

Regional TEC mapping with Random Field Priors and Kriging

I. Sayin,¹ F. Arikan,¹ and O. Arikan²

Received 23 November 2007; revised 1 July 2008; accepted 21 July 2008; published 28 October 2008.

[1] Total Electron Content (TEC) is an important parameter in monitoring ionospheric variability. In a given region, TEC can be obtained only by interpolation of measurements due to sparsity of the useful data. The lack of a complete ionospheric model hinders the choice of the optimum interpolation algorithm. A plausible solution to this problem can be found by investigating the performance of alternative interpolation algorithms on synthetically generated TEC surfaces for various sampling scenarios. The synthetic TEC data should represent the possible trends and variations of ionosphere. In this study, the performance of Random Field Priors (RFP) and Kriging interpolation algorithms are investigated over the parameter set of spatially correlated synthetic TEC data for various variance, range and trend options. Synthetic TEC data are sampled with regular and random sampling patterns, for number of samples from sparse to dense samplings. Interpolation scenarios are generated to investigate the improvement of the interpolation accuracy of the methods for each parameter. It is observed that for the random sampling patterns, when the trend is not modeled correctly, the errors of the algorithms increase and when the trend is modeled correctly, the reconstruction errors decrease. For the regular sampling patterns, the trend model does not affect the accuracy of the methods, and the reconstruction errors are close to lower bound error values. An example reconstruction is also provided over GPS-derived TEC, and error variances are compared over Kriging and Random Field Prior algorithms.

Citation: Sayin, I., F. Arikan, and O. Arikan (2008), Regional TEC mapping with Random Field Priors and Kriging, *Radio Sci.*, 43, RS5012, doi:10.1029/2007RS003786.

1. Introduction

[2] Ionosphere is the layer of the Earth's atmosphere that is dominated by a dispersive, temporally and spatially varying ionized plasma. Robust, reliable, and accurate imaging and monitoring of the ionosphere is a major challenge in performance improvement of communication and navigation systems. Total Electron Content (TEC) is a key parameter in investigation of the ionospheric variability. TEC is defined as the line integral of electron density along a ray path or as a measure of the total number of electrons along a path of the radio wave [Budden, 1985]. The unit of TEC is given in TECU where $1 \text{ TECU} = 10^{16} \text{ el/m}^2$. TEC is a derived quantity

and it is a function of electron density and the chosen ray path. TEC can be computed using the measurements and recordings of the vertical ionosondes both bottom-side and top-side, Faraday Rotation of satellite signals such as GLONASS and EISCAT, TOPEX/POSEIDON and JASON satellites with double frequency altimeters, GPS phase and delay recordings and incoherent backscatter radar signals [Komjathy, 1997]. In recent years, Global Positioning System (GPS) dual frequency signals are widely used to estimate both regional and global TEC values [Schaer, 1999]. In GPS-TEC computation, TEC on the slant ray path from the satellite to the receiver is called the slant TEC (*STEC*). When the *STEC* values are projected to the local zenith at the ionospheric pierce point, assuming the thin shell model of the ionosphere with a mapping function, the computed TEC value is called the vertical TEC (*VTEC*) [Arikan *et al.*, 2003, 2004, 2007; Nayir *et al.*, 2007]. In this study, the total electron content over the receiver station in the local zenith direction is denoted by the term TEC. Owing to the sparsity of TEC values obtained from the above listed

¹Department of Electrical and Electronics Engineering, Hacettepe University, Ankara, Turkey.

²Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey.

sources in time and space, accurate and robust interpolation of TEC for every desired location in a given region is a very challenging problem. The estimation of TEC values on a dense grid using the limited set of computed TEC is called TEC mapping [Hanbaba, 1999; Sayin, 2008].

[3] In the literature, there are various interpolation methods applied to ionospheric TEC mapping. A brief list includes Inverse Distance Weighting (IDW) [Jin *et al.*, 2004], Multiquadric function fitting (MQ) [Wielgosz *et al.*, 2003], Thin Plate Spline (TPS) [Moon, 2004], Spherical Harmonics (SH), Kriging (KG) and Neural Networks (NN). All of the above listed methods have both advantages and disadvantages. For example, IDW, MQ and TPS methods have low computational complexity and they are easy to implement. Yet, these methods do not include any statistical or background information on the nature of the ionosphere. SH is widely used in global mapping. The basis functions of SH span the whole surface of the globe and they have to be modified for regional TEC mapping [Schaer, 1999; Moon, 2004; Lazo *et al.*, 2004]. NN has higher computational complexity than the other above mentioned methods, and it is known as universal function approximator [Gurun, 2007]. When extrapolation is considered, NN provides a very powerful alternative [Leandro, 2004]. NN is applied to mapping of various ionospheric parameters, yet it is very difficult to decide on the optimum neural structure and its appropriate learning algorithm when it is applied to real ionospheric data [Gurun, 2007]. Also, NN requires a preprocessing stage to prepare the training and testing data sets, a process that increases the computational and implementation complexity.

[4] KG is a widely used spatial interpolation method in environmental sciences [Cressie, 1993]. It is known as the Best Linear Unbiased Estimator (BLUE). Kriging extracts the variability and spatial correlation structure of the surface from the measurement data [Sayin, 2008; Sayin *et al.*, 2007]. Where IDW, MQ and TPS uses only the geometric location of the samples to estimate the desired variable, KG also makes use of both the parameter value at the sample points and the covariance structure of the sample points [Chiles and Delfiner, 1999]. Kriging gives the estimation error variance from which confidence bounds can be obtained [Blanch, 2002]. KG has been applied to TEC mapping in the literature. In the study of Wielgosz *et al.* [2003], the TEC maps obtained with KG and MQ for the day 29 April 2003 and from the measurements of five stations in Ohio, are compared with the Global Ionospheric Maps (GIM) of International GPS Service for Geodynamics (IGS) centers (<ftp://cddisa.gsfc.nasa.gov/gps/products/ionex/>). It is noted that both MQ and KG maps are similar to each other and TEC estimates from these two methods are generally less than those of GIM in value. In the

work of Wielgosz *et al.* [2003], both MQ and KG methods are found to be suitable for sparse samplings, and MQ method is concluded to be more suitable for real time applications due to lower complexity compared to KG. In the work of Orús *et al.* [2005], accuracy improvement of Technical University of Catalonia (UPC) GIM with Kriging is proved by comparing the standard deviation and root mean square of the difference from the TOPEX/Poseidon-derived TEC for December 2002, and 0.3 TECU improvement in the standard deviation of the difference is noted. In the study of Stanislawska *et al.* [2002], Kriging is applied to real ionospheric data for quiet and disturbed days and the difference in the structure of the ionosphere is observed. Yet, there is no detailed study in the literature on the optimum performance parameters and models of KG for TEC mapping.

[5] Random Field (RF) model provides a space-time stochastic description of “disturbed disordered system” as given in the study of Vanmarcke [1983]. In the work of Vanmarcke [1983], various applications of RF to fields such as geostatistics, environmental sciences, economic risk management are discussed over the statistical model of the random variable. With the assumption that the measurement and estimation data sets have joint Gaussian distribution, Random Field Priors (RFP) method described by Vanmarcke [1983] provides unbiased and linear estimates under the minimum variance criterion. Also, RFP is capable of modeling the space-time structure of the random field and produce accurate interpolations, if a priori correlation matrix structures of the measurement and the field are provided. If the method is applied only for spatial interpolation, the a priori information corresponds to the mean or the trend in the covariance matrices. The RFP model has never been applied to ionospheric mapping before, yet with the strong underlying stochastic structure and the ease of implementation for both space and time estimation, makes it a very strong alternative to other mapping algorithms.

