North and South Hemispheric Solar Activity for Cycles 21-23: Asymmetry and Conditional Volatility of Areas of Plage Regions

E. Gonçalves[1,3] · **N. Mendes-Lopes**[1,3] · **I. Dorotovič**[2] · **J. M. Fernandes**[2,3] · **A. Garcia**[2,3] ·

Abstract In this work we study the dynamic evolution of the time series describing the areas of plage regions observed daily at the Observatório Astronómico da Universidade de Coimbra, in each one of the solar hemispheres during solar cycles 21-23. The classical ARMA model turned out to be insufficient to describe the time variations seen in the data, due to the presence of strong conditional variability. We found that the data are well fitted by ARMA mixed with power-δ TGARCH error models. The power index δ is non-integer; this property has recently been introduced in the literature on time series analysis and indicates the presence of strong volatility and long memory in the data series. We also detected the presence of dynamic asymmetry in the plage region areas observed in the two hemispheres, as two different temporal models are obtained to fit them. This conclusion is also supported by the dynamic evolution of the daily difference (north−south) time series that is significantly different from a white noise. This statistical modeling of time series, taking into account new and different characteristics of the solar activity, will be of great usefulness in subsequent forecast developments.

Keywords: Conditional volatility · North-south asymmetry · Sun: activity · Sun: chromosphere

1. Introduction

The activity of the Sun and the solar cycle is, presently, among the hot topics in astrophysics, not only for studies about the Sun itself but also for terrestrial and human connections, namely, the Sun's influence on the Earth climate changes

[1]CMUC, Department of Mathematics, University of Coimbra
email: esmerald@mat.uc.pt email: nazare@mat.uc.pt
[2]GCUC, Astronomical Observatory of the University of Coimbra email: ivan@mat.uc.pt email: jmfernan@mat.uc.pt
email: adriana@mat.uc.pt
[3]Department of Mathematics, University of Coimbra

and on the life time of artificial satellites (Hathaway, 2010). The study of solar activity and its evolution is based on the regular observations of a variety of solar features such as sunspots (number and area), plage regions, prominences, total radio flux, flares, *etc.*, including the geomagnetic activity. It is very well known that solar activity shows an asymmetry between the northern and southern hemispheres. Newcomb (1901) was probably the first one to point out this asymmetry by the analysis: "There seems to have been an abnormal delay in the increase of the spottedness of the southern hemisphere during the years 1880-1881." After this seminal paper, the asymmetry between the northern and southern hemispheres regarding the sunspot number was definitively established by Newton and Milsom (1955). Since then, many papers have been published showing that the claimed asymmetry is also observed in other solar activity indices, and numerous properties of the north-south (N–S) asymmetry have been revealed during the last several decades. Hereafter we briefly summarize some of these results obtained during the last decade. A general overview on the N–S asymmetry studies can be found *e.g.* in Duchlev (2001), Temmer *et al.* (2006), Javaraiah (2008), Sýkora and Rybák (2010), Bankoti *et al.* (2010, 2011), Gigolashvili *et al.* (2011), Li *et al.* (2010, 2011), Badalyan (2012), Deng *et al.* (2013), and references therein.

Many authors studied particularly the N–S asymmetry in the solar rotation rate, in relation to solar cycle predictions or Earth's climate studies. For example, Temmer *et al.* (2002) investigated the N–S asymmetry and differences in the rotation rate of the two hemispheres, based on the hemispheric sunspot numbers obtained at Kanzelhöhe Solar Observatory, Austria, and Skalnaté Pleso Observatory, Slovakia. Gigolashvili *et al.* (2011) performed a statistical study of the N–S asymmetry in solar differential rotation based on various solar features. Xie *et al.* (2012) investigated temporal variation of the solar cycle length using the hemispheric sunspot activity data through a method of continuous wavelet transformation. McIntosh *et al.* (2013) studied the hemispheric asymmetry in photospheric magnetic activity in relation to the abnormality of cycle 24. Vats and Chandra (2011) showed an asymmetry in the solar coronal rotation using radio (the Nobeyama Radio Heliograph) and X-ray (the *Yohkoh* Soft X-ray Telescope (SXT)) data. Javaraiah (2008) and Sen (2009) applied the study of the N–S asymmetry to solar cycle predictions. Georgieva *et al.* (2007) also found a connection between the N–S asymmetry and the Earth's atmospheric circulation. Furthermore, the magnetic synoptic maps from Kitt Peak between 1975 and 2003 show clearly periodic oscillations in the hemispheric asymmetry in magnetic flux (Knaack, 2004). Bzowski *et al.* (2003) studied the data taken during 1996–2002 from the *Solar Wind Anisotropies* (SWAN) instrument on the SOHO mission and found a N–S asymmetry in the solar wind mass flux throughout the solar cycle. Zhang *et al.* (2002), using data from the *Solar Wind Ion Composition Spectrometer* (SWICS) of *Ulysses* and synoptic charts derived from Kitt Peak magnetograms, suggested a N–S asymmetry in the temperature of the polar coronal holes.

