13.81	65.89	NETR5	TRM55971.00
Bo	boec	2	14.47	68.68	NETR5	TRM55971.00
Bodo	bod3	121	14.43	67.29	NETR8	TRM55971.00
Bomlo	bmlc	142	5.20	59.80	NETR5	TRM55971.00
Breivikbotn	brec	134	22.30	70.59	NETR5	TRM57971.00
Bygland	bygc	28	7.80	58.83	NETR5	TRM55971.00
Dagali	dgls	160	8.50	60.42	NETR8	TRM55971.00
Dombas	doms	140	9.11	62.07	NETR8	TRM55971.00
Donna	donc	90	12.47	66.10	NETR5	TRM55971.00
Drevsjo	drec	144	12.03	61.88	NETR5	TRM55971.00
Eidsvoll	eids	176	11.26	60.33	NETR9	TRM57971.00
Etne	etne	178	5.93	59.66	NETR9	TRM57971.00
Fauske	faus	183	15.39	67.26	NETR9	TRM57971.00
Fet	fetc	173	11.17	59.92	NETR9	TRM57971.00
Finnoy	fioy	180	5.87	59.17	NETR9	TRM57971.00
Finnsnes	finc	65	17.99	69.23	NETR5	TRM55971.00
Flisa	flic	98	12.01	60.61	NETR5	TRM55971.00
Floro	floc	42	5.04	61.60	NETR5	TRM55971.00
Folling	folc	91	11.62	64.12	NETR5	TRM59800.00
Forde	forc	168	5.89	61.46	NETR9	TRM57971.00
Fredrikstad	frec	143	10.94	59.21	NETR9	TRM57971.00
Froya	froc	60	8.66	63.87	NETR5	TRM55971.00
Gjora	gjoc	12	9.11	62.55	NETR9	TRM57971.00
Gloppen	gloc	41	6.19	61.77	NETR5	TRM55971.00
Grong	groc	3	12.31	64.46	NETR5	TRM55971.00
Hammerfest	hamc	137	23.66	70.67	NETR5	TRM57971.00
Hansnes	hanc	72	19.63	69.97	NETR5	TRM55971.00
Hardbakke	harc	38	4.84	61.08	NETR9	TPSCR3_GGD
Haukeli	hauc	79	7.20	59.81	NETR5	TRM55971.00
Havoysund	havc	141	24.66	71.00	NETR5	TRM57971.00
Hedalen	hedc	31	9.74	60.61	NETR5	TRM55971.00
Heggenes	hegc	50	9.07	61.15	NETR5	TRM55971.00
Hellesylt	helc	18	6.76	62.04	NETR5	TRM55971.00
Hemne	hemc	59	9.08	63.29	NETR5	TRM55971.00
Holasen	holc	45	11.18	63.70	NETR5	TRM55971.00
Honefoss	hfs4	1	10.25	60.14	NETR5	TRM59800.00

TABLE 5: Continued.

