COST Action: TU1302
Action Title: Satellite Positioning Performance Assessment for Road Transport – SaPPART

GENERATION OF SIMULATED DEGRADED TRAJECTORIES BY USING ERROR MODELS

COST is supported by the EU Framework Programme

ESF provides the COST Office through a European Commission contract

Document Profile

Document status:	
Deliverable title:	Generation of simulated degraded trajectories by using PVT error models
Work Group/Task Force:	SaPPART COST action
Preparation date:	26/11/2014
Submission date:	26/11/2014
Total pages:	24
Dissemination level:	
Authors:	Carlos Moriana-Varo
Contributors:	

Revision History

Rev.	Date	Authors/Company	Comments
1.0	26/11/2014	Carlos Moriana Varo / GMV	First draft

Table of Contents

List of Figures

List of Tables

1. Introduction

The purpose of this document is to analyze different algorithms for the generation of degraded trajectories, based on specific PVT error models, in the frame of the SaPPART COST action. These models correspond to simulation algorithms developed by IFSTTAR, although an alternative method, whose performance overcomes the challenges of the problem, is also introduced herein. This last method has been proposed by the author of the present document.

Besides, this memorandum summarizes the work the author has developed during his Short-Term Scientific Mission (STSM), during a period of two weeks, in IFSTTAR. Therefore, it is the purpose of the author to adapt the content of this document to the scheduled work plan, which was presented to the COST authority. This work plan consisted on the following main tasks:

- Review of the PVT developed so far within the SaPPART framework.
- Familiarization with the existing SW tools and the algorithms.
- Automated generation of simulated tracks.
- Analysis of the global and local behavior of the models under different configurations.

The structure of this document will be as follows: first, a brief description of the existing algorithms shall be performed. Second, a new method proposed by the author will be detailed. Third, experimental results for each PVT error model will be presented to the reader, taking into consideration typical target environments, such as open sky areas or urban canyons.

2. Background

In this paragraph, the different algorithms developed so far by IFSTTAR will be evaluated from a theoretical point of view. Nevertheless, it is worth to mention that an additional PVT model, apart from those included in the present document, has been implemented. This model is based on the automatic pattern detection by means of the Chopin method [1].

The main reason for not including this promising model is that, at the time of writing this memorandum the algorithm had not been fully tested.

Therefore, two of the PVT error models developed by IFSTTAR shall be explained in this paragraph. First, we will introduce the model based on Cauchy noise. Second, we will introduce a second model which includes an additional laplacian noise source.

All the methods proposed in this document generate PVT error signals (in our particular case, along-track and cross-track errors), which are compared with some reference error signals, from the statistical point of view (i.e. autocorrelation and probability distribution). These reference error signals have been generated by means of computing the real position errors, in the desired components, for several commercial receivers (in particular, the results of this document relates to u-blox LEA-6T receiver). The reference trajectography system Novatel SPAN CPT has been also used for this purpose.

Moreover, these error signals correspond to different types of environment, such as open sky or urban canyon areas.

2.1 PVT Model with Cauchy Noise

This paragraph describes briefly the PVT model based on Cauchy noise (see [2]). This model has its origin in the evaluation of the behavior of the raw data (i.e. along-track and cross track errors).

The Figure 1 represents the along-track and cross-track errors for real PVT signal recorded in the city of Nantes with a u-blox LEA-6T receiver. From the previous figure, we can extract some patterns concerning the structure of the signal: first, the presence of sudden peaks in the error signal, which seems to be fully uncorrelated with other parts of the signal; second, these sudden peaks last over time during a certain number of samples, generating step patterns; and third, that the signal is, with relatively high probability, included within some error bounds.

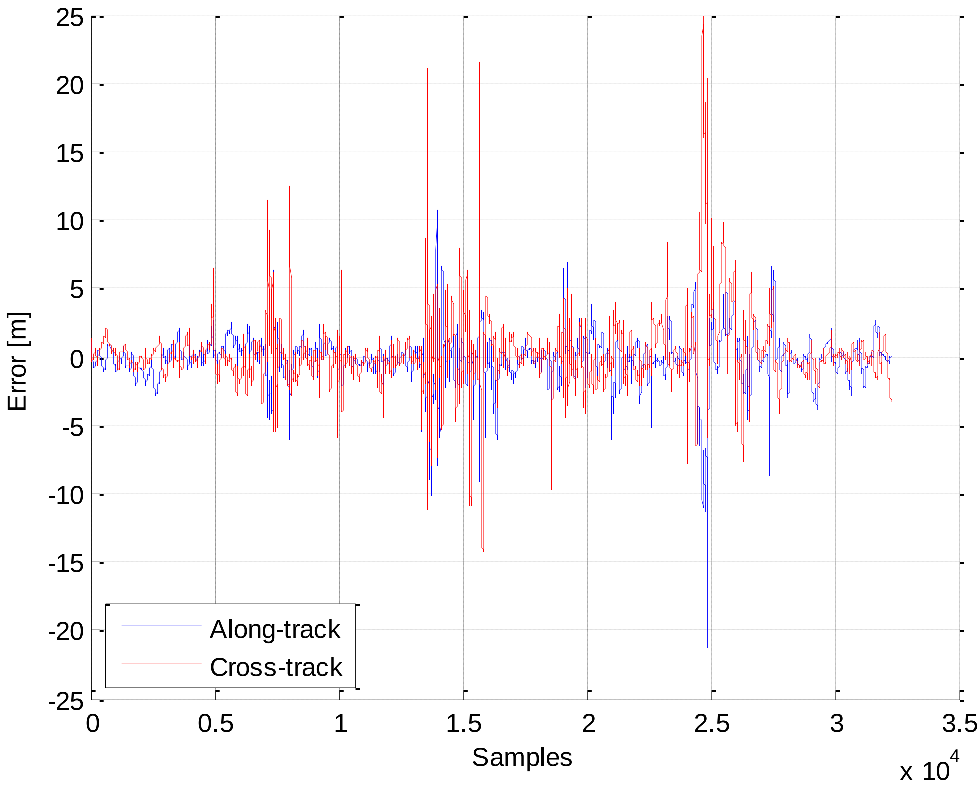

Figure 1 Along-track and cross-track errors for real u-blox LEA-6T position signal

For the sake of simplicity, it was decided to work separately with a radius signal (which is the norm of both along and cross track components), and an angle signal. These signals are generated independently.

The Cauchy model generates the magnitude of these steps by means of a truncated Cauchy distribution. Their frequency is controlled probabilistically by means of a uniform distribution. It is, if we denote A as the amplitude of the generated radius signal for a sample n:

$$A[n+1] = \begin{cases} A[n] & with\ probability\ (1-p) \\ C_r[n] & with\ probability\ p \end{cases}$$

where $C_r[n]$ is a realization of a Cauchy distribution, and p the probability of varying the size of the current step. The values of the parameters to generate $C_r[n]$ samples properly are estimated by means of Bayesian methods.

At this point, we have a square signal with different amplitudes, which does not correspond at all with the actual observed error signal, visually speaking (although both share similar statistic properties). An additional processing is performed, estimating the coefficients of an auto-regressive filter which relates both signals, such that the artificial square signal can be filtered and look much more alike a random noisy signal. Bayesian methods are also used for this purpose.

On the other hand, the angle signal is generated by a random walk based also in Cauchy noise:

$$\phi[n+1] = \phi[n] + C_\phi[n]$$

where $C_\phi[n]$ is a realization of a Cauchy distribution. The values of the parameters to generate $S_\phi[n]$ samples properly are estimated experimentally.

The computation of the along-track and cross-track error components is straightforward:

$$r_a[n] = \omega_a A[n] cos\phi[n]$$
$$r_c[n] = \omega_c A[n] sin\phi[n]$$

being ω_a and ω_c normalization constants. These constants model the fact that the error vector does not distribute equally in each analyzed component.

For a deeper analysis of the method, the reader is invited to consult [3].

COST is supported by the EU Framework Programme
European Science Foundation ESF provides the COST Office through a European Commission contract

2.2 PVT Model combining Laplace and Cauchy Noise

Although the results obtained with the Cauchy error model provided good results for the data analyzed, a new model was proposed, based again on the generation of a radius signal, but where the steps were generated by other distribution rather than the Cauchy one. The reasons to do so were mainly twofold:

- The Bayesian methods used for the automatic calibration of the Cauchy model (noise characteristics, filter coefficients) are difficult to work with the Cauchy distribution, due to its characteristics (neither mean nor variance are available).
- The magnitude of the generated steps in the first model was too high with respect to the observed error magnitudes in the reference data[1]. This was solved in the first method by truncating the Cauchy distribution.

The input data set used to describe this method is equivalent to the one in the first model, exemplified in Figure 1.
In this method, the radius signal $A[n]$ is generated analogously to the first model, where the step variation was controlled by a uniform distribution, but using a Laplace distribution (see [4]) instead of Cauchy, such that:

$$A'[n+1] = \begin{cases} A[n] & with\ probability\ (1-p) \\ L_r[n] & with\ probability\ p \end{cases}$$

where $A'[n+1]$ is the free of noise radius signal, $L_r[n]$ is a realization of a Laplace distribution, and p the probability of varying the size of the current step.
This model also adds to the radius signal an independent source of Cauchy noise, but much more limited in power than in the previous model. In particular, the relationship between the free of noise radius signal and the final radius signal is:

$$A[n+1] = A'[n+1] + C_r[n]$$

where $C_r[n]$ is again a realization of a Cauchy distribution. The values of the parameters to generate $L_r[n]$ and $C_r[n]$ samples properly are estimated by means of Bayesian methods.
As in the previous model, we estimate the coefficients of an auto-regressive filter which relates both the expected error signal with the simulated signal to smooth the final result.
Finally, the angle signal is generated by a random walk based again in Cauchy noise:

$$\phi[n+1] = \phi[n] + L_\phi[n]$$

where $S_\phi[n]$ is a realization of a Cauchy distribution. The values of the parameters to generate $S_\phi[n]$ samples properly are estimated experimentally.
The computation of the along-track and cross-track error components is straightforward:

$$r_a[n] = \omega_a A[n] cos\phi[n]$$
$$r_c[n] = \omega_c A[n] sin\phi[n]$$

being ω_a and ω_c normalization constants. These constants model the fact that the error vector does not distribute equally in each analyzed component.