Magnetic activity in polar regions and their N–S asymmetry have been extensively studied using polar faculae. Deng *et al.* (2013) analyzed the data of polar faculae during the time interval of 1952-1998 and studied the hemispheric

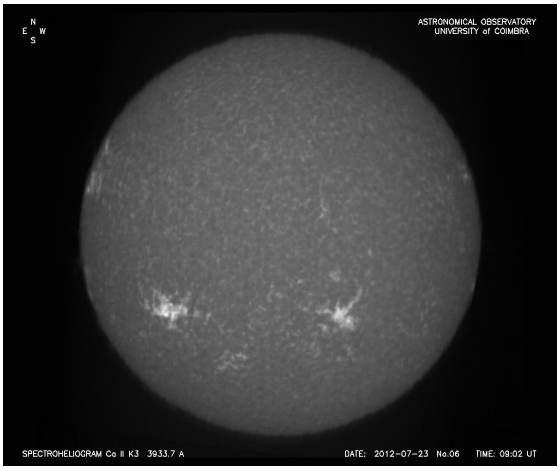

Figure 1. Calcium K-line spectroheliogram taken at the Observatório Astronómico da Universidade de Coimbra (OAUC) on 23 July 2012.

asynchrony of polar faculae and its relationship with the low-latitude solar activity. Li *et al.* (2009) studied the N–S asymmetry of high-latitude solar activity including polar faculae of cycles 19 to 23, and found that the N–S asymmetry in polar faculae counts is not simply related to the N–S asymmetry in sunspot numbers.

Dorotovič *et al.* (2010) presented a statistical descriptive analysis of the areas of plage regions (see Figure 1) between 1996 and 2006 measured on the Ca II K3 spectroheliograms obtained at the Observatório Astronómico da Universidade de Coimbra (OAUC), and concluded the existence of the N–S asymmetry. However, their study did not explore the dynamic properties of data in time. In order to analyze the dynamic nature of N–S asymmetry, it is essential to study the temporal evolution of data series. With this aim in mind, we use the same Ca II spectroheliograms of OAUC and analyzed the areas of plage regions in the two hemispheres during the period from March 1976 to December 2006 (covering the solar cycles 21, 22, and 23).

Time series modeling has undergone an important development in recent years. The linear formulation of the autoregressive moving average (ARMA) models has been found to be insufficient to describe adequately some data, like financial or physical ones. In fact, this kind of time series presents features of non-linearity behavior; particularly its conditional volatility, that is, its instantaneous variability, depends strongly on the past. In order to best take into account this property, several models describing the evolution of the conditional variance of the corresponding process appeared in the literature following the seminal paper of Engle (1982). This kind of models is known as autoregressive conditional heteroscedastic (ARCH) models (variance depending on time t) and their generalization (generalized ARCH, or GARCH, models).

Another property often found in those time series is the asymmetrical reaction of the volatility according to the sign of past observations, namely the realization of different behavior during a rising or falling period. This typical behavior is

taken into account in the threshold ARCH (TARCH) models (for a recent survey, see Francq and Zakoian, 2010) in which the conditional standard deviation of the process at time t is a piecewise linear function of negative and positive values of past observations. Similarly, the presence of long memory in the conditional variance after a shock as input has led to the proposal of power-law GARCH models (Ding, Granger, and Engle, 1993; Pan, Wang, and Tong, 2008).

A natural extension of the threshold ARCH processes that allows us to take into account both long memory property and asymmetry in the stochastic volatility (Gonçalves, Leite, and Mendes-Lopes, 2012) is the TGARCH model with non-integer power-law index δ. Moreover, the models of this general class are also well adapted for heavy tail data (as in the present case), since they are in general leptokurtic stochastic processes, that is, its kurtosis is greater than the standard value (3) of normal distribution. The combination of ARMA models with an error process following conditional heteroscedastic models greatly improves the adequacy of the fitting and it has been extensively studied in the stochastic literature and applied to real data (Weiss, 1984; Gonçalves and Mendes-Lopes, 2008).

In this study, the Calcium plage data of OAUC in the northern and southern hemisphere series are analyzed using the classical Box-Jenkins methodology (Box and Jenkins, 1976) and its generalization, in order to take into account the features of conditional volatility if they are detected in the residual series of the model firstly deduced by the classical procedure. The dynamic evolution of the daily difference of the observations in the northern and southern hemisphere data is analyzed by the same procedure.

The data studied here are generally leptokurtic, which means the presence of heavy tail in the data. Thus, the presence of another feature often observed in leptokurtic data, which is called the Taylor effect, is also analyzed in the same data series.

All statistical analysis are performed using the statistical software Eviews. Gouriéroux and Monfort (1995) is a general reference for statistical analysis and related parameters, and for a recent survey on time series we refer to Francq and Zakoian (2010).

2. Plage Region Areas: Estimation of Missing Observations and Their Validation

Usually an analysis of time series requires observations equally spaced in time. However, in the long series under study some observations are missing due to weather and instrument conditions. In order to interpolate missing values, a random average methodology was implemented (Gouriéroux and Monfort, 1990, Chapter 3). The study is undertaken for daily data on the northern and southern hemispheres of the Sun. In Figures 2 and 3 we present the observed plage region areas (in blue) and the corresponding interpolated (completed) series (in red), for the northern and southern hemispheres, respectively.

Figure 4 gives descriptive summaries (histogram and numerical parameters) and comparison tests of means, medians, and variances (Gouriéroux and Monfort, 1995) of the original and completed series for the northern hemisphere.

