Monte Carlo filtering for the detection of phase-coherent solar signal in the low-frequency domain.

G. Grec*and C. Renaud*

October 8, 2002

Abstract

The detection of the global solar modes is mostly made thanks to the associated local maximum in the power spectrum of the signal. This analysis has not been very successful below 1.3 mHz, due probably to the random signals generated by the surface perturbations emerging from the convective zone. We propose an heuristic description of an alternate analysis using a mean phase evaluation over a set of cross-correlation of 2 randomly generated complementary functions, themselves extracted from the helioseismic signal. This analysis is used to explore the frequency range below 1.5 mHz, where we expect the gravity modes together with the low-order acoustic modes. The first preliminary results obtained between 100 μHz and 300 μHz for GOLF [3] are briefly shown.

1 Introduction.

One of the major objectives of the helioseismic instruments aboard the SoHO space solar observatory [1] is to determine the frequency of the low radial order acoustic modes and of the gravity modes. These modes are all related to the deep layers of the sun, down to the core. Then the measurement of their frequency will constrain more tightly models of the internal structure. During the study of the instrumental specifications, the detection of low-frequency solar oscillations (below 1 mHz) was supposed to be a simple problem of the detection of periodic signals and to be achievable as the result of a long set of observations. With the assumption of a solar noise composed of random components coming from all possible other sources, a simple periodic signal should grows up above the background in the observed power spectrum as the length of the time series grows.

Data are now available almost continuously since Feb. 1996, but little major progress concerning solar mode detection at low frequencies has resulted. The initial statement is in fact related to the simple hypothesis of solar modes being detectable as perfectly coherent periodic oscillations. After 5 years of SoHO operations, the apparent failure of this simple approach is questionable. We can then take it for granted that the spectral analysis has reached a limit: no stable frequency can be identified apart from the acoustic modes, for which the level of excitation is high enough to produce a broad energy peak above the noise. At low frequency, the energy of a mode decreases, the life time increases and finally the p-mode sequence is lost below the background noise. In this part of the spectrum, below 1.2 mHz, an improved method of detection is needed.

2 The helioseismic signal.

We use here 5 years of data coming from GOLF [3], covering the period from Apr. 1996 to Apr. 2001, with 2 significative gaps in 1998. The photometric signal is normalized for the energy in the frequency domain of the "5 min" acoustical modes. This calculation gives a function $S(t)$, corrected for the solar distance, the radial velocity and the instrumental drifts. No assumption is made about the physical origin of the signal [6]. A relation of

*Département Cassini, UMR 6529 du CNRS, Observatoire de la côte d'azur, 06304 Nice, France.

this signal to the disk-averaged velocity can only be made for the p modes. The other components are related to any source of variation of the monochromatic intensity, to the instrumental noise, and to the photon noise.

3 Method of analysis.

Let first consider the benefit of an hypothetical stereoscopic observation made from 2 planets on opposite sides of the Sun: the signal of the global oscillations like the acoustical modes would be correlated, while the signal from the convection or other local stochastic sources should be independent. The correct way to detect the components common to the 2 observations is to compute the cross-correlation of the 2 signals, noted $F(t)$ and $G(t)$. In Fourier space, the related function is the cross-spectrum (CS), i.e. the complex product of the 2 Fourier transforms (FT). When 2 signals at a given frequency are related in time, the time lag corresponds to a stable relative phase for the corresponding spectral components, the phase being measurable in the CS. We extend this notion to the case of the single signal from the observable solar disk.

In order to make an analysis on several independent observations, the function $S(t)$ is split into a series $S_i(t)$ on successive 101 d long time intervals, then $1 \leq i \leq 17$. The 101 d duration of a single observation results in an acceptable frequency resolution.

Each function $S_i(t)$ is then used independently to make a series of N experiments using the functions $F_{i,n}(t)$ and $G_{i,n}(t)$, complementary in time, as defined in Eq. 1 and Eq. 2, where $0 \leq n \leq N$.

$$F_{i,n}(t) + G_{i,n}(t) = S_i(t) \tag{1}$$

$$F_{i,n}(t) \,.\, G_{i,n}(t) = 0 \tag{2}$$

To be free of systematic effects, as the obvious components related to a periodic window, we use random numbers to produce a set of functions $\text{Ш}_n(t)$ randomly distributed in time to compute $F_{i,n}(t)$ and $G_{i,n}(t)$ (Eq. 3 and Eq. 4).

