Ionosphere-Corrected Range Estimation in Dual Frequency GNSS Receivers

Danai Skournetou and Elena-Simona Lohan

Department of Communications Engineering

Tampere University of Technology

Tampere, Finland, 33720.

E-mail: see http://www.cs.tut.fi/tlt/pos/

Abstract

In Global Navigation Satellite Systems (GNSSs), the measurement of the satellite-receiver pseudorange requires the estimation of signal's total delay. Because the accuracy of the latter affects significantly the accuracy of the final position, it is essential to consider the effect of various error sources. Ionosphere is commonly regarded as one of the most influential sources due to the fact that it can significantly delay the signal; therefore, it is of paramount importance to mitigate its effects.

In single-frequency GNSS receivers, the ionospheric delay is typically found with the use of mathematical models. Because their accuracy is usu-

ally determined by their complexity, mass-market receivers employ relatively computationally simple models at the expense of limited accuracy. On the other hand, in dual-frequency receivers, we can virtually eliminate the ionospheric effects if higher order effects are ignored [1, 2]. While such an advantage has been widely recognised in the literature, the effect of the tracking error in the ionospheric correction and inherently on the range estimation is yet to be studied.

In this paper, we investigate the effect of tracking error on the ionosphere-corrected range in dual-frequency receivers. We statistically analyse the performance of Least Square (LS) method and we compare it with the simulation-based. Moreover, we compare the performance of the traditional approach with LS and Constrained LS methods, as well as with a new method for range estimation, proposed by the authors. The results showed that the traditional and LS methods perform well only under the restriction of zero tracking error, while our method has the best average performance.

1 Introduction

In Global Navigation Satellite Systems (GNSSs), the accuracy of the computed user position is a quality measure of paramount importance. In order to locate a user, the distance between its GNSS receiver and at least four satellites (also known as geometric range) needs to be estimated. In an inherent manner, the accuracy of the user's position depends on the accuracy of each range measurement. Ideally (i.e., in absence of errors), the unknown geometric range can be calcu-

lated with the knowledge of the signal's propagation time. For this purpose the receiver has to acquire the signal's transmission and reception times; the former is included in the transmitted navigation message and the latter (reception time) is found with the help of the receiver's internal clock. However, in realistic scenarios the presence of errors (both channel- and receiver-specific) "distorts" the propagation time, the measurement of which results in what is commonly denoted as pseudorange (as opposed to the true geometric range).

Among the various error sources, ionosphere accounts for the biggest part of signal's total delay [1], therefore, it is essential to alleviate its effects. Typically, it is considered to start at 50 km from the earth surface and to end at 1000 km. Unlike the lower layers of atmosphere (e.g. troposphere, stratosphere, etc.), ionosphere contains charged particles (electrons and ions), the content of which depends on various spatial and temporal parameters (e.g. altitude, season, time of the day, etc.), as well as on the occurrence of natural phenomena (e.g. electromagnetic storms and travelling ionospheric disturbances). The presence of the charged particles makes ionosphere a dispersive medium, thus, signals transmitted at different carrier frequencies have different phase advances and time delays.

When the signal propagates through ionosphere, its velocity changes due to the interaction with particles present in it. As a result, the signal's code is delayed and its phase is advanced. In particular, the signal is delayed almost by as much as the carrier phase is advanced, thus, it is sufficient to estimate one of the two parameters (if higher order and bending effects are ignored, then the values of code delay and carrier phase advance are exactly the same [3]). The interested

reader may refer to [1–4] for a detailed description of the ionospheric effects.

In order to mitigate the refraction effects, the knowledge of the involved refractive indices and signal's frequency is required. However, because ionosphere is an heterogeneous medium, meaning that the density of the ionised particles within it is not uniform (from now on we will consider only electrons since ions are much heavier [2]), its refractive index is defined by the electron density. Appleton and Lassen have derived a complex formula for computing the ionospheric refractive index [5], with which the ionospheric delay can be defined as the sum of first, second, third order and bending effects [3]. These effects are a function of the Total Electron Content (TEC), which is a space-time varying parameter to be estimated. It can be shown that for E1 signal, the second and third order effects contribute to the total ionospheric delay by a sub-metre and centimetre level, respectively (we remark, that the contribution of these two effects is similar for the rest of GNSS signals). Therefore, when mass-market receivers are considered, it suffices to consider only the first-order effect which accounts for almost 99% of the total delay [3]. For this reason, we ignore higher order terms in our model and whenever ionospheric effects are mentioned, they shall be associated only with first-order ones.

Most of the methods for ionospheric delay estimation have been proposed for single-frequency receivers since this is the dominant design when mass-market production is regarded. However, the performance of a single-frequency method can be useful also for receivers which operate at more than one frequency. For example, if signals from other frequencies are lost and the time needed to re-

acquire/re-track a lost signal is more than what can be afforded, the single-frequency method could be employed as backup option [6–9].

As mentioned earlier, the ionospheric delay depends on two parameters: the total electron density and the carrier frequency. While the latter is a known constant, the former needs to be estimated in order to further estimate the ionospheric delay. In single-frequency receivers, TEC is found with the help of an appropriately chosen model which shows the ionosphere status (i.e. TEC levels) for different locations and at different time periods. Moreover, such models are also responsible for making the necessary corrections for the ionospheric delay to a good degree of accuracy [10].

Unlike single-frequency receivers, no modelling of the ionosphere is needed when more than one carrier frequencies are available. For example, a dual-frequency receiver measures the pseudorange for each of the two received signals, both of which are contaminated by the same ionospheric effect. In theory (i.e., error-free scenario), proper combination of the available measurements allows the receiver to completely remove the ionospheric delay caused by first order effects [11] and this is one of the main advantages of dual-frequency receivers over single-frequency ones.

While the deployment of two frequencies is enough to compensate first-order ionoshperic effects, the choice of the two frequencies is of paramount important when considering the accuracy of the estimated range. Navstar Global Positioning System (GPS) -based dual frequency receivers utilise the L1 and L2 frequencies since these are the two signals currently transmitted form GPS satellites; how-

ever, with the advent of the new modernised signals the designers will have the flexibility to choose a better combination. Considering the future Galileo system, the research on dual-frequency receivers is in its infancy. In [12], a Galileo receiver which uses E1-E5a combination is tested but there is no explanation on the choice of the frequencies. An interesting study on the optimal frequency pair for future Galileo mass market receivers can be found in [13]. According to the study, the E1-E5 combination is much more difficult to implement than E1-E5a and E1-E5b combinations. The E1-E5a and E1-E5b pairs are shown to have similar performance in terms of acquisition, tracking and interference level. However, the use of the former is advised because it has the advantage of overlapping with the existing L1 and L5 bands, thus facilitating the design of a joint GPS/Galileo receiver [13].

