Paper presented at the

10th Australian International Aerospace Congress

incorporating the

14th National Space Engineering Symposium 2003

29 July – 1 August 2003

Brisbane, Queensland, Australia

Paper number
AIAC 2003–091

Multipath Mitigation by Neural Network for Spacecraft Attitude Determination from GPS Carrier Phase Observables

Lopes, R. V. F (1), Carrara, V. (1), Enderle, W. (2)
(1) Instituto Nacional de Pesquisas Espaciais - INPE, São José dos Campos, Brazil
(2) Queensland University of Technology, Brisbane, Australia

Abstract

A previous neural network algorithm for multipath mitigation on GPS-based attitude determination is adapted to space applications, where the attitude is unknown in the training process. One considers a user spacecraft with at least three GPS antennas and attitude determined from single differences of the GPS carrier phase L1. A recipe inspired on the on-board calibration problem for star sensors is used to cope with the lack of observability that arises when attitude is unknown. The algorithm was tested with simulated data based on a ground experiment and was able to reduce the multipath effect in more than 50%.

Nomenclature

λ	Wavelength of carrier phase L1.
$\boldsymbol{b}_i$	Baseline vector from the master antenna to the i^{th} slave antenna, in the antenna frame.
$\boldsymbol{V}_k^p$	Line of sight unit vector of p^{th} GPS satellite in the antenna frame at k^{th} sample time.
$\boldsymbol{v}_k^p$	Line of sight vector of p^{th} GPS satellite in stereographic coordinates at k^{th} sample time.
$\delta\boldsymbol{v}$	Distortion vector in stereographic coordinates.
$\phi_k^{i,p}$	Single-difference of carrier phase of p^{th} GPS satellite and i^{th} baseline at at k^{th} sample time.
$\bullet^T$	Transpose of a matrix.
$\langle\bullet\rangle_{\text{LPh}}$	Learning Phase sample average.
$\langle\bullet\rangle_{\text{EPh}}$	Evaluation Phase sample average.
$\langle\bullet\rangle_{\text{O}}$	Sample average over the whole antenna field of view.
$\langle\bullet\rangle_{*}$	Sample average over a neighborhood of 5^{o} from the antenna zenith.
$\tilde{\bullet}$	Calibrated value.

1. Introduction

In this paper, one addresses the spacecraft attitude determination problem from GPS interferometry only, envisaging future applications in the Brazilian space program. The aim of the work is to reduce attitude determination errors by means of multipath mitigation. One considers a set of at least three antennas linked to a single GPS receiver.

In a previous work, the authors proposed an algorithm for multipath mitigation using neural network [1]. That algorithm has two operating modes: the calibration mode and the correction mode. In the calibration mode the single difference of carrier phases are processed during a period long enough to assure that the GPS line-of-sights fill the effective field of view of the antenna set. In this mode attitude is supposed to be known and the neural network learns the distortion pattern in the GPS line-of-sight vectors due to multipath, as estimated by the observables in the antenna frame. In the correction mode, the neural network predicts and corrects the multipath effects. As a consequence, attitude determination accuracy is improved.

This paper adapts that algorithm to space applications where attitude is unknown in the calibration mode also. In this case there is a loss of observability of three degrees of freedom due to the fact that part of the multipath effect can be modeled as a three-axis rotation of the antenna frame. The proposed solution is based on the Shuster´s recipe for the polynomial calibration model that was originally stated in the context of on-board star sensor calibration [2] and afterwards applied to on-board multipath mitigation with a polynomial model [3].

The algorithm is tested with simulation data from a LEO Earth pointed satellite, based on ground experiment results. The results are compared with those obtained with a priori attitude knowledge.

2. Attitude Determination from GPS Interferometry and the Need for Multipath Mitigation

The between-antennas single difference of the GPS carrier phase L1 is related with the line of sight of the GPS satellite by the interferometry equation:

$$\lambda \, \phi_{i,k}^{p} = \boldsymbol{b}_i \cdot \boldsymbol{V}_k^{p} + \text{ observation errors,} \qquad (1)$$

with both $\boldsymbol{b}_i$ and $\boldsymbol{V}_k^{p}$ written in the antenna frame.

This set of equations can be solved for the line of sight unit-vector $\boldsymbol{V}_k^{p}$ provided there are at least two non-parallel antenna baselines. The coordinates of the line of sight of a GPS satellite in the reference frame are also known, from the ephemeris of the GPS satellites and the user spacecraft position. In this case, any well-known algorithm that solves the Wahba problem [4] can be used to evaluate the attitude of the user spacecraft. For instance, the QUEST algorithm [5] based on the q-method or the profile matrix algorithm [6] based on single value decomposition. The last one is used in this paper.

Accuracy of attitude determination from GPS observations depends mainly on the baseline lengths and the error present in the observations like random noise, baseline uncertainty (including the antenna phase center), and multipath delay. Baseline length is limited by the spacecraft mechanical architecture, but observation errors that cause apparent distortions on the GPS line of sights can be mitigated. In space applications, the multipath delay is mainly a function of the direction of the line of sight in the antenna frame and present a constant distortion pattern that can be estimated as far as the satellite architecture around the antennas remains unchanged.