For a deeper analysis of the method, the reader is invited to consult [3].

3. An Alternative PVT Error Model

This paragraph describes an alternative PVT error model, proposed by the author of the present memorandum during his STSM.
One of the problems that have the two previous PVT error models is that they shall identify a relatively high number of parameters. In order to do so, Bayesian estimation models have been implemented.
Nevertheless, three main drawbacks arise at this point: first, these parameters are not unique, and depend strongly on the data set analyzed; second, the number of parameters to estimate (7 in the first method, 11 in the second method); and third, the method is not automatic (this problem can be overcome in the future when the Chopin model is available).

[1] Although the Cauchy model evolved into the second model, due to the lack of control in the magnitude of the signal steps, new data recorded in Paris urban canyon areas have proved that this first method can be taken into consideration for the generation of PVT degraded trajectories.

The author proposes an algorithm which only needs to model a single parameter, once the input data set (along-track and cross-track real errors) is provided. This model tries to take advantage of the knowledge about the error distribution underlying into the input data.
In particular, the cumulative distribution functions (CDF) for each error (i.e. along-track and cross-track) are computed. It is worth to mention that these probability functions can be related to the general data set (regardless of the type of environment) or calculated for each type of scenario (i.e. specific distributions for open sky or urban canyon), which gives flexibility in order to describe how the model should be designed.
Once these functions are known, the generation of independent draws by means of a uniformly distributed random variable is straightforward, through the inverse transform technique, which claims the following:

Let be U a uniform variable in the range $[0,1]$. If $X = F^{-1}(U)$, then X is a random variable with CDF $F_X(X) = F$.

Therefore, for the experimental CDFs obtained, F_A and F_C, corresponding to the along-track and cross-track errors respectively, we are capable of generating independent samples (X_A, X_C), that fit the initial distribution.
Nevertheless, it has been observed that the along-track and cross track error components are actually correlated (in magnitude). To take into account this correlation, the final equations for the error signal generation are:

$$X_A = F_A^{-1}(U)$$
$$X_C = \lambda \cdot sign(1-2S) \cdot F_A^{-1}(U) + (1-\lambda) \cdot F_C^{-1}(V)$$

where S, U, V are independent uniform variables in the interval $[0,1]$, λ a correlation factor (experimental values are around $0.4 - 0.5$), and $sign(\cdot)$ the sign function.
Up to now, the method relies exclusively on the input data and no additional parameters have been estimated. However, something is missing at this point: the correlation through time of the two error signals.
This correlation is introduced through a modified Gauss-Markov process, in which the additive Gaussian noise is replaced by the pair of independent samples (X_A, X_C) generated in the previous step:

$$X_A^{(f)}[n+1] = e^{-\frac{\Delta t}{\tau}} \cdot X_A^{(f)}[n] + \left(1 - e^{-\frac{2\Delta t}{\tau}}\right)^{\frac{1}{2}} \cdot X_A$$
$$X_C^{(f)}[n+1] = e^{-\frac{\Delta t}{\tau}} \cdot X_C^{(f)}[n] + \left(1 - e^{-\frac{2\Delta t}{\tau}}\right)^{\frac{1}{2}} \cdot X_C$$

being $X_A^{(f)}[n]$ and $X_C^{(f)}[n]$ the output along-track and cross-track error signals generated by the model, Δt the time between samples (in the u-blox LEA-6T, $\Delta t = 0.2s$), and τ the correlation time between samples.
In the present method, all the modelling problems concentrate in the proper definition of the value of this correlation time, instead of playing with a huge set of parameters, which makes much easier the process required for characterizing a specific scenario.
One may identify the value of this τ parameter as the convergence state of the navigation filter (internal to the receiver) which is generating the true error signal. This assumption makes sense in a certain way: we could expect a noisier behavior of the error signal under transitory periods, and a smoother error evolution when the filter is working under steady-state regime. This convergence state may be given by the type of scenario under which the receiver is working (e.g. low satellite visibility in urban areas or open sky), poor geometries, or vehicle speed (e.g. high multipath in urban canyon at low velocities), among others.
A very simple and preliminary model has been generated to test this method, based mainly on the velocity of the vehicle. In particular, we have considered a high correlation time when the velocity is low, and vice versa. Notice that this is not fully realistic, since it is considering low correlation times when, for instance, the car is in a highway under open sky conditions. Thus, more variables shall be included in the definition of this correlation time. Nevertheless, from the statistical point of view, the proposed test is

enough to validate the statistical properties of the generated signal. The set of values for the τ parameter is as shown in Table 1:

Velocity condition [m/s]	Correlation time [s]
$v \geq 0$ && $v < 3$	120
$v \geq 3$ && $v < 5$	60
$v \geq 5$ && $v < 15$	30
$v \geq 15$	15

Table 1 Correlation time values in terms of the velocity

An additional condition has been added, concerning the way the algorithm handles the outliers of the distribution: if an outlier for the current pair (X_A, X_C) is produced (i.e. anyone of the uniform variables used to generate the pair are below or above certain limit probability), the correlation time is set to 1 second, with a transitory (configurable) period during which the τ parameter distributes uniformly between 0 and 5 seconds. The purpose of this modification is to consider both in the simulated signal sudden peaks with fast and low return to zero, just as observed in the true error signal.
Although this algorithm seems very different from the two previous methods, they all three share a common idea: estimate the CDF of the final distribution. In the first two methods, the CDF is parameterized to fit into a certain distribution, while in the third method the CDF is obtained experimentally. As an example of this, Figure 2 shows the CDF estimated by the Cauchy method with respect to the true one in a generic scenario.

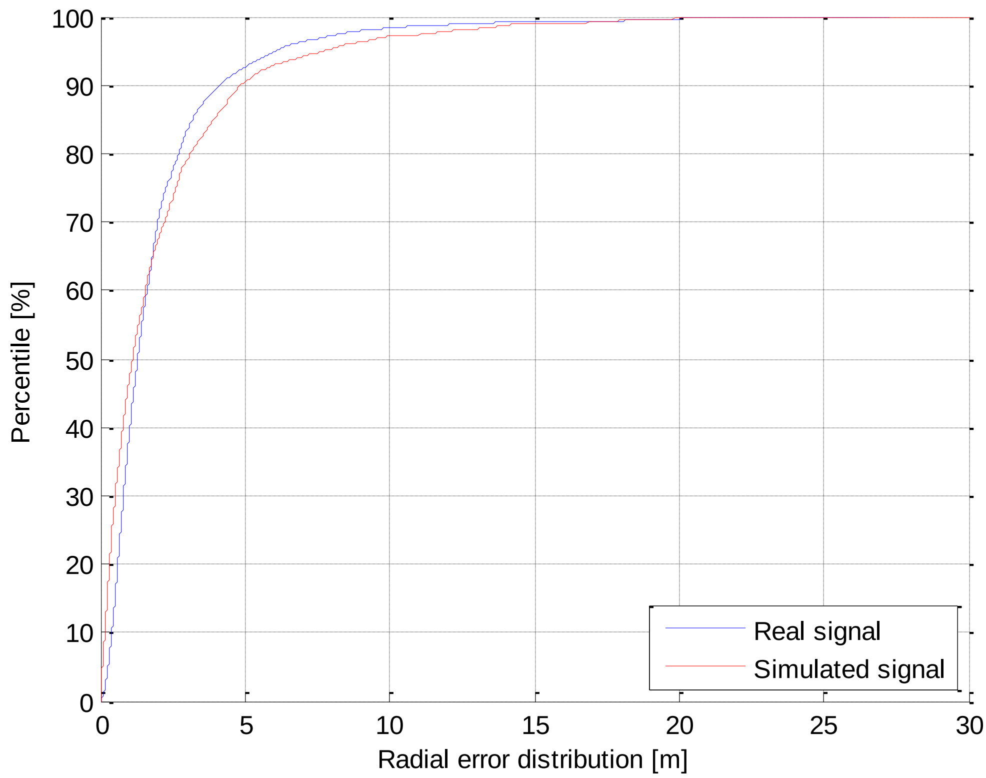

Figure 2 Cumulative distribution for radial error (data from Nantes area)

4. Experimental Results

This paragraph summarizes part of the experimentation process carried out by the author, including relevant results of the work done during the STSM, concerning the three PVT error models described before in this document.
All the algorithms herein included shall be evaluated under different sky conditions, in particular, in open sky and urban canyon areas. We have chosen to proceed this way in order to better evaluate the performance and adaptability of each model to very different situations: in open sky areas the error signal shall be kept within a very low margins (e.g. ±10m), while in urban canyon areas, the undesired

effect of non-line of sight (NLOS) effects, multipath and lack of sky visibility may lead to huge divergences in the trajectory, involving errors up to hundreds of meters.
The metrics considered to evaluate the performance of each method concerns mainly three different aspects: the CDF distribution, the autocorrelation function and the visual aspect of the generated PVT error signal.

4.1 Performance in Open Sky Environment

As a reference data for the evaluation of open sky environment, we have combined error signals corresponding to different scenarios, since there is not enough samples to generate a representative output in a single continuous data set. The combined data set contains error signal recorded in the periphery of the cities of Nantes and Paris (40000 samples). The evolution in time of the error signals (both along-track and cross-track components) is represented in the Figure 3.