[6] In the interpolation problem, the sparse values of TEC, taken over time and space samples, are connected with certain algorithms and missing values are reconstructed. In order to decide on the optimum interpolation method, it is a necessity to compare the performance of the mapping algorithms as a function of their respective parameters for reliable, robust and accurate TEC maps. The performance of the interpolation algorithms can not be determined accurately over real measurement data due to the fact that exact value of TEC for a point in space and time is not known precisely. TEC is a derived quantity and a proper space-time model of the ionosphere does not exist. It is very important to compare the performance of interpolation algorithms on synthetic TEC surfaces where all the parameters and sample values

are controlled and analyzed. This way, under the pre-determined error criterion, the reconstruction errors for various mapping algorithms can be determined a priori as a function of trend and variability in TEC, number of samples, and sampling pattern. Although, TEC maps can be produced as listed above, the degree of accuracy of those maps are not thoroughly investigated.

[7] The major goal of this study is to analyze the performance of Kriging and RFP algorithms over synthetic TEC surfaces in space for a given time. The synthetic TEC surfaces are generated by investigating the spatial trends of TEC in Global Ionospheric Maps (GIM). The performance parameters of KG and RFP algorithms are chosen to be the deterministic surface trend plus a statistical small scale disturbance, the sampling pattern type, and the number of samples. In this study, TEC maps are obtained by using GPS-TEC as the input to KG and RFP reconstruction algorithms. A comparison over the reconstructions and error variances are also provided. With this study, the optimum interpolation algorithm for a given surface, number of samples and sampling pattern can be chosen and can be compared with the sub-optimum algorithms. The performance improvement over varying parameters can be determined and cost-effective measurement schemes can be developed.

[8] In section 2, the synthetic TEC surface generation and interpolation scenarios are provided along with a general review of KG and RFP algorithms. In section 3 the results of the synthetic TEC interpolation simulations are given and the Interpolation methods are applied to real ionospheric data. The conclusions are in section 4.

2. Regional Synthetic TEC Interpolation With KG and RFP

[9] Ionosphere varies as a complex function of solar and geomagnetic activity, latitude, longitude, and time. These variations in space and time are generally made up of a deterministic trend and a variable scale statistical model. For example, ionosphere has an underlying diurnal and sunspot cycle periodicity in time, and east-to-west and north-to-south gradients in space. Ionospheric disturbances due to geomagnetic and solar activities can alter the underlying trend with an additive random variation. TEC can be modeled as a random field in space and time, $Z(\mathbf{x}, t)$, composed of a zero mean stationary random function $Y(\mathbf{x}, t)$ and a deterministic trend $\mu(\mathbf{x}, t)$ as

$$Z(\mathbf{x}, t) = \mu(\mathbf{x}, t) + Y(\mathbf{x}, t) \quad (1)$$

In the above equation $\mathbf{x} = [\theta \ \phi]^T$, where θ denotes latitude, ϕ denotes longitude and $(\cdot)^T$ is the transpose operator. $\mu(\mathbf{x}, t) = E\{Z(\mathbf{x}, t)\}$ represents the trend in TEC,

and it is generally modeled as a linear combination of known functions such as monomials. $E\{\cdot\}$ is the expectation operator. $Y(\mathbf{x}, t)$ represents the variation above $\mu(\mathbf{x}, t)$. Since distinct measurements related to $\mu(\mathbf{x}, t)$ and $Y(\mathbf{x}, t)$ are usually not available separately, observations in ionospheric physics can be used to estimate the structure of the trend functions. The space-time structure of the ionosphere can be divided into separate space and time functions due to the fact that these variations tend to have distinct trends of their own. In this study, we will concentrate on interpolation of TEC over space. Joint space-time estimations will be covered in follow-up study.

2.1. Interpolation Scenarios

[10] The synthetic TEC surfaces, \mathbf{g}_s , are generated on a dense grid defined in a region from $\theta_i \leq \theta \leq \theta_f$ and $\phi_i \leq \phi \leq \phi_f$ divided into N_θ points in latitude and N_ϕ points in longitude. The lexicographical index $l = n_\theta + (n_\phi - 1)N_\theta$ provides a mathematical ease to handle the spatially distributed TEC values by ordering them in a one dimensional vector as $\mathbf{g}_{s_{N_\theta N_\phi \times 1}}$, where \mathbf{g}_s is defined as:

$$\mathbf{g}_{s_{N_\theta N_\phi \times 1}} = [g_s(1) \ \dots \ g_s(l) \ \dots \ g_s(N_\theta N_\phi)]_{1 \times N_\theta N_\phi}^T \quad (2)$$

where $g_s(l)$ denotes the TEC value at the l^{th} grid point.

[11] For N_a measurements or sample points, the measurement vector is defined as

$$\mathbf{d}_{N_a \times 1} = [d(1) \ \dots \ d(n_a) \ \dots \ d(N_a)]_{1 \times N_a}^T \quad (3)$$

where $d(n_a)$ is the measurement value at the n_a^{th} measurement point. In this study, the measurement points are taken either from \mathbf{g}_s or GPS-derived vertical TEC. Mapping or interpolation can be considered as the problem of estimating \mathbf{g}_s , from \mathbf{d} . TEC estimates from the KG and RFP algorithms at the grid points can be given by

$$\hat{\mathbf{z}}_s = [\hat{Z}_s(1) \ \dots \ \hat{Z}_s(l) \ \dots \ \hat{Z}_s(N_\theta N_\phi)]_{1 \times N_\theta N_\phi}^T \quad (4)$$

where $\hat{Z}_s(l)$ is the estimate for the l^{th} grid point. In this study, the regional TEC map grid is formed for $N_\theta = 11$ and $N_\phi = 24$. The spacings in latitude and longitude are chosen to be 1° . Such a grid is large enough to represent Central Europe. The chosen spatial resolution is higher than those of IGS-GIM, which is 5° in longitude and 2.5° in latitude.

[12] The trend structure for the synthetic TEC surfaces are formed by the following six functions

$$\mu_1(\mathbf{x}) = 18 \quad (5)$$

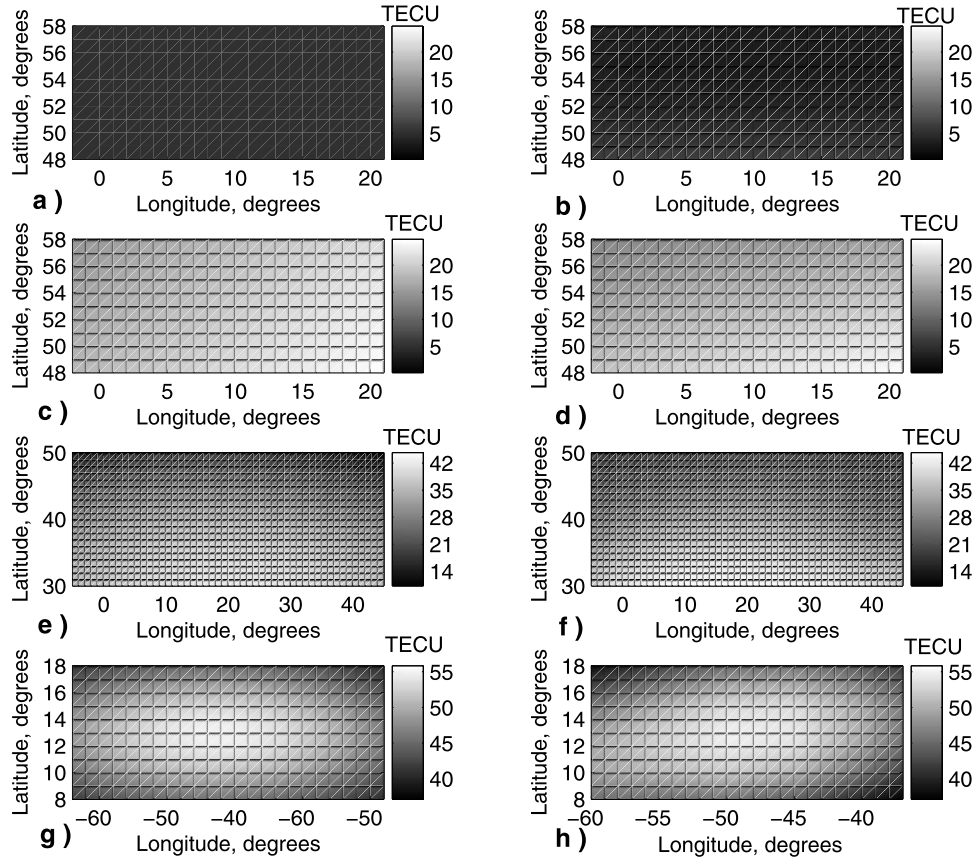


Figure 1. Trends used in the generation of synthetic TEC surfaces (a) constant, (c) linear, (e) second order, and (g) Gaussian are compared to their GIM equivalents in (b) 1 April 2003 0200 UT, (d) 1 April 2004 1000 UT, (f) 23 May 2004 1200 UT, and (h) 1 April 2004 0000 UT. Latitude is in degrees.