Although the distortion vector is a three-dimensional vector, it has only two degrees of freedom, since the line of sights are constrained to be unit vectors. Therefore, a distortion model needs to represent it in planar coordinates. Now, GPS antennas are omni-directional, but the observation errors are stronger in signals with low elevations with respect to the antenna plane than in signals coming from the antenna zenith. For this reason, the orthogonal projection of the distortion vector in the antenna plane is not suitable to the problem because it masks the radial component of distortions close to the antenna plane. The stereographic projection (see Figure 1) avoids this hindrance and is adopted here:

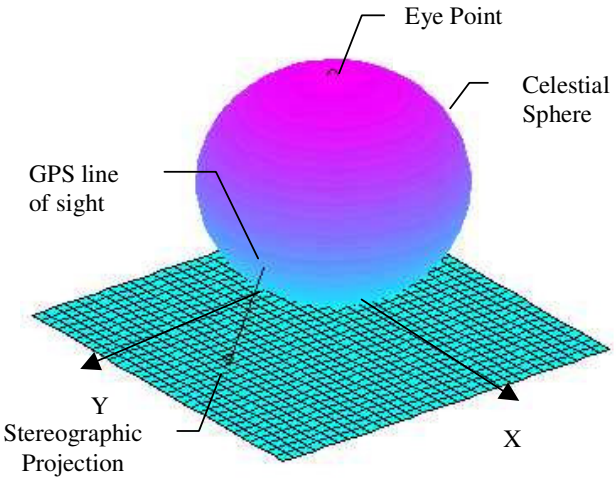

Figure 1. The Stereographic Projection

$$\boldsymbol{v} = \begin{Bmatrix} V_x \\ V_y \end{Bmatrix} \frac{2}{1+V_z} \quad \Leftrightarrow \quad V = \begin{Bmatrix} 4\boldsymbol{v} \\ 4 - \boldsymbol{v}\cdot\boldsymbol{v} \end{Bmatrix} \frac{1}{4+\boldsymbol{v}\cdot\boldsymbol{v}}. \tag{2}$$

3. Neural Networks

A Neural Network is an entity composed of individual processing units called neurons grouped in layers. The network is structured in such a way that the output of one neuron is the input of one or more neurons. If there is at least one closed path in the neuron's connections, meaning that the input of a given neuron comes from the combination of its output and the output of other neurons, then the network is said to be recurrent. In case that there is no such closed path, then the neural network is called a feedforward network (Figure 2). Feedfoward networks have, in general, one input layer composed of one neuron for each network input and an output layer, which also presents one neuron for each network output. One or more hidden layers separate the network input from its output. In such feedforward network the neuron in a given layer applies each an activation function *f* to the sum of the weighted outputs of the previous layer. For hidden and input layers, the activation function generally is a biased nonlinear differentiable function like the sigmoid or hyperbolic tangent, for instance, while the output layer can be a linear function [7]. The process to obtaining the network weights is called supervised training, once the computational method calculates the weights increment at each step based on the error generated by the network output (compared with the expected values) and in some optimization rule. Training consists of an interactive process in which the weights are adjusted by propagating the output error through the network layers. Nonlinear continuous functions can be approximated with a given accuracy by a 2 layer neural net with linear function in the output and the sigmoid activation function in the input layer [8,9]:

$$f(x) = \frac{1-e^{-x}}{1+e^{-x}} \tag{3}$$

A feedforward network composed by l layers, as shown in Figure 2, can be seen as a mapping function with n_0 input elements and n_l output parameters. If x_i^k is the output of the i^{th} neuron of layer k, w_{ij}^k is the weight of the j^{th} input (coming from the j^{th} neuron of the preceding layer) and f^k is the activation function of layer k, then:

$$x_i^k = f^k\left(\bar{x}_i^k + b_i^k\right) = f^k\left(\sum_{j=1}^{n_{k-1}} w_{ij}^k\, x_j^{k-1} + b_i^k\right), \tag{4}$$

where b_i^k is the neuron bias that allows the neuron to present a non-null output for a null input.

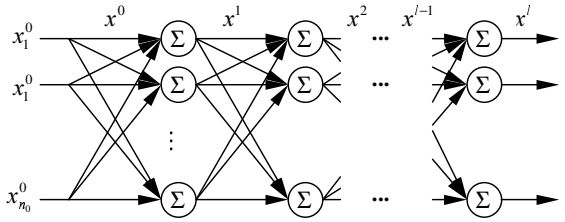

Figure 2. A feedforward neural network

Generally the neuron bias and the weights can be obtained at the same time, by assuming the inclusion of new unit input for the bias. In a vector-matrix representation form the preceding equation yields

$$\boldsymbol{x}^k = \boldsymbol{f}^k\left(\bar{\boldsymbol{x}}^k\right) = \boldsymbol{f}^k\left(\boldsymbol{W}^k \boldsymbol{x}^{k-1}\right) \quad , \tag{5}$$

where the weight matrix of layer k, $\boldsymbol{W}^k$, includes the neuron bias:

$$\boldsymbol{W}^k = \begin{bmatrix} w_{11}^k & \cdots & w_{1n_{k-1}}^k & b_1^k \\ w_{21}^k & \cdots & w_{2n_{k-1}}^k & b_2^k \\ \vdots & & \vdots & \vdots \\ w_{n_k 1}^k & \cdots & w_{n_k n_{k-1}}^k & b_{n_k 1}^k \end{bmatrix}. \tag{6}$$

The dimensions of the output vector $\boldsymbol{x}^k$ and the weight matrix $\boldsymbol{W}^k$ are now n_k+1 and $n_k \times n_{k-1}+1$, respectively.