Site name	4chars	ID	Lon	Lat	Receiver	Antenna
Honningsvag	hons	111	25.96	70.98	NETR8	TRM59800.00
Horten	horc	158	10.47	59.41	NETR9	TRM57971.00
Hustad	husc	15	7.14	62.97	NETR5	TRM55971.00
Ibestad	ibec	153	17.18	68.78	NETR5	TRM55971.00
Innfjorden	innc	17	7.55	62.50	NETR5	TRM55971.00
Jostedalen	josc	167	7.28	61.66	NETR9	TRM57971.00
Karasjok	karc	139	25.52	69.47	NETR5	TRM57971.00
Kautokeino	kaus	112	23.02	69.02	NETR8	TRM59800.00
Kirkenes	kirc	129	30.04	69.73	NETR5	TRM57971.00
Kjopsvik	kjoc	82	16.39	68.10	NETR5	TRM55971.00
Kobbelv	kobc	83	15.89	67.58	NETR5	TRM55971.00
Kongsvinger	konc	172	12.00	60.19	NETR5	TRM57971.00
Koppang	kopc	52	11.04	61.57	NETR5	TRM55971.00
Kristiansand	krss	185	7.91	58.08	NETR5	TRM55971.00
Kristiansund	krsu	162	7.73	63.11	NETR9	TRM57971.00
Kvanangsbotn	kvac	73	22.06	69.72	NETR5	TRM55971.00
Kvikne	kvic	122	10.32	62.56	NETR9	TRM57971.00
Kyrkjebo	kyrc	58	5.90	61.16	NETR9	TRM57971.00
Lakselv	lakc	138	24.96	70.05	NETR9	TRM57971.00
Lauvsnes	lauc	14	10.90	64.50	NETR5	TRM55971.00
Leikanger	leic	6	6.86	61.18	NETR5	TRM55971.00
Leknes	lekc	136	13.61	68.15	NETR5	TRM57971.00
Leksvik	leks	186	10.63	63.67	NETR9	TRM57971.00
Lierne	liec	30	13.65	64.40	NETR5	TRM59800.00
Lillehammer	lilc	32	10.44	61.14	NETR5	TRM55971.00
Lindas	linc	37	5.16	60.73	NETR9	TPSCR3_GGD
Lista	lstc	24	6.69	58.09	NETR9	TPSCR3_GGD
Lodingen	lodc	62	15.99	68.41	NETR9	TRM57971.00
Lofoten	lofs	119	13.04	67.89	NETR8	TRM55971.00
Lom	lomc	126	8.57	61.84	NETR8	TRM57971.00
Lonsdal	lonc	88	15.46	66.74	NETR5	TRM55971.00
Loppa	lopc	74	22.35	70.24	NETR5	TRM55971.00
Loten	lotc	163	11.35	60.82	NETR5	TRM57971.00
Luroy	lurc	48	13.01	66.51	NETR5	TRM55971.00
Lysefjorden	lysc	9	6.39	59.03	NETR5	TRM55971.00
Mare	mare	184	11.43	63.93	NETR5	TRM57971.00
Maurset	mauc	124	7.33	60.41	NETR9	TRM57971.00
Maze	mazc	149	23.67	69.45	NETR9	TRM57971.00
Mebonden	mebc	46	11.03	63.23	NETR5	TRM55971.00
Mehamn	mehc	133	27.85	71.04	NETR5	TRM57971.00
Mo i Rana	moic	49	14.14	66.31	NETR5	TRM55971.00
Moelv	moec	174	10.70	60.93	NETR9	TRM57971.00
Molde	mldc	159	7.15	62.73	NETR9	TRM57971.00
Moldjord	molc	84	14.57	67.01	NETR5	TRM55971.00
Mosjoen	mosj	188	13.20	65.84	NETR5	TRM55971.00
Myre	myrc	67	15.09	68.91	NETR5	TRM55971.00
Mysen	mysc	97	11.33	59.57	NETR5	TRM55971.00
Namsos	nams	187	11.51	64.47	NETR9	TRM57971.00
Narvik	narc	66	17.43	68.44	NETR5	TRM55971.00
Olderdalen	oldc	75	20.53	69.60	NETR5	TRM55971.00
Opera	opec	4	10.75	59.91	NETR5	TRM55971.00
Ornes	ornc	47	13.73	66.86	NETR5	TRM55971.00
Oslo	osls	175	10.37	59.74	NETR8	TRM55971.00
Osterbo	ostc	81	7.51	60.83	NETR5	TRM55971.00
Overbygd	ovec	152	19.29	69.03	NETR9	TRM55971.00

Table 5: Continued.