As mentioned earlier, theoretically it can be shown that a dual-frequency receiver can successfully mitigate the ionospheric effects and thus, determine the user's position more accurately, in reality, this is valid only under certain constraints. More precisely, the authors' research led to the observation that a dual-frequency receiver is able to remove the ionospheric delay only when the pseudorange measurement errors are minimal, i.e., below a certain threshold. Such observation acted as the motivation for this manuscript, the targets of which can be summarised as follows: first, we show what is the impact of the measurement error on the accuracy of the range estimation in the presence of ionospheric errors. Secondly, we implement two alternative methods for estimating the receiver-satellite range, namely linear Least Square (LS) and Constrained linear Least

Square (CLS). Thirdly, we introduce a new method for improving the range estimation in dual-frequency receivers, called Brute Force Constraint (BFC) and compare the performance of these three methods with the traditional one. Finally, we present simulation results which suggest the best frequency combination among dual-frequency Galileo receivers.

The remainder of this manuscript is organised as follows: Section 2.1 describes the traditional approach for estimating the pseudorange, provides a statistical analysis of the LS solution and includes a performance comparison between the traditional and LS methods. Section 3 presents two new enhanced methods for pseudorange estimation. Section 4 includes the results of our research and their interpretation. Finally, Section 6 concludes the most important findings of this work.

2 Theoretical Background

In code-based GNSS receivers, we can model the measured pseudorange in units of length as [8, 14]

$$\rho_i = \rho + E + I_i + \varepsilon_i \tag{1}$$

where ρ is the true satellite-receiver range, E encompasses all the error sources which are common to all received signals (e.g. clock bias, tropospheric delay) and I_i is the ionospheric delay corresponding to the signal transmitted in f_i carrier

frequency. More precisely, the first-order ionospheric delay is modelled as [3]

$$I_i = \frac{40.3}{f_i^2} TEC \tag{2}$$

where TEC is the total electron content measured in TEC Units (TECUs) with 1 TECU=10^{16} electrons/m^2. The measurement error, ε_i, is a residue of the processing done in the code tracking stage and is equal to $\rho_i - \rho$ or equivalently to $c(\hat{\tau}_i - \tau)$, where c is the speed of light, $\hat{\tau}_i$ and τ are the estimated and the true code delay, respectively, both given in units of time. We notice that the code tracking error is different for different signals because it depends on signal-specific characteristics such as type (i.e. data or pilot), modulation, frequency, etc. and it represents mostly the effects of noise and multipath propagation.

2.1 Traditional Approach

In order to reduce the complexity of the range estimation problem, we ignore the common error sources (such as errors due to nonsynchronism of the satellite and receiver clocks, ephemerides and the tropospheric refraction) since their effects contaminate both frequencies in the same way and do not affect the validity of the results [15, 16]. In what follows, we assume that the range measurements are contaminated only by errors caused due to noise, ionospheric and multipath effects.

Starting from Eq. (1), we can form the following system of linear equations

for a dual-frequency receiver (i.e. $i = 1, 2$)

$$\begin{cases} \rho_1 = \rho + \dfrac{40.3}{f_1^2} TEC + \varepsilon_1 \\ \rho_2 = \rho + \dfrac{40.3}{f_2^2} TEC + \varepsilon_2 \end{cases} \tag{3}$$

In the traditional approach, we linearly combine the available pseudorange measurements in order to estimate the two unknown parameters. Here, the parameter of interest is the satellite-receiver range which can be computed as

$$\hat{\rho} = \frac{f_1^2}{f_1^2 - f_2^2}\rho_1 - \frac{f_2^2}{f_1^2 - f_2^2}\rho_2 \tag{4}$$

In the ideal case, i.e. when $\varepsilon_i = 0$ for $i = 1, 2$, we can completely remove the ionosheric effect and thus, calculate the true range successfully. Although the linear combining provides us with ionosphere-free measurements, it also has the side-effect of increasing the noise [1]. Assuming that the pseudorange measurements are uncorrelated ([2]), we can calculate the noise variance of the range measurement which has undergone Ionospheric Correction (IC) as

$$\sigma^2_{\rho^{IC}} = \frac{f_1^2}{f_1^2 - f_2^2}\sigma^2_{\rho_1} + \frac{f_2^2}{f_1^2 - f_2^2}\sigma^2_{\rho_2} \tag{5}$$

For example, if we assume that $\sigma^2_{\rho_1} = \sigma^2_{\rho_2} = 1$, the noise variance of the ionosphere-free range is 6.7 times higher for E1-E5a frequency combination. Using Eq. (5), it can be shown that in Galileo system the bigger the frequency separation ($|f_1 - f_2|$) is, the smaller the noise variance of the ionosphere-free range is. Considering the

future Galileo signals (E1, E5, E5a and E5b), it can be also shown, that the E1-E5a frequency combination is the best in terms of noise variance (see Table 1 for the numerical values of the frequency separation for all combinations). We notice that because the E5-E5a frequency combination is the worst and thus not a likely option, it will not be included in the following simulation results.

2.2 Least Square Method

With the help of vector notations, we can write the system given in (3) in a compact form as

$$\begin{aligned} \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} &= \begin{bmatrix} 1 & \frac{40.3}{f_1^2} \\ 1 & \frac{40.3}{f_2^2} \end{bmatrix} \begin{bmatrix} \rho \\ TEC \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \\ \mathbf{r} &= \mathbf{A}\mathbf{x} + \mathbf{e} \end{aligned} \tag{6}$$

where $\mathbf{r}$ is the observation vector that contains the pseudorange measurements, $\mathbf{A}$ is a 2×2 matrix, $\mathbf{x}$ is the unknown parameter vector to be estimated and $\mathbf{e}$ is the measurement error vector.

One of the most popular methods of solving a system of linear equations is the one that tries to minimise the squared difference between the observed data and $\mathbf{Ax}$, known as ordinary linear Least Square (LS) method. In particular, the LS solution is

$$\hat{\mathbf{x}}_{LS} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{r} \tag{7}$$

where T denotes the operation of transposition. Now, our target is to analyse the

LS performance by statistically characterising the estimation residual. For this purpose, we derive the variance-covariance matrix of the LS solution (denoted as $\Sigma_{\hat{\mathbf{x}}_{LS}}$). By definition, we have

$$\Sigma_{\hat{\mathbf{x}}_{LS}} \coloneqq E\left[(\hat{\mathbf{x}}_{LS} - \mathbf{x}_{LS})(\hat{\mathbf{x}}_{LS} - \mathbf{x}_{LS})^T\right] \tag{8}$$

where $E[\bullet]$ denotes the expectation operation. For computing the difference, $\hat{\mathbf{x}}_{LS} - \mathbf{x}_{LS}$, we substitute Eq. (6) into Eq. (7) and we get

$$\begin{aligned} \hat{\mathbf{x}}_{LS} &= (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{x} + \mathbf{e}) \\ &= (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A}\mathbf{x} + (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{e} \\ &= \mathbf{x} + (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{e} \end{aligned} \tag{9}$$

In order to calculate the variance-covariance matrix, we substitute Eq. (9) into Eq. (8) and after some mathematical manipulations we get

$$\Sigma_{\hat{\mathbf{x}}_{LS}} = \mathbf{B}\mathbf{S}\mathbf{B}^T \tag{10}$$

where $\mathbf{B} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ and $\mathbf{S}$ is a 2×2 matrix defined as

$$\begin{aligned} \mathbf{S} &= E\left[\mathbf{e}\mathbf{e}^T\right] \\ &= \begin{bmatrix} \sigma_1^2 + \mu_1^2 & E\left[\varepsilon_1, \varepsilon_2\right] \\ E\left[\varepsilon_2, \varepsilon_1\right] & \sigma_2^2 + \mu_2^2 \end{bmatrix} \end{aligned} \tag{11}$$

where σ_i^2 and μ_i are the measurement error variance and mean, respectively, for the signal transmitted in f_i for $i = 1, 2$. The non-diagonal terms of $\mathbf{S}$ are both equal to $\lambda(\sigma_1 + \mu_1)(\sigma_2 + \mu_2)$, where λ is the correlation factor.