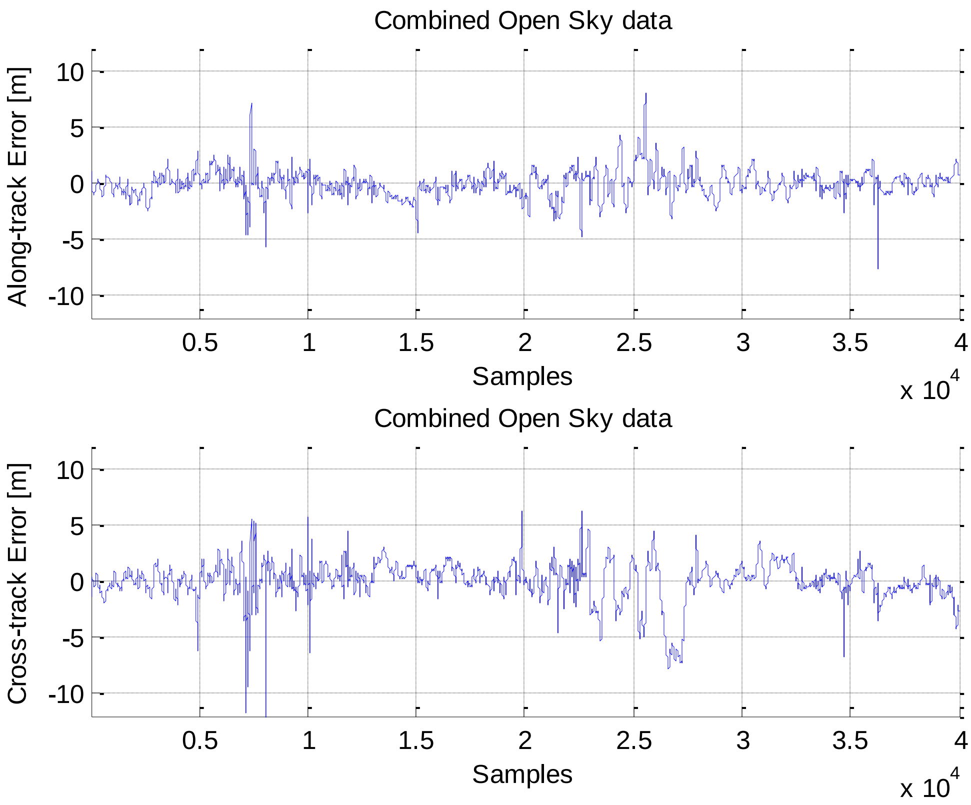

Figure 3 Along-track and cross-track errors in open sky environment

4.1.1 CDF Distribution Analysis

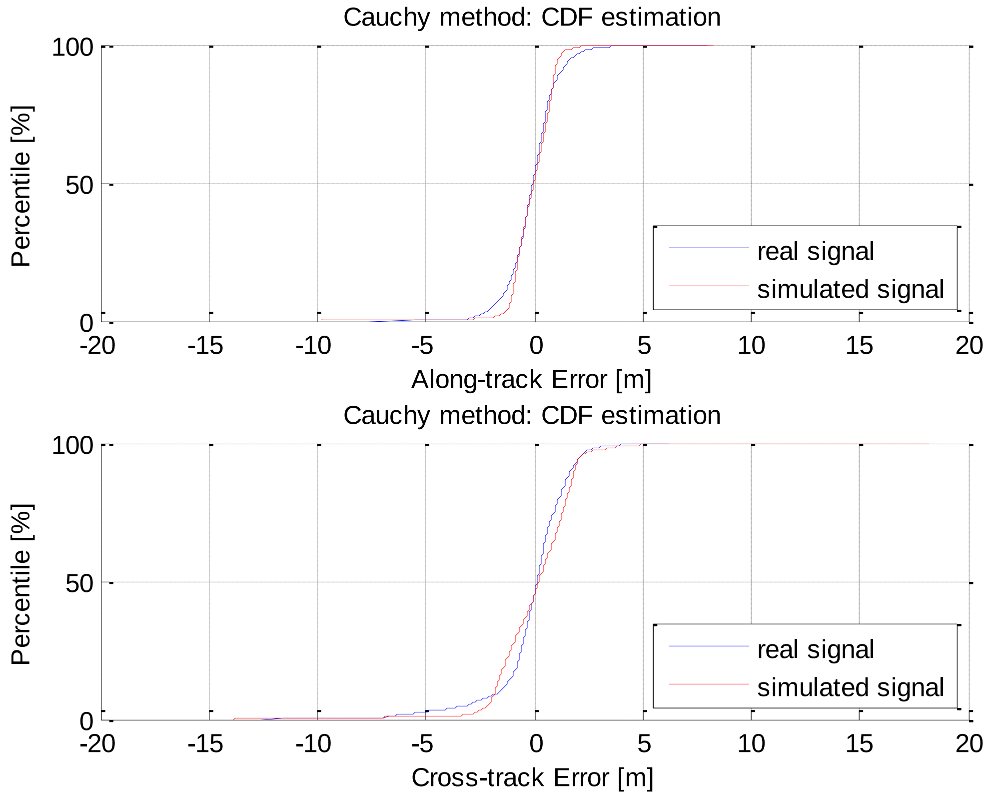

Figure 4 Cumulative distributions of signals in Cauchy method (open sky)

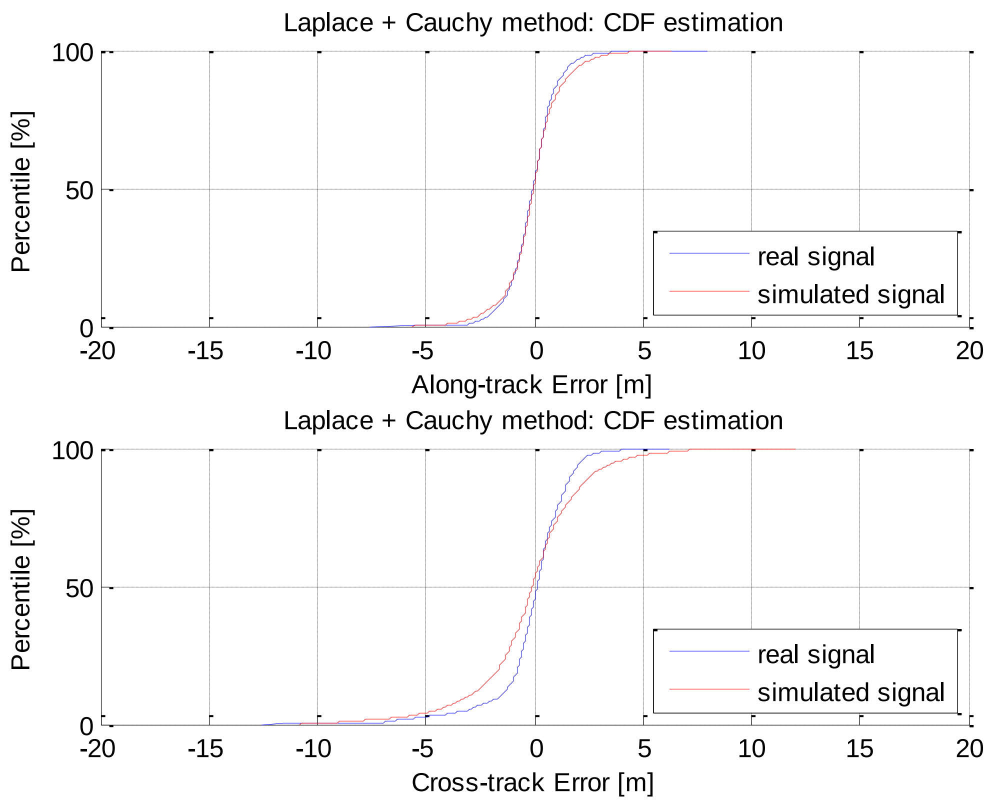

Figure 5 Cumulative distributions of signals in Laplace-Cauchy method (open sky)

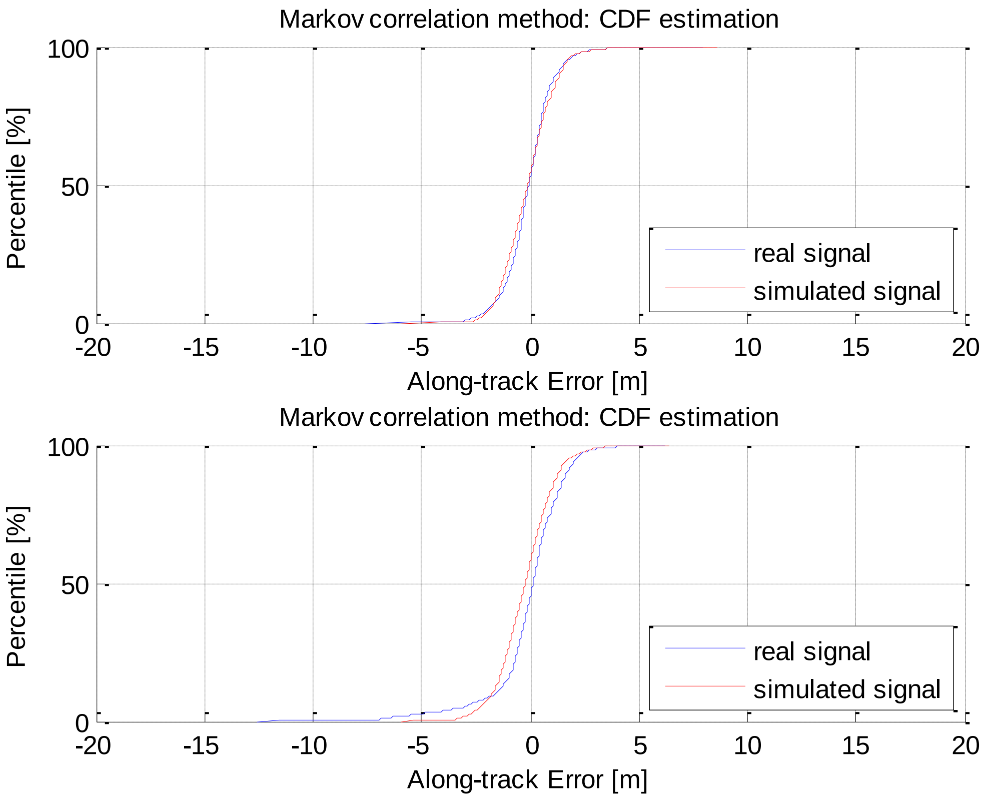

Figure 6 Cumulative distributions of signals in Markov correlation method (open sky)