$$F_{i,n}(t) = S_i(t) \,.\, (\Pi_P(t) * \text{Ш}_n(t)) \tag{3}$$

$$G_{i,n}(t) = S_i(t) \,.\, (1 - \Pi_P(t) * \text{Ш}_n(t)) \tag{4}$$

where $\Pi_P(t) = 1$ for $0 \leq t \leq P$ and $\Pi_P(t) = 0$ elsewhere, $\text{Ш}_n(t) = \sum_{k=1}^{\infty} \epsilon_{k,n}\, \delta(t - kP)$ where $\delta(t)$ is the Dirac distribution and the random series $\epsilon_{k,n} = 1$ or $\epsilon_{k,n} = 0$ with the probability distribution $\psi(0) = \psi(1) = 0.5$.

Are the $F_{i,n}(t)$ and $G_{i,n}(t)$ statistically independent? The heuristic answer is "likely yes" for the stochastic components (in the simple example of a system of fringes due to a pair of random pulses, the signal still exists in 1 case over 4 when the 2 pulses are in a given window, but in this case they disappear from the complementary signal and then the fringes disappear in the CS. On the other hand, for any periodic component, the signals in $F_{i,n}(t)$ and $G_{i,n}(t)$ remain identical.

As $F_{i,n}(t)$ and $G_{i,n}(t)$ are almost simultaneous, the CS $\mathcal{P}_{i,n}(\nu)$ (Eq. 5) is a real number for any frequency related to a periodic component.

$$\mathcal{P}_{i,n}(\nu) = \mathcal{F}_{i,n}(\nu) \,.\, \mathcal{G}^*_{i,n}(\nu) \tag{5}$$

where $\mathcal{F}_{i,n}(\nu)$ is the TF of $F_{i,n}(t)$ and $\mathcal{G}^*_{i,n}(\nu)$ is the conjugate of the TF of $G_{i,n}(t)$.

Practically a small imaginary component in $P_{i,n}(\nu)$ should be expected, due to the noise in the signal at the same frequency and to the small random difference of the time averages over the 2 windows. We then search for any component following Eq. 6 and Eq. 7.

$$\Re(\mathcal{P}_{i,n}(\nu)) \;> 0 \tag{6}$$

$$\mid \Im(\mathcal{P}_{i,n}(\nu)) \;/\; \Re(\mathcal{P}_{i,n}(\nu)) \mid < d \tag{7}$$

In our present work, the duration of the elementary window is $P = 6\,\text{h}$, the number of experiments for each 101 d sample is $N = 4500$, the maximum selected for $\Im/\Re$ is $d = 0.1$. The spectral resolution of the calculation is $\rho = 3.6\text{nHz}$.

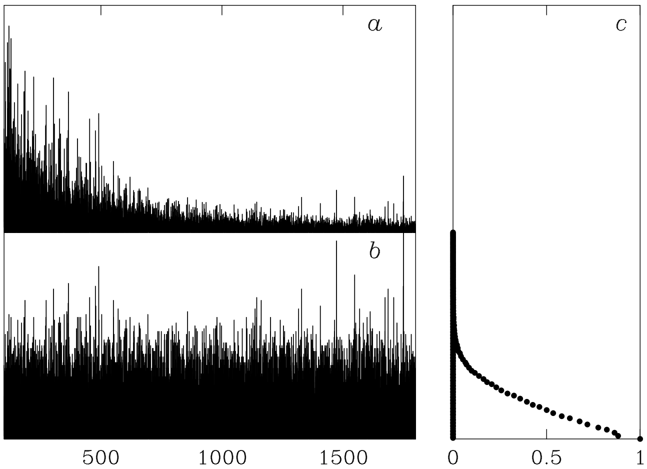

Figure 1: Example of detection for a 101 days long data segment, 1500 random experiments with a 6 h elementary window are used. • a : $A_{i,N}(\nu)$ signal amplitude • b : $R_{i,N}(\nu)$ number of detected events • c : amplitude distribution of $R_{i,N}(\nu)$. Scales are arbitrary except b horizontal scale, frequency in μHz.

4 Comments on artifacts.

An FFT algorithm gives a sampled FT, the sampling in frequency being related to the length of the input file. For a periodic component, the width of the line depends on the real duration of the signal. There is obviously no relation between any frequency we are looking for and the duration of the observation defining the frequency resolution. We have then to over sample the calculation of the FT, as we do in the simple case of the calculation of a spectral density.