$$\mu_2(\mathbf{x}) = 33.18 - 0.3\theta + 0.3\phi \quad (6)$$

$$\mu_3(\mathbf{x}) = -44.61 + 2.96\theta + 0.66\phi - 0.033\theta^2 - 0.033\phi^2 \quad (7)$$

$$\mu_4(\mathbf{x}) = 38.06 - 0.43\theta + 8.66 \exp\left(-\left(\frac{\theta - 53}{15}\right)^2 - \left(\frac{\phi - 9.5}{10}\right)^2\right) \quad (8)$$

$$\mu_5(\mathbf{x}) = 1 + 5 \exp\left(-\left(\frac{\theta - 53}{7}\right)^2 - \left(\frac{\phi - 9.5}{10}\right)^2\right) \quad (9)$$

$$\mu_6(\mathbf{x}) = 21.09 + 6.01 \sqrt{\cos^2(\theta - 53) + \sin^2(\phi - 9.5)} - 6.01 \exp[0.25(\cos(\theta - 53) + \cos(\phi - 9.5))] \quad (10)$$

The coefficients of the trend functions given by equations (5) to (8) are chosen to represent the trends in the TEC distributions for a quiet day of the ionosphere for different time instants in a day. Trend functions given by equations (9) and (10) represent a severe disturbance or variation of TEC. An example of these synthetic trend functions and GIM maps is provided in Figure 1. Figures 1a, 1c, 1e, and 1g are the chosen trend functions to represent constant, linear, second order, and gaussian trends. Figures 1b, 1d, 1f, and 1h are extracted from GIM maps corresponding to those trends in the left column for different days and hours.

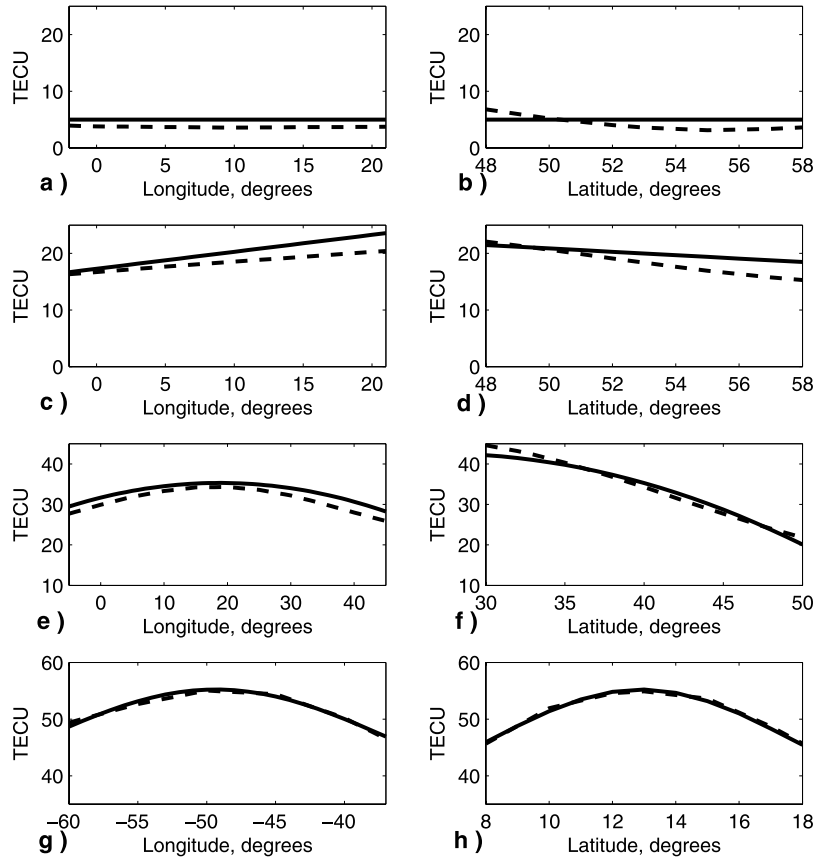


Figure 2. Comparison of synthetic TEC surfaces (solid lines) and GIMs (dashed lines) from Figure 1 for a fixed latitude in Figures 2a, 2c, 2e, and 2g and for a fixed longitude in Figures 2b, 2d, 2f, and 2h.

[13] For a better comparison of the agreement of the synthetic TEC surfaces with the trends in GIMs, TEC values for a fixed latitude and longitude are provided in Figure 2. Figures 2a, 2c, 2e, and 2g are TEC values for a constant latitude, corresponding to the synthetic TEC and GIM values in Figure 1. The solid line represents synthetic TEC and the dashed line represents GIM-TEC. In Figures 2b, 2d, 2f, and 2h are the TEC values corresponding to the synthetic TEC and GIM values in Figure 1, for a fixed longitude. Figures 2a and 2b correspond to the trend in Figures 1a and 1b. Figures 2c and 2d correspond to the trend in Figures 1c and 1d. Figures 2e and 2f correspond to the trend in Figures 1e and 1f. Figures 2g and 2h correspond to the trend in Figures 1g and 1h. As it can be observed from Figure 2, the chosen synthetic trend surfaces are in excellent agreement with those from GIM.

[14] The small scale variations on the trends are generated using Cholesky Decomposition method [Sayin, 2008; Cressie, 1993]. An exponential covariance

function is used for generating the small-scale variability of the TEC surface as

$$cov_Y(h) = \sigma^2 \exp\left(\frac{-|h|}{a}\right) \quad (11)$$

where $cov_Y(h)$ is the covariance of the TEC values at two points, separated with the distance lag h . The variance σ^2 determines the dispersion of the values and the range a determines the correlation between the values of different points. To represent different states of the ionosphere, the variance σ^2 is chosen from the set 0.64, 1.44, and 2.56 as the residual variance of GPS-TEC when IRI model is assumed as the trend. For the range a denoting the correlation distance of $Y(\mathbf{x})$, 5, 10 or 15 are chosen to represent lowest to highest correlation between the points. On a very quiet day, ionosphere can be modeled by using $\sigma^2 = 0.64$ and $a = 15$. A disturbed ionosphere on the other hand may have a large σ^2 , and small a . The assessment of the performance with various interpolation algorithms over the generated surfaces require the

mapping algorithms to be run over various realization of surfaces having the same trend $\mu(\mathbf{x})$, variance σ^2 and range a .

[15] For the sampling of synthetic TEC surfaces, regular and random patterns are used. Regular sampling structure use square, triangular and hexagonal distributed sample points. Random sampling patterns are chosen as uniformly distributed, inhibited and clustered as in the works of *Sayin* [2008], *Yfantis et al.* [1987], and *Zimmermann et al.* [1999]. The sparse GPS-TEC measurement points are often in the structure of inhibited or clustered. For example, there is a large cluster of GPS stations over land regions such as Central Europe and there are vast spaces over seas and oceans where there are no stations.