In Fig. 1, we see the variances of the range error for LS method in the case of uncorrelated pseudorange measurements (i.e., $E[\varepsilon_1, \varepsilon_2] = E[\varepsilon_2, \varepsilon_1] = 0$). More precisely, we notice that the simulation-based curve (represented by a solid line) matches well with the theory-based one (represented by a filled circle) for all frequency combinations. The theory-based variance is found from the variance-covariance matrix as $\Sigma_{\hat{\mathbf{x}}_{LS}}(1, 1)$. For the simulation-based, we have used the simulation profile described in Section 4.1, where we used zero mean measurement errors ($\mu_i = 0$ for $i = 1, 2$).

Fig. 2 shows how the theoretical variance of the range error varies for different correlation factors when the pseudorange measurement errors have zero mean and five different values of standard deviation (we notice that only this figure is plotted using logarithmic scale for the sake of good visibility). The curve for the case of $\sigma = 0$ is not visible because the resulting variance of estimation error is zero (regardless the value of the correlation factor). For $\sigma \neq 0$ and when the correlation factor is -1 (i.e., perfect negative correlation), we have the highest error variance, while for increasing correlation factor, the variance decreases. For example, in the case of $\sigma = 0.01$ the variance for correlation factors -1 and 1 is about 49 and 39 meters, respectively. However, the variance difference between smallest and highest correlation factor increases with increasing error variance. In other words, when the measurements are correlated in the negative direction, the

impact of error is smaller than when they are correlated in the positive direction. In order to interpret this behaviour, we need to understand the physical meaning of the correlation factor. Negative correlation means that when one error variable increases the other one decreases. If for example we assume that the second measurement is subtracted from the first one (recall that in the case of zero mean error, LS method is mathematically equivalent to the traditional one) and that when ε_1 increases, ε_2 decreases, the difference $\varepsilon_1 - \varepsilon_2$ increases. If however, we have positive correlation (i.e., when ε_1 increases, ε_2 increases as well), the error difference decreases (due to partial error cancelling) and thus, the variance of the estimation error becomes smaller.

3 Enhanced Methods

The main disadvantage of both traditional and LS methods is that their solution is unrestricted. Consequently, the estimate of the unknown vector $\mathbf{x}$ may violate certain physical limitations, associated with the unknown parameters.

3.1 Constrained Least Square

One way to avoid the aforementioned problem is to impose certain constraints in the solution of the ordinary LS, leading to what is commonly known as Constrained Least Square method (CLS). More precisely, the idea is to minimise the squared difference between the observed data and $\mathbf{Ax}$, subject to the linear inequality constraint $\mathbf{A}\hat{\mathbf{x}}_{CLS} \geq \mathbf{b}$. CLS is expected to provide a more accurate solu-

tion than LS at the expense of increased computational burden.

In our model, we have set $\mathbf{b} = [0\ 0]$ which means that both range and TEC estimates are forced to be non-negative. Concerning the constraint on range, the zero value represents model-specific theoretical minimum (see Sec. 4.1 for a detailed model description). Simulation results show that this minimum value was not attained (see right middle plot of Fig. 4), thus, the zero value is a rather relaxed constraint. If any range-specific auxiliary information is available at the receiver, then the lower bound of range can be adjusted accordingly. The zero value for TEC constraint represents a physical theoretical minimum bound (i.e., not model-specific). While this minimum bound would occur in vacuum, low TEC values are possible to occur in certain time/space dependent conditions. We notice that we have tested other constraint vectors as well but they did not provide any better results than the ones produced with the non-negative constraint (the empirical results are omitted for the sake of clarity).

3.2 Brute Force Constraint

In order to reduce the computational complexity of CLS method and retain the advantage of increased accuracy via constrained solution, we have designed a simple method for estimating the range and the TEC parameters, called Brute Force Constraint (BFC). The pseudocode of the method is given below:

Calculate the mean of the pseudorange measurements as

$\hat{\rho} = M[\rho_1\ \rho_2] = \frac{1}{2}\sum_{i=1}^{2}\rho_i$, where $M[\cdot]$ denotes the arithmetic average

while (true)

$OK = 1$

for $k = 1$ till 2

Estimate TEC as $\widehat{TEC}_k = \frac{f_k^2}{40.3}(\rho_k - \hat{\rho})$

if $\widehat{TEC}_k \leq 0$

Re-estimate range as $\hat{\rho} = 2\rho_k - \hat{\rho}$

$OK = 0$

break

end

end

if $OK == 1$

break

end

end

$\widehat{TEC} = M[\widehat{TEC}_1\ \widehat{TEC}_2]$

In the first step of the algorithm, we make an estimate of the true range by calculating the average of the two available pseudorange measurements. In this step, it is assumed that no a-priori information is available on the quality of the measurements, thus both are equally treated (if any auxiliary information exists, we could weight the measurements according to their credibility). Then, we apply an iterative procedure which terminates if and only if both estimated TEC measurements are physically valid (i.e., positive). More precisely, for each received

signal we estimate its TEC using the averaged pseudorange measurement. If the estimated value of TEC is less or equal to zero then we re-estimate the range in such a way, that we force the re-estimated TEC value to be positive (this is the step from which our algorithm is named). Notice that estimated TEC is negative if $\rho_k < \hat{\rho}$ thus, when redefining $\hat{\rho}$ we have to ensure that the resulting TEC estimate will not be negative. This is achieved by setting $\hat{\rho} = 2\rho_k - \hat{\rho}$. As soon as both estimated TEC values are valid, the algorithm returns the estimated parameters. From the description above, it becomes evident that BFC requires at most two iterations (one for each frequency) in order to converge to a physically valid solution and this is what makes it computationally light.

4 Simulation results

The target of this section is two-fold. First, we perform Monte Carlo type of simulations for comparing the performance of LS, CLS and BFC methods with the traditional solution given in Eq. (4). Secondly, we assess the results in order to find what frequency combination is the best for a dual-frequency Galileo receiver. We notice, that for the sake of clarity we present only results related to range estimation, since this is the parameter of interest.

4.1 Simulation profile

In our research we consider the carrier frequencies E1 (1575.42 MHz), E5 (1189 MHz), E5a (1176.45 MHz) and E5b (1207.14 MHz), assigned to Galileo system,

which are dedicated to open and commercial services (OS and CS) [17]. As a result, we have six possible frequency combinations: E1-E5, E1-E5a, E1-E5b, E5-E5a, E5-E5b and E5a-E5b, from which the E5-E5a is not considered here since in theory it was shown to perform by far the worst (see Section 2.1 for more information). The simulation setup is as follows: we generate 2000 random realisations of the signal and of the measurement errors. More precisely, the true range is uniformly distributed between 18000 and 25000 km and the TEC is uniformly distributed between 1 and 250 TECU. We would like to emphasise that the effects associated with the satellite/receiver clock are not considered for the sake of reduced modelling complexity (in the presence of such errors, satellite clock bias can be removed using the time parameters included in the navigation message and the receiver bias can be found by linear combination of at least four pseudoranges). This implies that negative pseudorange measurements cannot occur, although in reality they are likely to appear (e.g., if the receiver clock drift is large, above a certain threshold). The limits of the TEC parameter have been chosen in such a way that typical values encountered in various latitudes are included [18–20].