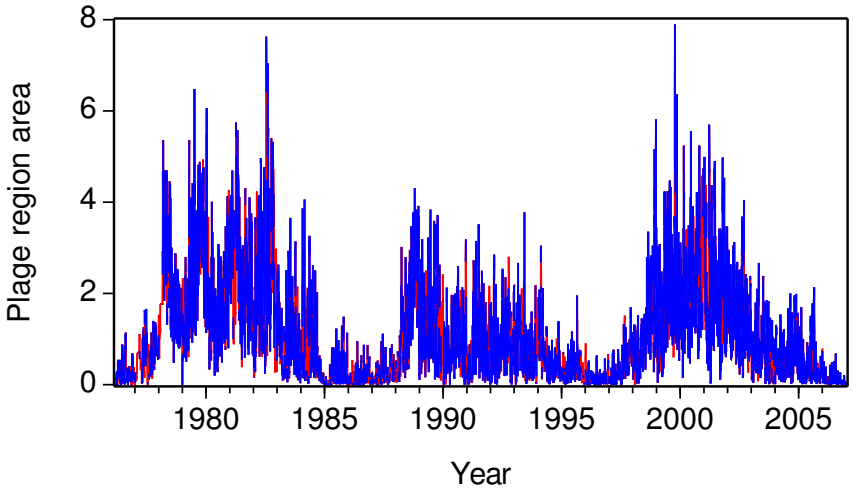

Figure 2. OAUC plage region areas of the northern hemisphere, in percent of the total area of the hemisphere. The blue curve shows the original observation while the red curve shows the completed time series.

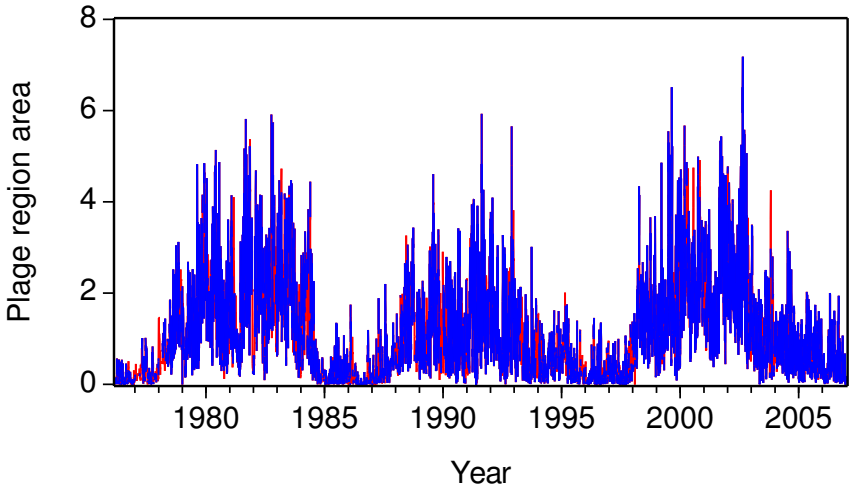

Figure 3. The same as Figure 2, but for the southern hemisphere.

We remark that the p-value of the Jarque-Bera test ($< 10^{-6}$) and the kurtosis value greater than 3 indicate clear non-Gaussianity of data as well as its leptokurtosis. This hypothesis is also supported by the estimated probability densities obtained by the nonparametric kernel method (Silverman, 1998) using the Epanechnikov kernel, which is optimal in a minimum variance sense, with

bandwidth $h = 0.3401$ [1], and is illustrated for the northern hemisphere in Figure 5. (Figures for the southern hemisphere data are not shown because they are similar to Figures 4 and 5).

The chi-squared tests (Gouriéroux and Monfort, 1995) confirm the probabilistic equivalence between the observed and the completed series. For the series of the northern hemisphere data, the null hypothesis of equality of the two distributions is clearly accepted (χ^2-test, p-value 0.665 in Table 1); usually the null hypothesis is rejected if the p-value is as small as 0.01 or 0.05. Similarly, for the southern hemisphere the equality of the distributions is also accepted with p -value equal to 0.32 (Table 2).

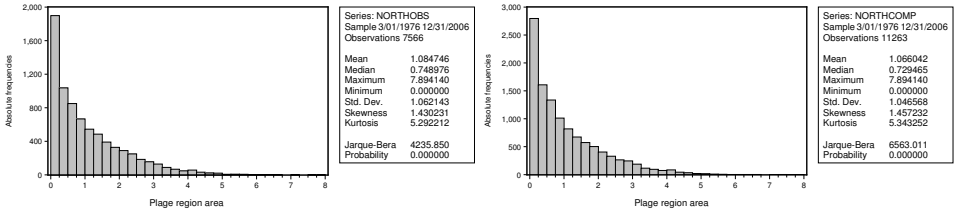

Figure 4. Descriptive summaries of the observed and completed series of plage region areas (northern hemisphere). The histograms show the number of data points per every 0.25 of the plage region areas in percent of the hemisphere.

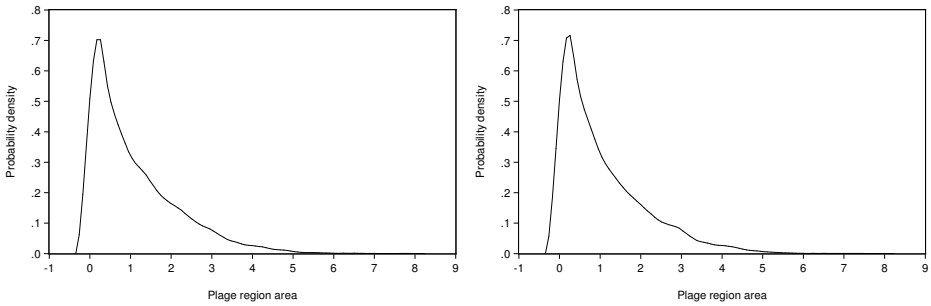

Figure 5. Estimated probability densities for the observed (left) and completed (right) series of plage region areas (northern hemisphere).

To discard a possible loss of information by aggregation, we repeated all these analysis for the recorded and estimated observations, in each year. This study has led to similar conclusions, namely that the random average methodology used to

[1]This parameter has the same units as the area of plage regions.

