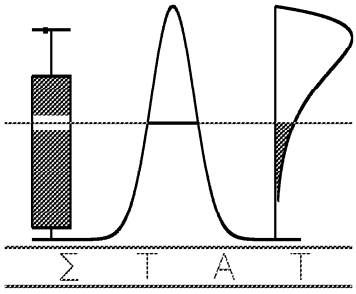
T E C H N I C A L

R E P O R T

0239

Revisiting the moment method as a mode identification technique

J. De Ridder, G. Molenberghs, and C. Aerts



I A P S T A T I S T I C S

N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Revisiting the Moment Method as a Mode Identification Technique

Joris De Ridder, Geert Molenberghs and Conny Aerts *

January 23, 2003

Abstract

The moment method is a well known mode identification technique in asteroseismology, which uses a time series of the first 3 moments of a spectral line to estimate the discrete mode parameters ℓ and m. The method, contrary to many other mode identification techniques, also yields other important real-valued parameters such as the inclination angle i, and the rotational velocity v_e. We developed a statistical formalism for the moment method based on so-called generalized estimating equations (GEE). This formalism allows the estimation of the uncertainty of the real-valued parameters taking into account that the different moments of a line profile are correlated and that

*Joris De Ridder is post-doctoral researcher, Institute of Astronomy, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium (email: joris@ster.kuleuven.ac.be). Geert Molenberghs is Professor, Center for Statistics, Limburgs Universitair Centrum, Universitaire Campus, B-3590 Diepenbeek, Belgium (email: geert.molenberghs@luc.ac.be). Conny Aerts is Associate Professor, Institute of Astronomy, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium (email: conny@ster.kuleuven.ac.be).

the uncertainty of the observed moments also depends on the model parameters. Furthermore, we set up a procedure to take into account the mode uncertainty, i.e., the fact that often several modes (ℓ, m) can describe the data. We applied our method to the star HD181558, from which we learned that numerically solving the estimating equations is an intensive task. We report on the numerical methods we use and we end with the results for HD181558 together with our conclusions.

Keywords: Generalized estimating equations, time series, sandwich estimator, astrostatistics, discriminant function

1 Introduction

Asteroseismology is the exploration of the stellar interior with stellar ocillations. One uses the fact that most stars oscillate, in a regular way, and with several frequencies simultaneously. These oscillations manifest themselves at the surface of the star through brightness variations, temperature variations, and surface velocity variations, which can be observed. A star can only oscillate in one or several of its "natural" frequencies which are determined by the internal structure of the star. With suitable inversion techniques it is possible to use the observed frequencies to derive information about the internal structure of the star.

To do so, however, the characteristics of the oscillations need to be known first. That is, a *mode identification* has to be carried out, where one estimates the parameters characterising the oscillations from observational data. There are only few mode identification techniques, and the properties of their estimators are hardly if at all studied. Statistical uncertainties of the estimates, for example, are never reported.

Nevertheless, from an astrophysical point of view, uncertainties are important because wrong mode indentifications can mislead inversion techniques. It is therefore necessary to know a priori exactly how wrong one can be with the estimates.

In this paper, we present an in-depth study of the statistical properties of one particular mode identification technique, namely the so-called *moment method*. For examples of applications of this method we refer to, e.g., Aerts et al. (1998), Uytterhoeven et al. (2001), Aerts & Kaye (2001), and Chadid, De Ridder, Aerts & Mathias (2001).

2 Astrophysical Background

As in any inferential method, the moment method uses a theoretical model to describe the observations. To understand the statistically relevant properties of this theoretical model and its parameters, we first need to briefly discuss some of the physics of stellar oscillations and how they can be observed.

In Figure 1 a diagrammatic illustration of the surface of an oscillating star is given. For our application, the most important aspect of stellar oscillation is the surface velocity. The lighter parts of the stellar surface have an inward velocity while the darker parts have an outward velocity. The figure is only a snapshot: the star is varying periodically and half an oscillation cycle later the situation is reversed with the lighter parts moving outward and the darker parts moving inward. For slowly-rotating oscillating stars, each of the oscillation modes can be described with a spherical harmonic Y_ℓ^m which is what we actually drew in Figure 1. In reality, the motion of a surface element is more complex because it not only moves vertically but also

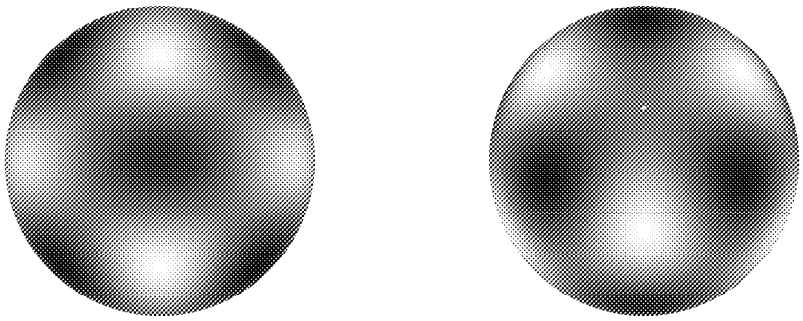


Figure 1: A diagrammatic illustration of the surface of an oscillating stars for the mode $(\ell, m) = (5, 3)$. In the left picture we look at the equator, and in the right picture we almost look at the pole of the star.

horizontally.

In terms of model parameters we have to estimate 3 unknown parameters per oscillation mode: 2 discrete parameters and 1 real-valued parameter. The 2 discrete parameters are the so-called mode numbers ℓ and m of the spherical harmonic, which describe the configuration of the inward and outward going regions. To describe the 3-dimensional motion of the stellar matter, only one parameter is needed: the amplitude v_p of the vertical motion. The reason is that there exists a theoretical linear relation between the amplitude of the vertical motion and the amplitude of the horizontal motion. To compute the constant of proportion K, however, the mass and the radius of the star is required and these quantities are often not very accurately known. Nevertheless, in what follows we will assume, as a first approximation, that K is known, to considerably simplify the treatment.