Furthermore, the measurement errors are assumed to be uncorrelated [2], unless otherwise stated. Statistically, the error for each carrier frequency (here, f_1 to f_4 frequencies correspond to E1, E5, E5a and E5b, respectively) is modelled as a random variable that follows the normal distribution (the assumption of normal distribution has been commonly encountered in the literature [2, 21–23]). More precisely, the errors are distributed according to $\varepsilon_i \sim \mathcal{N}(\mu_{\varepsilon_i}, \sigma_{\varepsilon_i})$, where $\mu_{\varepsilon_i} = 1/f_{s,i}$

with $f_{s,i}$ being the sampling frequency and σ_{ε_i} takes values from 0 to 0.2 chips with a step of 0.01 chip. Because E1 signal has smaller chip rate than the other three, a standard deviation error of 0.01 chips translates into 2.932 m. of error for E1 signal and into 0.293 for the rest (see Table 2 for chips-to-meters mapping). We notice that the choice of the mean range error represents a worst-case scenario, since a typical implementation for tracking the code delay (such as the commonly used narrow Early-Minus-Late discriminator) can provide estimates with resolution within two consecutive samples (e.g., in a three-correlator tracking unit, the delay estimates can fall anywhere between early and late correlators). However, critical environments which are characterised by heavy multipath propagation effects, such as densely-built areas, can cause high tracking errors (e.g., for the GPS C/A and Galileo Open Service signals, these errors may be several decades of metres, depending the on the channel profile and the receiver) [24–26]. Furthermore, the sampling frequency, $f_{s,i}$, was taken equal to $2W_i$, where W_i is the signal bandwidth (in order to fulfill Nyquist rule). Because the main spectral lobes of the Galileo signals fall within the range $[-2N_{B,i}f_{chip,i}\ 2N_{B,i}f_{chip,i}]$, where $N_{B,i}$ and $f_{chip,i}$ are the modulation order and chip rate of the signal transmitted in f_i frequency, respectively, it can be easily shown that $W_i = 4N_{B,i}f_{chip,i}$ [27]. We recall that the modulation order is defined as the ratio between the sub-carrier rate and the chip rate and it depends on the used modulation [27]. For example, for a BPSK-modulated signal on frequency f_i, $N_{B,i} = 1$; for a CBOC-modulated signal on f_i, $N_{B,i} = 12$ [27]. In other words, the mean error depends on two parameters: the modulation order and the chip rate of the signal in use (see Table 3 for mean

measurement errors in metres).

The description of the included methods are given in Sections 2.2 and 3. CLS method is implemented with the help of quadratic programming. Also, we recall that we have chosen $\mathbf{b} = [0\ 0]^T$ as the inequality constraint vector. Moreover, for reference purpose, we also include the cases in which no ionospheric correction is done in any of the two signals. The performance measure we used is the Root Mean Square Error (RMSE) expressed in units of meters.

4.2 Assessment of results

The simulation results are divided into two categories: In the first one, the mean error is zero (i.e. $\mu_{\varepsilon_1} = \mu_{\varepsilon_2} = 0$) and in the second one we have $1/(8N_{B,i}f_{chip,i})$ (recall that $\mu_{\varepsilon_i} = 1/f_{s,i} = 1/(2W_i) = 1/(2 * 4N_{B,i}f_{chip,i})$). In Fig. 3, we see the RMSE for E1-E5a pair versus the standard deviation. We chose this frequency combination because from theoretical point of view it was shown to result in the smallest noise variance (see Section 2.1). In the case of zero mean (left plot), the poor performance of LS can be explained by the fact that no constraints are set, which consequently allows for physically invalid solutions. We notice that because the mean measurement errors are zero, the traditional method is mathematically equivalent to the LS. Therefore, the performance curve of traditional method is omitted from the left plot of Fig. 3 in order to avoid redundancy. It is important to observe that in the absence of tracking error (i.e. when the mean and deviation of error are zero), both the traditional and the LS approaches result in zero RMSE, however, their overall performance indicates how sensitive these

methods are in the presence of even small errors.

Regarding the CLS method, we see that it performs considerably better than LS, which is as expected, since CLS restricts the solution to valid values. The method we proposed based on brute force constraint performs the best of all methods for error deviation higher than 0.03. However, when the error deviation is higher than 0.13 chips, the BFC does not offer any advantage if E5a signal is used. The results for non-zero mean are shown in the right plot of Fig. 3 from which it is obvious that introducing a mean tracking error caused all methods to perform worse than with zero mean measurement errors. Nevertheless, in practise it is more likely to have non-zero mean measurement errors due, for example, to multipath bias. We see that CLS method performs the best for small error deviations (i.e. up to 0.03 chips), while BFC is the best when error deviation is higher than 0.03 chips.leme. Moreover, comparison between the two plots shows that CLS and BFC methods perform better in cases of non-zero mean error. Nonetheless, we would like to emphasise that BFC method is the least computationally demanding, while CLS method is the most complex.

Fig. 5 shows the histograms of TEC and range estimation errors for BFC (top plots), LS (middle plots) and CLS (bottom plots) methods. We see that in the BFC method the estimation error has the smallest dispersion both for TEC and range estimates, while in LS method we have the biggest spreading. Moreover, BFC and CLS methods seem to overestimate TEC (longer tail on the left side of the histogram) and underestimate range (longer tail on the right side of the histogram) more often, while the estimation errors for LS method have a more

bell-like shape.

In order to investigate further the non-symmetric histograms in the case of CLS and BFC methods, we plot the Absolute Mean Error (AME) and the Variance of Error (VoE) in the case of zero mean. From the left plot of Fig. 6 we see that LS method is unbiased (zero mean), while in the case of CLS and BFC methods some bias is present. On the other hand, comparison based on the variance of error shows that BFC method has the smallest variance while LS the highest. So, by choosing a method that imposes certain constraints we trade variance for mean (a lower mean and higher variance in LS case is counter-balanced by a much lower variance, but slightly higher bias in the other algorithms, explained by the presence of the constraints

The impact of correlated noise when mean is zero is shown in Fig. 7. Specifically, we see that when $\sigma = 0.01$ chips (left plot) LS and CLS methods benefit from positively correlated noise since RMSE decreases with increasing correlation factor. Such behaviour is in agreement with the theoretical results shown in Fig. 2. Regarding BFC method, we see that RMSE is not much affected by the presence of correlated errors. In the case of $\sigma = 0.1$ chips (left plot), CLS method performs better than LS, unless in the case of strong positive correlation (i.e., when $\lambda \geq 0.8$) where they perform similarly.