Site name	4chars	ID	Lon	Lat	Receiver	Antenna
Pasvik	pasc	148	29.69	69.37	NETR9	TRM57971.00
Porsgrunn	prgc	171	9.66	59.14	NETR9	TRM57971.00
Portor	porc	7	9.43	58.80	NETR5	TPSCR3_GGD
Prestasen	prec	10	6.25	59.49	NETR9	TPSCR3_GGD
Rauland	rauc	55	8.05	59.70	NETR9	TRM55971.00
Rena	renc	33	11.37	61.14	NETR5	TRM55971.00
Roan	roac	92	10.30	64.21	NETR5	TRM55971.00
Roros	rorc	123	11.39	62.58	NETR5	TRM55971.00
Rosendal	rosc	35	6.01	59.98	NETR9	TPSCR3_GGD
Royrvik	royc	19	13.53	64.90	NETR5	TRM59800.00
Sandvika	svic	151	10.52	59.89	NETR9	TRM55971.00
Seljord	selc	43	8.63	59.49	NETR5	TRM55971.00
Sirevag	sirc	11	5.79	58.50	NETR9	TPSCR3_GGD
Skaland	skac	154	17.30	69.45	NETR9	TRM57971.00
Stjordal	stjc	170	10.92	63.47	NETR9	TRM57971.00
Sulitjelma	sulc	86	16.08	67.12	NETR9	TRM55971.00
Sveindal	svec	26	7.47	58.49	NETR5	TRM55971.00
Svolvar	svoc	87	14.56	68.23	NETR5	TRM55971.00
Tana	tanc	131	28.19	70.20	NETR5	TRM55971.00
Terrak	terc	57	12.38	65.09	NETR5	TRM55971.00
Tingvoll	tinc	16	8.21	62.91	NETR5	TRM55971.00
Tjome	tjmc	68	10.40	59.13	NETR5	TRM55971.00
Tonstad	tnsc	34	6.71	58.67	NETR5	TRM55971.00
Tregde	tgde	96	7.55	58.01	NETR9	AOAD/M_T
Treungen	trec	25	8.52	59.02	NETR5	TRM55971.00
Trofors	troc	39	13.39	65.54	NETR5	TRM55971.00
Tromso	tro1	115	18.94	69.66	NETR8	TRM59800.00
Trondheim	trds	125	10.32	63.37	NETR8	TRM55971.00
Trysil	trys	145	12.38	61.42	NETR8	TRM55971.00
Tyin	tyic	89	8.23	61.18	NETR5	TRM55971.00
Tysvar	tysv	179	5.40	59.43	NETR9	TRM57971.00
Ulefoss	ulec	8	9.28	59.28	NETR5	TPSCR3_GGD
Ulsak	ulsc	54	8.62	60.84	NETR5	TRM55971.00
Vadso	vadc	130	29.74	70.08	NETR5	TRM57971.00
Valle	valc	78	7.48	59.25	NETR5	TRM55971.00
Vardo	vars	110	31.03	70.34	NETR8	TRM59800.00
Vega	vegs	118	11.96	65.67	NETR8	TRM59800.00
Veggli	vegc	44	9.17	60.04	NETR9	TRM55971.00
Vikna	vikc	56	11.24	64.86	NETR5	TRM55971.00
Vinstra	vinc	53	9.75	61.60	NETR5	TRM55971.00
Volda	vold	181	6.08	62.14	NETR9	TRM57971.00
Voss	vssc	128	6.42	60.63	NETR9	TPSCR3_GGD
Skjervoy	skjc	76	20.98	70.03	NETR5	TRM55971.00
Skollenborg	skoc	5	9.68	59.62	NETR9	TPSCR3_GGD
Skreia	skrc	29	10.93	60.65	NETR5	TRM55971.00
Smola	smol	182	7.96	63.51	NETR5	TRM55971.00
Smorfjord	smrc	146	24.95	70.52	NETR9	TRM57971.00
Stadt	stac	20	5.32	62.12	NETR5	TRM55971.00
Stavanger	stas	165	5.60	59.02	NETR8	TRM55971.00
Steigen	stgc	164	15.02	67.78	NETR5	TRM55971.00

requires less memory, and produces good results as TPS and OK.