The increasing number of neurons or hidden layers normally makes the neural net to better approximate the target data. Nevertheless, the capacity of generalization, i. e. the ability to interpolate between points where the neural net was not trained is more accentuated on nets with few neurons or hidden layer [10]. The number of neurons in the hidden layers is important under the approximation point of view: few neurons tend to decrease the stability and result bad approximation, too much neurons cause oscillation on the output between the trained points [11].

The training process normally minimizes the output error through the application of an optimization method, like the steepest descent. These methods need to know with some extent how the net output varies with respect to a given neuron weight. This can be achieved with the backpropagation algorithm developed by Werbos [7] which obtains the partial derivative of the output elements in a recursive way. In matrix form, the derivative of the output vector with respect to the j^{th} weight of the i^{th} neuron of the k^{th} layer results the expression:

$$\frac{\partial x^l}{\partial w_{ij}^k} = \Delta^k \begin{bmatrix} 0 & \cdots & x_j^{k-1} & \cdots & 0 \end{bmatrix}^T, \tag{7}$$

where Δ^k is the backpropagation matrix, obtained from:

$$\Delta^k = \Delta^{k+1} \boldsymbol{W}^{k+1} \boldsymbol{F}^k, \tag{8}$$

with initial condition $\Delta^l = \boldsymbol{F}^l$, where $\boldsymbol{F}^k$ is a diagonal matrix with the derivatives of the activation function f^k:

$$\boldsymbol{F}^k = \frac{d\boldsymbol{f}^k(\bar{\boldsymbol{x}}^k)}{d\bar{\boldsymbol{x}}^k} = \begin{bmatrix} f^{k\prime}(\bar{x}_1^k) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & f^{k\prime}(\bar{x}_{n_k}^k) \end{bmatrix}. \tag{9}$$

The steepest descent method (sometimes misunderstood with the backpropagation algorithm), shows, as the backpropagation, a high degree of parallelism and simplicity. The weights are corrected based on the minimization of the neural net output error. Weight updating starts at the net output layer and then the error is backpropagated to the preceding layer in order to compute its weight corrections. The minimization criterion uses the network output quadratic error as the performance index:

$$J(t) = \frac{1}{2} \boldsymbol{\varepsilon}(t)^{\mathrm{T}} \, \boldsymbol{\varepsilon}(t), \tag{10}$$

where $\boldsymbol{\varepsilon}(t)$ is the network output error at time t, defined by:

$$\boldsymbol{\varepsilon}(t) = \boldsymbol{y}^d(t) - \boldsymbol{y}(t), \tag{11}$$

where $\boldsymbol{y}^d(t)$ and $\boldsymbol{y}(t)$ are expected and actual network output. Weight updates of layer k are performed using:

$$\boldsymbol{W}^k(t+1) = \boldsymbol{W}^k(t) - \alpha \nabla J^k , \tag{12}$$

where the gradient of the square backpropagated error ∇J^k comes from the backpropagation matrix:

$$\nabla J^k = -\Delta^{k^T} \boldsymbol{\varepsilon} \boldsymbol{x}^{k-1\mathrm{T}} . \tag{13}$$

Convergence of the weights depends on the adjusting of the learning rate coefficient α ranging from 0 to 1. Small values cause the learning process to be too slowly, while high gain tends to unstable the weight updates.

Steepest descent or gradient method is today the most common training procedure. It is ease to implement in computers, it is very fast but it converges in a strongly slow rate. This certainly is the main reason of the extremely long training times in most neural net applications. Nevertheless, there are other training methods that show improved learning speed, as the least square [12] or the Levenberg-Marquardt algorithm [13].

The neural net error after the training process depends strongly of the training set. Care must be taken in choosing the training data in order to evenly fulfill the input-output state space, so the weights can be adjusted globally. Unfortunately it is difficult, not to say impossible, to identify such a global training set in some problems [14] as the one presented here. On the other side, at least in theory, the training set do not need to be very large, by taking into consideration the network ability in acquiring enough system information by generalizing and interpolating the input data.

The neural network implemented in this paper has one hidden layer and activating function of hyperbolic tangent type. The training method is the steepest descent with an adaptive α value, as follows. One starts with α=1. Each time a step increases the performance index, it is canceled and α is divided by 2. If a step decreases the performance index, in the next step α is doubled, except if it is already 1.