5 Receiver design implications

When designing a dual-frequency GNSS receiver, it is important to decide what are the two frequencies to be used. Because different combinations may have different advantages over others, a designer is advised to choose and prioritise the performance metrics based on which a choice is made. In this section, we suggest what is the best frequency combination to be used, assuming the channel scenario described in Section 4.1 and a certain method.

As it was shown earlier, when the mean measurement error is zero and the standard deviation is small, the LS or equivalently the traditional method performs the best and the frequency pair to be chosen is the one with the largest separation. The choice of this pair is supported not only from the theoretical analysis in Section 2.1 but also from the results shown in the left plot of Fig. 8. On the other hand, when non-zero mean error is expected, CLS method performs the best for small standard deviation (however, in this case the performance difference between CLS and BFC is marginal). In the right plot of Fig. 8 we see the performance of CLS method for different frequency combinations and for small error deviations, the E5a-E5b pair appears to be the best option.

If the measurement errors are characterised by high variance, the suggested BFC method was shown to perform the best in both cases of zero and non-zero mean error (see Fig. 3). The performance of BFC for the various frequency combination can be seen in Fig. 9. We see that if $\sigma \geq 0.15$, then any combination of E1 signal with the rest is the best choice, while for $\sigma < 0.15$, E5-E5b and

E5a-E5b pairs give the best results. At this point, it is important to emphasise the finding that unlike in the case of zero mean error, in the case of non-zero mean error the best pair does not depend on the frequency separation. For example, for CLS method we see that the theoretical order is preserved only for error deviation higher than 0.06 chips (see left plot of Fig. 8).

6 Conclusions and future work

This paper focuses on the estimation of the ionosphere-corrected satellite-receiver range in dual-frequency receivers. First, we described the traditional method for estimating the geometric range. Then, we modelled our estimation problem in such a way, that linear methods can be applied. More precisely, we statistically analysed the performance of the ordinary linear Least Square (LS) method and compared it with the traditional one. In order to avoid physically-invalid solutions, we employed the Constraint Least Square (CLS) method, in which linear inequality constraints are imposed. Moreover, a new method with reduced complexity, called Brute Force Constraint (BFC), is described. Next, we examined the effect of code tracking error in the performance of the above mentioned methods and provided comparative discussion of them.

The simulation results showed that the LS method has the same performance with the traditional approach. Moreover, it was shown that both of them are able to successfully estimate the true range (i.e. resulting Root Mean Square Error (RMSE) was zero) when there is no measurement error (e.g., tracking gives per-

fect delay estimates). However, in the realistic case when even a small measurement error occurs (due for example to incomplete multipath mitigation in the delay tracking), their performance deteriorates significantly.

Moreover, the results showed that CLS method performs better than LS at the expense of increased computational complexity. Nevertheless, in the case of zero mean error, LS method was the only unbiased estimator but also the one characterised with the highest variance. The proposed BFC method and CLS were shown to perform better in the case where the tracking error variable was characterised by non-zero mean. Finally, the results indicated that the E1-E5a frequency pair is the best combination for a dual-frequency receiver.

Based on these observations, and on the fact that in severe multipath propagation scenarios, such as in urban areas, where measurement errors are highly likely, we would recommend that dual-frequency ionospheric correction should be made based on the proposed BFC algorithm, instead of using the classical solutions (linear combination or LS). BFC algorithm offers a low complexity solution with significant advantage over the classical solutions when measurement errors are present.

Priorities for future research into the ionosphere-corrected range estimation include further optimisation of BFC algorithm and performance evaluation using real satellite data.

Acknowledgement

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. [227890]. The work has also been supported by TISE graduate school and by the Academy of Finland.

References

[1] Kaplan, E.: Understanding GPS: Principles and Applications. Artech House (1996)

[2] Misra, P. and Enge, P.: Global Positioning System: Signals, Measurements, and Performance. Ganga-Jumuna Press, Lincoln, Massachusetts (2001)

[3] Odijk, D.: Fast Precise GPS Positioning in the Presence of Ionospheric Delays. PhD thesis, Publications on Geodesy 52, Netherlands Geodetic Commission, Delft, Netherlands (2002)

[4] Klobuchar, J. A.: 'Ionospheric effects on GPS'. GPS World,. 1991 pp. 48–51

[5] Appleton, E. V. U. R. S. I. papers (1927) Washington.

[6] Bidaine, B., Prieto-Cerdeira, R., and Orus, R.: 'Nequick: In-depth analysis and new developments'. In proc. of the 3rd ESA Workshop on Satellite Navigation User Equipment Technologies NAVITEC 2006 [CD-Rom], Noordwijk, The Netherlands, 2006

[7] Seo, J., Walter, T., and Enge, P.: 'Availability benefit of future dual frequency gps avionics under strong ionospheric scintillation'. ION GPS/GNSS, Savannah, GA, 2009

[8] Yang, Y., Sharpe, R., and Hatch, R.: 'L1 backup navigation for dual frequency gps receiver'. ION GPS/GNSS, Portland, Oregon, 2003 pp. 1258–1263

[9] Batchelor, A., Fleming, P., and Morgan-Owen, G.: 'Ionospheric delay estimation in the european global navigation overlay service'. Remote Sensing of the Propagation Environment (Digest No: 1996/221), IEE Colloquium on,. 1996 pp. 3/1 –3/6

[10] Rao, G. S. B. and Goswami, R.: 'Ionospheric delay estimation for improving the global positioning system position accuracy'. IETE Journal of Research, 2008, 54, (1), pp. 23–29

[11] Engel, U.: 'A theoretical performance analysis of the modernized GPS signals'. Position, Location and Navigation Symposium, 2008 IEEE/ION,. 2008 pp. 1067–1078

[12] Gerein, N., Olynik, M., Clayton, M., Auld, J., and Murfin, T.: 'A dual-frequency l1/e5a galileo test receiver'. In Proc. of of ENC-GNSS, Munich, Germany, 2005

[13] Hurskainen, H., Lohan, E.-S., Nurmi, J., Sand, S., Mensing, C., and Detratti, M.: 'Optimal dual frequency combination for galileo mass market receiver

baseband'. Signal Processing Systems, 2009. SiPS 2009. IEEE Workshop on,. 2009 pp. 261–266

[14] Farrell, J. A. and Barth, M.: The Global Positioning System and Inertial Navigation. McGraw-Hill (1999) (ISBN-0-07-022045-X).

[15] De Oliveria Camargo, P., Monico, J. F. G., and Ferreira, L. D. D.: 'Application of ionospheric corrections in the equatorial region for l1 gps users'. Earth, Planets Space, 2000, 52, (11), pp. 1083–1089

[16] Kim, B. C. and Tinin, M. V.: 'Effect of ionospheric irregularities on accuracy of dual-frequency gps systems'. Geomagnetism and Aeronomy, 2007, 47, (2), pp. 238–243

[17] Hein, G. W., Godet, J., Issler, J. L., Martin, J. C., Erhard, P., Rodriguez, R. L., and Pratt, T.: 'Status of Galileo Frequency and Signal Design'. CDROM Proceedings of ION GPS,. 2002

[18] Gao, G. X., Datta-Barua, S., Walter, T., and Enge, P.: 'Ionosphere effects for wideband GNSS signals'. ION Annual Meeting, Cambridge, Massachusetts, 2007

[19] Hein, G. W. and Avila-Rodriguez, J. A.: 'Combining Galileo PRS and GPS M-Code'. Inside GNSS, 2006, 1, (1), pp. 48–56

[20] Mannucci, A. J.: Ionosphere and ionospheric delay. www.cosmic.ucar.edu/summercamp_2005/presentations/Mannucci_Anthony_

20050602.pdf (2005) [presentation available online; accessed 5-January-2010].