For local methods the covariance structure is in general not conserved. For gentle variation in data, the Akima with
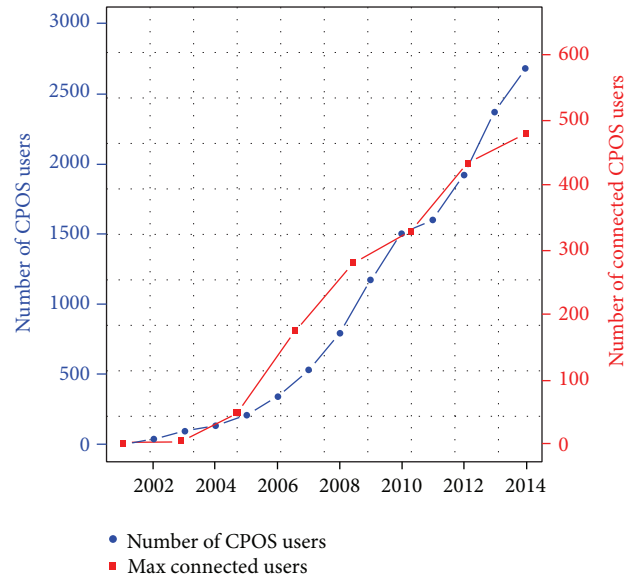

Figure 13: Number of Norwegian CPOS RTK users (blue curve) increased from 1 to ~3000 and the users connected to the system at the same time (red curve) increased from 1 to ~500.

bicubic method is an appropriate choice because it is the real spline method. While IDW is stable, it is inaccurate and in addition does not conserve the covariance structure of the process under study.

One major benefit of these techniques is that there is no need for any prior estimation of the spatial dependency, as in the case of Bayesian analysis (e.g., Kalman filter).

10. Discussions

(1) As we mentioned in the Introduction, processing large data sets is a challenge of the future, and our suggestion for how to handle this is formulated as follows. First of all, we already have enough mathematical tools to do the job, so we do not need to develop new ones. These tools can be considered as elementary building blocks in the hands of the data analyst/modeler. The main challenge is to know that the right tools exist, what they can do for us, what their strengths and weaknesses are, and how to combine them in appropriate ways to describe the observed variations as well as possible. Figure 13 shows the number of users connected at the same time and the historical data of the users using the CPOS services in Norway. Both curves increase exponentially in a period of one decade, and if the development continues to follow the same pattern, the existing tools will not be sufficient to process the large data sets.

(2) Data quality and quantity are important to perform reliable statistical analysis, and elementary checks are necessary before starting analysis.

(3) In geodesy and geophysical data analysis, the Gauss-Markov model and Kalman Filter are often considered when modeling and when state estimation is necessary. Since new navigation satellite system (e.g., Galileo, Beidou) in addition to the old GPS and GLONASS becomes operation, massive data sets need to be processed in real-time, so we are experiencing a computational paradigm shift.

(4) To avoid information loss between the correction and interpolation segments, picking the right algorithm for the job is essential for improving the user position errors.

Appendix

CPOS Station Characteristics

See Table 5.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements, and Performance*, Ganga-Jamuna Press, 2006.

[2] H. Landau, U. Vollath, and X. Chen, "Virtual reference station systems," *Journal of Global Positioning Systems*, vol. 1, no. 2, pp. 137–143, 2002.

[3] F. Takac and O. Zelzer, "The relationship between network RTK solutions MAC, VRS, PRS, FKP and i-MAX," in *Proceedings of the 21st International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS '08)*, pp. 348–355, 2008.

[4] N. A. C. Cressie, *Statistics for Spatial Data*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York, NY, USA, 1993.

[5] J. Dennis and J. J. More, "Quasi-Newton methods, motivation and theory," *SIAM Review*, vol. 19, no. 1, pp. 46–89, 1977.

[6] M. A. Luersen and R. le Riche, "Globalized nelder-mead method for engineering optimization," *Computers and Structures*, vol. 82, no. 23–26, pp. 2251–2260, 2004.