4. The Multipath Mitigation Algorithm and On-Board Application Aspects

The proposed algorithm to mitigate the multipath delay effect has two modes. The first one is the calibration mode that takes advantage from the spacecraft relative motion with respect to the GPS constellation to store GPS line of sights coming from all possible directions during a period called learning phase. The stored observations are then compared with their respective predicted values. The stereographic coordinates of the residuals are the observed distortions. The neural network is then trained to learn the observed distortion pattern.

Once the calibration mode is finished the algorithm enter in the correction mode where the distortions on the observed line of sights are predicted by the neural network and subtracted from them before the attitude evaluation.

If the spacecraft attitude is accurately known by other means than the GPS observations during the learning phase, the line of sight observations can be easily predicted from its coordinates in the reference frame. In this case the algorithm works in the aided calibration mode. Otherwise, the attitude matrix must be first estimated at every time based on the distorted observations that one is trying to calibrate, on an iterative algorithm. In this case, the algorithm works in the self-calibration mode. It starts estimating the attitude with the non-calibrated observations. Then the neural network is trained with the obtained residuals. Once it learns the distortion pattern, the observations are corrected and the attitude is refined. The process cycle repeats until the solution stabilizes.

5. The Shuster's Recipe

There is a subtle question related with the possibility of the algorithm to diverge in the self-calibration mode. Indeed it may be shown that the problem of observability arises and may cause divergence. The problem is due to the fact that usually in space applications, one cannot detect post-launch misalignment of the antenna frame with respect to the spacecraft. It is indistinguishable from a rotation of the spacecraft as a whole. Therefore, the algorithm is unable to determine correctly a distortion pattern equivalent to a misalignment due to an intrinsic lack of observability of three degrees of freedom.

The way to cope with this question is the Shuster's recipe adapted to the multipath mitigation problem: the distortion part equivalent to a misalignment must vanish around the antenna zenith. The recipe is implemented as follows:

a) In the self-calibration mode, each time the neural network learns a given distortion pattern, one evaluate the distortions in a circle with radius 5°;

b) The equivalent misalignment vector $\boldsymbol{\theta}$ in the antenna zenith is given by:

$$\boldsymbol{\theta}_x = -\left\langle \delta \boldsymbol{v}_y \right\rangle_* , \quad \boldsymbol{\theta}_y = \left\langle \delta \boldsymbol{v}_x \right\rangle_* , \quad \boldsymbol{\theta}_z = \left\langle \boldsymbol{v}_x \delta \boldsymbol{v}_y - \boldsymbol{v}_y \delta \boldsymbol{v}_x \right\rangle_* \left\langle \left| \boldsymbol{v} \right|^2 \right\rangle_*^{-1} ; \tag{14}$$

c) Before attitude is refined by means of the corrected line of sights, they are all rotated by the negative of the equivalent misalignment vector.

6. Numerical Results

The algorithm was tested with digital simulation of a LEO Earth pointed satellite in a Sun synchronous orbit. The satellite was equipped with four GPS antennas on a plane normal to the satellite yaw axis, in the corners of a square with side 0.80m long.

In the learning phase, the satellite attitude was such that the x-axis of the antenna frame was in the roll direction and the y-axis in the pitch direction. In the evaluation phase the satellite was rotated by 90 degrees around the yaw axis in order to emphasize that the observed line of sights then came from other directions than those observed in the learning phase. Figure 3 shows the path of the GPS satellites with respect to the antenna frame in stereographic coordinates in both learning phase and evaluation phase, respectively, while Table 1 presents the input parameters of the simulation. The evolution of the number of GPS satellites at sight can be seen in Figure 4.