[21] Kuusniemi, H., Lachapelle, G., and Takala, J.: 'Position and velocity reliability testing in degraded gps signal environments'. GPS Solutions, 2004, 8, (4), pp. 226 – 237

[22] Liu, G. and Liu, Y.: 'An methodology of automatic guidance system'. Networking, Sensing and Control, 2006. ICNSC '06. Proceedings of the 2006 IEEE International Conference on,. 2006 pp. 568 –571

[23] O' Keefe, K., Julien, O., Cannon, M., and Lachapelle, G.: 'Availability, accuracy, reliability, and carrier-phase ambiguity resolution with galileo and gps'. Acta Astronautica, 2006, 58, (8), pp. 422–434

[24] Lohan, E. S., Lakhzouri, A., and Renfors, M.: 'Feedforward delay estimators in adverse multipath propagation for Galileo and modernized GPS signals'. EURASIP Journal on Applied Signal Processing, 2006, 2006,, 19 Article ID 50971.

[25] Skournetou, D. and Lohan, E. S.: 'Discontinuity-based code delay estimator for GNSS signals'. 4th Advanced Satellite Mobile Systems (ASMS'08), Bologna, Italy, 2008 pp. 213–222

[26] Skournetou, D., Sayed, A. H., and Lohan, E. S.: 'A deconvolution algorithm for estimating jointly the Line-Of-Sight code delay and carrier phase

of GNSS signals'. European Navigation Conference in GNSS (ENC-GNSS), Naples, Italy, 2009 (in press).

[27] Lohan, E. S., Lakhzouri, A., and Renfors, M.: 'Binary-offset-carrier modulation techniques with applications in satellite navigation systems'. Wiley Journal of Wireless Communications and Mobile Computing, DOI: 10.1002/wcm.407, published on-line, 2006, 7, (6), pp. 767–779

Signals & carrier frequencies, kHz	E5 (1189)	E5a (1176.45)	E5b (1207.14)
E1 (1575.42)	386.42	398.97	368.28
E5		12.55	18.14
E5a			30.69

Table 1: Absolute frequency separation (in KHz) for different freq. combinations.

σ, chips / Signals	0.01	0.05	0.1	0.15	0.2
E1	2.932	14.662	29.325	43.988	58.651
E5, E5a, E5b	0.293	1.466	2.932	4.398	5.865

Table 2: Standard deviation of measurement range error: Chips-to-meters mapping.

Signals	μ, metres
E1	18.3284
E5	0.9164
E5a, E5b	3.6657

Table 3: Non-zero mean measurement range error in units of metres.

List of Figures

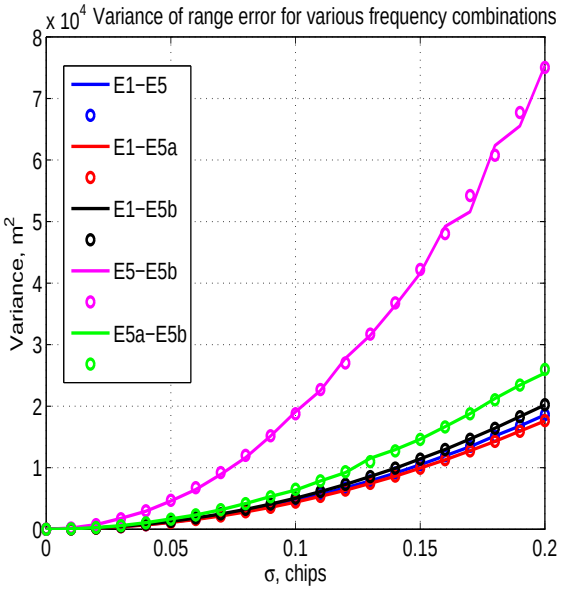

Figure 1: Variance of range error for various frequency combinations and zero mean measurement errors (solid line: simulation-based, filled circle: theory-based)

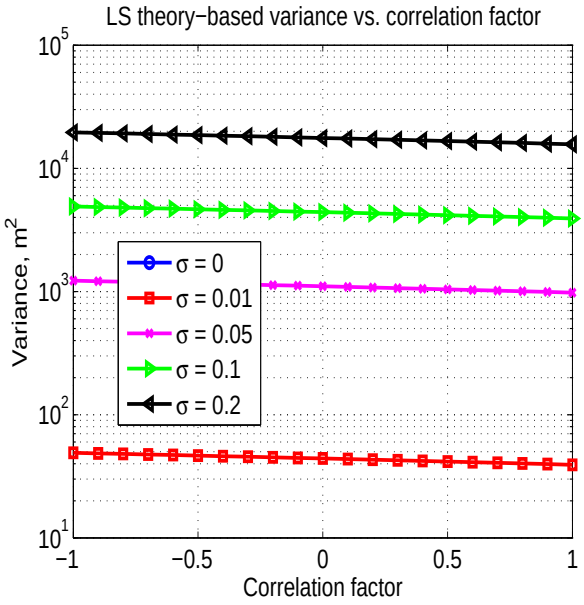

Figure 2: Theory-based variance of LS method versus correlation factor (λ), for different standard deviations of measurement error (σ) and mean equal to zero ($\mu = 0$).

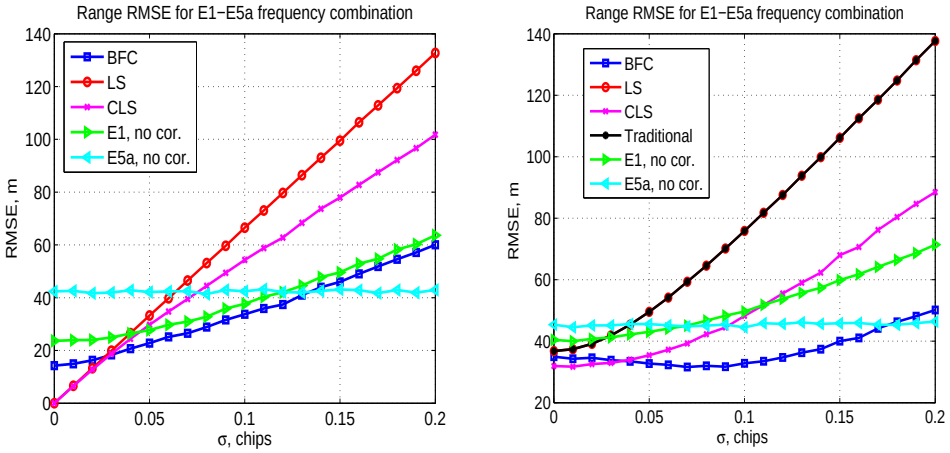

Figure 3: Range RMSE for E1-E5a frequency combination, zero (left plot) and non-zero (right plot) mean.