The calculation of $F_{i,n}(t)$ and $G_{i,n}(t)$ from $S_i(t)$ (Eq. 1 and Eq. 2) obviously produces sidelobs in $\mathcal{F}_{i,n}(\nu)$ and $\mathcal{G}_{i,n}(\nu)$. As $\mathcal{F}_{i,n}(\nu) + \mathcal{G}_{i,n}(\nu) = \mathcal{S}_i(\nu)$ those sidelobs are opposite, resulting in the CS in a negative number rejected in the Eq.6. Then the sidelobs can not produce a false detection. Nevertheless, blind intervals associated to the strong spectral components in the signal could be present in the CS, accordingly to the duration P of the elementary window.

5 Signal detection.

For each function $\mathcal{P}_{i,n}(\nu)$ we select the bins for which Eq. 7 is satisfied and count $R_{i,N}(\nu)$ the total number of selected bin for the N experiments. For the same selected bins, we calculate $A_{i,N} = \sum_{n=1}^{N} \Re(\mathcal{P}_{i,n}(\nu))$. A typical result is shown Fig.1: $R_{i,N}(\nu)$ shows several lines above a random distribution. $A_{i,N}(\nu)$ increases at low frequency, similarly to the plain PS of $S_i(t)$.

In order to identify the frequency intervals where the signal goes above the noise, we define for each section a cumulative function $S_j(f_j) = \sum_{n=1}^{N} \sum_{\nu=f_j}^{f_j+w} R_{i,N}(\nu)$ on successive frequency intervals of width $w = 0.2\,\mu$Hz (corresponding to a 50 d lifetime for the signal). After the selection of the higher values of $S_j(f_j)$, we obtain a short set of discrete values Φ_j for the frequency f_j.

Finally, a solar signal related to a resonant mode is suspected if the same frequency Φ_j exists in the result of the calculation for several temporal sections $S_i(t)$ of the original data (Tab.1). We can then calculate a weighted frequency of the detected lines, as

$$\Phi_j = \sum_{n=1}^{N} \sum_{\nu=f_j}^{f_j+w} \nu R_{i,N}(\nu) \Bigg/ \sum_{n=1}^{N} \sum_{\nu=f_j}^{f_j+w} R_{i,N}(\nu) \tag{8}$$

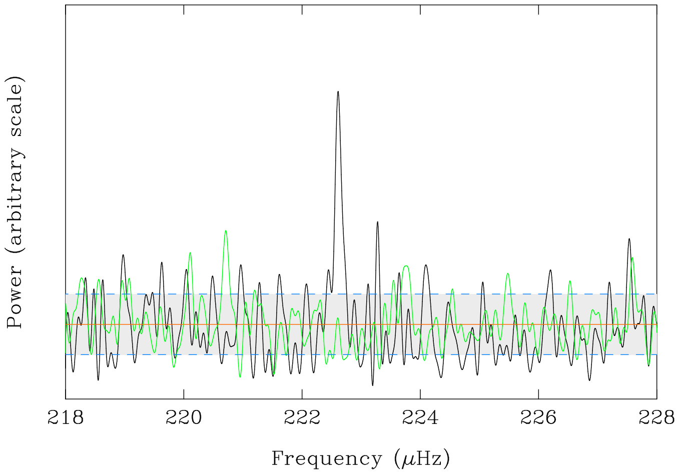

Figure 2: Typical result deduced from Tab. 1 for the solar g mode $l = 2$, $n = 3$. We calculate the averaged PS $f_1(\nu)$ for 5 time intervals(black solid line). It is then compared to the counterpart average $f_2(\nu)$ for the other time intervals (green line). A sharp line is visible in $f_1(\nu)$. The mean (solid horizontal) and rms (blue dashed horizontals) are estimated on the plotted interval for the PS outside the line. The max. of the line is above 7 σ.

6 Use for power spectrum averaging.

At this point, we can simply go back to the standard PS. The additional information we have now extracted from the phase of the CS indicates the segments $S_i(t)$ (defined in Sec. 3) in which a given periodic signal is detected. We can then compare 2 averaged PS: Firstly we select the sections of signal on which a given solar period is detected and compute the mean PS. Secondly we compute the mean PS for the remaining observations.