A last real-valued parameter related to the oscillation is the oscillation period. However, for good datasets, this oscillation period can often be quite accurately determined from the data with other methods so that it is usually not counted as another unknown real-valued parameter during mode identification.

In the model, 2 additional unknown parameters not connected to the oscillations are present. A first one is the rotational velocity at the equator of the star, usually denoted as v_e. The second one is the inclination angle i under which we observe the star. This is illustrated in Figure 1. Both pictures show the same Y_ℓ^m, but on the left hand side we are looking on the equator, while on the right hand side we are looking almost on the pole. Clearly, i has a large impact on how the surface velocity field is observed.

A last unknown model parameter is specifically related to the kind of observational data we use. In the case of the moment method it concerns high-resolution spectroscopic data. The gathered star light is decomposed into its colours so that a detailed spectrum can be constructed, i.e., received light flux as a function of the wavelength of the light. At certain wavelenghts, such a spectrum contains absorption lines where the light has been partially blocked by certain chemical elements at the surface of the star. An example of the Si^+ absorption line at $\lambda = 412.805$ nm for the non-radially oscillating star HD181558 is shown in the left hand panel of Figure 2. Here, an observational time series of 30 high-quality spectra gathered by De Cat and Aerts (2002) is shown. The oscillations in the star cause the absorption line to change its position and shape in time. Precisely these line profile variations are used to estimate the parameters mentioned above. To model them, another unknown parameter is needed, denoted by σ, which is related to the width the line profile would have in the absence of pulsation. From an astrophysical point of view this parameter is unimportant and can be considered a nuisance.

Modeling the line profiles themselves turns out to be *very* computationally expensive. That is why Balona (1986) deviced the moment method, which replaces each line

profile by three numbers: the first, the second, and the third moment denoted by y_1, y_2, and y_3 respectively. These quantities are measures for the average position, the width and the skewness of the line profile. In practice, no moments higher than y_3 are considered because the higher-order observed moments are often too noisy and they unduly complicate the calculations. It is common practice to express the moments with the unit $(\mathrm{km/s})^n$, $n = 1, 2, 3$. A time series of theoretical moments can be computed much faster than a time series of theoretical line profiles. The nuisance parameter σ, however, remains. In the right hand panels of Figure 2 we show a time series of the three moments for the star HD181558.

A theoretical moment at one point of time is computed by integrating over the contributions of all points on the visible stellar surface. The expressions for the moments can be written in closed form (Aerts et al. 1992), but they are quite lengthy and hardly of any practical use to computing derivatives. We preferred to use different expressions which are computationally more advantageous, but which involve an integration with integration bounds that depend on the inclination angle i.

Since each of the model parameters has a physical meaning, there are restrictions on their values. In order of importance we have: $\ell \in \{0, 1, 2, \cdots\}$, $m \in \{-\ell, \cdots, 0, \cdots, +\ell\}$, $i \in [0^\circ, 360^\circ[$, $v_e \geq 0$, $v_p > 0$, and $\sigma > 0$.

3 Current Statistical Status

The moment method is a multi-response problem where a time series of 3 responses is used to extract 6 (i.e., 2 discrete + 4 real-valued) parameters. In what follows we will

use the notation

$$\mathbf{Y}_i \equiv (y_1(t_i),\ y_2(t_i),\ y_3(t_i))' \tag{1}$$

and

$$\boldsymbol{\mu}_i \equiv (\mu_1(t_i,\ell,m,\boldsymbol{\beta}),\ \mu_2(t_i,\ell,m,\boldsymbol{\beta}),\ \mu_3(t_i,\ell,m,\boldsymbol{\beta}))' \tag{2}$$

for the first three observed and theoretical moments respectively, at time point t_i ($i = 1, \cdots, n$), where we grouped the real-valued parameters in the vector $\boldsymbol{\beta} \equiv (v_p, \sigma, v_e, i)'$.

It is important to understand how the moment method is currently used. Theoretically it can be shown that for a monoperiodic star, the time dependence of the moments takes the following form:

$$\mu_1 = a_1\ \sin(2\pi\nu t + \alpha_1), \tag{3}$$

$$\mu_2 = b_0 + b_1\ \sin(2\pi\nu t + \delta_1) + b_2\ \sin(4\pi\nu t + \delta_2), \tag{4}$$

$$\mu_3 = c_1\ \sin(2\pi\nu t + \gamma_1) + c_2\ \sin(4\pi\nu t + \gamma_2) + c_3\ \sin(6\pi\nu t + \gamma_3), \tag{5}$$

where ν is the oscillation frequency. The phases α_1, δ_i, γ_j are constants, while the amplitudes a_1, b_i, c_j depend on the parameters $(\ell, m, \boldsymbol{\beta})$. A *discriminant* $\Gamma_l^m(\boldsymbol{\beta})$ is constructed to estimate these parameters by comparing the observed amplitudes with their theoretical counterparts:

$$\Gamma_\ell^m = \sqrt{\left(f_{\tilde{a}_1}\ |\tilde{a}_1 - a_1|\right)^2 + \sum_{i=0}^{2}\left(f_{\tilde{b}_i}\ \sqrt{|\tilde{b}_i - b_i|}\right)^2 + \sum_{i=1}^{3}\left(f_{\tilde{c}_i}\ \sqrt[3]{|\tilde{c}_i - c_i|}\right)^2}, \tag{6}$$

where the tilde denotes observed quantities, and where the weights f are introduced to incorporate the estimated standard errors $\Delta\tilde{a}_1$, $\Delta\tilde{b}_i$ and $\Delta\tilde{c}_j$ of the corresponding observed amplitudes:

$$f_{\tilde{a}_1} \equiv W^{-1}\ \frac{\tilde{a}_1}{\Delta\tilde{a}_1}, \qquad f_{\tilde{b}_i} \equiv W^{-1}\ \frac{\tilde{b}_i}{\Delta\tilde{b}_i}, \qquad f_{\tilde{c}_i} \equiv W^{-1}\ \frac{\tilde{c}_i}{\Delta\tilde{c}_i},$$

$$W \equiv \frac{\tilde{a}_1}{\Delta\tilde{a}_1} + \frac{\tilde{b}_0}{\Delta\tilde{b}_0} + \frac{\tilde{b}_1}{\Delta\tilde{b}_1} + \frac{\tilde{b}_2}{\Delta\tilde{b}_2} + \frac{\tilde{c}_1}{\Delta\tilde{c}_1} + \frac{\tilde{c}_2}{\Delta\tilde{c}_2} + \frac{\tilde{c}_3}{\Delta\tilde{c}_3}$$