Table 1. Chi-squared test of equality of the observed and completed distributions of plage region areas (northern hemisphere)[2]. Zero cells (0%) have expected frequencies less than 5. The minimum expected cell frequency is 264.8.

Bin	Observed N	Expected N	Residual
$[0.0, 0.5]$	2935	2958.3	−23.3
$[0.5, 1.5]$	2545	2580.0	−35.0
$[1.5, 2.5]$	1256	1218.1	37.9
$[2.5, 3.5]$	556	544.8	11.2
$[3.5, 8]$	274	264.8	9.2
Total	7566		
Test statistics			
	Chi-squared		2.387
	Degrees of freedom		4
	p-value		0.665

Table 2. Chi-squared test of equality of the observed and completed distributions of plage region areas (southern hemisphere). Zero cells (0%) have expected frequencies less than 5. The minimum expected cell frequency is 151.3.

Bin	Observed N	Expected N	Residual
$[0.0, 0.5]$	2591	2663.2	−72.2
$[0.5, 1.0]$	1478	1505.6	−27.6
$[1.0, 1.5]$	1208	1180.3	27.7
$[1.5, 2.0]$	824	802.0	22.0
$[2.0, 3.0]$	937	923.1	13.9
$[3.0, 4.0]$	359	340.5	18.5
$[4.0, 6.0]$	169	151.3	17.7
Total	7566		
Test statistics			
	Chi-squared		7.005
	Degrees of freedom		6
	p-value		0.320

complete the series maintains the original probabilistic structure. For instance, the equality of the observed and completed distributions of plage region areas, in each year and each hemisphere, is always accepted with a p-value greater than 0.6, using the χ^2-test.

[2]This test is an asymptotic one and the approximations are not valid if some cells have expected frequency less than 5. The same footnote applies to Table 2 and Table 6.

3. Temporal Evolution Modeling, Asymmetry, and Taylor Effect

In order to analyze the temporal evolution of the plage region areas, we consider the general class of ARMA models with power-law index δ in the TGARCH errors. Let us consider a real stochastic process $\eta = (\eta_t,\, t \in \mathbb{Z})$ and, for any integer t, let [3] $\eta_t^+ = \eta_t \mathbb{I}_{\{\eta_t \geq 0\}}$, $\eta_t^- = -\eta_t \mathbb{I}_{\{\eta_t < 0\}}$ and let $\underline{\eta}_t$ be the $\sigma-$ field generated by $(\eta_{t-i}, i \geq 0)$.

The stochastic process $\eta = (\eta_t,\, t \in \mathbb{Z})$ is said to follow a δ power threshold generalized autoregressive conditional heteroscedastic (δ-TGARCH) model with orders p and q (positive integers) if, for every $t \in \mathbb{Z}$, we have

$$\begin{cases} \eta_t = \sigma_t \varepsilon_t \\ \sigma_t^\delta = \omega + \sum_{i=1}^p \left[\alpha_i \left(\eta_{t-i}^+\right)^\delta + \beta_i \left(\eta_{t-i}^-\right)^\delta\right] + \sum_{j=1}^q \gamma_j \sigma_{t-j}^\delta \end{cases} \tag{1}$$

for some real constants $\delta \neq 0$, $\omega > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$ $(i = 1, ..., p)$, $\gamma_j \geq 0$ $(j = 1, ..., q)$, and where $\varepsilon = (\varepsilon_t,\, t \in \mathbb{Z})$ is a sequence of independent and identically distributed real random variables such that ε_t is independent of $\underline{\eta}_{t-1}$, for every $t \in \mathbb{Z}$. If $\delta < 0$ we consider the following convention:

$$\begin{aligned} \left(\eta_t^+\right)^\delta &= 0 \quad \text{if} \quad \eta_t \leq 0, \\ \left(\eta_t^-\right)^\delta &= 0 \quad \text{if} \quad \eta_t \geq 0, \end{aligned}$$

for every $t \in \mathbb{Z}$. If $\gamma_j = 0$ $(j = 1, ..., q)$, the δ-TGARCH(p, q) model is simply denoted δ-TARCH(p).

Moreover, we recall that the stochastic process $V = (V_t,\, t \in \mathbb{Z})$ follows an ARMA(r,s) model with error process η if

$$V_t = \sum_{i=1}^r \varphi_i V_{t-i} + \eta_t + \sum_{j=1}^s \theta_j \eta_{t-j} \tag{2}$$

where φ_i $(i = 1, ..., r)$ and $\theta_j (j = 1, ..., s)$ are real numbers. The case of $s = 0$ is the AR(p) model.

3.1. Temporal Evolution of the Two Series

To characterize the temporal evolution of several series in study, the Box-Jenkins methodology as well as its generalization are often used; therefore, we begin by analyzing the autocorrelation and partial autocorrelation [4] functions of the series (Francq and Zakoian, 2010).

Figure 6 shows the autocorrelation and partial autocorrelation functions of the northern hemisphere series. As we can see, the autocorrelation function decreases

[3] The indicatrix function is defined as $\mathbb{I}_A = \begin{cases} 1, & \text{if } A \text{ happens} \\ 0, & \text{otherwise.} \end{cases}$

[4] The partial autocorrelation of lag k is the correlation between observations X_t and X_{t-k} after removing the linear relations among $X_{t-1}, \ldots, X_{t-k+1}$.

exponentially and the partial autocorrelation is significantly null for lags greater than 2 (these values are, with 95% of confidence, in the neighborhood of zero). Therefore, we may estimate that the data are represented by an AR(2) model. A similar conclusion was obtained for the southern hemisphere series.