4.1.2 Autocorrelation Function Analysis

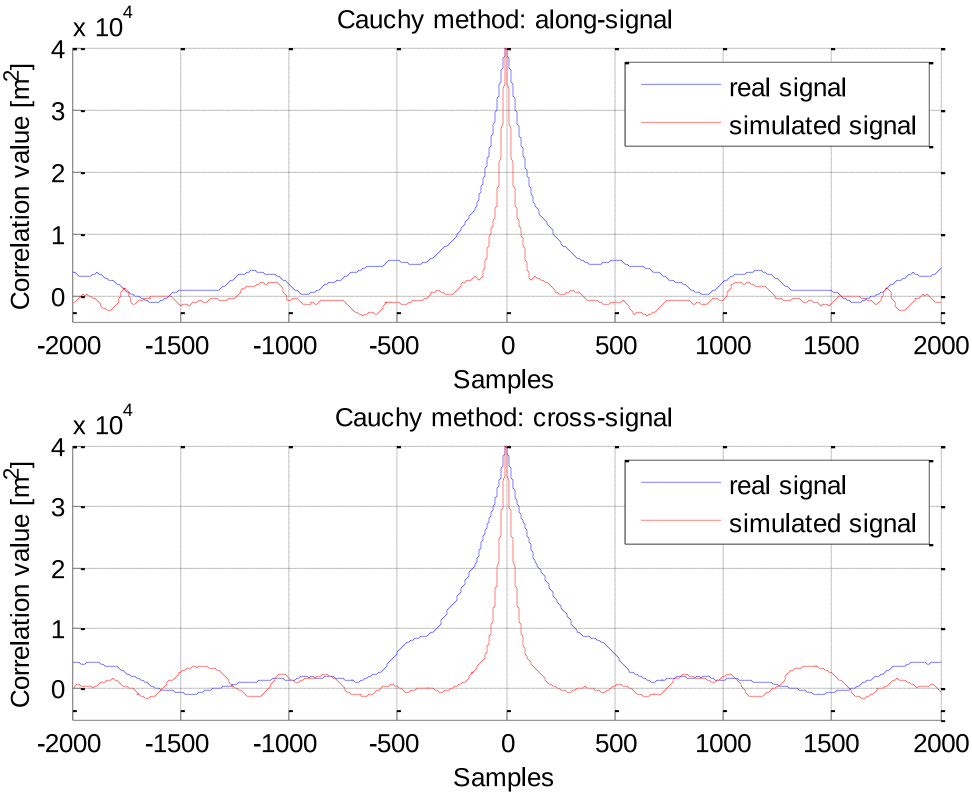

Figure 7 Autocorrelation function of the signals in Cauchy method (open sky)

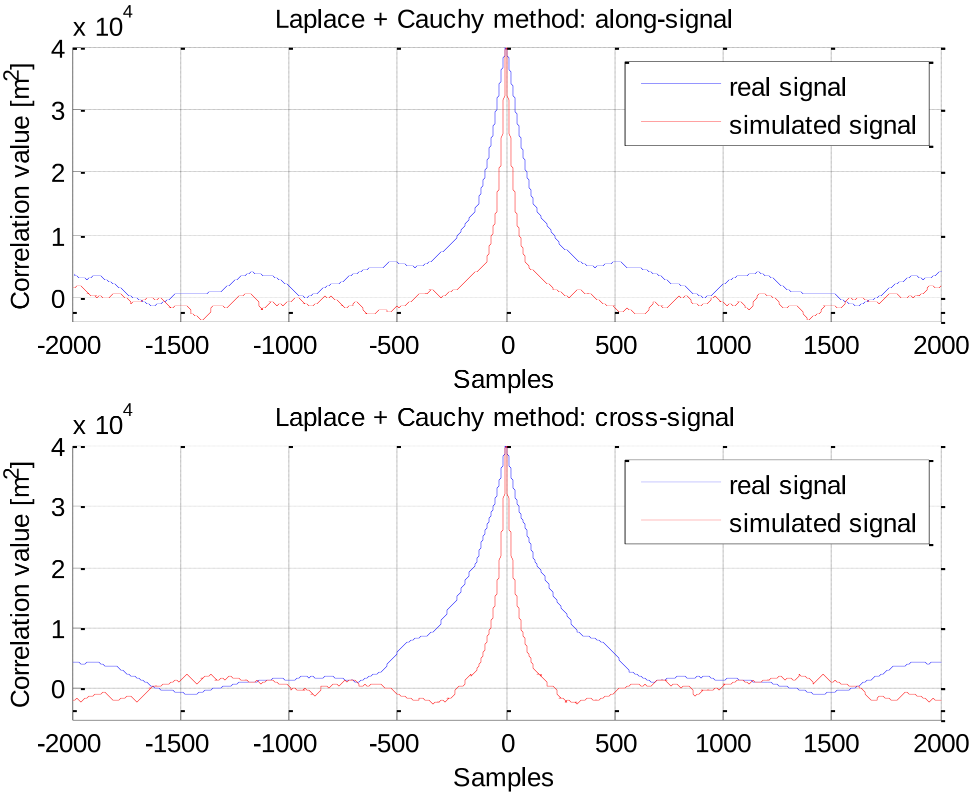

Figure 8 Autocorrelation function of signals in Laplace-Cauchy method (open sky)

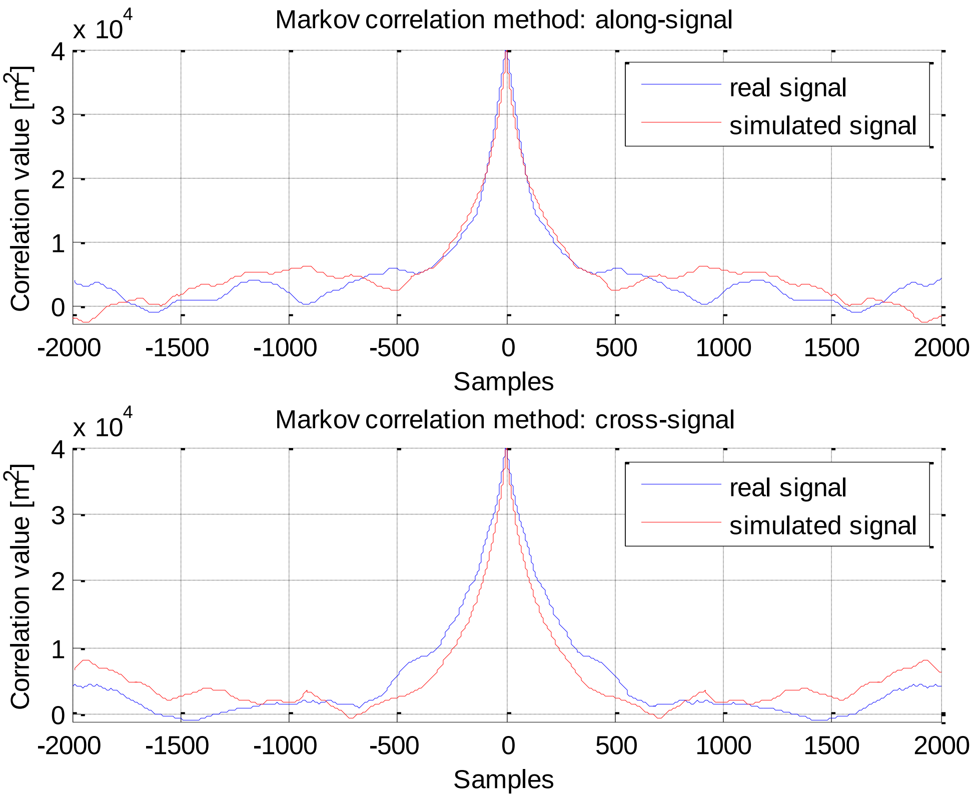

Figure 9 Autocorrelation function of signals in Markov correlation method (open sky)

