Tema

Harmonic Analysis of Multipath Index Time Series in GPS Stations[†]

E.M. SOUZA[1*], D.B.M. ALVES[2] and F.L. SCHUMACHER[3]

ABSTRACT. The identification of the cyclical and seasonal variations can be very important in time series. In this paper, the aim is to identify the presence of cyclical or seasonal variations in the indices of the multipath effect on continuous GPS (*Global Positioning System*) stations. Due to the model used to obtain these indices, there should not have cyclical variations in these series, at least due to the multipath effect. In order to identify the presence of cyclical variations in these series, correlograms and Fourier periodograms were analyzed. The Fisher test for seasonality was applied to confirm the presence of statistical significant seasonality. In addition, harmonic models were adjusted to check in which months of the year the cyclical effects are occurring in the multipath indices. The possible causes of these effects are pointed out, which will direct the upcoming investigations, as well as the analysis and correlations of other series. The importance of this analysis is mainly due to the fact that errors in the collected signals of these stations will directly influence the accuracy of the results of the whole community that directly or indirectly uses GPS data.

Keywords: harmonic analysis, multipath index, time series.

1 INTRODUCTION

Data from GNSS (Global Navigation Satellite Systems), or specially, GPS system are affected by several errors, which may occur on satellite, signal propagation, and receiver or ground station. Most of these errors can be eliminated, reduced or modeled, depending on the applied positioning method or desired accuracy. One of the most used positioning methods is the relative one, where is possible to estimate a receiver position in relation to a reference station (generally this reference station belongs to a continuous monitoring network). Applying relative positioning, most of the errors are eliminated, depending on receiver's distance. But, there is a kind of error that is not eliminated, the multipath effect. This error occurs because reflected signals arrive to

---

*Corresponding author: Eniuce Menezes de Souza
[†]Grateful to CNPq for the processes 473973/2012-0 and 303079/2011-8 and FAPESP 2012/19906-7.
[1]Department of Statistics, UEM – Maringá State University, Av. Colombo, 5790, 87020-900 Maringá, PR, Brazil. E-mail: emsouza@uem.br
[2]Cartography Department, UNESP – São Paulo University, Rua Roberto Simonsen, 305, 19060-900 Presidente Prudente, SP, Brazil. E-mail: danibarroca@fct.unesp.br
[3]Undergraduate Student of Statistics, UEM – Maringá State University, Av. Colombo, 5790, 87020-900 Maringá, PR, Brazil. E-mail: fernandalschumacher@gmail.com

the receiver antenna together the direct one. Antenna can receive signals reflected in neighboring surfaces (indirect), like buildings, cars, trees, water etc. Therefore, the received signals can be composed by direct and indirect signals and introduce distortions in signal and its modulation [8]. Multipath modeling is very difficult because it depends on physical conditions of each receiver's location. So the ideal is to collect the data avoiding such effect, but in many times this is not possible.

After data collection, we can estimate some indices that can represent multipath effect, indicating its presence in the data. One of these is the Multipath (MP) index, which is obtained from linear combinations of GPS signals. The development and details concerning this index can be found in [8] and [6].

When data from continuous monitoring stations are used, mainly as base station in relative positioning, it is expected a minimum of multipath effect, because these stations are, theoretically, well located in places without obstructions. Otherwise, if there is multipath presence in a reference station, this error will affect directly the user positioning. Therefore, in this paper, the main goal is to evaluate the existence and magnitude of multipath in continuous monitoring stations from temporal series of MP index, which also provides information about data quality. Furthermore, we aim to verify which events can influence the index values, and to investigate the presence of cyclic variations and seasonality statistically significant in the series. In this investigation, the multipath repeatability will be very useful to prove the multipath presence. This repeatability occurs because if the receiver remains stopped and the environment nearby remains unchanged, the multipath effect should repeat every day according to the satellite movement. Thus, behaviors or periodicities beyond the multipath repeatability might indicates other effects influencing the MP calculus, or in other words, been taken into account together with multipath effect.

An important effect that can influence the MP index series and motivates this investigation is related to ionospheric irregularities, called ionospheric scintillation ([3], [4]). This effect can be described as a fast change in phase and amplitude of GPS signal, caused by irregularities of electron density. Scintillation can degrade or cause the GPS signal lost. The Brazilian territory is one of the most affected regions of the Earth, with high scintillation activity in sunset hours. Ionospheric scintillation occurs mainly around the peak of the 11-year ionosphere solar cycle and the most affected months are March, April, September and October, near equinox.