[16] The interpolation accuracy is expected to increase as the number of samples increase, yet as the number of samples (corresponding to measurements) increase, the computational complexity of the algorithms also increase. Thus, it is important to find a threshold after which a significant improvement in the interpolation performance do not occur with increasing number of samples and keeping computational complexity to a minimum. The number of sampling points used in this study out of 246 grid points are 20, 30, 40, 50, 60 and 70, corresponding to 7.6%, 11.4%, 15.2%, 19.0%, 22.7% and 26.5% of the total number of grid points, respectively.

2.2. Kriging and Random Field Prior Mapping Algorithms

[17] Kriging is a widely used interpolation technique that linearly estimates the process by minimizing the error variance with respect to an unbiasedness condition. Kriging estimate is the linear combination of the values at measurement points as

$$\hat{Z}_s(l) = \sum_{n_a=1}^{N_a} w_{l;n_a} Z(\mathbf{x}_{n_a}) \quad (12)$$

where $w_{l;n_a}$, for $1 \leq n_a \leq N_a$, are the Kriging coefficients for the l^{th} grid point. Kriging algorithms, generally, assume the trend $\mu(\mathbf{x})$ in the random function as a linear combination of known deterministic functions, $f_{n_k}(\mathbf{x})$ as

$$\mu(\mathbf{x}) = E\{Z(\mathbf{x})\} = \sum_{n_k=1}^{N_k} a_{n_k} f_{n_k}(\mathbf{x}), \quad f_1(\mathbf{x}) = 1 \quad (13)$$

where a_{n_k} , $1 \leq n_k \leq N_k$, is the unknown trend coefficient. For unbiased estimation, the coefficients w have to satisfy the unbiasedness condition as

$$\sum_{n_a=1}^{N_a} w_{l;n_a} f_{n_k}(\mathbf{x}_{n_a}) = f_{n_k}(\mathbf{x}_l), \quad n_k = 1, \dots, N_k \quad (14)$$

The Kriging interpolation algorithm can be applied on various assumptions of the trend $\mu(\mathbf{x})$. When the trend is assumed to be constant, corresponding to $N_k = 1$, the interpolation is called Ordinary Kriging (OK). When $N_k > 1$ the method is called Universal Kriging. In this study, the Universal Kriging algorithm which assumes a linear trend, corresponding to $N_k = 3$ (UK1) and the Universal Kriging which assumes a second order trend corresponding to $N_k = 6$ (UK2) are chosen to be the mapping algorithms along with OK.

[18] Random Field Priors (RFP) interpolation model allows multivariate and/or multidimensional estimation, given that the first and second order statistics are known. For TEC mapping, the first order statistic corresponds to the mean or trend function and the second order statistic corresponds to the covariance function of the multivariate random field [*Sayin*, 2008]. In our implementation of RFP for spatial interpolation of TEC at a given time instant, the mean TEC surface, the covariance and cross-covariance matrices of the \mathbf{g}_s and \mathbf{d} vectors are assumed to be known. For interpolation of real data, these covariance and cross-covariance matrices can be obtained from an empirical ionospheric model such as IRI or from past data on the ionospheric parameters such as GIM. When a constant or a low order variation in the trend is chosen, the performance of RFP in this study will be similar to Simple Kriging. When a highly varying trend is tried, this similarity will cease to exist yet for all cases RFP will form a minimum error bound for OK, UK1 and UK2 [*Sayin*, 2008; *Olea*, 1999].

[19] The exponential covariance function given in equation (11) is used in both KG and RFP algorithms. The semivariogram function that is necessary for reconstruction using Kriging and in evaluating error variances is computed as

$$\gamma(h) = cov(0) - cov(h) \quad (15)$$

as given by *Olea* [1999] for isotropic random functions. $cov(\cdot)$ is the covariance function in equation (11). h is the distance between two sample points in space. In the next section, we will explain the results of interpolation scenarios.

3. Results and Discussion

[20] In this section, spatial interpolation performances of RFP, OK, UK1 and UK2 methods are compared first on synthetic TEC surfaces over the important performance parameters and then, these interpolation algorithms are applied to GPS-TEC data.

3.1. TEC Mapping Over Synthetic Surfaces

[21] The synthetic TEC surfaces are generated for six trend functions in equations (5) to (10). Small scale

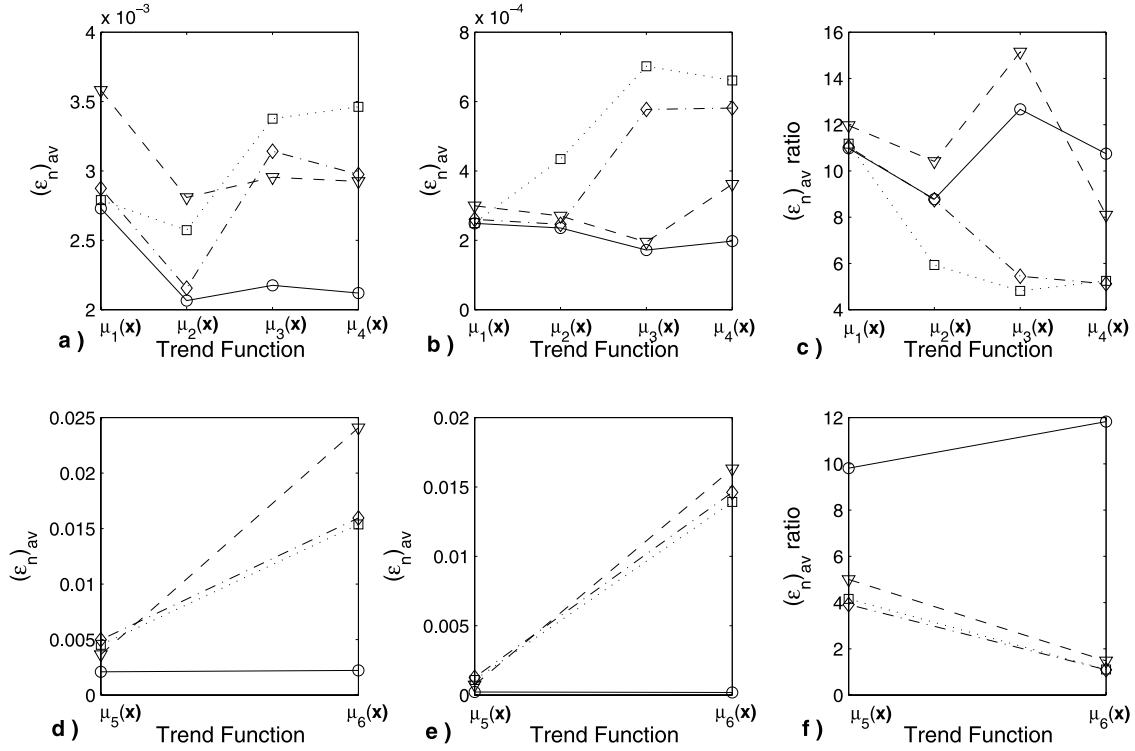


Figure 3. Average normalized error versus trend function: (a) and (d) high level of scale disturbance; (b) and (e) low level of scale disturbance. (c) The ANE ratio of Figures 3a to 3b; (f) the ANE ratio of Figures 3d to 3e. Circles, squares, diamonds and triangles represent RFP, OK, UK1, and UK2, respectively.