(Aerts 1996). Expression (6) for Γ_ℓ^m avoids that the third moment y_3 with its large values overshadows the first moment y_1, but it has the disadvantage that it cannot discriminate the sign of the mode number m. The parameters are estimated by searching for the minimum of Γ_ℓ^m in a rectangular grid in the parameter space. For the real-valued parameters, it is hoped for that the grid is fine enough in order not to miss the global minimum. Finally, a table is produced with the top 5 or 6 best fitting $(\ell, m, \boldsymbol{\beta})$ parameter sets.

The best strategy to obtain the final estimate for (ℓ, m) and $\boldsymbol{\beta}$ together with their uncertainties, is currently open to debate. Despite the usefulness of the table with the best parameter sets, it was never attempted to estimate the uncertainties of the parameters obtained with the moment method, because of the severe difficulties involved. In this work, we set a first step by trying to estimate the uncertainties of the real-valued parameters $\boldsymbol{\beta}$. Estimating the uncertainties of the discrete parameters ℓ and m is a more challenging problem, and will be left for future research.

Even for a given (ℓ, m) value, it is currently unknown how precise the real-valued parameters are estimated. For example, is the uncertainty of the inclination angle as small as 5°, or is perhaps 30° a more typical value? Moreover, very often several (ℓ, m) couples give almost equally good fits. The question is raised as to how should we take this into account for our best estimate of $\boldsymbol{\beta}$ and its uncertainty? In what follows, we will try to answer these questions.

4 New Statistical Approach

We need a new estimating method which not only produces point estimates, but in addition interval estimates.

A first step is to explore the use of the least squares method. Of course the three responses y_1, y_2 and y_3 are dependent, and their covariance matrix $\mathbf{V}$ is unknown. As Seber and Wild (1989) show, the multi-response least-squares method allows one to compute consistent estimates of the covariance matrix of the responses, the model parameters, and the covariance matrix of the model parameters. However, this is only true if $\mathbf{V}$ does not depend on the parameters $\boldsymbol{\beta}$, which is not the case for the moment method. For example, the uncertainty of the first moment of a line profile (y_1) can be estimated with the second moment (y_2), and the latter depends on ℓ, m, and $\boldsymbol{\beta}$. Or, with an astrophysical example, the faster the star rotates (larger v_e) the broader the line profile, and the less precise we know the position or the first moment of the line profile. Therefore, in the case of the moment method, minimizing the weighted sum of squares is not appropriate. In particular, ignoring the $\boldsymbol{\beta}$ dependence of $\mathbf{V}$ deteriorates the efficiency of the estimator.

The *generalized estimating equations* (GEE) methodology, as developed by Liang and Zeger (1986), is better suited for the purpose of the moment method. We recall that this method does not assume a particular *joint* probability density for the responses y_1, y_2 and y_3, nor that they are i.i.d. The theory neither assumes that the theoretical model is linear in its parameters, and the exact covariance matrix $\mathbf{V}$ of the responses does not need to be known in full detail. The method *does* assume, however, that the different observations $\mathbf{Y}_i$ ($i = 1, \cdots, n$) are independent, that a working approximation of the covariance matrix of the responses is available, and that the expectation values

$E[\mathbf{Y}_i] \equiv \boldsymbol{\mu}_i(\ell, m, \boldsymbol{\beta})$ $(i = 1, \cdots, n)$ are correctly specified.

We will use GEE to estimate the uncertainties of the real-valued parameters $\boldsymbol{\beta}$. We recall that in the GEE method, the parameters are estimated by locating the root of the quasi-score function $\mathbf{U}(\boldsymbol{\beta})$:

$$\mathbf{U}(\boldsymbol{\beta}) \equiv \sum_{i=1}^{n} \mathbf{D}_i^t \cdot \mathbf{W}_i^{-1} \cdot (\mathbf{Y}_i - \boldsymbol{\mu}_i), \tag{7}$$

where n is the size of the time series. The 3×4 matrix $\mathbf{D}$ is defined by $D_{rs} \equiv \partial \mu_r / \partial \beta_s$, and the 3×3 symmetric matrix $\mathbf{W}_i$ is a working approximation of the true covariance matrix $\mathbf{V}_i$ of the quantities $\mathbf{Y}_i$:

$$\mathbf{V}_i \equiv E[(\mathbf{Y}_i - \boldsymbol{\mu}_i(\boldsymbol{\beta}))(\mathbf{Y}_i - \boldsymbol{\mu}_i(\boldsymbol{\beta}))^t], \tag{8}$$

where $\boldsymbol{\beta}$ are the true (but unknown) parameters. It can be shown (e.g., Liang and Zeger 1986; Zeger and Liang 1986; Diggle, Heagerty, Liang and Zeger 2002) that the root $\hat{\boldsymbol{\beta}}$ is a consistent and asymptotically normal estimate of the true $\boldsymbol{\beta}$, with sandwich covariance matrix

$$\mathrm{Cov}[\hat{\boldsymbol{\beta}}] = \mathbf{I}_0^{-1}\ \mathbf{I}_1\ \mathbf{I}_0^{-1}, \tag{9}$$

where we used the abbreviations

$$\mathbf{I}_0 \equiv -E\left[\frac{\partial \mathbf{U}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right] = \sum_{i=1}^{n} \mathbf{D}_i^t\ \mathbf{W}_i^{-1}\ \mathbf{D}_i \tag{10}$$

and

$$\mathbf{I}_1 \equiv \mathrm{Cov}[\mathbf{U}(\boldsymbol{\beta})] = \sum_{i=1}^{n} \mathbf{D}_i^t\ \mathbf{W}_i^{-1}\ \mathbf{V}_i\ \mathbf{W}_i^{-1}\ \mathbf{D}_i. \tag{11}$$