Fig. 2 is an example of a peak in the mean PS of the selected sections compared to a flat power for the other time section. The statistical significance of this result is the next point to study: as we use now an averaged PS, the distribution of amplitude of a detected line is a gaussian (central limit theorem). Then this simple analysis allows to detect lines significantly above the noise after the selection of the PS. Fig. 3, 4 and 5 give the mean PS for frequency domains of observed or expected values of 3 p modes. The detection of the p-mode doublet shows that our fully automated analysis cross-check the result resulting from a careful selection of time sections made by observers following the frequency recurrence of the solar p modes [7].

7 Frequency of the detected oscillations.

The value ϕ_j, given Tab. 1, is not a true measurement for the solar mode frequency. Several choices exist to define a mean frequency for the detected signal, involving $R_{i,N}(\nu)$ and $A_{i,N}(\nu)$ defined Sec. 5 or not.

The averaging of PS (Sec. 6) is a simple application of the CS analysis, giving results directly comparable with the classical determination of the frequency of a solar mode. In most analysis the frequency determination results from the fit of a Lorentz profile to the PS (or in complex form the FT). The multiplet structure of non-radial modes is included in the fit, giving the central frequency and the rotational splitting of the solar made. This approach assumes that successive points are statistically independent, excluding any averaging or over sampling of the FT.

On opposite, $R_{i,N}(\nu) = 0$ and $A_{i,N}(\nu) = 0$ with a very hight probability, apart from the detected lines. The frequencies in Tab.1 are calculated with Eq. 8. They are given as an early indication, the exact values should come from a more detailed analysis.

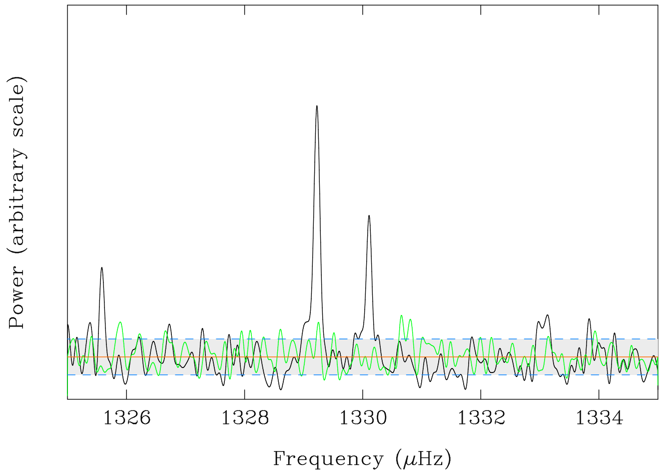

Figure 3: Same as Fig. 2 for the solar p mode $l = 1$, $n = 8$ (theoretical value of the central frequency of the mode 1329.5 μHz [9]). The 2 components due to the rotationnal splitting of this solar mode are extracted from the background noise.

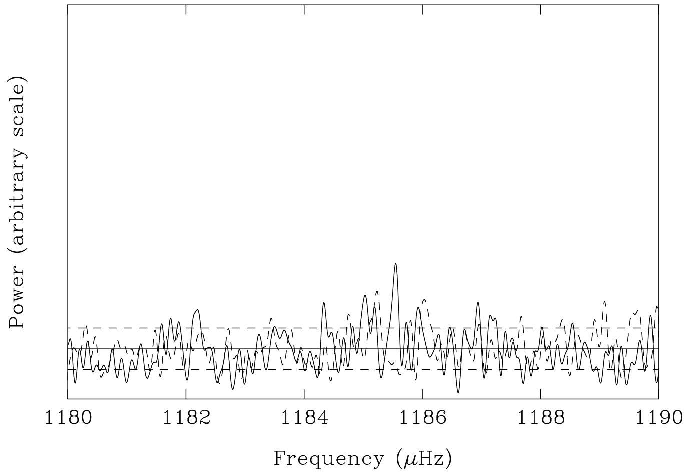

Figure 4: Result for the solar p mode $l = 1$, $n = 7$ (theoretical value of the central frequency of the mode 1185.4 μHz [9]).