4.1.3 Visual Aspect of the PVT Error Signal

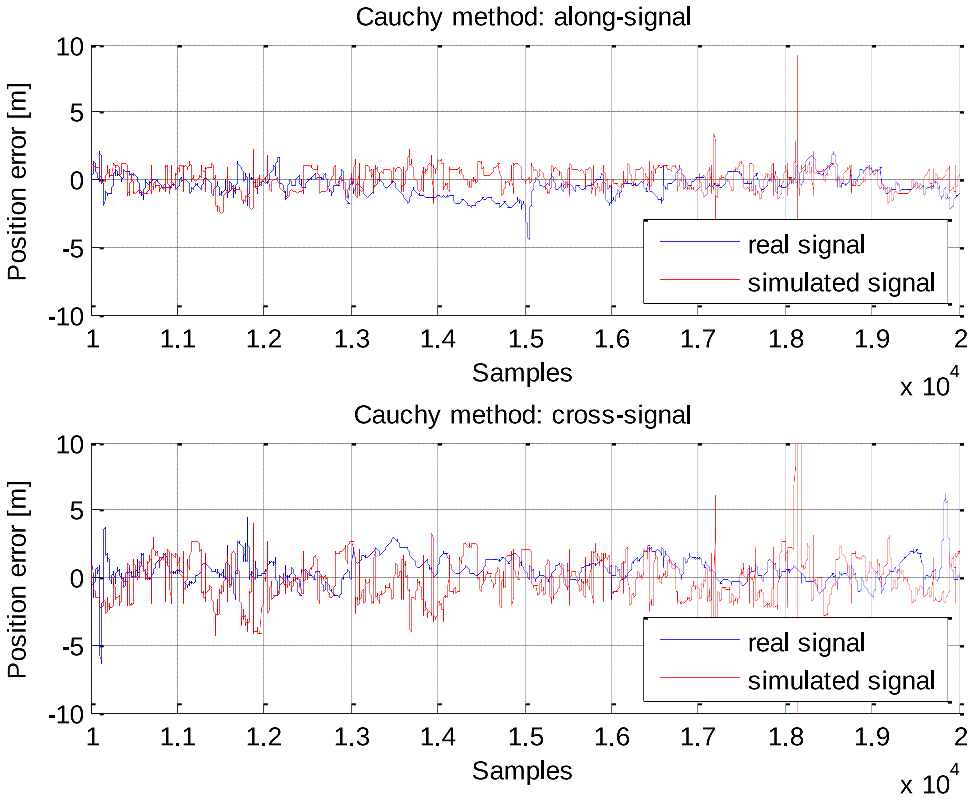

Figure 10 Time evolution of signals with Cauchy method (open sky)

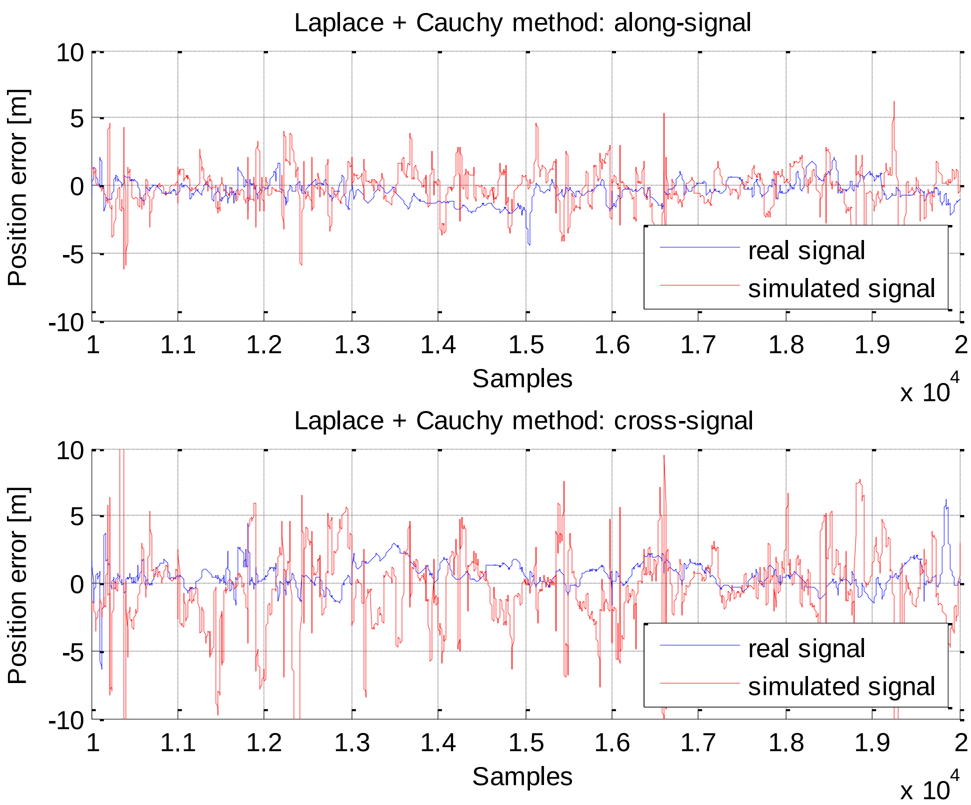

Figure 11 Time evolution of signals with Laplace-Cauchy method (open sky)

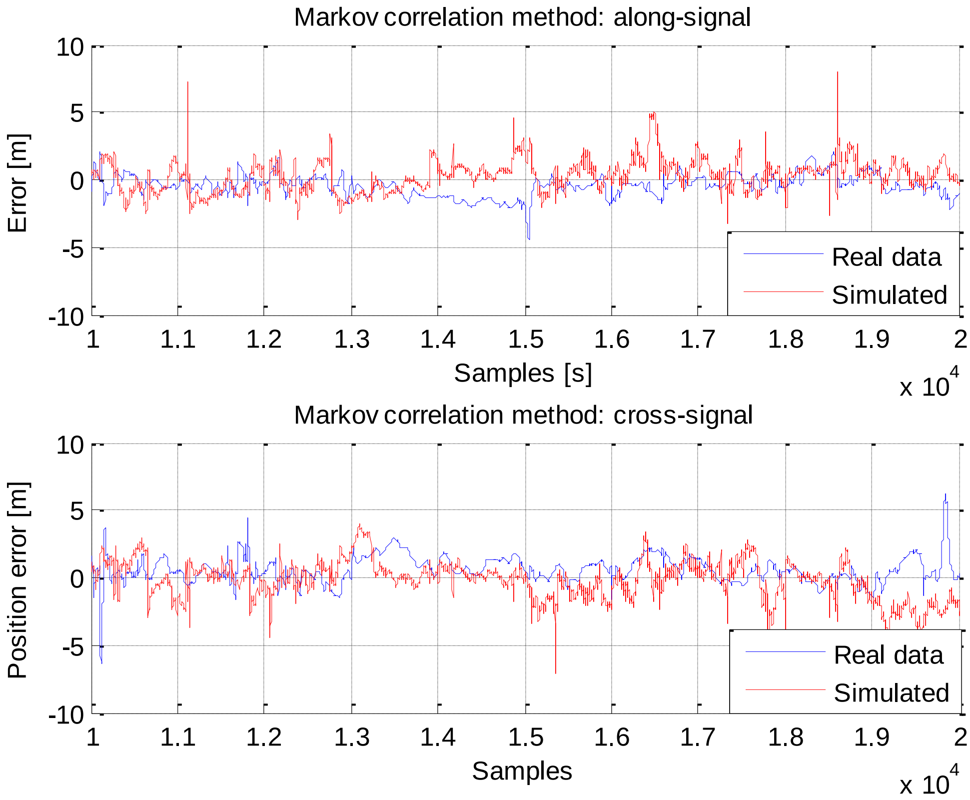

Figure 12 Time evolution of signals with Markov correlation method (open sky)

4.2 Performance in Urban Canyon Environment

As a reference data for the evaluation of urban canyon environment, we have selected the scenarios corresponding to the 29/04/2014, recorded in the Defense and Mairie XII Paris areas with the u-blox LEA-6T receiver (46355 samples in total). The evolution in time of the error signals (both along-track and cross-track components) for each data set is represented in Figure 13 and Figure 14.

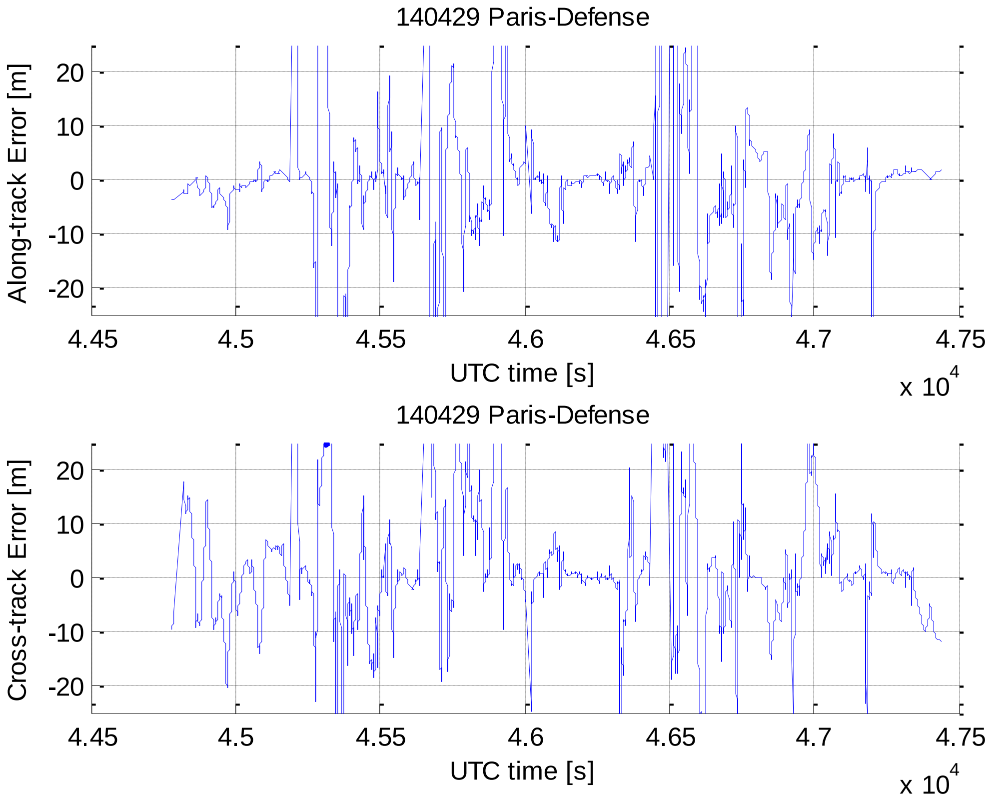

Figure 13 Along-track and cross-track errors in Defense area (Paris)