The unknown covariance matrices $\mathbf{V}_i$ in the expression for $\mathbf{I}_1$ are estimated by

$$\hat{\mathbf{V}}_i = (\mathbf{Y}_i - \boldsymbol{\mu}_i(\hat{\boldsymbol{\beta}})) \cdot (\mathbf{Y}_i - \boldsymbol{\mu}_i(\hat{\boldsymbol{\beta}}))^t. \tag{12}$$

The so-called sandwich estimator in expression (9) makes the estimate of the covariance matrix of $\boldsymbol{\beta}$ more robust against misspecification of the covariance matrix of the responses.

For the working approximation $\mathbf{W}$ for the covariance matrix $\mathbf{V}$, we suggest the following idea. Consider the mirror image $w(v,t) = 1 - p(v,t)$ of the normalized spectral line $p(v,t)$ (where the abscissa values are expressed in the unit km/s) as a distribution function for the variable v, and compute, for a given time t_i:

$$\begin{aligned} W_{rs} &= E[(y_r - \mu_r)(y_s - \mu_s)] \\ &= E\left[\left(\frac{\sum w_j\ v_j^r}{\sum w_j} - \mu_r\right)\left(\frac{\sum w_j\ v_j^s}{\sum w_j} - \mu_s\right)\right] \\ &= \frac{1}{(\sum w_j)^2}\ E\left[\sum_j w_j(v_j^r - \mu_r) \cdot \sum_j w_j(v_j^s - \mu_s)\right] \\ &= \frac{1}{(\sum w_j)^2}\ \sum_j w_j^2\ (E[v_j^r\ v_j^s] - \mu_r\ \mu_s) \\ &= \Gamma \cdot (\mu_{r+s} - \mu_r\ \mu_s), \end{aligned} \tag{13}$$

where the sum over the index j runs over all the pixels of the spectral line, and where we define

$$\Gamma \equiv \sum_j w_j^2 \Big/ \Big(\sum_j w_j\Big)^2 . \tag{14}$$

Here, we assume that the different observed points of the line profile are uncorrelated. In summary, to estimate the uncertainty of the first three moments of the line profile we use the higher theoretical moments of the line profile, up to the sixth moment $\mu_6(\ell, m, \boldsymbol{\beta})$.

Having derived an estimator for $\boldsymbol{\beta}$ and its uncertainty, given a couple (ℓ, m), we should take into account that we do not actually know the correct (ℓ, m) values. Given that (ℓ, m) is a couple of *discrete* parameters, this problem is notoriously hard. We

can do this by "weighting" each mode (ℓ, m) with a goodness-of-fit function. The best guess for both $\boldsymbol{\beta}$ and its uncertainty is then computed with a weighted mean over all relevant modes (ℓ, m). The entire estimation procedure can be sumarized as follows:

1. Set up a set of couples of the degree ℓ and the azimuthal number m: $\{(\ell_j, m_j)\}$.

2. For each of the couples (ℓ_j, m_j), solve the quasi-score equations and estimate the real-valued parameters $\hat{\boldsymbol{\beta}}_j$ and their covariance matrix $\mathrm{Cov}[\hat{\boldsymbol{\beta}}_j]$.

3. Compute for each of the modes (ℓ_j, m_j), the weight G_j^2 which indicates how well the theoretical moments $\boldsymbol{\mu}(\hat{\boldsymbol{\beta}})$ fit the observational moments $\mathbf{y}$:

$$G_j^2 = \sum_{k=1}^{3} \sum_{i=1}^{n} \frac{(y_k(t_i) - \mu_k(\hat{\boldsymbol{\beta}}_j, t_i))^2}{\mu_{2k}(\hat{\boldsymbol{\beta}}_j, t_i) - \mu_k^2(\hat{\boldsymbol{\beta}}_j, t_i)}. \tag{15}$$

4. The best estimate for the degree and the azimuthal number $(\tilde{\ell}, \tilde{m})$ is the (ℓ_j, m_j) that has the lowest weight G_j^2. The corresponding best estimate for the real-valued parameters $\tilde{\boldsymbol{\beta}}$ can be computed with

$$\tilde{\boldsymbol{\beta}} = \frac{\sum\limits_{\{(\ell_j, m_j)\}} \hat{\boldsymbol{\beta}}_j \; G_j^{-2}}{\sum\limits_{\{(\ell_j, m_j)\}} G_j^{-2}} \tag{16}$$

and the corresponding covariance matrix is the sum of the intra-mode variance and the inter-mode variance:

$$\mathrm{Cov}[\tilde{\boldsymbol{\beta}}] = \frac{\sum\limits_{\{(\ell_j, m_j)\}} \mathrm{Cov}[\hat{\boldsymbol{\beta}}_j] \; G_j^{-2}}{\sum\limits_{\{(\ell_j, m_j)\}} G_j^{-2}} + \frac{\sum\limits_{\{(\ell_j, m_j)\}} (\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_j) \cdot (\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_j)^t \; G_j^{-2}}{\sum\limits_{\{(\ell_j, m_j)\}} G_j^{-2}}. \tag{17}$$

For both practical and astrophysical reasons, only modes with a degree ℓ up to a certain limit (e.g. $\ell \leq 4$) are considered.