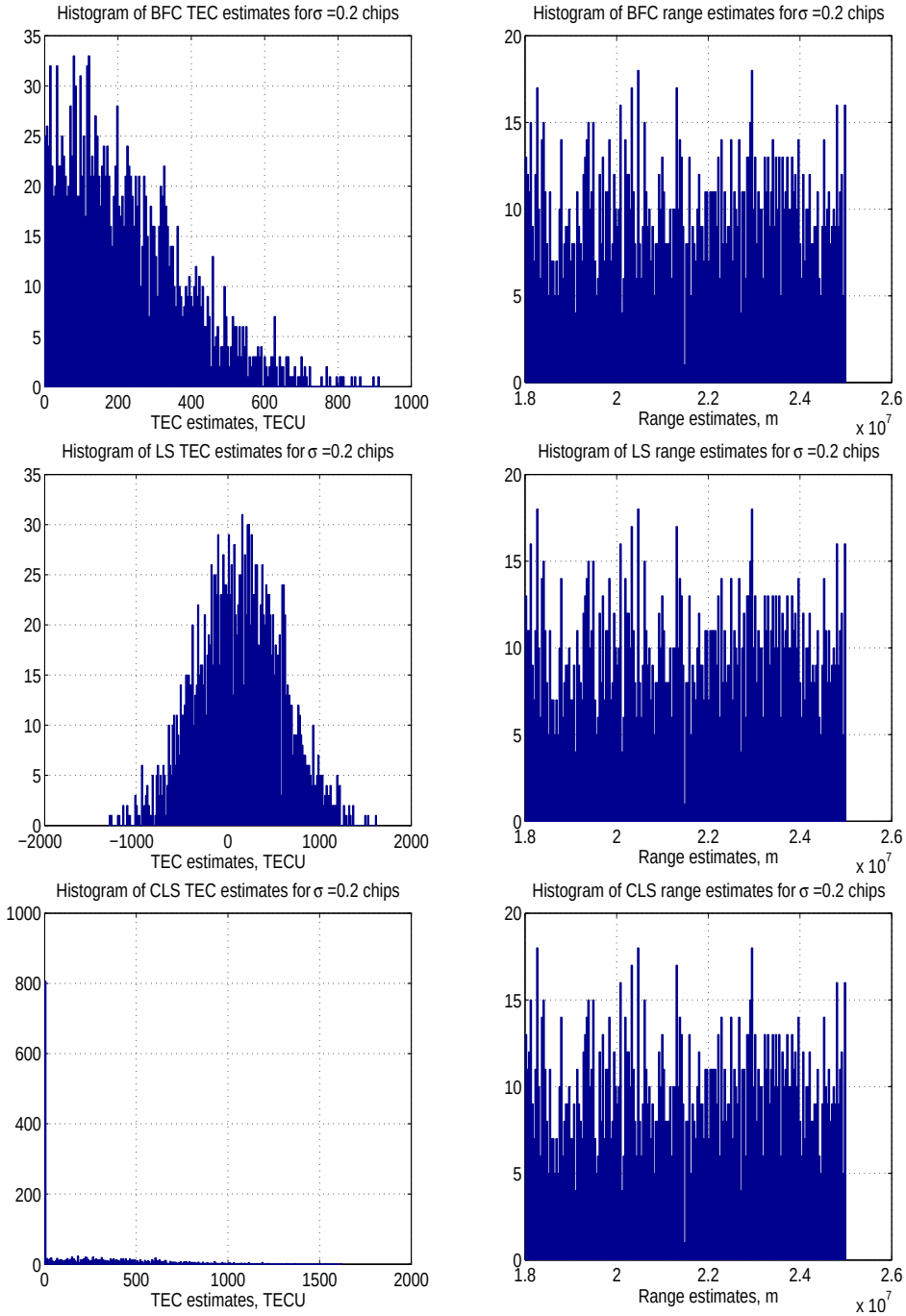

Figure 4: Histograms of TEC and range estimates of BFC (top plots), LS (middle plots) and CLS (bottom plots), for E1-E5a frequency combination, zero mean and $\sigma = 0.2$ chips.

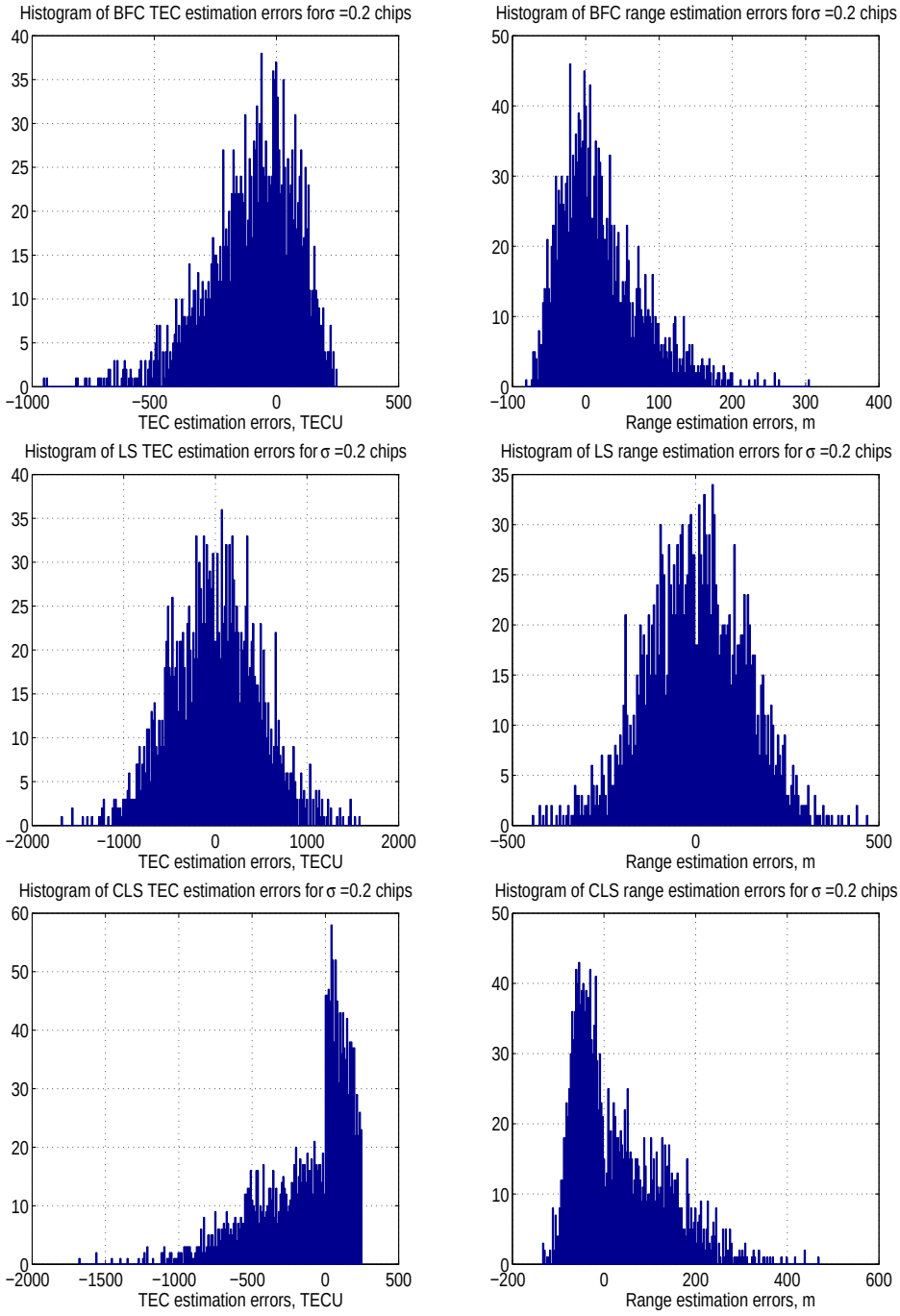

Figure 5: Histograms of TEC and range estimation errors of BFC (top plots), LS (middle plots) and CLS (bottom plots), for E1-E5a frequency combination, zero mean and σ = 0.2 chips.