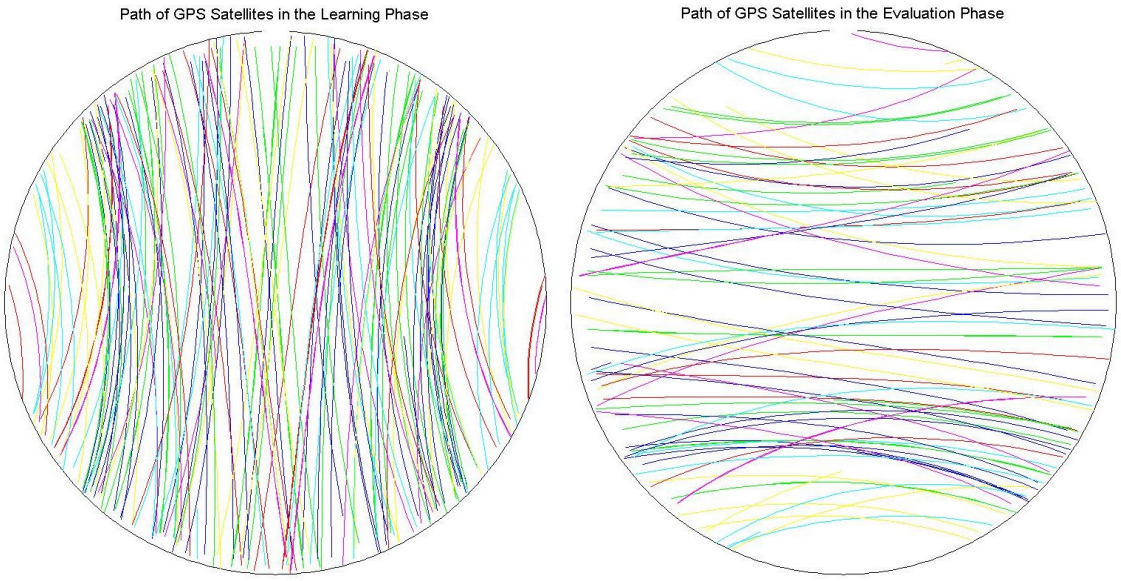

Figure 3. Path of GPS satellites over the field of view of the GPS antennas during: a) the learning phase (12 hours); b) the evaluation phase (6 hours).

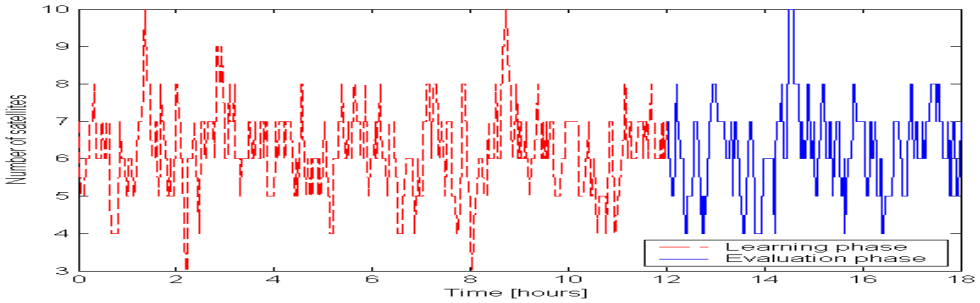

Figure 4. Number of GPS satellites at sight during the simulation

Table 1. Simulation parameters

Orbit altitude:	650 km
Orbit inclination:	98°
Attitude:	Earth oriented, antenna frame facing zenith
Number and geometry of GPS antennas:	4 antennas in the corners of a square with side 0.80m long
Range of noise level on L1 carrier phase:	0.35mm at the antenna zenith; 3.5mm at antenna plane.
RMS of multipath delay:	6mm
Antenna mask angle:	15°
Sampling interval:	60s
Sample period in the learning phase:	12 hours
Sample period in the evaluation phase:	6 hours
Number of neurons in the hidden layer:	60

An empirical model based on real data taken at DLR on a turntable, under the effect of strong, intentional multipath, was used to simulate the multipath. In the experiment, two antennas formed a baseline one meter long, where two metal plates were inserted, one in between the antennas and the other outwards (see Figure 5). As the turntable rotated around the local vertical, the multipath delay was observed from all azimuth angles and from eight different elevations corresponding to the elevation of the GPS satellites at sight at that time, as exemplified in Figure 6.

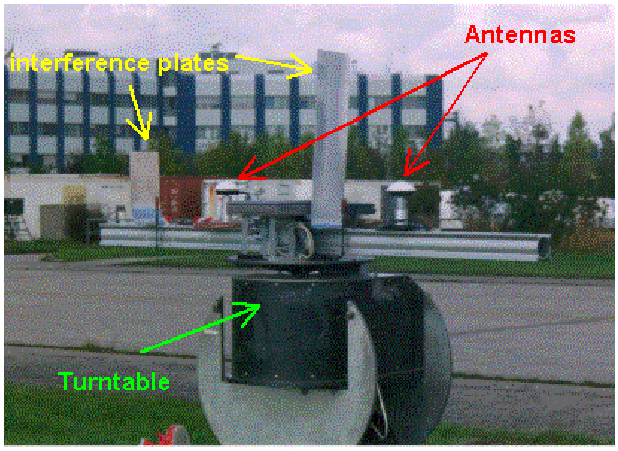

Figure 5. Two GPS antennas on a turntable with the interference of two metal plates. DLR, 1999.

A series of surface spherical harmonics of 30th order was then fitted to those data and used as a model to simulate the multipath delay on the master antenna of the satellite. This empiric model was also used to simulate the multipath delay of the three slave antennas, with the same zonal and tesseral amplitudes of the first one, but all randomly phased. The results can be seen in Figure 7 in stereographic coordinates in the antenna frame.