variations are generated and added on the trend functions for all nine combinations of σ^2 and a . Six sampling patterns, and six different number of samples are used in generating scenarios. For each surface scenario, four interpolation algorithms are applied. The total of $7,776 = [6 \text{ (trend functions)} \times 3 \text{ } (\sigma^2 \text{ values)} \times 3 \text{ } (a \text{ values)} \times 6 \text{ (sampling patterns)} \times 6 \text{ (number of samples)} \times 4 \text{ (interpolation algorithms)}]$ interpolation scenarios are generated. The generation of $Y(\mathbf{x})$ using one set of σ^2 and a corresponds to one realization of the small scale variations over the trend. For statistically meaningful reconstructions, for each set of σ^2 and a , $N_r = 10$ realizations are generated. The reconstruction errors are computed using the average normalized error (ANE) criterion. First, the normalized error for each realization is computed as

$$\epsilon_n = \frac{\|\hat{\mathbf{z}}_s - \mathbf{g}_s\|^2}{\|\mathbf{g}_s\|^2} \quad (16)$$

where $\|\cdot\|$ is the \mathcal{L}_2 norm defined by $\|\mathbf{g}_s\|^2 = \sum_{i=1}^N [g_s(i)]^2$. The ANE, $(\epsilon_n)_{av}$, of one scenario is the average of the

normalized errors over N_r realizations of that scenario, as

$$(\epsilon_n)_{av} = \frac{1}{N_r} \sum_{i=1}^{N_r} \epsilon_n(i) \quad (17)$$

where $\epsilon_n(i)$ is the normalized error obtained from the i^{th} realization.

[22] The ANE from each interpolation scenario is investigated and analyzed by keeping one parameter constant at a time and observing the variation of ANE for other parameters. This way, both the dominant parameters and performance improvement for reconstruction algorithms are compared. The following discussion is a summary of the performance analysis. Some examples to support the conclusions are also provided.

[23] By examining the ANE for interpolation scenarios, it is observed that the most dominant factor in accuracy of the reconstructions is the spatial distribution of ionosphere in the form of trend and small scale variations. This is a very important conclusion in judging the accuracy of the reconstructed maps using KG and

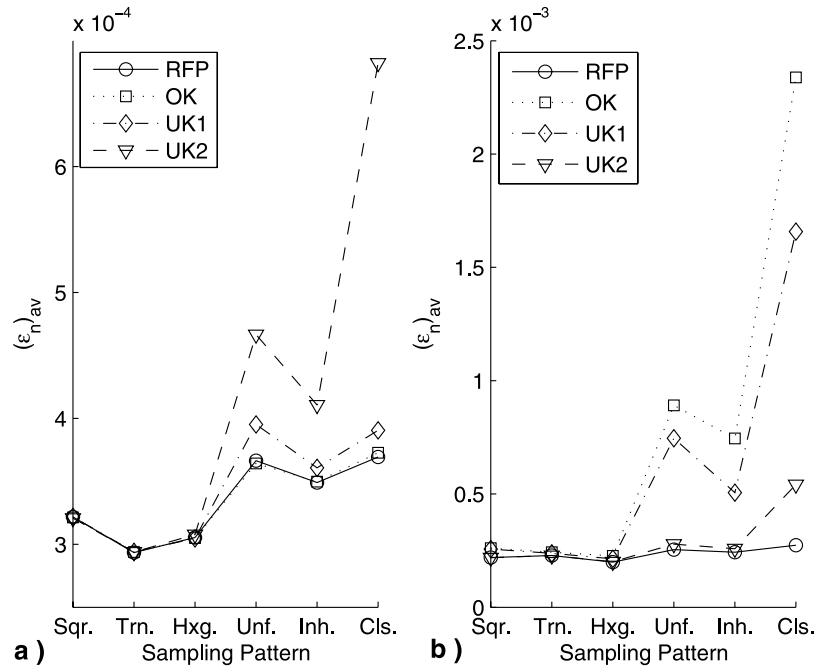


Figure 4. Average normalized error versus sampling pattern and average small scale disturbance on the trend (a) constant trend $\mu_1(\mathbf{x})$; (b) second order trend $\mu_3(\mathbf{x})$.

RFP algorithms over real data for quiet and disturbed days such as in the works of *Wielgosz et al.* [2003], *Blanch* [2002], *Orús et al.* [2005], and *Stanislawski et al.* [2002]. In order to demonstrate the effect of the ionospheric surface in reconstructions, an example is provided in Figure 3 for uniformly distributed sampling pattern with 11.4% sampling rate over all the trend functions and reconstruction algorithms. In Figures 3a and 3b, the ANE for trends $\mu_1(\mathbf{x})$, $\mu_2(\mathbf{x})$, $\mu_3(\mathbf{x})$, and $\mu_4(\mathbf{x})$ are given. In Figures 3d and 3e, the ANE plots for the trends $\mu_5(\mathbf{x})$ and $\mu_6(\mathbf{x})$ are given. In Figures 3a and 3d, the parameters of small scale disturbance are chosen as $\sigma^2 = 2.56$ and $a = 5$ corresponding to higher level of small scale variability. In Figures 3b and 3e, the parameters of small scale disturbance are chosen as $\sigma^2 = 0.64$ and $a = 15$ corresponding to lower level of small scale variability. Figure 3c shows the ratio of the ANEs in Figure 3a to the ANEs in Figure 3b, for each reconstruction algorithm and each trend. The ratio of the ANEs in Figure 3d to the ANEs in Figure 3e, for each reconstruction algorithm and each trend is provided in Figure 3f. When Figures 3a and 3d are compared, it is observed that the ANEs in Figure 3d are higher than those in Figure 3a. Similarly, when Figures 3b and 3e are compared, it is observed that the ANEs in Figure 3e are higher than those in Figure 3b. When the ANEs in Figures 3a and 3b are compared with their ratios in Figure 3c, it is observed

that the effect of higher level of small scale variability is significant for trend functions $\mu_1(\mathbf{x})$ to $\mu_4(\mathbf{x})$ for RFP and UK2. The ratio of ANE for trends $\mu_5(\mathbf{x})$ and $\mu_6(\mathbf{x})$ in Figure 3f varies depending on the type of the reconstruction algorithm.

[24] From the analysis of ANE for the large number of interpolation scenarios, we have observed that the second most dominant parameter in the performance of reconstruction algorithms is the sampling pattern of the surface. For regular sampling patterns tried in this study, regardless of the number of samples, the trend or level of small scale disturbance, the ANE for all interpolation algorithms are very close to each other. The ANE increases significantly for random sampling patterns. Although ANE for uniformly distributed sampling pattern with a high number of samples is closer in value when compared to those from regular sampling patterns, clustered sampling pattern generally causes the highest ANE over all trends, levels of small scale disturbance and interpolation algorithms. Examples on the effect of sampling pattern on the performance criterion are provided in Figures 4, 5, and 6. In Figure 4, the ANE versus all sampling patterns are provided for all KG and RFP algorithms. In Figure 4a, the trend is chosen as constant. In Figure 4b, the trend has second-order variation. For both cases, the level of small scale disturbance is average with $\sigma^2 = 0.64$ and $a = 10$. The sampling rate for both

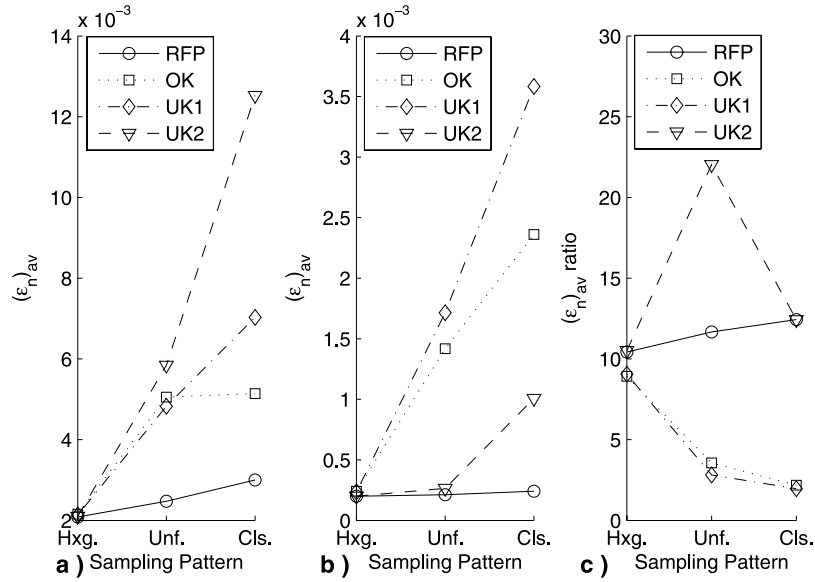


Figure 5. Average normalized error versus sampling pattern types for the second order trend, $\mu_3(\mathbf{x})$, with (a) highest small scale disturbance and (b) lowest small scale disturbance, for a sampling rate of 7.6%. (c) The ratio of Figures 5a to 5b.