Thus, Section 2 presents how to build the harmonic models to investigate the cyclic variations of this time series. The results and analysis are discussed in Section 3. Finally, the concluding remarks are presented in Section 4.

2   INVESTIGATING CYCLIC VARIATIONS

Considering that the MP index is generated daily, the time series have 365 observations per year. Although the daily repeatability and the variations over the day can not be verified by this index, it is very useful to investigate behaviors that are repeated with larger intervals, such as months and

years. The analysis of the autocorrelation function (ACF) or correlogram is useful in identifying the presence of autocorrelation, however its correct interpretation is one of the most difficult tasks in time series [2]. In that sense, particularly for investigation of periodicities, harmonic analysis and Fourier periodogram can be very helpful. The harmonic analysis is historically known for allowing the investigation of "hidden" periodicities in time series. In usual regression models in the form:

$$Y_t = \mu_t + \varepsilon_t, \tag{2.1}$$

where μ_t is a deterministic function and $\{\varepsilon_t\}$ is a sequence of independent white noise of Y_t, with mean zero and variance σ_ε^2 [5], seasonality or cyclical effects can be incorporated via cosine functions:

$$\mu_t = \delta \cos(2\pi\omega t + \phi), \tag{2.2}$$

where $\delta > 0$ is the amplitude, ω is the frequency and ϕ, which follows uniform distribuition $\phi \sim U(-\pi, \pi)$, is the curve phase. As t varies, the curve oscillates between the maximum δ and minimum $-\delta$. Since the curve repeats itself exactly every $1/\omega$ time units, $1/\omega$ is called the period of the cosine curve. However, equation (2.2) is not appropriate because the nonlinearity of the parameters δ and ϕ. Hence it is convenient to use the reparametrization:

$$\delta \cos(2\pi\omega t + \phi) = \alpha_1 \cos(2\pi\omega t) + \beta_1 \sin(2\pi\omega t), \tag{2.3}$$

where $\delta = \sqrt{\alpha_1^2 + \beta_1^2}$, $\Phi = \arctan(-\beta_1/\alpha_1)$, $\alpha_1 = \delta \cos\Phi$ and $\beta_1 = -\delta \sin\Phi$.

Then for a fixed frequency ω we can use $\cos(2\pi\omega t)$ and $\sin(2\pi\omega t)$ as predictor variables and adjust the α_j and β_j by the least squares method (LSM). The model can thus be written as:

$$\mu_t = \beta_0 + \alpha_1 \cos(2\pi\omega t) + \beta_1 \sin(2\pi\omega t). \tag{2.4}$$

where the constant term β_0 can be interpreted as a cosine with frequency equal to zero.

In this sense we could write the decomposition of Y_t in terms of periodic functions. Thus the general linear combination of m cosine curves with arbitrary amplitudes, frequencies and phases can be written as:

$$Y_t = \beta_0 + \sum_{j=1}^{m} \left(\alpha_j \cos(2\pi\omega_j t) + \beta_j \sin(2\pi\omega_j t) \right). \tag{2.5}$$

Although the α_j and β_j can be estimated by the LSM, for some frequencies the procedure is simplified. If n is odd we have $n = 2k + 1$ and the sine and cosine predictor variables with frequency $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{k}{n}$, where $\frac{k}{n}$ can be written as $\frac{1}{2} - \frac{1}{2n}$, are orthogonal, beyond $\omega = 0$. These frequencies are called Fourier frequencies and

$$\hat{\beta}_0 = \bar{Y} = \frac{1}{n}\sum_{t=0}^{n} Y_t, \hat{\alpha}_j = \frac{2}{n}\sum_{t=1}^{n} Y_t \cos\left(\frac{2\pi j t}{n}\right) \text{ and } \hat{\beta}_j = \frac{2}{n}\sum_{t=1}^{n} Y_t \sin\left(\frac{2\pi j t}{n}\right). \tag{2.6}$$

where $\bar{Y}$ is the arithmetic mean. If n is even, $n = 2k$ and the equations in (2.6) are still valid for $j = 1, \ldots, k - 1$ and $\omega_k = k/n = 1/2$. Moreover, it is possible to show that the contribution

of the harmonic δ_j in (2.3) is given by $n\delta_j^2/2$, which is the known Fourier periodogram [2, 5]. Therefore:

$$I\left(\frac{j}{n}\right) = \frac{n\left(\hat{\alpha}_j^2 + \hat{\beta}_j^2\right)}{2}, \tag{2.7}$$

for $j = 1, \ldots, k$ if n is even and $j = 1, \ldots, k - 1$ if n is odd. Actually, other representations for the periodogram can be constructed, but the meaning is the same.