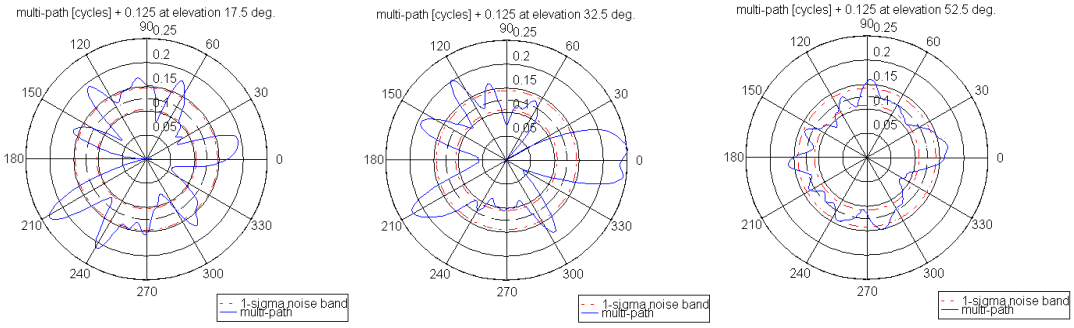

Figure 6. Multipath delay in L1-cycles from some representative elevations: 17.5°; 32.5°; and 52.5°.

When the line of sight of a GPS satellite is determined from the single differences of carrier phase under the effect of such multipath delays, they present distortions with different amplitudes and in different directions. The distortion pattern is quite irregular and more intense close to the edges of the antenna field of view where the metal plates were placed in the experiment. Figure 8a shows the pattern of the distortion arrows in stereographic coordinates in the antenna frame, the contour lines

corresponding to the different levels of distortion intensity. Figure 8b shows the same pattern as learned by the neural network.

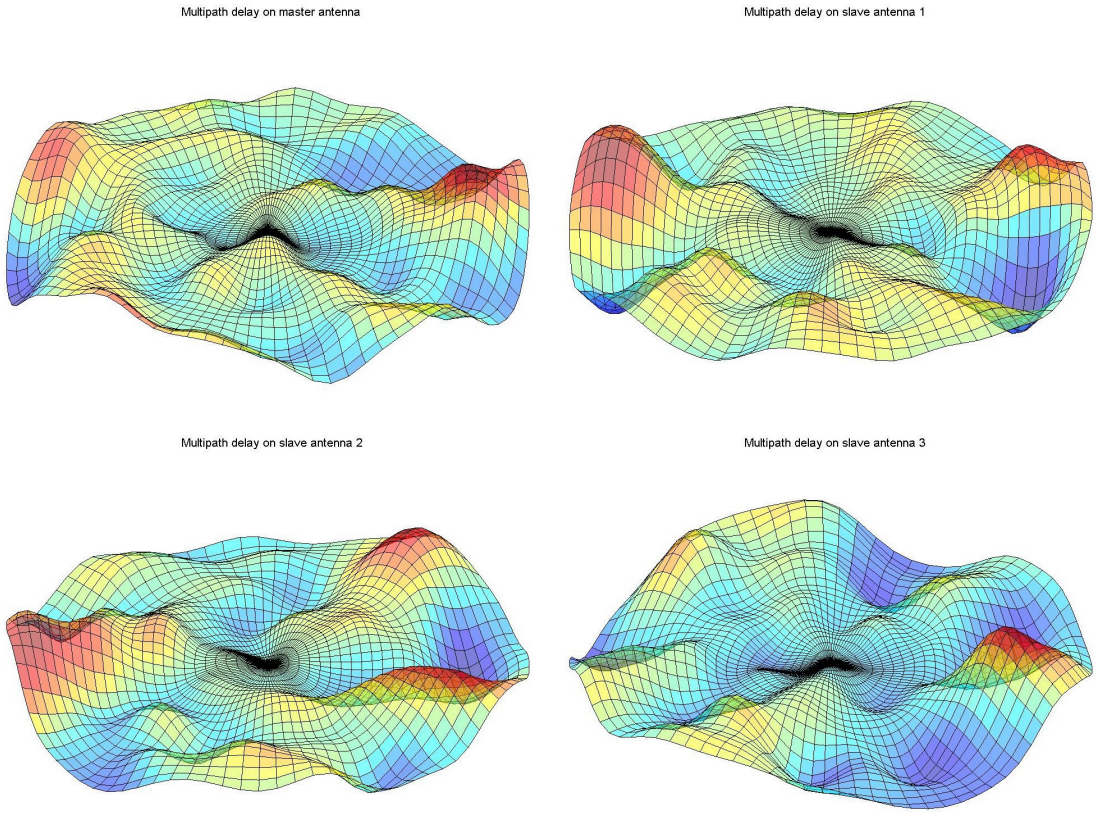

Figure 7. Multipath delay pattern in stereographic coordinates in the antenna frame for the master antenna (a) and the slave antennas (b), (c) and (d).