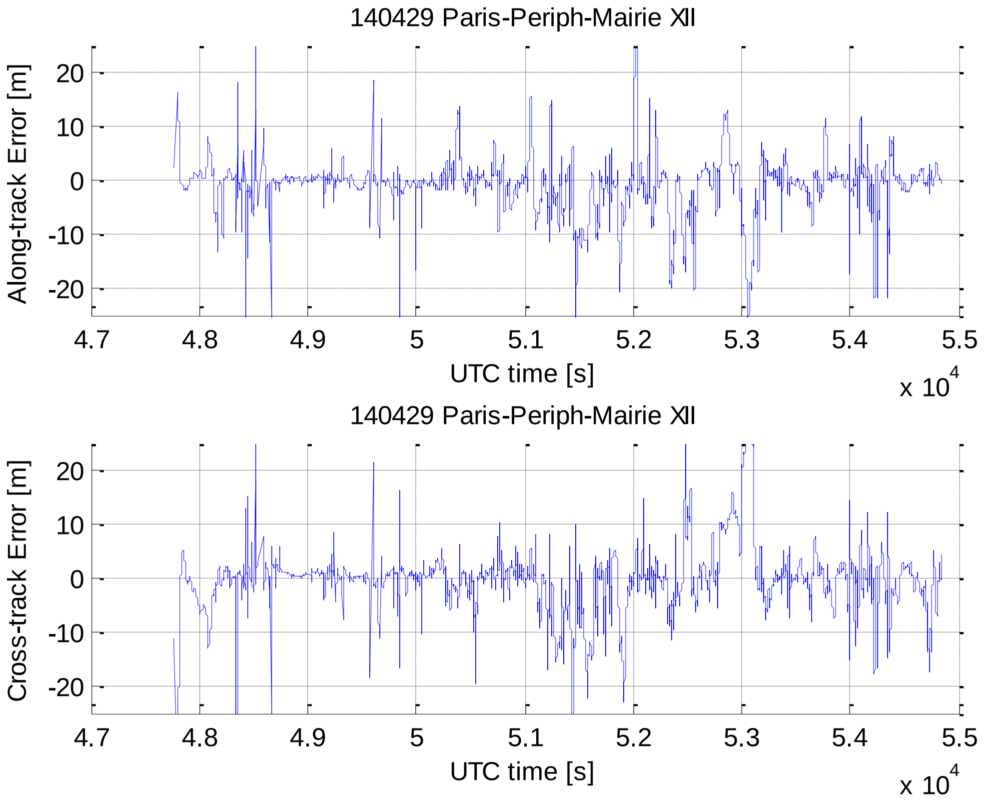

Figure 14 Along-track and cross-track errors in Mairie XII area (Paris)

4.2.1 CDF Distribution Analysis

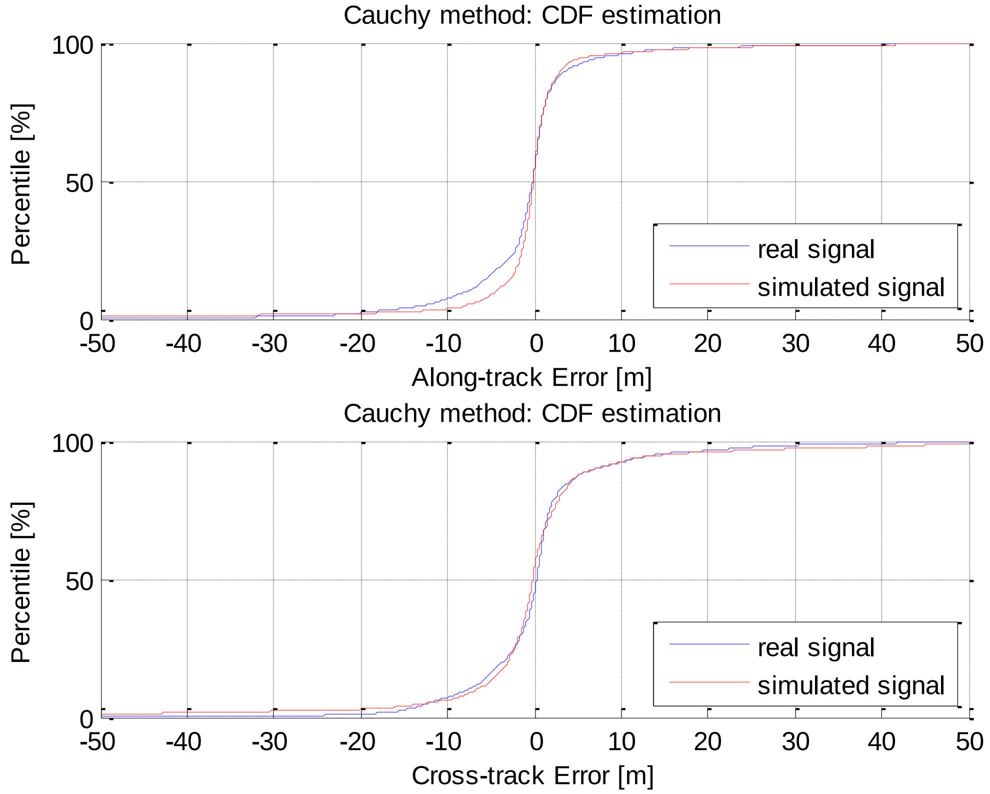

Figure 15 Cumulative distributions of signals in Cauchy method (urban canyon)

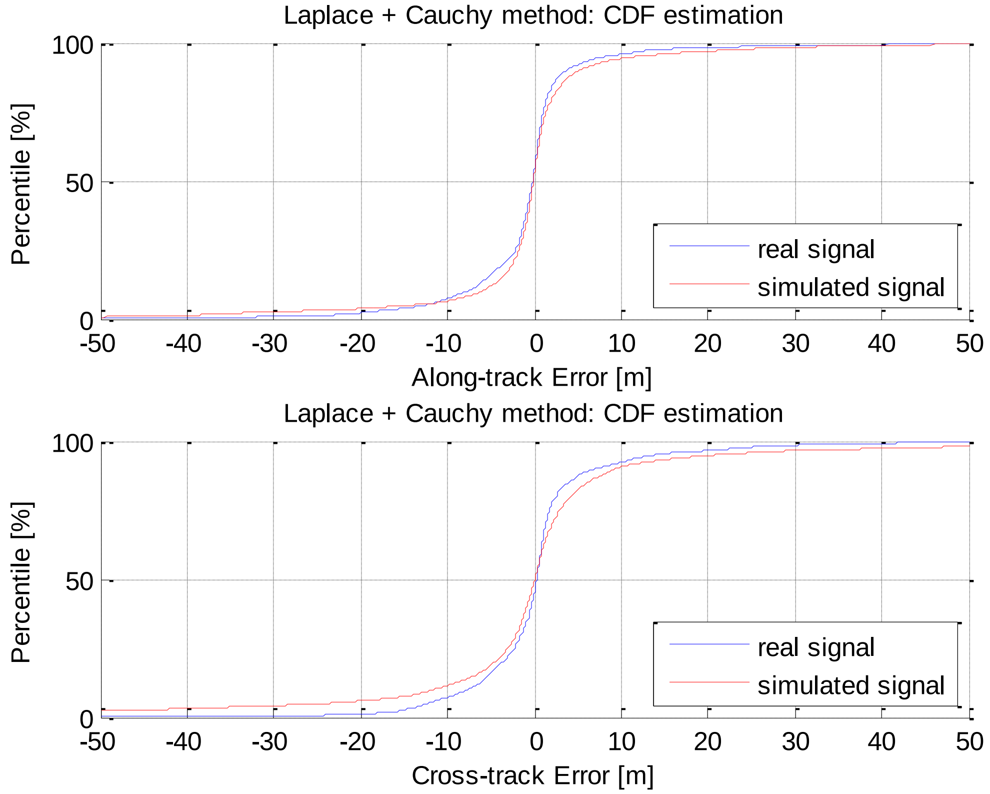

Figure 16 Cumulative distributions of signals in Laplace-Cauchy method (urban canyon)

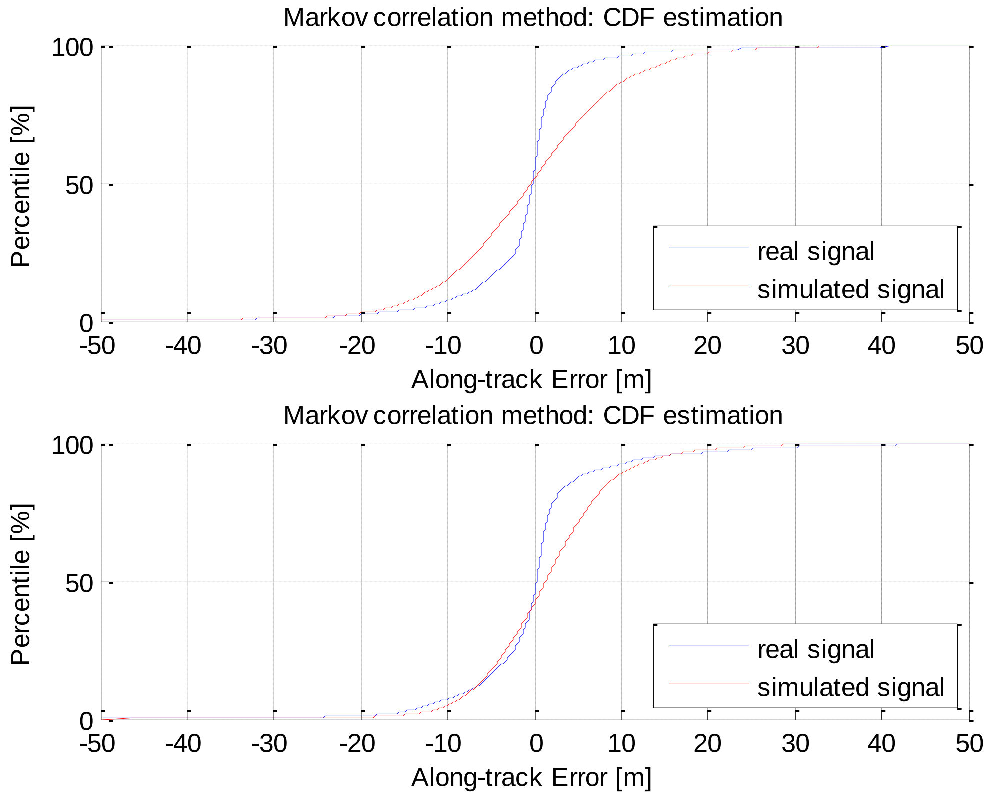

Figure 17 Cumulative distributions of signals in Markov correlation method (urban canyon)