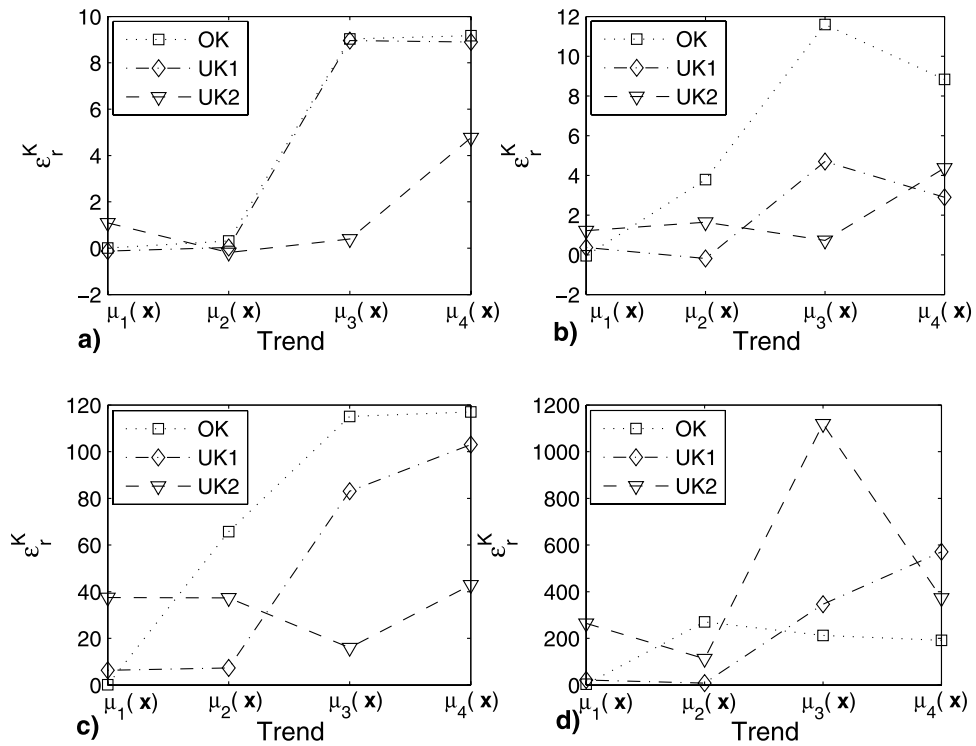


Figure 6. Relative error versus the trend with a high level of scale variability for (a) square, (b) hexagonal, (c) uniformly distributed, and (d) clustered sampling patterns.

figures is 11.4%. The highest error in both Figures 4a and 4b is observed for the clustered sampling pattern. The minimum error bound is again obtained for RFP. The second lowest error is due to the reconstruction algorithm that assumes the correct trend. It is OK for Figure 4a and UK2 for Figure 4b.

[25] In Figures 5a and 5b, the ANE variation against the sampling patterns is presented for the second order trend with the highest ($\sigma^2 = 2.56$, $a = 5$) and the lowest ($\sigma^2 = 0.64$, $a = 15$) small scale disturbance, respectively, for a sampling rate of 7.6%. Figure 5c shows the ratio of the ANEs in Figure 5a to the ANEs in Figure 5b, for each reconstruction algorithm and each trend. It is observed from Figure 5c that the ANE for the surfaces with lowest and highest variability are similar to each other for the regular sampling pattern, for all reconstruction algorithms. The performance for the random samplings can differ in ANE depending on the type of the reconstruction algorithm.

[26] The variation of relative ANE with respect to trends and sampling patterns are presented in Figure 6 with 11.4% sampling rate. Owing to the fact that the RFP method provides a minimum error bound on the reconstruction errors, the relative ANE of Kriging algorithms with respect to RFP are calculated as:

$$\epsilon_r^K = 100 \times \frac{(\epsilon_n)_{av}^K - (\epsilon_n)_{av}^{RFP}}{(\epsilon_n)_{av}^{RFP}} \quad (18)$$

In equation (18), $(\epsilon_n)_{av}^K$ is the average normalized error. K represents one of OK, UK1 or UK2. Figures 6a, 6b, 6c, and 6d represent the variation of $(\epsilon_n)_{av}^K$ with respect to trend with a high level of small scale variability for square, hexagonal, uniformly distributed and clustered sampling patterns, respectively. In this figure, we observe the importance of the trend and sampling pattern in the performance of the interpolation algorithms. As discussed before, the relative errors increase significantly for the surfaces with a high variability trend and for random sampling pattern. The reconstruction algorithm which assumes the correct trend function has the lowest relative error. The regular sampling patterns provide similar relative errors. The errors increase approximately for 10 and 100 times for uniformly distributed and clustered sampling patterns, respectively.

[27] The third most important factor in the performance evaluation of the reconstruction algorithms is the number of samples in sampling pattern. From our analysis, it is observed that for any trend and small scale variability, and sampling pattern, ANE decreases with increasing number of samples for all interpolation methods. An example is provided in Figure 7 that shows the variation of ANE versus number of samples for constant, linear, and second order trends, for uniformly distributed sampling pattern. The interpolation algorithms that

assume the correct trend function converge faster to the minimum error bound limit as number of samples increase. Although the convergence slows down for higher variability trend surfaces as in Figure 7c, 11.4% sampling rate corresponding to 30 samples for the generated grid forms a break point in the performance improvement for all interpolation algorithms, surface trends and small scale variability.

[28] The performance analysis of KG and RFP algorithms provides a reliable basis for generating accurate GPS-TEC maps for tomography, vertical profile generation and ionospheric earthquake precursor investigation. The specific and general conclusions from this study are applied to a pilot project in Turkey for the distribution and number of GPS stations and accuracy of the instantaneous TEC maps for quiet and disturbed ionosphere. Currently, the GPS stations are clustered around Marmara Sea region and western Anatolia. The optimum number and distribution of additional GPS stations can be determined by considering the performance improvement in synthetic TEC reconstructions.