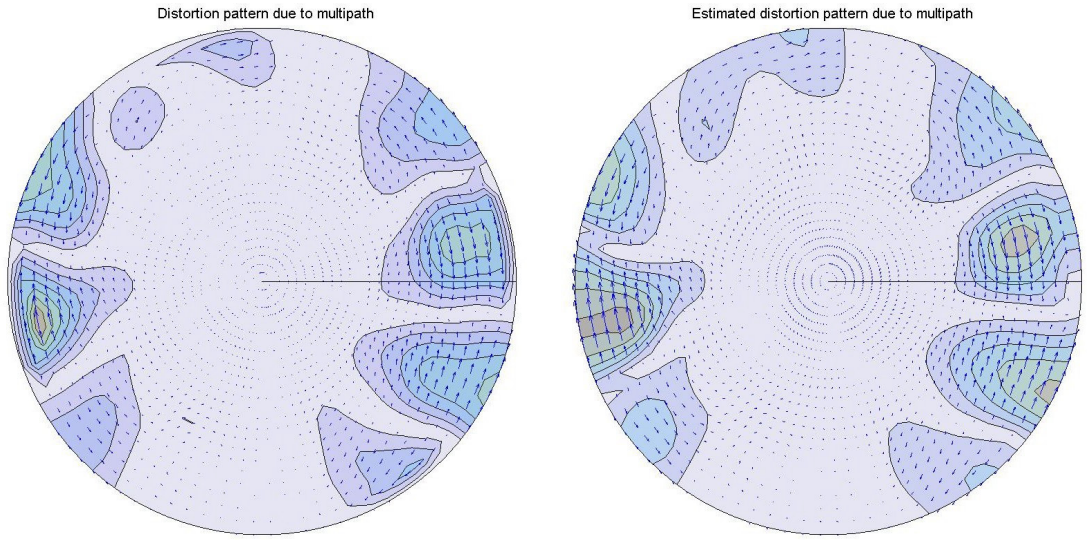

Figure 8. The distortion pattern due to multipath, in stereographic coordinates in the antenna frame: a) before calibration; b) after calibration.

Figures 9a and 9b show respectively the effect of the sample period and the number of neurons in the hidden layer of the neural network on the algorithm performance. The efficiency of distortion prediction η is defined here as the ratio between the accuracy gain in the evaluation phase and the learning

phase in terms of the direction of the observed line of sights. The distortion reduction factor γ is the accuracy gain for hypothetic noise-free observations of the line of sights from the whole antenna field of view. The performance parameters η and γ are respectively given by:

$$\eta \equiv \left(\frac{\left\langle \left| \delta \tilde{V}_k^p \right|^2 \right\rangle_{\mathrm{EPh}}}{\left\langle \left| \delta V_k^p \right|^2 \right\rangle_{\mathrm{EPh}}} \frac{\left\langle \left| \delta V_k^p \right|^2 \right\rangle_{\mathrm{LPh}}}{\left\langle \left| \delta \tilde{V}_k^p \right|^2 \right\rangle_{\mathrm{LPh}}} \right)^{1/2}, \quad \gamma \equiv \left(\frac{\left\langle \left| \delta \tilde{V} \right|^2 \right\rangle_{\mathrm{O}}}{\left\langle \left| \delta V \right|^2 \right\rangle_{\mathrm{O}}} \right)^{1/2}. \tag{15}$$

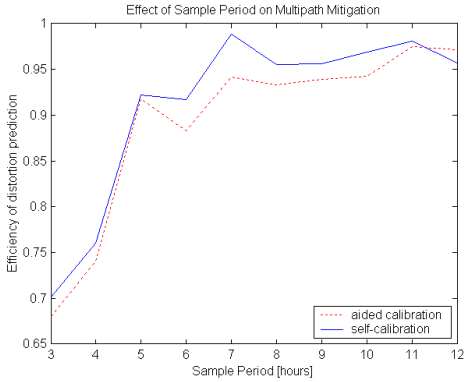

Figure 9a. Effect of the sample period on multipath mitigation.

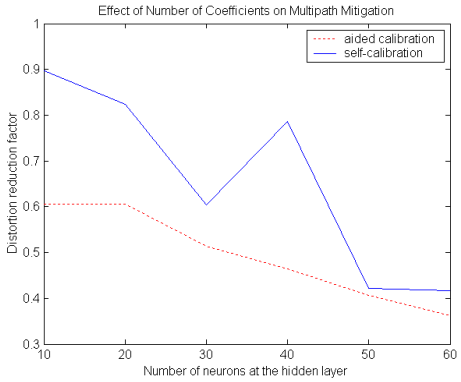

Figure 9b. Effect of the number of neurons on multipath mitigation.

From those results, one may see that if the learning phase is too short, the ability to predict the distortion is poor. Also, if the number of neurons in the hidden layer is too small, the accuracy gain is not as good as it could be. Thus, it was selected a sample period of 12 hours and 60 neurons at the hidden layer. Figure 10a shows the sampled distribution function of the distortion intensity in the evaluation phase in three different cases: before calibration; after calibration when a priori attitude knowledge is available (aided calibration); and after a calibration based on the GPS observations only (self-calibration). Analogously, Figure 10b shows the sample distribution function of the total attitude error (module of the three axis error vector) in the evaluation phase for the same cases.

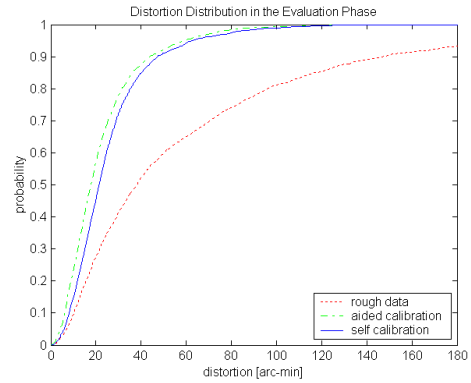

Figure 10a. Distribution of distortion in the observed line-of-sights.