4.2.2 Autocorrelation Function Analysis

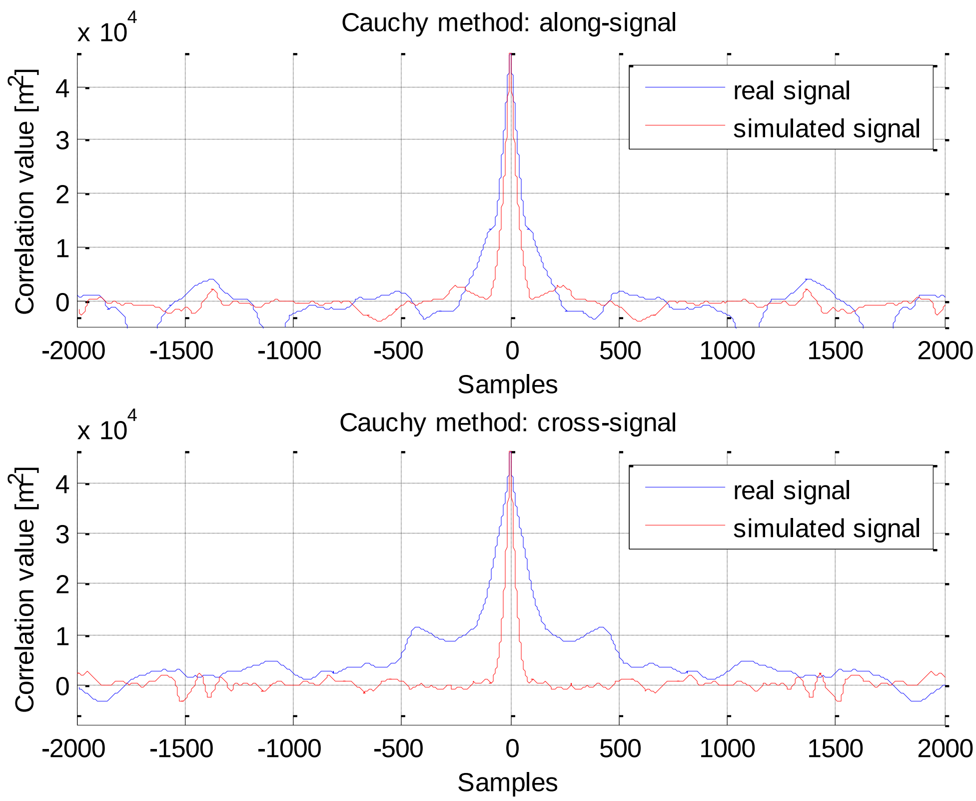

Figure 18 Autocorrelation function of signals in Cauchy method (urban canyon)

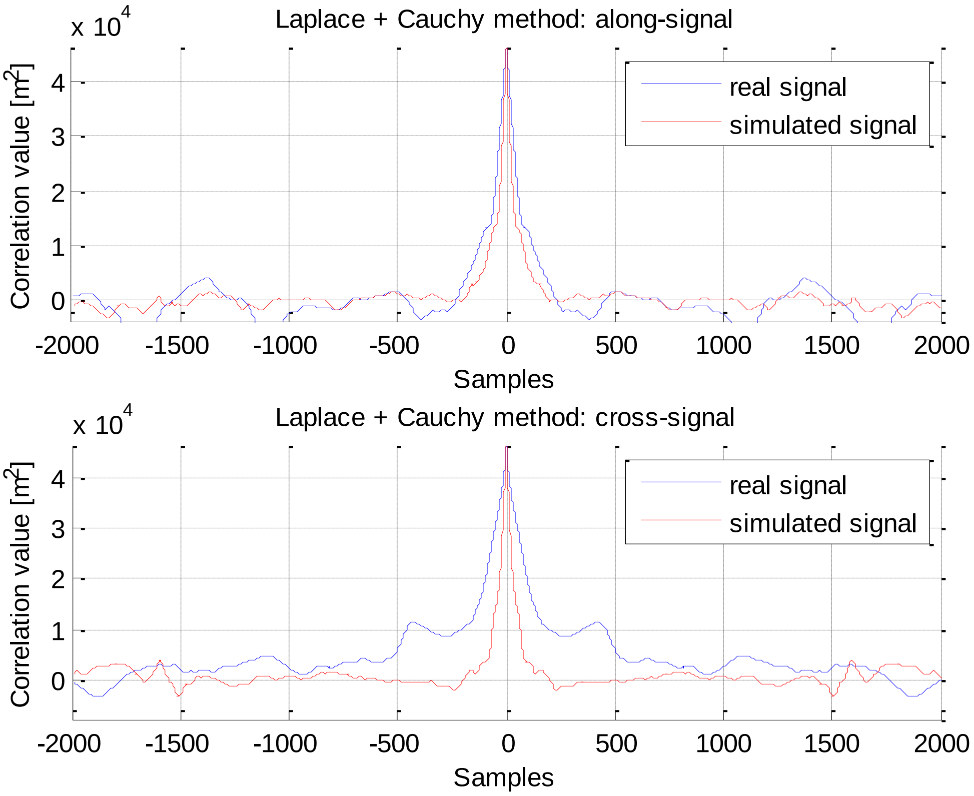

Figure 19 Autocorrelation function of signals in Laplace-Cauchy method (urban canyon)

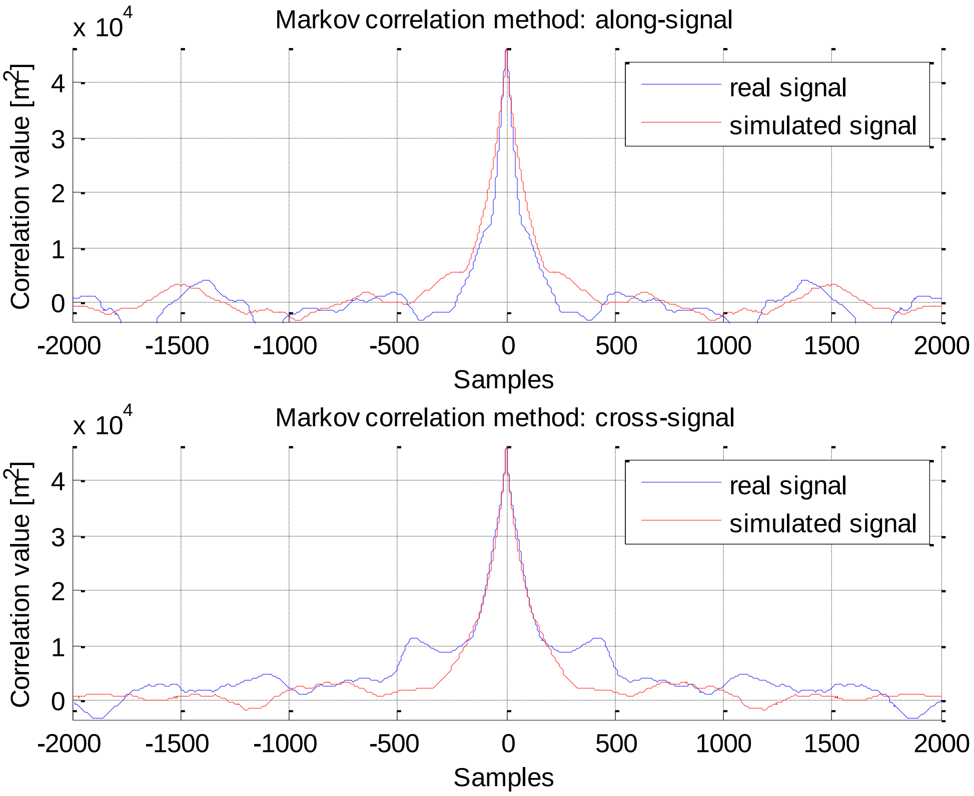

Figure 20 Autocorrelation function of signals in Markov correlation method (urban canyon)