[7] W. J. Gordon and J. A. Wixom, "Shepard's method of 'metric interpolation' to bivariate and multivariate interpolation," *Mathematics of Computation*, vol. 32, no. 141, pp. 253–264, 1978.

[8] R. J. Renka and R. Brown, "Algorithm 792: accuracy tests of ACM algorithms for interpolation of scattered data in the plane," *ACM Transactions on Mathematical Software*, vol. 25, no. 1, pp. 78–94, 1999.

[9] H. Akima, "A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points," *ACM Transactions on Mathematical Software*, vol. 4, no. 2, pp. 148–159, 1978.

[10] H. Akima, "Algorithm 761: scattered-data surface fitting that has the accuracy of a cubic polynomial," *ACM Transactions on Mathematical Software*, vol. 22, no. 3, pp. 362–371, 1996.

[11] C. H. Reinsch, "Smoothing by spline functions," *Numerische Mathematik*, vol. 10, pp. 177–183, 1967.

[12] J. Duchon, "Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces," *Analyse Numerique*, vol. 10, pp. 5–12, 1976.

[13] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, USA, 2010.

[14] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society Series B: Methodological*, vol. 48, no. 3, pp. 259–302, 1986.

[15] H. Rue and L. Held, *Gaussian Markov Random Fields: Theory and Applications*, CRC Monographs on Statistics & Applied Probability, CRC Press, London, UK, 2005.

[16] M. Abramowitz, I. A. Stegun, and D. Miller, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, vol. 32, Dover Publications, 1965.

[17] M. L. Stein, *Interpolation of Spatial Data*, Springer Series in Statistics, Springer, New York, NY, USA, 1999.

[18] J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models. Part I. A discussion of principles," *Journal of Hydrology*, vol. 10, no. 3, pp. 282–290, 1970.

[19] H. V. Gupta, H. Kling, K. K. Yilmaz, and G. F. Martinez, "Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling," *Journal of Hydrology*, vol. 377, no. 1-2, pp. 80–91, 2009.

[20] J. M. Dow, R. E. Neilan, and C. Rizos, "The international GNSS service in a changing landscape of global navigation satellite systems," *Journal of Geodesy*, vol. 83, no. 3-4, pp. 191–198, 2009.

[21] N. Kjorsvik, O. Ovstedal, and J. G. Svendsen, "Estimating local ionosphere parameters in Norway," in *Proceedings of the 15th International Technical Meeting of the Satellite Division of the Institute Of Navigation (ION GPS '02)*, pp. 2304–2310, 2002.

[22] P. Guttorp and P. D. Sampson, "Methods for estimating heterogeneous spatial covariance functions with environmental applications," in *Handbook of Statistics XII: Environmental Statistics*, pp. 663–690, 1994.

[23] N. D. Le and J. V. Zidek, *Statistical Analysis of Environmental Space-Time Processes*, Springer Series in Statistics, Springer, New York, NY, USA, 1st edition, 2006.

[24] C. M. Jarque and A. K. Bera, "A test for normality of observations and regression residuals," *International Statistical Review*, vol. 55, no. 2, pp. 163–172, 1987.

The Scientific
World Journal

International Journal of
Rotating
Machinery

Journal of
Engineering

Journal of
Sensors

Advances in
Mechanical
Engineering

International Journal of
Distributed
Sensor Networks

Advances in
Civil Engineering

Advances in
OptoElectronics

Journal of
Robotics

International Journal of
Chemical Engineering

VLSI Design

International Journal of
Navigation and
Observation

Modelling &
Simulation
in Engineering

Advances in
Acoustics and Vibration

Active and Passive
Electronic Components

International Journal of
Antennas and
Propagation

Journal of
Control Science
and Engineering

Shock and Vibration

Journal of
Electrical and Computer
Engineering

Hindawi

Submit your manuscripts at
http://www.hindawi.com