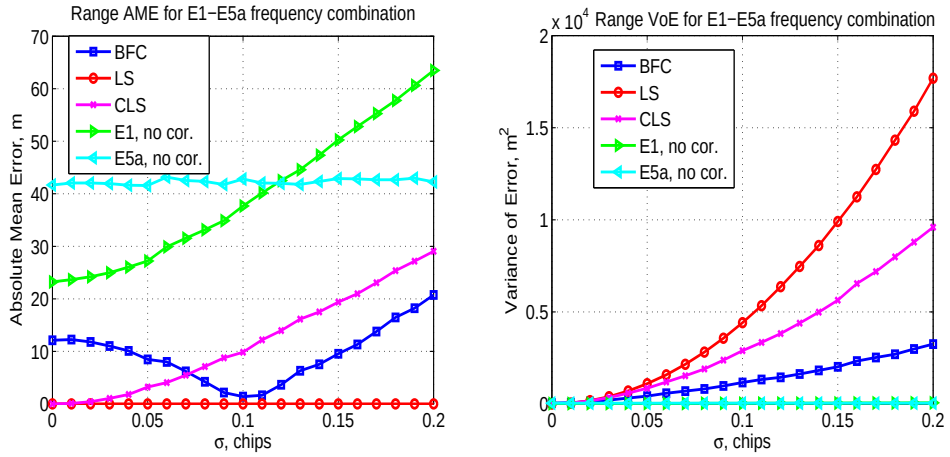

Figure 6: AME (left plot) and VoE (right plot) for E1-E5a frequency combination and zero mean.

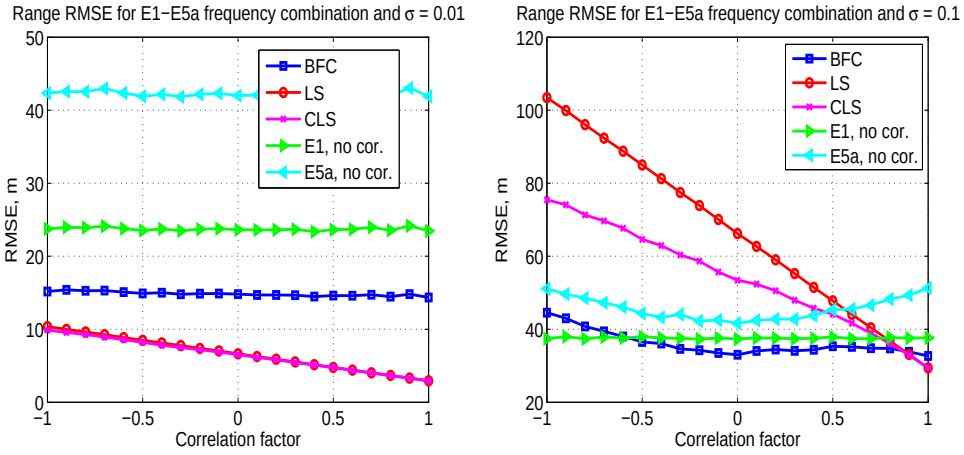

Figure 7: Range RMSE versus correlation factor for E1-E5a frequency combination, zero mean, $\sigma = 0.01$ (left plot) and $\sigma = 0.1$ (right plot).

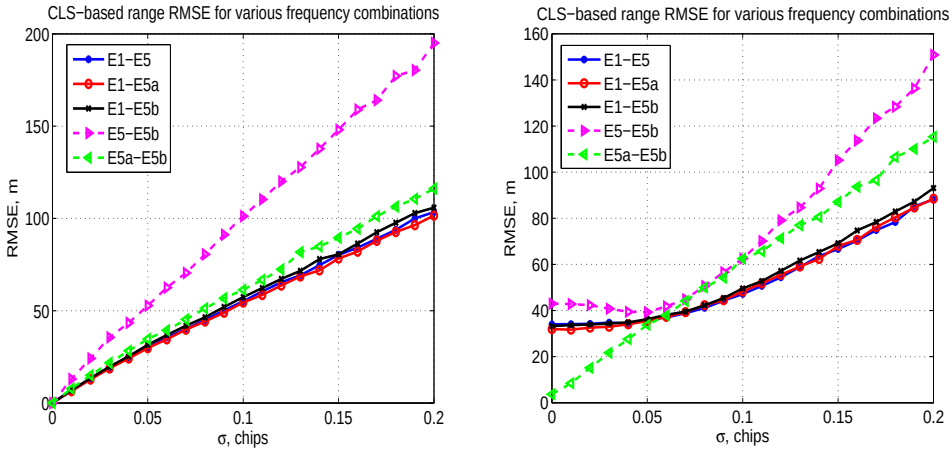

Figure 8: Range RMSE of CLS method for various frequency combinations, zero (left plot) and non-zero mean (right plot).

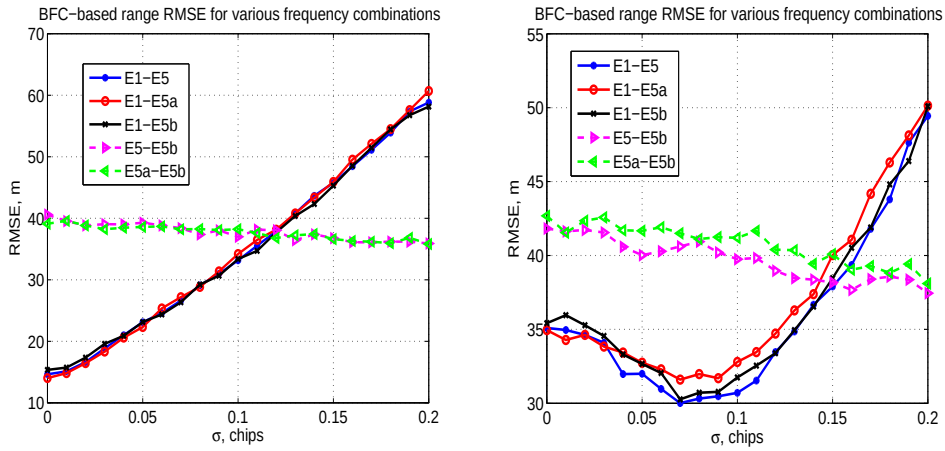

Figure 9: Range RMSE of BFC method for various frequency combinations, zero (left plot) and non-zero mean (right plot).