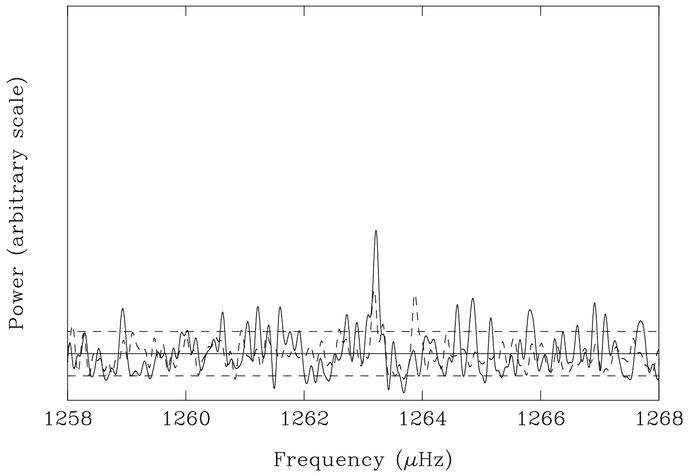

Figure 5: Result for the solar p mode $l = 0$, $n = 8$ (theoretical value of the central frequency of the mode 1263.7 μHz [9]).

Table 1: Frequency table of the detected oscillations (plain synodic measurement). The heading integers are the order of the temporal section $S_i(t)$, each section in which the oscillation is detected is marked with a number in the range 1 to 9, proportional to the number $R_{i,N}(\nu)$ of detected events (see Sec. 5). The last columns give the closest frequency and mode for degrees $l \leq 3$ from [9]. The double lines are more likely broad lines rather than split, taking into account the frequency window $w = 200\,\mathrm{nHz}$ (Eq. 8).

μHz	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	μHz	mode
128.87					5				2	2					1			128.13	$g4$, $l=1$
147.29	5		1				2							2				148.88	$g9$, $l=3$
166.10	6													1		3	3		
171.69				4			6	3						4		2		171.06	$g5$, $l=2$
171.94	2		1							1			5					*id.*	(broad line)
174.94	1					2		1							2				
178.53		1		4		2											3	178.12	$g7$, $l=3$
187.33							1			5		1			2				
212.13				3				3							2	2	3		
214.76		3	1	1												1			
222.65					3	2						1	1				2	222.7	$g3$, $l=2$
229.46	1					1	1						4		3				
238.70						9					1	1					2	238.78	$g4$, $l=3$
243.90										3	2		4				5		
262.14				4						1		5				1		262.08	$g3$, $l=3$
264.94	3									2			5			1		263.54	$g1$, $l=1$
265.30	4				1	2								4				*id.*	(broad line)
287.09								4				4	2				2	285.47	$p1$, $l=1$
295.90			1					3			1						3	297.01	$g1$, $l=2$
303.49							1	2	4		2								

8 Comparison with analytic models.

The identification of a solar signal with a solar mode must be supported by an approximate knowledge of a set of frequencies, as was the case the first time a regular spectrum corresponding to p-modes was detected [5]. That is at least true for GOLF, where the sun is observed as a star and no information on the geometrical pattern is available. We use the results from [9] to relate our present results with the calculated frequencies in Tab.1 (periods in Fig.6). Several close coincidences exist[1].

9 Discussion and conclusion.

We there detect a periodic signal using his phase coherence. In comparison the power spectrum used for the classical analysis gives only an information about the energy in a given bandpass: for a periodic signal that is a sharp line, exactly like any fringe due to the random background.

It is clear from the results from 300 μHz to 1000 μHz that the convective background contributes to our filtered signal. Then, the most we can expect for the resonant modes is to detect them with an higher probability than the other pseudo-periodic events, as the resonance may extend the life time and the level of excitation of the related signal. Our first results favour this hypothesis.

The lines in $R_N(\nu)$ have a different amplitude distribution than in a PS. The noise in $R_N(\nu)$ is not frequency dependent, allowing a direct statistical study, free of observer's bias and model fitting.

[1]The margin is not discussed here. The authors [9] estimate the uncertainty related to the solar model used for the calculation at 3 μHz, the rotational splitting extends this range dependently on the degree of the mode.