The estimation of this AR model leads to a heteroscedastic residual series. In fact, applying the Lagrange multiplier ARCH test to this residual, the null hypothesis of homoscedasticity is rejected with p-value $< 10^{-4}$ (Table 3). A generalization of the Box-Jenkins methodology must be used.

Therefore, we reanalyze the northern hemisphere series considering the class of AR(2) models with general δ-TGARCH error processes. We assume that the temporal evolution of the northern hemisphere series is described by the process $N = (N_t, t \in \mathbb{Z})$ such that

$$N_t = 0.073 + 0.765N_{t-1} + 0.226N_{t-2} + X_t \tag{3}$$

with a TARCH(1) error process $X = (X_t, t \in \mathbb{Z})$ defined by

$$\begin{cases} X_t = \sigma_t \varepsilon_t, \\ \sigma_t^{0.108} = 0.659 + 0.304 \left(\varepsilon_{t-1}^{+}\right)^{0.108} + 0.289 \left(-\varepsilon_{t-1}^{-}\right)^{0.108}, \end{cases} \tag{4}$$

where $(\varepsilon_t, t \in \mathbb{Z})$ are independent real random variables with a centered and reduced Gaussian distribution [5].

The associated residual series has properties of homoscedastic error process, that is, conditional variance independent of t. In fact, the residual correlogram (Figure 7) is compatible with that of a white noise as all the autocorrelations are significantly null (with 95% of confidence in a neighborhood of zero). Moreover, the Lagrange multiplier ARCH test applied to that series accepts the null hypothesis of homoscedasticity with p-value 0.81 (Table 4). These facts validate the previous model, and therefore we conclude that the northern hemisphere series is well fitted by a non-centered AR(2) model with a TARCH(1) error process with power index $\delta = 0.108$.

Table 3. Output of ARCH test after model AR(2) estimation for the northern hemisphere series.

Heteroscedasticity test: ARCH			
F-statistics	487.5138	Prob. $F(1, 11258)$	$< 10^{-4}$
Obs*R-squared[6]	467.3619	Prob. chi-square (1)	$< 10^{-4}$

[5] We note that the AR parameter values are estimated by the classical least squares method, while in the δ-TGARCH parameters the maximum likelihood estimation is used (Francq and Zakoian, 2010).

[6] The statistics labeled Obs*R-squared is the Lagrange multiplier test statistic for the null hypothesis of non serial correlation. The same footnote applies to Table 4 and Table 5.

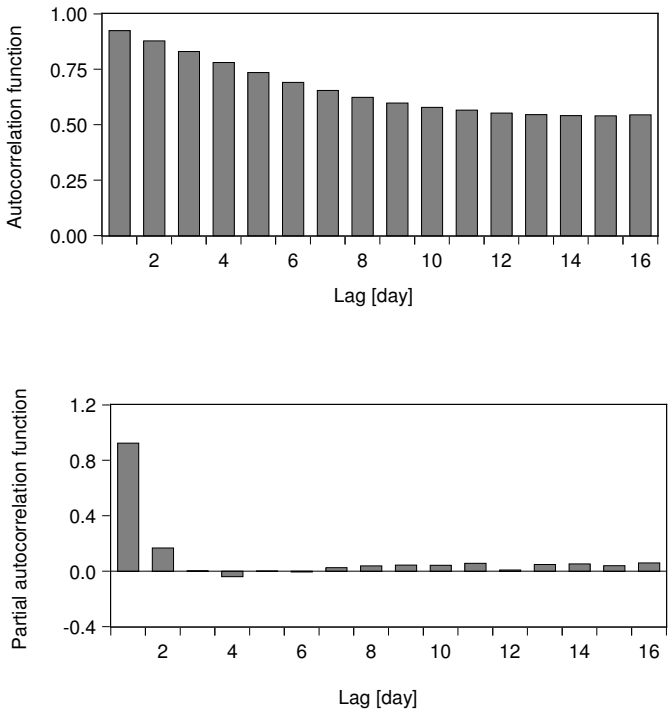

Figure 6. Autocorrelation and partial autocorrelation coefficients for the plage region area series from the northern hemisphere.

Figure 8 shows our results on the plage region areas of the northern hemisphere. The completed series is in red, the series estimated by the model mentioned above is in green, and the corresponding residues are in blue. We observe that the temporal model captures well the evolutionary characteristics of the observed series.

A similar analysis on the southern hemisphere series led to the process $S = (S_t, t \in \mathbb{Z})$ with the form

$$S_t = 0.211 + 0.721 S_{t-1} + 0.262 S_{t-2} + Y_t \tag{5}$$

with the TARCH(1) error process Y with power index $\delta = 0.081$ given by

$$\begin{cases} Y_t = \sigma_t \varepsilon_t, \\ \sigma_t^{0.081} = 0.68 + 0.2932 \left(\varepsilon_{t-1}^{+}\right)^{0.081} + 0.2814 \left(-\varepsilon_{t-1}^{-}\right)^{0.081}, \end{cases} \tag{6}$$

where $(\varepsilon_t, t \in \mathbb{Z})$ are independent real random variables with a centered and reduced Gaussian distribution. The residual is also compatible with a homoscedastic error process; the Lagrange multiplier ARCH test accepts the null hypothesis of homoscedasticity with $p-$value 0.80.