3.2. Generation of TEC Maps With GPS-TEC Data

[29] The KG and RFP algorithms are applied to generate TEC maps using GPS-TEC over a region containing 39 stations encircled by Yerevan, Armenia (40.23°N, 44.50°E), Robledo, Spain (40.43°N, 4.25°W), Tromsø, Norway (69.66°N, 18.94°E) and Nicosia, South Cyprus (35.14°N, 33.39°E). The sampling rate corresponds to approximately 2.5%. When the distribution of GPS stations are considered, it seems to be made up of several clusters and there are regions where no sampling can be obtained. The vertical TEC values are obtained by IONOLAB-TEC from www.ionolab.org using Reg-Est algorithm discussed by *Arikan et al.* [2003, 2004, 2007] and *Nayir et al.* [2007]. The experimental semivariogram is obtained by using

$$\gamma^*(h) = \frac{1}{2N(h)} \sum_{\substack{i,j \text{ pair} \\ i \neq j}}^{N(h)} [Z(\mathbf{x}_i) - Z(\mathbf{x}_j)]^2 \quad (19)$$

where h shows the distance lag between GPS stations located at \mathbf{x}_i and \mathbf{x}_j . $Z(\mathbf{x}_i)$ and $Z(\mathbf{x}_j)$ are the TEC values at points \mathbf{x}_i and \mathbf{x}_j , respectively. $N(h)$ is the number of point pairs with the same distance lag h . The experimental semivariogram are fitted to the theoretical semivariogram function $\gamma(h)$ defined in equation (15) and (11). An example of generated TEC maps is presented in Figure 8 for 16 October 2004 at 0600 UT. This day is listed as the fifth negatively disturbed day of October 2004 in the web site of the Ionospheric Dispatch Center in Europe (IDCE) (<http://www.cbk.waw.pl/rwc/idce.html>). Figures 8a, 8c, 8e, and 8g are the TEC maps generated by RFP, OK,

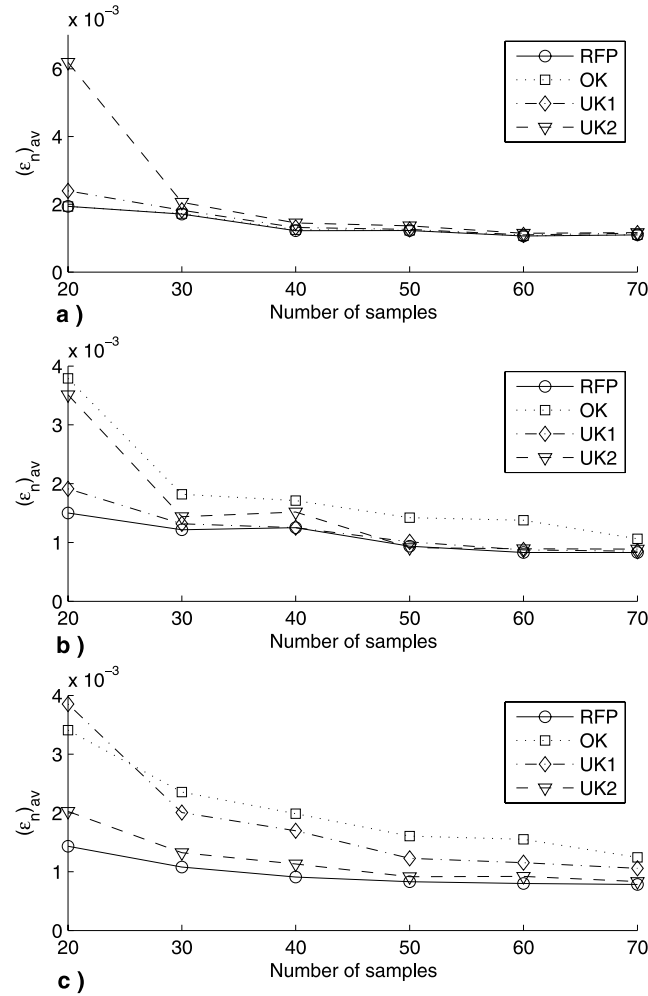


Figure 7. Average normalized error versus number of samples for surfaces with (a) constant $\mu_1(\mathbf{x})$, (b) linear $\mu_2(\mathbf{x})$, and (c) second order $\mu_3(\mathbf{x})$ trends.

UK1, and UK2 interpolation algorithms, respectively. Figures 8b, 8d, 8f, and 8h are the error variances of the TEC maps generated by RFP, OK, UK1, and UK2 interpolation algorithms, respectively. The computation of error variances are provided by *Sayin* [2008], *Blanch* [2002], *Cressie* [1993], and *Olea* [1999]. At 0600 UT, the Sun rises from the east causing an increased ionization in the ionosphere. This gradient of TEC in the ionosphere from north to south and east to west can easily be observed from Figure 8. The lowest error variance is obtained by RFP, and the average error variance (AEV) is 0.3398 TECU^2 . As the constraint on the interpolation algorithms increases, the error variances also increase. For OK, AEV is 0.3414 TECU^2 ; for UK1, AEV takes the value of 0.3452 TECU^2 ; and finally for

UK2, AEV is equal to 0.3584 TECU^2 . The minimum error variance points in Figures 8b, 8d, 8f, and 8h correspond to the locations of GPS stations.

4. Conclusion

[30] In this study, the performance of Random Field Priors, Ordinary Kriging and Universal Kriging assuming linear and second order trends are compared over synthetically generated surfaces with respect to their performance parameters of surface trend, small scale surface variability, sampling pattern and number of samples. The performance criterion is chosen as average normalized error (ANE). The performance improvement or degradation between the interpolation scenarios including the interpolation algorithms are

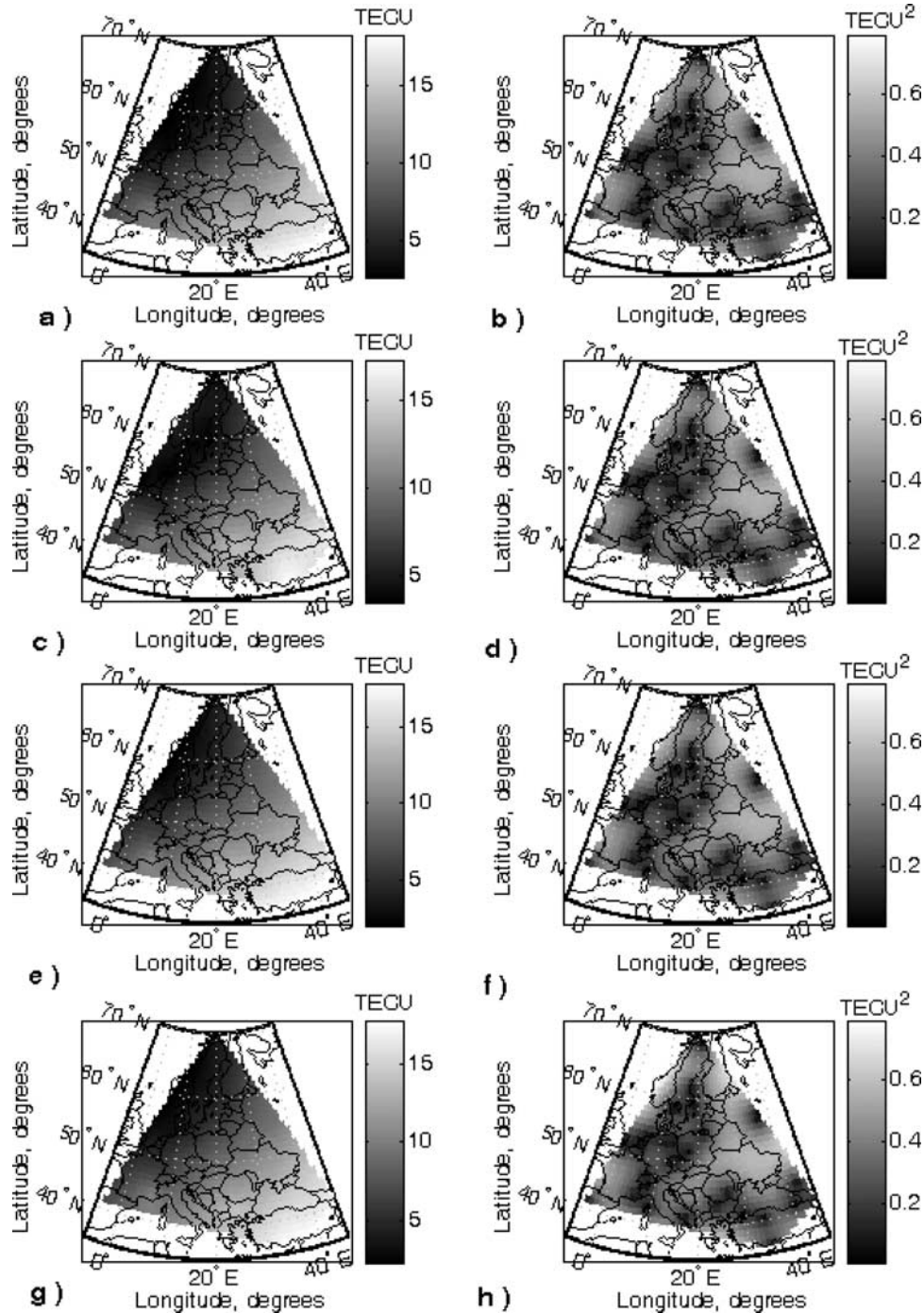


Figure 8. The GPS-TEC maps by (a) RFP, (c) OK, (e) UK1, (g) UK2 and corresponding error variances (EV) (b) EV-RFP, (d) EV-OK, (f) EV-UK1, (h) EV-UK2 for 16 October 2004 at 0600 UT.