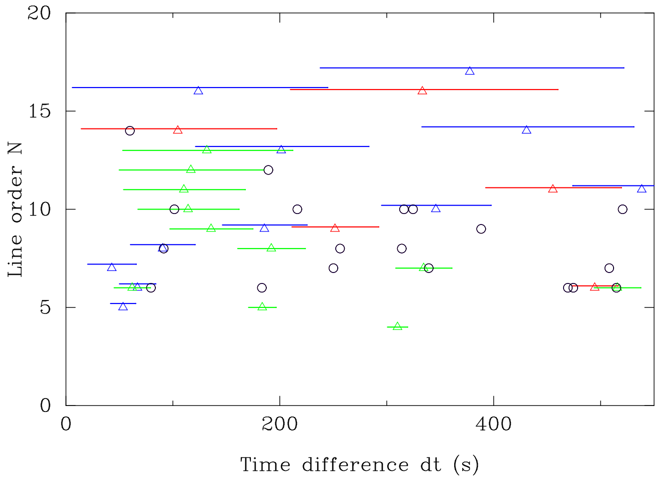

Figure 6: The periods P calculated for the signals detected in the range from 100 μHz to 300 μHz (Tab. 1) are plotted $\circ$. The periods P predicted for the degree l=1 (red), 2 (blue) and 3 (green) solar g modes are plotted $\triangle$, the uncertainty bars correspond to $\pm 1.5\,\mu$Hz. We use a folded periodogram adapted to the asymptotic behaviour of degree l=3 g modes. (The horizontal scale is the difference of periods (dt), the vertical scale is the line order N. The period of a mode is $P = 550\,N + dt$ s.)

At the present level of analysis, we conclude that we have detected of sporadic quasi-periodic signals. Several oscillations at same frequency are detected in different sections of the solar signal. Several but not all of those frequencies correspond to those calculated for the g modes and the low-order p modes.

There is no evidence for a permanent periodic solar signal. The search for oscillations with a very long coherence time, an up to now expected property of the solar modes, seems then not to be an adequate approach. The detection of solar g modes has been somewhat suspected in the PS of the signal [4]. We can now explain why the corresponding peak does not become higher when the observation time increases: our present result indicates that a given mode is hardly visible continuously for as long as half a year.

The physical nature of the GOLF signal is unknown, apart from the p-mode oscillations for which the velocity component has been identified [8] and [10]. A forthcoming analysis of the measurements from SOI [11] and VIRGO [2] should give complementary informations, as they give data on the surface velocity and the brightness variations averaged over the solar disk.

Acknowledgements: The GOLF instrument is due to a team issued from a consortium of institutes. SOHO is a space solar observatory issued from a common project from ESA and NASA. This work has been performed using the computing facilities provided by the SIVAM program for medium power computer from the Observatoire de Nice.

References

[1] Domingo, V., Fleck, B., Poland, A.I., 1995 *Sol. Phys.*, **162**, 1.

[2] Fröhlich, C. and the VIRGO team, 1995, *Sol. Phys.*, **162**, 101.

[3] Gabriel, A.H. and the GOLF team, 1995, *Sol. Phys.*, **162**, 61.

[4] Gabriel, A. H., Turck-Chièze, S., García, R. A., Pallé, P. L. Boumier, P., Thiery, S., Baudin, F., Grec, G., Ulrich, R. K., Bertello, L., Roca Cortés, T., Robillot, J.-M., 1999, *Advances in Space Research*, **24**, 147.

[5] Grec, G., Fossat, E., Pomerantz, M., 1980, *Nature*, **288**, 441.

[6] Grec, G. & Renaud, C., "Reduction of the Golf signal",1999, unpublished.

[7] Lazrek, M., Baudin, F., Bertello, L., Boumier, P., Charra, J., Fierry-Fraillon, D., Fossat, E., Gabriel, A.H., García, R.A., Gelly, B., Gouiffes, C., Grec, G., Pallé, P.L., Pérez Hernández, F., Régulo, C., Renaud, C., Robillot, J.M., Roca Cortés, T., Turck-Chièze, S., Ulrich, R.K., 1997, *Sol. Phys.* **175-2**, 227.

[8] Pallé, P.L., Régulo, C., Roca Cortés, T., Garcia, R.A., Jiménez, A., Boumier, P., Gabriel, A.H., Grec, G., Robillot, Turck-Chièze, S., Ulrich, R.K., & the SOI Team: 1999, *Astron. Astrophys.*, **341**, p625.

[9] Provost, J., Berthomieu, G., Morel, P., 2000, *A.&A.*, **353**, 775.

[10] Renaud C., Grec, G., Boumier, P., Gabriel, A.H. ,Robillot, J.M., Roca Cortès, T., Turck-Chièze S., Ulrich, R.K. 1999, *Astron. Astrophys.*, **345**, 1019.

[11] Scherrer, P.H. and the SOI team, 1995, *Sol. Phys.*, **162**, 129.