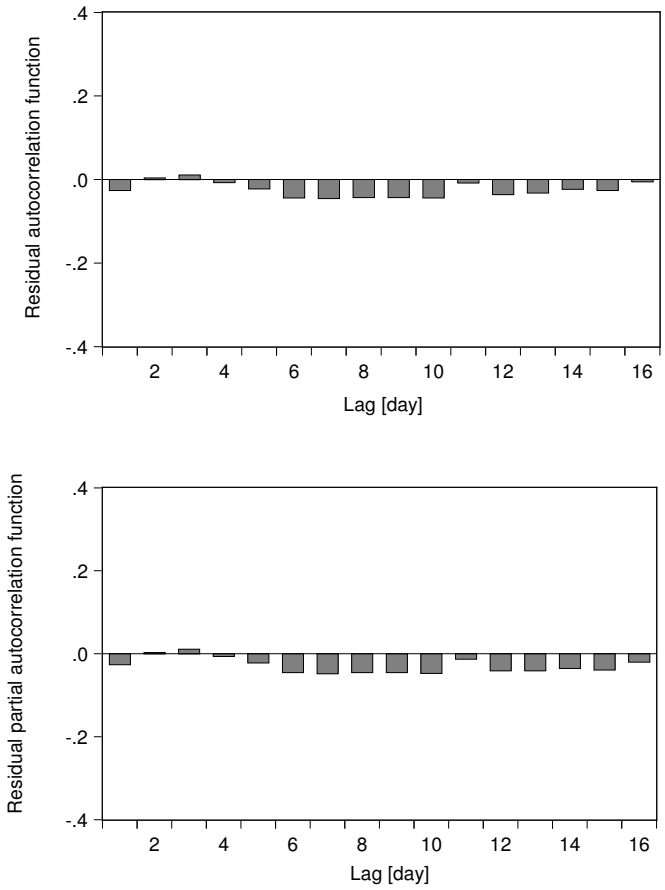

Figure 7. Residual correlogram for the northern hemisphere series after model AR(2) with δ-TGARCH error.

We emphasize that ARMA-TGARCH models with integer power δ do not capture well the heteroscedasticity of the residual series, which indicates a long memory property in the volatility (Ding, Granger, and Engle, 1993) of the plage region area data.

3.2. N-S Asymmetry

The study developed in the previous section also reveals an asymmetry of activity in the northern and southern solar hemispheres, in line with Dorotovič *et al.* (2010), as two different models were obtained for the two hemispheres. In order to validate this conjecture we describe the dynamic nature of the series of differences (northern hemisphere − southern hemisphere) between the values obtained on

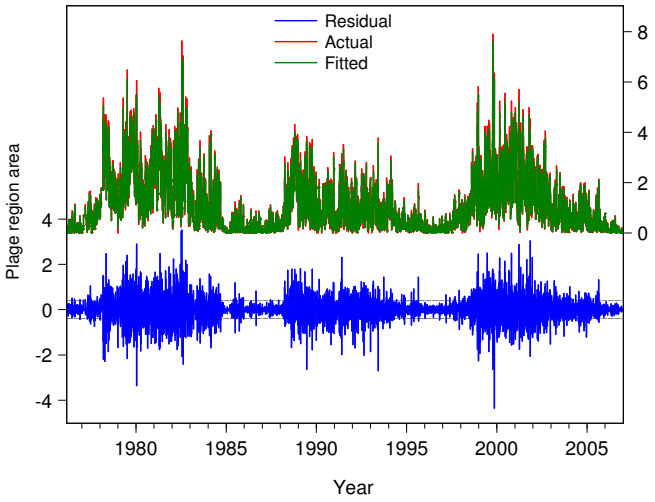

Figure 8. The daily northern hemisphere time series after model fitting. The red curve shows the completed data, while the green curve shows the fitted model. The blue indicates the residuals.

Table 4. Output of ARCH test after model AR(2) with δ-TGARCH error for the northern hemisphere series.

ARCH test			
F-statistics	0.057522	Prob. $F(1, 11258)$	0.810460
Obs*R-squared	0.057532	Prob. chi-square (1)	0.810440

the same day in each one of the hemispheres, through the process D_t, $t \in \mathbb{Z}$. The generalized Box-Jenkins methodology applied to this series leads to the following model

$$D_t = -0.03 + 0.736 D_{t-1} + 0.251 D_{t-2} + Z_t \tag{7}$$

with

$$\begin{cases} Z_t = \sigma_t \varepsilon_t, \\ \sigma_t^{0.44} = 0.481 + 0.377 \left(\varepsilon_{t-1}^{+}\right)^{0.44} + 0.377 \left(-\varepsilon_{t-1}^{-}\right)^{0.44}, \end{cases} \tag{8}$$

where $(\varepsilon_t, t \in \mathbb{Z})$ are independent real random variables with a centered and reduced Gaussian distribution. This model is clearly distinct from a white noise. In fact, the daily difference time series is modeled by a non-centered AR(2) process, D. Moreover, the error process follows an absolute value TARCH(1) model with power index 0.44 [Equation (8)].

The analysis of the residual series of this process D shows its compatibility with a homoscedastic error process; namely, the Lagrange multiplier ARCH test

Table 5. The ARCH test for the residual series of D_t (north−south differences).

ARCH test			
F-statistics	0.468926	Prob. $F(1, 11257)$	0.493496
Obs*R-squared	0.468990	Prob. chi-square (1)	0.493452

Table 6. Test of equality of the northern and southern hemisphere series. Zero cells (0%) have expected frequencies less than 5. The minimum expected cell frequency is 431.3 .