In the following section this estimation procedure is applied to a dataset of the star HD181558.

5 Application to HD181558

HD181558 belongs to the class of the Slowly Pulsating B stars (SPBs). Although the star is multi-periodic (De Cat and Aerts 2002) it has a very dominant (in amplitude) first mode, which allows us to use a mono-periodic model in a good approximation. In fact, the amplitude of this mode is the largest ever observed for an SPB. The dataset used for this GEE application has already been shown in Figure 2. In what follows we always assume $K = 21$.

Our first goal was to estimate $\boldsymbol{\beta}$ for each mode (ℓ, m) with $\ell \leq 4$, by solving the non-linear quasi-score equations. It turned out, however, that this was not just a technical detail of the procedure, but was in fact a major issue.

First, it turned out that the quasi-score function $\mathbf{U}(\boldsymbol{\beta})$ is computationally slow to evaluate, even though we programmed the computer codes in efficient C++. The reason is that 1 evaluation requires the computation 18 time series: 6 time series for the moments μ_1 up to μ_6 for the working approximation $\mathbf{W}$, and 12 time series for the moments μ_1 up to μ_3 for different parameters $\boldsymbol{\beta}$ to numerically compute (with forward differences) the derivatives in $\mathbf{D}$. For this reason we first determined a good initial guess for $\hat{\boldsymbol{\beta}}$ for the local search routine, by systematically making a rough scan of the 4D parameters space for each mode (ℓ, m) with a computationally less expensive goodness-of-fit function $g(\boldsymbol{\beta})$:

$$g(\boldsymbol{\beta}) \equiv \sum_{d=1}^{3} \frac{1}{d} \sqrt[d]{\frac{1}{n} \sum_{i=1}^{n} |y_d(t_i) - \mu_d(t_i, \boldsymbol{\beta})|}. \tag{18}$$

The construction with the d^{th} root and the division by d is simply to prevent the higher order moments from overshadowing the lower order moments. The sampling of the parameter space was done probabilistically and non-uniformly. For each parameter

β_i, a physical range was determined and this range was subdivided into intervals. After each set of 10000 sampled points, each interval of each parameter β_i was assigned a sampling probability according to the lowest $g(\boldsymbol{\beta})$ value recorded up to then, with the β_i component in the corresponding interval. The sum of probabilities over all intervals of a parameter β_i was set to one. For each couple (ℓ, m) a total of 200000 points was sampled. This procedure was set up to sample more the more promising regions of the parameter space.

In Table 1 we give for the star HD181558 the lowest $g(\boldsymbol{\beta})$ value recorded for each mode (ℓ, m). The couple $(\ell, m) = (0, 0)$ can be excluded on astrophysical grounds because such modes do not occur in SPBs. As can be seen, there is not just one mode (ℓ, m) that stands out, but there are several candidate modes that describe the data well. Our final estimate of $\boldsymbol{\beta}$ should take into account this mode uncertainty.

The 24 scans of a 4D parameter space with 200000 points each, was a rather time consuming but necessary task to find suitable initial guesses for $\hat{\boldsymbol{\beta}}$ for the local search algorithm. Concerning this local search algorithm, we explored several alternatives. First, we implemented the multi-dimensional Newton-Raphson method to find the root of $\mathbf{U}$, with and without Fisher scoring to approximate the Jacobian. However, despite the fact that we included a backtracking algorithm to counter too large Newton steps, we found that the Newton-Raphson method was rather easily led astray. Another option is to locate the minimum of $\|\mathbf{U}\|$. Although this is more vulnerable to local minima, the advantage is that there is a larger arsenal of function minimizing techniques than of root-finding techniques. An obvious candidate local descent method is the multi-dimensional Newton method. However, we decided not to implement this method, the main reason being that the Hessian is not available. Approximating it nu-

$g_{\min}$	$\ell = 1$	$\ell = 2$	$\ell = 3$	$\ell = 4$
$m = +4$				11.9
$m = +3$			7.52	11.0
$m = +2$		6.57	6.71	11.3
$m = +1$	4.72	4.74	5.85	10.8
$m = 0$	6.37	6.37	6.79	11.6
$m = -1$	4.79	6.57	7.05	10.7
$m = -2$		4.68	6.86	10.5
$m = -3$			6.92	11.3
$m = -4$				11.3

Table 1: For each couple (ℓ, m), the 4D parameter space was scanned with the goodness-of-fit function g defined by equation (18) and with the dataset of the star HD181558 shown in Figure 2. The minimum $g(\boldsymbol{\beta})$ value for each of the (ℓ, m) couples is given.

merically is likely to give unreliable results, given the fact that computing the function $\|\mathbf{U}\|$ itself already requires computing numerical derivatives of the moments. Instead we looked for more robust algorithms, and implemented two methods which need no derivatives at all, only function values: the conjugate-direction (Powell's) method (see, e.g., Press et al. 1992, p. 420) and the Torczon (1989) simplex method. Our best experience concerning efficiency and robustness was with the conjugate-direction method, which we subsequently used to locate the root $\hat{\boldsymbol{\beta}}$ of $\mathbf{U}$ for all modes (ℓ, m).

Even with the conjugate-direction method, the algorithm did not always converge. The reason, as it turns out, is that the quasi-score functions have "false" zeros, for example there are cases where the components of $\mathbf{U}$ approach zero for $\beta_2 \equiv \sigma \to \infty$. Quite often, the algorithm converged to a point outside the physically relevant range of the parameters, even when several different initial guesses for $\hat{\boldsymbol{\beta}}$ were tried. Although they did not occur for our dataset of the star HD181558, we should mention two other possible causes of numerical difficulties. First, it may be possible that the working approximation $\mathbf{W}$ is not invertible, for example if $\boldsymbol{\beta}$ approaches zero. Secondly, the matrix $\mathbf{I}_0$ may not be invertable, and hence no covariance matrix can be computed. This occurs, for example, for $\beta_4 = i \to 0°$ because $\beta_3 \equiv v_e$ appears only in $\beta_3 \sin\beta_4 = v_e \sin i$ in the equations, so that the third row and the third column of $\mathbf{I}_0$ are zero. We stress, however, that the latter example is a problem of intrinsic non-identifiability and is not specific for the GEE approach. One simply cannot derive the rotational velocity if the star is looked pole-on.