The idea of locating hidden frequencies or periodicities in time series comes from repeating the analysis in (2.3) for all frequencies $\frac{1}{n}, \ldots, \frac{k}{n}$, correlating them with the series for a intensity measure. In other words, quantifying the importance of each frequency to the series in question.

By adjusting a model for the Fourier frequencies we have the harmonic analysis. If the series can be considered deterministic with periodicity S, then the harmonic model in equation (2.5) can be adjusted for $j = 1, \ldots, S/2$ [7].

Even if α_j and β_j were null for all j, it is possible that peaks occur on the ordinate of the periodogram due to random fluctuations. Thus, to verify whether the periodicity is significant or not, assuming that the process Y_t is Gaussian, Fisher shows a procedure to test the hypothesis $H_0 : \alpha_j = \beta_j = 0 \, \forall j$ (there is no periodicity) based on the statistic:

$$g_{obs} = \frac{\max I\left(\frac{i}{n}\right)}{\sum_{j=1}^{[K]} I\left(\frac{j}{n}\right)}, \tag{2.8}$$

where $[K] = k$ if n is even and $[K] = k - 1$ if n is odd [7]. Fisher showed that for n odd the exact distribution of g_{obs}, under H_0, is given by:

$$P(g > a) = m(1 - a)^{m-1} - \binom{m}{2}(1 - 2a)^{m-1} + \ldots + (-1)^x \binom{m}{x}(1 - xa)^{m-1}, \tag{2.9}$$

where $m = [K]$ and x is the largest integer smaller than $\frac{1}{a}$ $\left(x = \left[\frac{1}{a}\right]\right)$. The expression given in (2.9) can be approximated using only the first term of the expansion, that is:

$$P(g > a) \cong m(1 - a)^{m-1}. \tag{2.10}$$

For a given significance level α, we can find the critical value $a(\alpha)$ such that $P(g > a(\alpha)) = \alpha$. If the observed value of the statistic (g_{obs}) is greater than $a(\alpha)$ we must reject H_0, which means that there is a periodicity. Substituting a for g_{obs} the approximate test p-value can be found.

3   RESULTS

For the application of this research, data from a continuous monitoring stations of the Brazilian Network for Continuous Monitoring (RBMC – http://www.ibge.gov.br/home/geociencias/geodesia/rbmc) were used. The station is located in Sao Paulo (POLI). This choice was made based on the amount of data available (at least six years of data) and because any antenna or

software change occured in this period for this station. More details and the time series for other RBMC stations can be seem in [1].

The TEQC software (http://facility.unavco.org/software/teqc/) was used to obtain the MP daily indices of the station, with data from 2006 to 2012. All the other analyzes were performed in R software (http://cran.r-project.org/).

Figure 1 shows the MP time series of the POLI station and their respective ACF.
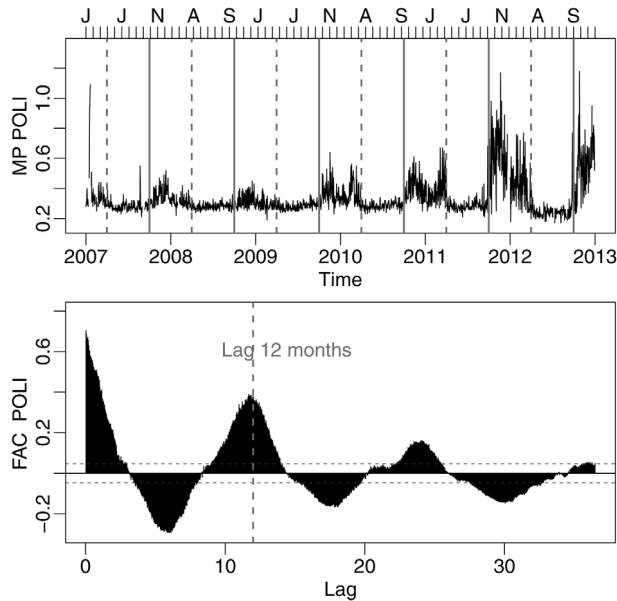

Figure 1: Top: MP time series of the POLI station. The red solid lines indicates the beginning of October. The red dashed lines indicates the end of March. Bottom: Autocorrelation function (ACF).