Bin	Observed N	Expected N	Residual
$[-2, -1]$	1201	870.1	330.9
$[-1, -0.5]$	1910	2360.6	−450.6
$[-0.5, 0]$	1398	1399.7	−1.7
$[0, 0.5]$	1085	998.7	86.3
$[0.5, 1]$	740	680.9	59.1
$[1, 2]$	783	824.7	−41.7
$[2, 8]$	449	431.3	17.7
Total	7566		
Test statistics			
	Chi-squared		227.277
	Degrees of freedom		6
	p-value		$< 10^{-3}$

applied to this residual series accepts its homoscedasticity with $p-$value equal to 0.49 (Table 5).

It is interesting to observe that these two series, north and south, not only have different evolution, but their probability distributions are also clearly different. In fact, the chi-squared test applied to the north and south series, centered in advance, clearly rejects the null hypothesis of equality (p-value $< 10^{-3}$, Table 6).

In conclusion, with this study we have shown that the evolution of plage region areas observed in the northern hemisphere is different from that of the southern hemisphere, having neither static nor dynamic symmetry.

3.3. Taylor Effect

The Taylor effect is a property detected in several empirical studies, often observed in real data sets with conditional volatility, in which the autocorrelations of the absolute value time series are larger than those of the squared time series (Taylor, 1986; Granger and Ding, 1995). This property is also a characteristic of

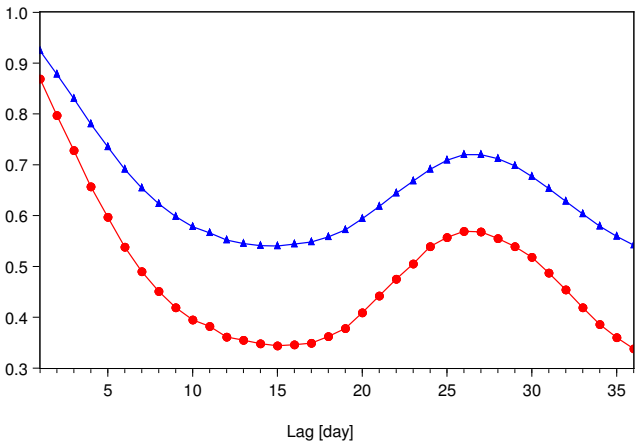

Figure 9. The Taylor effect in the northern hemisphere series. The upper blue curve shows the autocorrelation of the absolute value time series while the lower red curve corresponds to the squared time series.

leptocurtik data which are the case of the present topic (as the curtosis of our series is always greater than 3; see Figure 4). Figure 9 shows clear existence of the Taylor effect in the northern hemisphere data series. A similar plot can be obtained for the southern hemisphere series. Theoretical studies on this property have been carried out in some conditional heteroscedastic models, particularly in threshold ones (Gonçalves, Leite, and Mendes-Lopes, 2009).

4. Conclusions

In this paper we analyzed the plage region areas during solar cycles 21-23, using daily spectroheliograms obtained at the Observatório Astronómico da Universidade de Coimbra. As the analysis requires data equally spaced in time, we interpolated the missing observations (due to weather and instrument conditions), preserving the same statistical properties as the original series.

We found that the time series of plage region areas in the northern hemisphere is clearly different from that in the south hemisphere. This conclusion was obtained from the difference time series and also from the analysis of the temporal evolution of the two series. In fact, these studies led to quite different probabilistic properties in the northern and southern hemispheric data.

Moreover, the performed statistical analysis revealed the presence of heavy tail with high volatility. The residuals from the ARMA model fitted to these series showed strong conditional heteroscedasticity. Furthermore, we observed that TGARCH models with integer power δ do not capture well such volatility. On the contrary, we found that these series have a temporal evolution well characterized by AR models with errors following power TARCH formulations,

with non-integer power index δ. This allows us to conjecture that these series of observations have asymmetric stochastic volatility and long memory.

Further, we pointed out that the presence of the Taylor effect in these series is certainly justified by their leptokurtosis and strong conditional volatility. This feature was introduced by Taylor (1986) after observing its presence in several financial time series. With the present study we showed that this property usually associated with financial data is also clearly present in astronomical data. The results obtained here are in coherence with the conclusions deduced for solar activity from several studies of the classical sunspot series (R.Gonçalves, Pinto, and Stollenwerk, 2009). Recently, Pop (2012) and Noble and Whealtland (2013) pointed out that daily sunspot numbers follow a Laplace distribution, which is a heavy tail distribution. Moreover they discussed strong variability in that data series, justifying it from the complicated local processes associated with sunspot formation, evolution, and decay.

Finally, we would like to emphasize our results in the context of the solar cycle prediction. Although the prediction of solar cycles has been studied using different formalisms and approaches, it is clear that the absence of a full explanation (and a mechanism) to understand the solar cycle introduces a strong limitation in the prediction. The statistical approach remains of the most reliable ones (Hathaway, Wilson, and Reichmann, 1999), namely a time series model is pursued which is well-fitted to data. In addition, taking into account different behavior of solar activity in the northern and southern hemispheres, the development of models describing each one of these series is a positive contribution to prediction improvement. With this kind of models the precision of forecast can be expressed as a function of current and past states of the process. This approach will allow, in particular, the detection of sub-periods of stronger or weaker volatility. The estimation of the probability distribution of these processes, like in Pop (2012) for sunspots, is also an open question and can contribute, in particular, to measure the quality of the produced forecasts.

Acknowledgements We are grateful to an anonymous reviewer for advice that helped us to improve significantly this paper.