Making detailed 1D slices of the 4D function $\|\mathbf{U}\|$ would be too time consuming, but we did keep record of the minimal goodness-of-fit values $g_{\min}$ in each of the intervals of each parameter β_i (disregarding the values of the other components $\beta_{j\neq i}$). We

made plots of these records, and in Figure 3 we show four typical examples of them. Although the function $g(\boldsymbol{\beta})$ need not have exactly the same "surface" as the function $\|\mathbf{U}(\boldsymbol{\beta})\|$ (the difference is similar to the well-known difference between L_1-norm and L_2-norm minimization), we may assume that many of the features of the surface of $\|\mathbf{U}(\boldsymbol{\beta})\|$ are also visible in the surface of $g(\boldsymbol{\beta})$. We observe that the minimum in the upper left panel of Figure 3 for the well-fitting mode $(\ell, m) = (2, -2)$ is quite localized. This is much in contrast with the almost flat surface in the lower left panel for the badly-fitting mode $(\ell, m) = (4, -3)$. Intuitively, one can expect that the equivalent for the case of the $\|\mathbf{U}\|$ function hampers the iterations towards the minimum, and that this increases the chance to wander out of the physically relevant part of the parameter space. In fact, this is exactly what happened for this mode. More generally, we observe a quite strong correlation between how well a mode fits the data (with $g_{\min}$ as goodness-of-fit value) and the likelihood that the root finding algorithm does not converge. Concerning the lower right panel showing two minima in the plot of the inclination angle i, we need to mention that there exist theoretical symmetry relations that reveal that the moment method can only determine i mod 90°. Taking this fact into account, we never observed "real" multiple minima, although a minimum can often be a very wide valley.

The fact that we quite often do not seem to find the root of $\mathbf{U}$ does not need to contradict the theory outlined in Section 4. We solve for the root of the observed $\mathbf{U}$ function because we know that $E[\mathbf{U}] = \mathbf{0}$. However, the latter is only true if the model is correctly specified, i.e. if $E[\mathbf{Y}] = \boldsymbol{\mu}(\ell, m, \boldsymbol{\beta})$. Therefore, theoretically, the existence of a root in the 4D parameter space of the real-valued parameters $\boldsymbol{\beta}$ cannot be guaranteed for a "wrong" couple (ℓ, m), and this is exactly what we observe for

badly fitting modes. For this reason we interpreted a non-convergence (after repeatedly trying) as an indication that the candidate mode should be disregarded.

In Table 2, we list the roots of the quasi-score function for those modes for which there was convergence in the physically relevant part of the parameter space. The closeness of $\|\mathbf{U}\|$ to zero varies from mode to mode. Restarting the algorithm at the point were it stopped, did not bring $\|\mathbf{U}\|$ closer to zero. For some of the solutions the algorithm might have stranded in a local minimum, for example the modes $(\ell, m) = (3, -1)$ and $(\ell, m) = (4, 0)$. In addition, we are suspicious about the solution for the mode $(\ell, m) = (2, 2)$ since the inclination angle equals exactly 270°, and the parameter v_e has an unlikely small value.

In Figure 4 we show fits for the three best fitting modes, with the function G^2 (see Eq. 15) as a goodness-of-fit. As mentioned before, there is not just one, but several modes that can fit the observed data quite well. Note that the goodness-of-fit values in Table 1 provide us with an indication of the relative merits of wavenumber choice. Of course, at this point we lack knowledge about the reference distribution of these values, unlike in classical fit statistics (e.g., likelihood-ratio based). However, similar instances exist in both a frequentist (e.g., Akaike Information Criterion) and a Bayesian context (e.g., Bayes factors). Nevertheless, we assert that these numbers, especially when supported by careful graphical inspection, are useful to substantially narrow down our uncertainty about the wavenumbers, in spite of an intrinsically complicated modelling endeavor. To this end, the last column in Figure 4 displays mode $(\ell, m) = (3, 1)$ with a fit, substantially worse than the one in the first three columns of the same figure.

We used the modes in Table 2, to compute the *weighted* mean $\tilde{\boldsymbol{\beta}}$ and its standard error, with Eqs. (16) and (17). The results, including the intra-mode and inter-mode

(ℓ, m)	$\|\mathbf{U}\|_{\min}$	G^2	$\hat{v}_p$	$\hat{\sigma}$	$\hat{v}_e$	$\hat{\imath}$
(1,0)	0.15	2.7	2 (1)	9.0 (0.9)	13 (19)	320 (48)
(1,1)	$5.1\ 10^{-24}$	0.63	2.01 (0.08)	6.3 (0.2)	15.2 (0.7)	117 (2)
(1,-1)	$1.1\ 10^{-5}$	1.1	4.0 (0.2)	4.2 (0.6)	25 (2)	336 (1)
(2,0)	0.10	3.1	0.9 (0.4)	7 (2)	30 (24)	331 (13)
(2,1)	1.7	0.61	1.7 (0.1)	3.8 (0.8)	16 (1)	71 (1)
(2,2)	0.014	2.4	1 (1)	10 (2)	0.4 (29)	270 (360)
(2,-2)	0.0032	0.72	1.62 (0.06)	4.3 (0.4)	17.6 (0.7)	129 (1)
(3,1)	$1.1\ 10^{-25}$	2.3	1.00 (0.04)	6.8 (0.7)	18 (2)	145 (7)
(3,2)	0.40	3.1	1.1 (0.2)	6 (3)	17 (10)	49 (23)
(3,-1)	3.5	7.0	1.3 (0.3)	3 (12)	7 (12)	189 (5)
(4,0)	2.7	11	0.4 (0.3)	8 (12)	49 (188)	19 (96)
(4,-4)	0.047	8.7	0.6 (3)	7 (24)	20 (87)	295 (30)

Table 2: Roots of the quasi-score functions for those modes where there was convergence in the physically relevant part of the parameter space, for the star HD181558. The values between brackets are the standard errors obtained with the sandwich estimator (9). G^2 is the weight of the mode as defined by Eq. (15). v_p, σ and v_e are expressed in km/s, and the inclination angle i in degrees.