It is possible to see in Figure 1 that from beginning of October (red solid line) to the end of March(red dashed line), the effects are more intense. We can also note two peaks per year, which occurs in this period, around March and October. Furthermore, the magnitude of these peaks is increasing with time. The presence of effects that are repeated every year in the time series can also be seen in ACF (Fig. 1, bottom). To confirm such seasonalities, which refer to the presence of effects that are being repeated in periods of 12 months, the Fourier periodogram was estimated and the result is shown in Figure 2.

In Figure 2 one can notice a peak in blue, which occurs at the frequency 0.002734375 and represents a periodicity of 365 days ($= 1/0.002734375$). The red line is indicating exactly this frequency, which coincides with the estimated (in blue). So we can confirm the seasonality by Fourier periodogram. Concerning to the Fisher periodicities test, from equation (2.8) is obtained $g_{obs} = 0.25$. By equation (2.10) we have $P(g > g_{obs}) < 0.0001$. Therefore we reject H_0, i.e. there is a statistically significant periodicity, confirming the previous results.
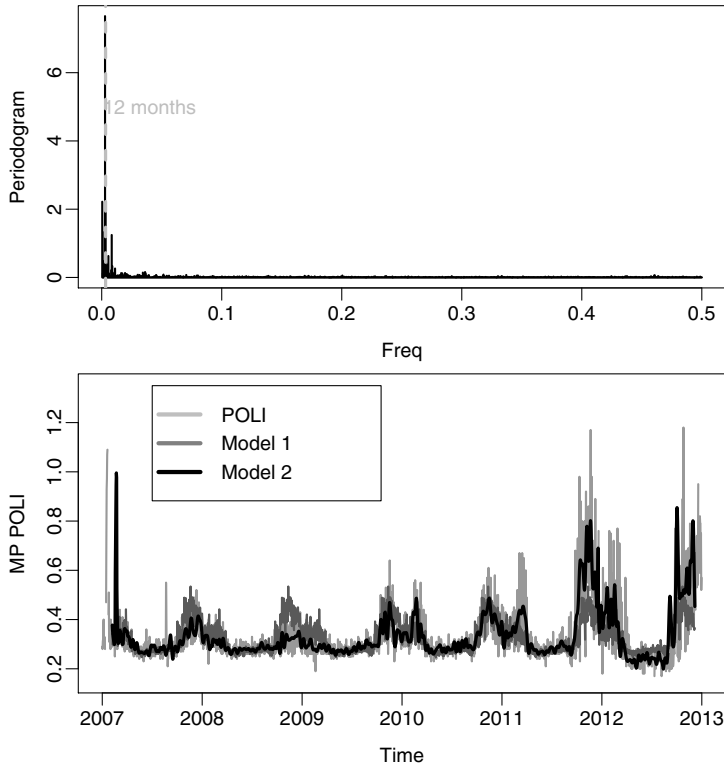
Figure 2: Top: Fourier periodogram for time series MP for POLI. Bottom: Time series of MP for station POLI and the estimated harmonic models.

Although the presence of seasonality has been confirmed, it is still necessary to identify its behavior, indicating the times when its effect is more expressive. In this sense, the Fourier harmonic modeling presented in Section 2 is quite useful. In Figure 2 adjustments of two harmonic models for the MP of POLI station are illustrated. Model 1 considers the series with deterministic seasonality of 365 days (green) and model 2 was adjusted for Fourier frequencies without considering deterministic seasonality (black).

The main harmonic in the model 1 occurs in $j = 1$, with frequency $1/365 = 0.002734375$ and amplitude $\delta_1 = \alpha_1^2 + \beta_1^2 = 0.04364 + 0.00544 = 0.04908$. This harmonic is responsible for 71.64% of the variance of the series. In model 2 various components of frequency contribute to the variance of the series. The most important harmonic in model 2 has amplitude $\delta = 0.04871 + 0.00588 = 0.05459$. This harmonic along with the annual harmonic contribute with 84.86% of the variance of the series.

Both modelings were statistically significant. The adjust of the second model was better, with σ_ε smaller and R^2 larger (Table 1). Even the model 1 showing an R^2 of 0.4485, one can see in Figure 2 that the harmonic model was able to detect the peaks near to October and March.

Table 1: Adjusts of the harmonic models for time series of the MP for the POLI station.