observed by analyzing the ANE for a large number of scenarios. It is observed that, the most important interpolation performance parameter is the trend of the

surface and the degree of small scale disturbance on it. For all scenarios, RFP assumes that trend of the interpolated TEC surface is known, and thus RFP provides a

minimum error bound. The interpolation algorithm that assumes the correct trend, performs similar to RFP. For a surface with a high variability of small scale disturbance or the assumption of wrong trend for interpolation algorithms, the errors can grow by 10 and 100 fold, respectively. The second most important parameter that effect the performance of interpolation algorithms is the sampling pattern. The error increases by three to five fold for the random sampling when compared to regular sampling patterns. The worst performance is obtained when the sampling points are clustered in an area. The third important parameter is the number of samples and for an approximately 10% sampling rate corresponds to a performance improvement of two to four fold even for highly varying surface trends. Although the performance of all KG and RFP algorithms can be compared over specific scenarios, it is observed that in generation of instantaneous TEC maps with GPS-TEC data, the average error variance increases for the interpolation algorithms that have to satisfy more constraints. In the application of the KG and RFP algorithms to GPS-TEC data, the lowest error variance is observed for RFP. Ordinary Kriging assuming lowest constraint for reconstruction performed similar to RFP. Therefore, we conclude that irrespective of the trend, small scale disturbance, sampling pattern and number of samples, OK is an optimum choice for instantaneous TEC maps where a priori information over the region is not available or the exact trend in the ionosphere is not known. In future studies, the spatio-temporal interpolation and prediction methods for the ionospheric parameters will be investigated.

[31] **Acknowledgment.** This study is supported by TUBITAK EEEAG grant 105E171.

References

- Arikan, F., C. B. Erol, and O. Arikan (2003), Regularized estimation of vertical total electron content from Global Positioning System data, *J. Geophys. Res.*, *108*(A12), 1469, doi:10.1029/2002JA009605.
- Arikan, F., C. B. Erol, and O. Arikan (2004), Regularized estimation of vertical total electron content from GPS data for a desired time period, *Radio Sci.*, *39*, RS6012, doi:10.1029/2004RS003061.
- Arikan, F., O. Arikan, and C. B. Erol (2007), Regularized estimation of TEC from GPS data for certain mid-latitude stations and comparison with the IRI model, *Adv. Space Res.*, *39*, 867–874, doi:10.1016/j.asr.2007.01.082.
- Blanch, J. (2002), An ionosphere estimation algorithm for WAAS based on Kriging, paper presented at ION GPS 2002, Portland, Oreg., 24–27 Sept.
- Budden, K. G. (1985), *The Propagation of Radio Waves: The Theory of Radio Waves of Low Power in the Ionosphere and Magnetosphere*, Cambridge Univ. Press, New York.
- Chiles, J. P., and P. Delfiner (1999), *Geostatistics: Modelling Spatial Uncertainty*, John Wiley, New York.
- Cressie, N. A. C. (1993), *Statistics for Spatial Data*, John Wiley, New York.
- Gurun, M. (2007), Static and dynamic regional ionosphere mapping with TEC data by using neural network based methods (in Turkish), M.S. thesis, Hacettepe Univ., Ankara, Turkey.
- Hanbaba, R. (Ed.) (1999), The final report of COST 251, Space Res. Cent., Warsaw, Poland.
- Jin, S., J. Wang, H. Zhang, and W. Zhu (2004), Real-time monitoring and prediction of ionospheric electron content by means of GPS, *Chin. Astron. Astrophys.*, *28*, 331–337.
- Komjathy, A. (1997), Global ionospheric Total Electron Content mapping using the Global Positioning System, Ph.D. thesis, Dep. of Geod. and Geomat. Eng., Univ. of New Brunswick, Fredericton, New Brunswick, Canada.
- Lazo, B., A. Calzadilla, K. Alazo, M. Rodriguez, and J. S. Gonzalez (2004), Regional mapping of F2 peak plasma frequency by spherical harmonic expansion, *Adv. Space Res.*, *33*, 880–883.
- Leandro, R. F. (2004), A new technique to TEC regional modelling using a neural network, paper presented at ION GNSS 2004, Long Beach, Calif., 20–21 Sept.
- Moon, Y. (2004), Evaluation of two-dimensional ionosphere models for national and regional GPS networks in Canada, M.S. thesis, Univ. of Calgary, Calgary, Alberta, Canada.
- Nayir, H., F. Arikan, O. Arikan, and C. B. Erol (2007), Total Electron Content estimation with Reg-Est, *J. Geophys. Res.*, *112*, A11313, doi:10.1029/2007JA012459.
- Olea, R. A. (1999), *Geostatistics For Engineers and Earth Scientists*, Kluwer Acad., Dordrecht, Netherlands.
- Orús, R., M. Hernandez-Pajares, J. M. Juan, and J. Sanz (2005), Improvement of global ionospheric VTEC maps by using Kriging interpolation technique, *J. Atmos. Sol. Terr. Phys.*, *67*, 1598–1609.
- Sayin, I. (2008), Total Electron Content mapping using Kriging and Random Field Priors, M.Sc. thesis, Hacettepe Univ., Ankara, Turkey.
- Sayin, I., F. Arikan, and O. Arikan (2007), Synthetic TEC mapping with ordinary and universal Kriging, paper presented at RAST-2007, Recent Advances in Space Research, Istanbul, Turkey, 14–16 June.
- Schaer, S. (1999), Mapping and predicting the Earth's ionosphere using the Global Positioning System, Ph.D. thesis, Univ. of Bern, Bern, Switzerland.
- Stanislawska, I., G. Juchnikowski, L. R. Cander, L. Ciraolo, P. A. Bradley, Z. Zbyszynski, and A. Swiatek (2002), The Kriging method of TEC instantaneous mapping, *Adv. Space Res.*, *29*, 945–948.
- Vanmarcke, E. (1983), *Random Fields: Analysis and Synthesis*, MIT Press, London.
- Wielgosz, P., D. Brzezinska, and I. Kashani (2003), Regional ionosphere mapping with Kriging and Multiquadric method, *J. Global Posit. Syst.*, *2*, 48–55.

Yfantis, E. A., G. T. Flatman, and J. V. Behar (1987), Efficiency of Kriging estimation for square, triangular, and hexagonal grids, *Math. Geol.*, 19, 183–205.

Zimmermann, D., C. Pavlik, A. Ruggles, and M. P. Armstrong (1999), An experimental comparison of ordinary and uni-

versal Kriging and inverse distance weighting, *Math. Geol.*, 31, 375–390.

F. Arikan and I. Sayin, Department of Electrical and Electronics Engineering, Hacettepe University, Beytepe, Ankara 06800, Turkey. (arikan@hacettepe.edu.tr; isiltan@ee.hacettepe.edu.tr)

O. Arikan, Department of Electrical and Electronics Engineering, Bilkent University, Bilkent, Ankara 06800, Turkey. (oarikan@ee.bilkent.edu.tr)