$\tilde{\beta}_i$	Weighted Mean	Intra-mode Variance	Inter-mode Variance
$\tilde{v}_p$	1.8 (1.0)	0.33	0.74
$\tilde{\sigma}$	5.5 (4.1)	14	3.1
$\tilde{v}_e$	17 (26)	612	46
$\tilde{\imath}$	164 (132)	7300	10079

Table 3: The weighted mean over all 12 modes mentioned in Table 2, computed with Eqs. (16) and (17). The values mentioned between brackets are standard errors. The intra-mode variance and inter-mode variance are computed with respectively the first and the second term of Eq. (17). $\tilde{v}_p$, $\tilde{\sigma}$, $\tilde{v}_e$ are expressed in km/s, and the inclination angle $\tilde{\imath}$ is expressed in degrees.

variance, are given in Table 3. In the specific case of HD181558, one could argue that the modes with $\ell = 3$ and $\ell = 4$ can be disregarded on astrophysical grounds. The reason is that these modes would require a very large oscillation amplitude at the surface of the star to cause the large observed amplitude of the first moment y_1. For this reason, we also computed $\boldsymbol{\beta}$ with the $\ell = 1$ and $\ell = 2$ modes of Table 2 only. The results are listed in Table 4. With Tables 3 and 4 we achieve the goal of this application of the revised version of the moment method: we have obtained a best guess for the real-valued parameters and their standard errors, where we took into account the mode uncertainty. An important result is that the uncertainties of the parameters can be large, in fact larger than we expected. Especially the rotational velocity v_e cannot be estimated precisely. The large inter-mode uncertainty of the inclination angle i is not surprising since the inclination angle is known to be largely dependent on the mode

$\tilde{\beta}_i$	Weighted Mean	Intra-mode Variance	Inter-mode Variance
$\tilde{v}_p$	2.0 (1.0)	0.19	0.71
$\tilde{\sigma}$	5.3 (2.0)	0.82	3.1
$\tilde{v}_e$	16 (12)	100	37
$\tilde{\imath}$	170 (137)	8342	10450

Table 4: The same information as in Table 3 is shown, except that the mean is computed over those 7 modes in Table 2 with $\ell = 1$ and $\ell = 2$.

numbers (ℓ, m). To conclude, we note that the values for the weighted means do not change much by excluding the $\ell = 3$ and $\ell = 4$ modes. The reason is that the latter modes have a lower weight anyway, as can be seen from the G^2 values in Table 2.

6 Summary and Conclusions

We revisited one particular mode identification technique, the moment method, to incorporate estimates of the uncertainties of the real-valued parameters. Because of the difficulties to overcome, this was never done for the moment method, nor for any other mode identification technique.

We found that, in the specific case of the moment method, the method of least-squares does not give consistent estimates of the real-valued parameters and for this reason we resorted to the GEE method as described by Liang and Zeger (1986) which does give a consistent normally distributed estimate. This method requires a working approximation of the covariance matrix of the 3 responses which we set up using the higher theoretical moments. An important source of uncertainty is the fact that often

not just one but several candidate modes can describe the data. We set up a procedure to weight each mode and to compute a weighted mean over all modes of the parameter vector and its uncertainty. To compute the latter we introduced the intra-mode and the inter-mode uncertainty.

Subsequently, we applied our procedure to the SPB star HD181558. The solution of the estimating equations turned out to be a tedious task, one of the reasons being that the quasi-score function $\mathbf{U}$ is computationally slow to evaluate. We needed to set up a strategy to systematically scan the parameter space for each mode (ℓ, m) to obtain good initial guesses for the local search method. In addition, we experimented with several local root finding methods of which we selected the method of conjugate directions to minimize $\|\mathbf{U}\|$ as the most robust and efficient method. We nevertheless experienced quite often convergence difficulties. In fact, one of the conclusions of this work is that GEE applied to the moment method is computationally rather demanding. This seems, however, intrinsic to the problem and not to GEE.

The scanning of the parameter space showed that several modes can explain the dataset of HD181558, and that it is therefore not useful to mention only the very best fitting mode. We retained 12 modes as candidate modes for which an estimate of the real-valued parameters $\hat{\boldsymbol{\beta}}$ can be computed, and used these estimates to obtain a best guess $\tilde{\boldsymbol{\beta}}$ for the real-valued parameters plus their uncertainties, taking into account the mode uncertainty. Doing so, we discovered that the parameter uncertainties can be large. It might be tempting for the multiperiodic case to use a two-stage approach where the inclination angle i and the rotational velocity v_e is determined with the dominant mode, and where these parameters are subsequently fixed while determining the mode parameters of the other modes, to have the dimension of the parameter space

reduced. Our results show that such an approach can be very dangerous: in the case of HD181558 it can hardly be justified because of the large uncertainty on i.

We finally mention that we proposed a new goodness-of-fit function, which is to be preferred above the previously used discriminant (6), as it works at least as good as the latter and allows to discriminate between positive and negative azimuthal numbers which was one of the shortcommings of the use of (6).

Acknowledgement

Research supported by a PAI program P5/24 of the Belgian Federal Government (Federal Office for Scientific, Technical, and Cultural Affairs).