Model	Par	Estimates	SE	t-values	$P(> \lvert t \rvert)$	σ_ε	R^2	P-value
1	α_1	0.2089	0.0069	29.86	< 0.0001	0.2266	0.4485	< 0.0001
	β_1	–0.0737	0.0069	–10.63	< 0.0001			
2	α_1	0.0371	0.0050	7.37	< 0.0001	0.1506	0.7566	< 0.0001
	β_1	–0.0912	0.0051	–17.75	< 0.0001			
	α_6	0.2207	0.0052	42.40	< 0.0001			
	β_6	–0.0767	0.0050	–15.47	< 0.0001			

4   CONCLUDING REMARKS

The positioning performed using data from continuous monitoring stations has been much employed in recent years. In order to evaluate the quality of data from these stations were generated and evaluated MP time series data of POLI station of the RBMC.

Only with the time series graph it was noticeable variations that repeat periodically. To evaluate this phenomenon more effectively, correlograms, periodograms and harmonic adjusts of the series were analyzed. With the results presented in Figures 1 and 2 the seasonality became clear. To confirm the seasonality, Fisher's exact test was performed, where it was possible to verify the presence of statistically significant seasonality. Moreover, by harmonic adjustments that consider or not deterministic seasonality, it was possible to identify peaks in March and October. The fact that these peaks are in the months of March and October coincides with recent research related to the effects of ionospheric scintillation, which has been more intense in those months. It is expected that non-modeled effects can be absorbed in the calculation of these indices, which explains the appearance of the seasonal effect on the index studied. These results will direct future research towards the correction of the effect of ionospheric scintillation.

RESUMO. A identificação de variações cíclicas e sazonais é muito importante em séries temporais. Neste artigo, objetiva-se identificar a presença de variações cíclicas e sazonais nos índices do efeito do multicaminho em estações GPS (*Global Positioning System*) de monitoramento contínuo. Devido ao modelo usado na obtenção de tais séries, não deveriam existir variações cíclicas nas mesmas, pelo menos devido ao multicaminho. No sentido de identificar a presença de variações cíclicas nestas séries, foram analisados correlogramas e periodogramas de Fourier. Para confirmar a presença de sazonalidade estatisticamente significativa, foi utilizado o teste de Fisher para sazonalidade. Além disso, modelos harmônicos foram ajustados para verificar em que meses do ano os picos dos efeitos cíclicos estão ocorrendo. As possíveis causas dos efeitos são apontadas, as quais direcionarão investigações futuras, bem como a análise e correlações com outras séries. A importância desta análise se deve principalmente ao fato de que erros nos sinais coletados afetarão a acurácia dos resultados de toda a comunidade que usam dados GPS diretamente ou indiretamente.

Palavras-chave: análise harmônica, índice de multicaminho, séries temporais.

REFERENCES

[1] D.B.M Alves, E.M. Souza, V.Y. Kaneshiro & J.S. Souza. Análise de séries temporais de multicaminho em estações de monitoramento contínuo. *Boletim de Ciências Geodésicas*, **19** (2013), 353–373.

[2] C. Chatfield. The Analysis of Times Series: An Introduction. Chapman and Hall: London (1995).

[3] R.S. Conker, M.B. El-Arini, C.J. Hegarty & T. Hsiao. "Modeling the Effects of Ionospheric Scintillation on GPS/Satellite-Based Augmentation System Availability". *Radio Science*, **38** (2003).

[4] R. Tiwari, S. Skone, S. Tiwari & H.J. Strangeways. 3WBMod Assisted PLL GPS Software Receiver for Mitigating Scintillation Affect in High Latitude Region. IEEE, 2011. Disponível em: <http://www.ursi.org/proceedings/procGA11/ursi/FG-4.pdf>. Acesso em: out. 2012.

[5] J.D. Cryer & K.S. Chan. Time Series Analysis: With Applications in R, Springer Texts in Statistics (2008).

[6] L.H. Estey & C.M. Meertens. TEQC: The multi-purpose toolkit for GPS/GLONASS Data. *GPS Solutions*, **3**(1) (1999), 42–49.

[7] P.A. Morettin & C.M.C. Toloi. Análise de Séries Temporais. Edgar Blucher: São Paulo (2004).

[8] E.M. Souza. Análise de Wavelets para Detecção e Correção do Multicaminho no Posicionamento Relativo GNSS Estático e Cinemático, 2008. Tese (Doutorado em Ciências Cartográficas) – Universidade Estadual Paulista, Presidente Prudente.