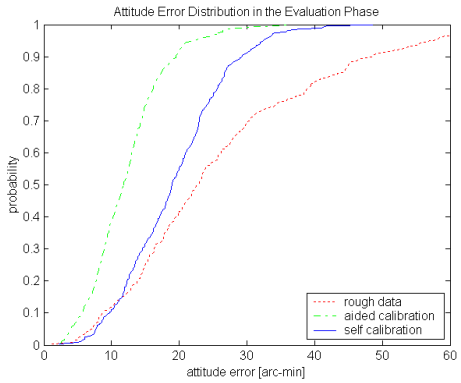

Figure 10b. Distribution of attitude total error.

Table 2 summarizes the algorithm performance in terms of accuracy in the simulation.

Table 2. Performance parameters in arc-min.

Performance Parameter	Before calibration	After aided calibration	After self-calibration
Global distortion (noise-free)	62.28	22.58	26.00
Distortion in the evaluation phase	84.73	29.13	32.22
Roll-Pitch error (1σ)	19.10	8.08	12.78
Yaw error (1σ)	15.56	7.13	10.26
Total attitude error (1σ)	31.17	13.47	20.78
Total attitude error (95%)	57.39	22.03	32.40

7. Conclusions and Recommendations

An algorithm for multipath mitigation by neural network has been developed and implemented. The algorithm improves the accuracy of spacecraft attitude determination from GPS observations only. In the specific test scenario, the algorithm was able to improve the attitude accuracy level from 1^o to 0.5^o. Also, the distortion level on the observed GPS line of sight due to multipath only decreased from 1^o to less than 0.5^o.

Besides the multipath scenario, the absolute accuracy level achieved in a given application depends on several other factors independent from the algorithm itself, as the baseline length, the noise level, the satellite attitude and the correspondent number and geometric distribution of the GPS satellites at sight. Nevertheless, the results indicate that the algorithm has a potential to substantially reduce the effect of multipath on board of a spacecraft so improving the accuracy of its attitude determination.

The Shuster's recipe applied to the self-calibration problem was proven to be efficient in avoiding the algorithm divergence that was otherwise present. Furthermore, the existence of abundant attitude GPS observations assured a considerable accuracy gain even in the absence of a priori attitude knowledge.

The algorithm represents an encouraging step envisaging future applications of attitude determination with GPS in the context of the Brazilian space activities. More investigations with data from a GPS simulator under more realistic flight conditions are foreseen.

Acknowledgments

The first author acknowledges his sponsorship by CNPq, process No. 300388/95-0(RN).

References

1. Lopes RVF, Carrara V, Enderle W, Arbinger C (2000) Mitigating Multipath by Neural Network. (AAS 00-207) Advances in the Astronautical Sciences, 105(2), 1639-1650.
2. Shuster MD, Lopes RVF (1994) Parameter Interference in Distortion and Calibration. (AA 94-186) Advances in The Astronautical Sciences, 87(1), 595-611.
3. Lopes RVF, Milani PG (2000) Consistent On-Board Multipath Calibration for GPS Based Spacecraft Attitude Determination. Proceedings of the ION GPS 2000, Salt Lake City, USA, pp. 2216-2226.
4. Wahba G. (1965) A Least Squares Estimate of Spacecraft Attitude, Problem 65-1. SIAM Review, 7(3), 409.
5. Shuster MD, Oh SD (1981) Three-Axis Attitude Determination from Vector Observations. Journal of Guidance, Control and Dynamics, 4(1), 70-77.
6. Markley FL (1988) Attitude determination using Vector Observations and the Single Value Decomposition. Journal of the Astronautical Sciences, 36(3), 245-258.
7. Hunt KJ, Sbarbaro D, Zbikowski R, Gawthrop PJ (1992) Neural networks for control systems - a survey. Automatica 28(6), 1083-1112.
8. Cybenko G (1989) Approximation by superposition of a sigmoidal function. Mathematics of Controls, Signals and Systems 2(4), 303-314.
9. Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. Neural Networks 2(5), 359-366.
10. Baffes PT, Shelton RO, Phillips TA (1991) *NETS,* a neural network development tool. Huston, Lyndon B. Johnson Space Center. (JSC-23366)
11. Billings SA, Jamaluddin HB, Chen S (1992) Properties of neural networks with applications to modelling non-linear dynamical systems. International Journal of Control 55(1),193-224.
12. Chen S, Billings SA (1992) Neural networks for nonlinear dynamic system modelling and identification. International Journal of Control 56(2), 319-346.
13. Hagan MT, Menhaj M (1994) Training feedforward networks with the Marquardt algorithm. IEEE Transactions on Neural Networks 5(6), 989-993.
14. Carrara V, Varotto SEC, Rios Neto A (1998) Satellite Attitude Control Using Multilayer Perceptron Neural Networks (98-345). Advances in the Astronautical Sciences. 100(1), 565-579.