References

Badalyan, O. G.: 2012, *Astron. Lett.* **38**, 51.
Bankoti, N. S., Joshi, N. C., Pande, S., Pande, B., Pandey, K.: 2010, *New Astron.* **15**, 561.
Bankoti, N. S., Joshi, N. C., Pande, B., Pande, S., Uddin, W., Pandey, K.: 2011, *New Astron.* **16**, 269.
Box, G.E.P., Jenkins, G.M.; 1976, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, 285.
Bzowski, M., Mäkinen, T., Kyrölä, E., Summanen, T., Quémerais, E.: 2003, *Astron. Astrophys.* **408**, 1165.
Deng, L. H., Qu, Z. Q., Liu, T., Huang, W. J.: 2013, *Adv. Space Res.* **51**, 87.
Ding, Z., Granger, C.W.J., Engle, R.F.: 1993, *J. Empir. Finance* **1**, 83.
Dorotovič, I., Rybak, J., Garcia, A., Journoud, P.: 2010, In: Dorotovič, I. (ed.), *Proceedings of the 20th National Solar Physics Meeting*, Slovak Central Observatory, Hurbanovo, Slovakia, 58.
Duchlev, P. I.: 2001, *Solar Phys.* **199**, 211.

Engle, R.F.:1982, *Econometrica* **50**, 987.
Engle, R.F., Paton, A.J.: 2001, *Quantitative Finance* **1**, 237.
Francq, Ch., Zakoian, J.M.: 2010, *GARCH Models*,Wiley, Chichester, UK, 93.
Georgieva, K., Kirov, B., Tonev, P., Guineva, V., Atanasov, D.: 2007, *Adv. Space Res.* **40**, 1152.
Gigolashvili, M. Sh., Japaridze, D. R., Mdzinarishvili, T. G.: 2011, *Astrophysics* **54**, 593.
Gonçalves, E., Mendes-Lopes, N.: 2008, *Séries temporais. Modelações lineares e não lineares*, Edições Sociedade Portuguesa de Estatística (SPE), Lisbon, 282.
Gonçalves, E., Leite, J., Mendes-Lopes, N.: 2009, *Stat. Prob. Lett.* **79**, 602.
Gonçalves, E., Leite, J., Mendes-Lopes, N: 2012, *Stat. Prob. Lett.* **82**, 1597.
Gonçalves, R., Pinto, A.A., Stollenwerk, N.: 2009, *Astrophys. J.* **691**, 1583.
Gouriéroux, Ch., Monfort, A.: 1990, *Series Temporelles et Modèles Dynamiques*, Economica, Paris, 63.
Gouriéroux, Ch., Monfort, A.: 1995, *Statistics and Econometric Models, Vols.1 & 2*, Cambridge University Press, Cambridge, Vol. 2, 17, 109.
Granger, C.W.J., Ding, Z.: 1995, *Ann. Econom. Statist.* **40**, 67.
Haig, J.D.: 2007, *Living Rev. Solar Phys.* **4** (2). http://solarphysics.livingreviews.org/Articles/lrsp-2007-2/.
Hathaway, D. H.: 2010, *Living Rev. Solar Phys.* **7** (1). http://solarphysics.livingreviews.org/Articles/lrsp-2010-1/.
Javaraiah, J.: 2008, *Solar Phys.* **252**, 419.
Knaack, R., Stenflo, J. O., Berdyugina, S. V.: 2004, *Astron. Astrophys.* **418**, L17.
Li, K. J., Wang, J. X., Xiong, S. Y., Liang, H. F., Yun, H. S., Gu, X. M.: 2002, *Astron. Astrophys.* **383**, 648.
Li, K. J., Gao, P. X., Zhan, L. S., Shi, X. J., Zhu, W. W.: 2009, *Mon. Not. Roy. Astron. Soc.* **394**, 231.
Li, K. J., Liu, X. H., Gao, P. X., Zhan, L. S.: 2010, *New Astron.* **15**, 346.
McIntosh, S. W., Leamon, R. J., Gurman, J. B., Olive, J.-P., Cirtain, J. W., Hathaway, D. H., Burkepile, J., Miesch, M, Markel, R. S., Sitongia, L.: 2013, *Astrophys. J.* **765**, 146.
Newcomb, S.: 1901, *Astrophys. J.* **13**, 1.
Newton, H. W., Milsom, A. S.: 1955, *Mon. Not. Roy. Astron. Soc.* **115**, 398.
Noble, P.L., Wheatland, M.S.: 2013, *Solar Phys.* **282**, 565.
Pan, J., Wang, H., Tong, H.: 2008, *J. Econom.* **142**, 352.
Pop, M.-I.: 2012, *Solar Phys.* **276**, 351.
Sen, A. K.: 2009, *Adv. Space Res.* **44**, 1321.
Silverman, B.: 1998, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London, p. 34.
Sýkora, J., Rybák, J.: 2010, *Solar Phys.* **261**, 321.
Taylor, S.:1986, *J. Econom. Dynam. Control* **18**, 931.
Temmer, M., Veronig, A., Hanslmeier, A.: 2002, *Astron. Astrophys.* **390**, 707.
Temmer, M., Rybák, J., Bendík, P., Veronig, A., Vogler, F., Otruba, W., Pöotzi, O., Hanslmeier, A.: 2006, *Astron. Astrophys.* **447**, 735.
Vats, H. O., Chandra, S.: 2011, *Mon. Not. Roy. Astron. Soc.* **413**, L29.
Weiss, A.A.: 1984, *J. Time Series Anal.* 5, 129.
Xie, J.-L., Shi, X.-J., Xu, J.-C.: 2012, *Res. Astron. Astrophys.* **12**, 187.
Zhang, J., Woch, J., Solanki, S. K., von Steiger, R.: 2002, *Geophys. Res. Lett.* **29**, 77.