4.2.3 Visual Aspect of the PVT Error Signal

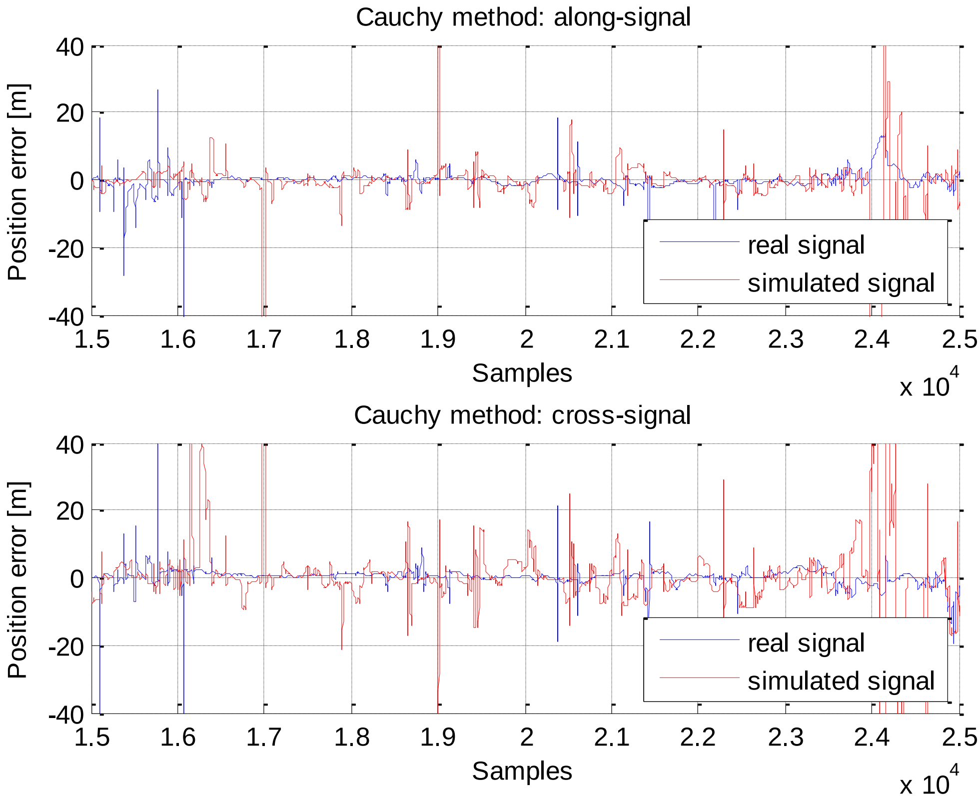

Figure 21 Time evolution of signals with Cauchy method (urban canyon)

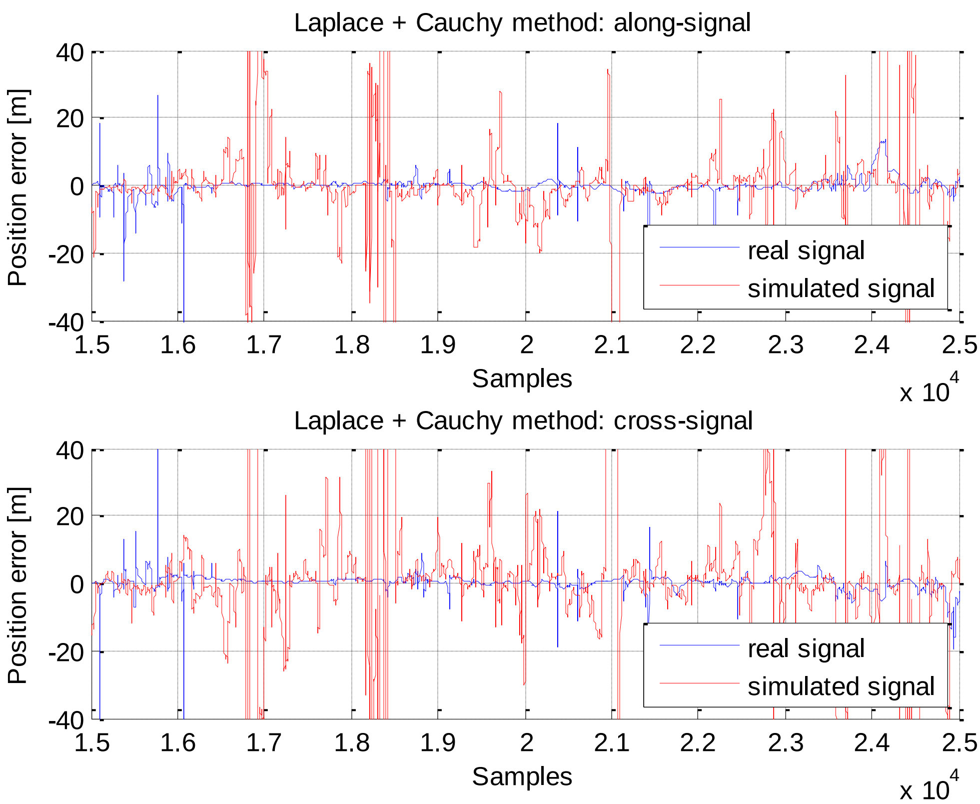

Figure 22 Time evolution of signals with Laplace-Cauchy method (urban canyon)

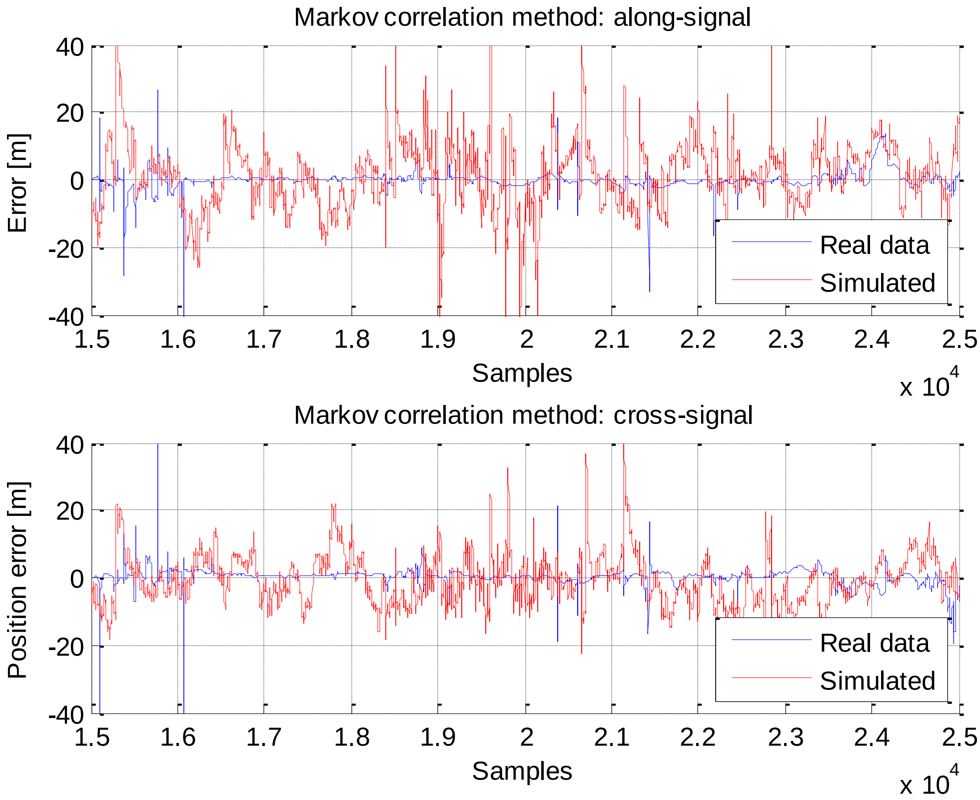

Figure 23 Time evolution of signals with Markov correlation method (urban canyon)

4.3 Analysis of the Results

Several conclusions can be extracted from the experimental results presented in the previous section, concerning the three types of metrics that have been defined (i.e. cumulative distribution function, autocorrelation function and visual aspect of the generated signal).

First, regarding the cumulative distribution function, different behaviors have been observed. In the case of open sky environments (see Figure 4, Figure 5 and Figure 6), we notice how the Markov correlation method (proposed by the author) obtains the best performance, although the CDF estimation is degraded in the case of urban canyon areas (see Figure 15, Figure 16 and Figure 17). In average, the Laplace-Cauchy method seems to be the one that fits better the global statistic, while the Cauchy method is the one that obtains the worse performance.

Second, regarding the autocorrelation function, the Markov correlation method obtains the best performance, achieving similar shapes to those corresponding to the true data (from u-blox LEA-6T). Moreover, it has been tested that the other two proposed methods, Cauchy and Cauchy-Laplace, have difficulties to obtain similar autocorrelation functions to the Markov method, tending to model much more uncorrelated distributions than the real one. This happens both for open sky environments (see Figure 7, Figure 8 and Figure 9) and urban canyon areas (see Figure 18, Figure 19 and Figure 20).

Third, concerning the visual aspect of the generated signal, the Cauchy method seems to produce a more artificial signal than the other two proposed methods, both in open sky (see Figure 10, Figure 11 and Figure 12) and in urban canyon (see Figure 21, Figure 22 and Figure 23).

At the light of the obtained results, and although the results are acceptable for all the algorithms presented herein, the author is inclined to use the Markov correlation method, at least until the model based on the very promising Chopin algorithm is ready. This is so because of several reasons:

- It is the method with the better autocorrelation properties.
- The generated output signals look less artificial than in the other proposed methods.
- The Markov correlation model only requires parameterizing a single parameter, instead of a set of them (more than seven).
- The configuration of the design parameter has not been changed for the generation of the open sky and urban canyon output signals, while for the other two methods have required optimization to fit into the different scenario statistics.

5. Conclusions

At the light of the obtained results, we can conclude that the three methods proposed herein could be considered in a future for the generation of PVT degraded trajectories, in the frame of the SaPPART COST action, although additional work is still required in order to optimize their performance.
The author considers that, after considering the experimental results, the Markov correlation model is the one that obtains best general performance among the others proposed in this memorandum, although it still needs to be improved in its CDF estimation performance. The main reasons to conclude that are:

- Better autocorrelation properties.
- Good visual aspect.
- One configuration valid for both open sky and urban areas.
- Only one parameter to model.

Finally, we remind that, at the time of writing this document, an additional implementation for a PVT error model based on the Chopin algorithm was being developed. This method might overcome all the problems presented in this memorandum and improve the final performance, concerning all the proposed metrics.

6. References

[1] N. Chopin, "Dynamic detection of change points in long time series", in *Annals of* the *Institute of Statistical Mathematics* 59 (2), 349-366, 2007.

[2] Cauchy distribution: http://en.wikipedia.org/wiki/Cauchy_distribution

[3] J. Monsifrot, "Modelling and simulation of position errors for the validation of a map-matching GNSS-based application", IFSTTAR*,* 2014

[4] Laplace distribution: http://en.wikipedia.org/wiki/Laplace_distribution