7 References

Aerts, C., de Pauw, M. and Waelkens, C. (1992), "Mode identification of pulsating stars from line profile variations with the moment method. An example - The Beta Cephei star Delta Ceti", *Astronomy and Astrophysics*, 266, 294-306.

Aerts, C. (1996), "Mode identification of pulsating stars from line-profile variations with the moment method: a more accurate discriminant", *Astronomy and Astrophysics*, 314, 115-122

Aerts, C., De Cat, P., Cuypers, J., Becker, S.R., Mathias, P., De Mey, K., Gillet, D. and Waelkens, C. (1998), "Evidence for binarity and multiperiodicity in the beta Cephei star beta Crucis", *Astronomy and Astrophysics*, 329, 137-146

Aerts, C. and Kaye, A.B. (2001), "A Spectroscopic Analysis of the gamma Doradus Star HD 207223 = HR 8330", *The Astrophysical Journal*, 553, Issue 2, 814-822

Balona, L.A. (1986), "Mode identification from line profile variations", *Monthly Notices of the Royal Astronomical Society*, 220, 647-656

Chadid, M., De Ridder, J., Aerts, C. and Mathias, P. (2001), "20 CVn: A monoperiodic radially pulsating delta Scuti star", *Astronomy and Astrophysics*, 375, 113-121

De Cat, P. and Aerts, C. (2002), "A study of bright southern slowly pulsating B stars. II. The intrinsic frequencies", *Astronomy and Astrophysics*, in press

Diggle, P.J., Heagerty, P., Liang, K-Y and Zeger, S.L. (2002), Analysis of Longitudinal Data, Oxford University Press, Oxford

Liang, K.Y. and Zeger, S. (1986), "Longitudinal data analysis using generalized linear models", *Biometrika*, 73, 13

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992), "Numerical Recipes in C. The art of scientific computing." (2nd ed.), Cambridge University Press

Seber, G.A.F. and Wild, C.J. (1989), "Nonlinear Regression", Wiley series in prob-

ability and mathematical statistics, John Wiley & Sons, Inc.

Torczon, V.J. (1989), "Multi-directional search: a direct search algorithm for parallel machines", PhD Thesis, Rice University, Houston (Texas, U.S.A.)

Uytterhoeven, K., Aerts, C., De Cat, P., De Mey, K., Telting, J.H., Schrijvers, C., De Ridder, J., Daems, K., Meeus, G. and Waelkens, C. (2001), "Line-profile variations of the double-lined spectroscopic binary kappa Scorpii", *Astronomy and Astrophysics*, 371, 1035-1047

Zeger, S.L., Liang, K-Y (1986), "Longitudinal data anlaysis for discrete and continuous outcomes", *Biometrics*, 42, 121-130

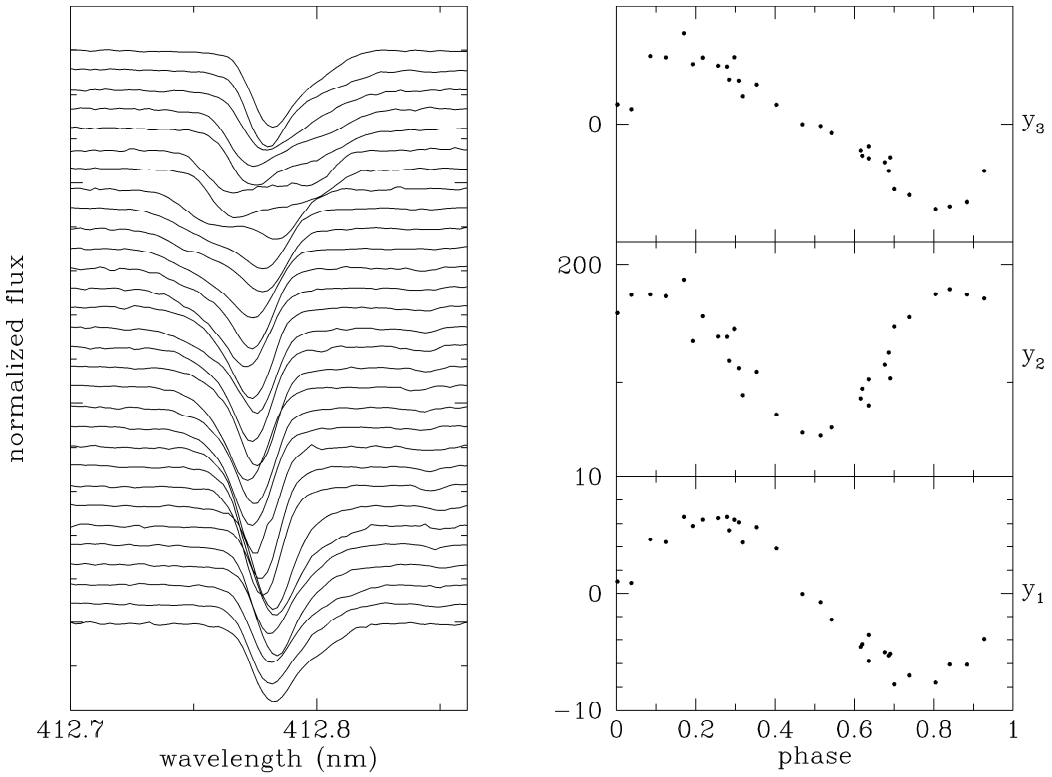


Figure 2: In the left panel, a time series of Si^+ (412.805 nm) absorption lines of the non-radially oscillating star HD181558 is shown. The line profiles are 'sorted' to nicely cover an entire oscillation cycle of the dominant mode, which has a period of about 29h42m. Each of the line profiles is vertically shifted to obtain a clear visual effect. In the right panels, the first moment y_1 (in km/s), the second moment y_2 (in km^2/s^2) and the third moment y_3 (in km^3/s^3) of all line profiles are shown as a function of the oscillation phase which is the fraction of the oscillation period.

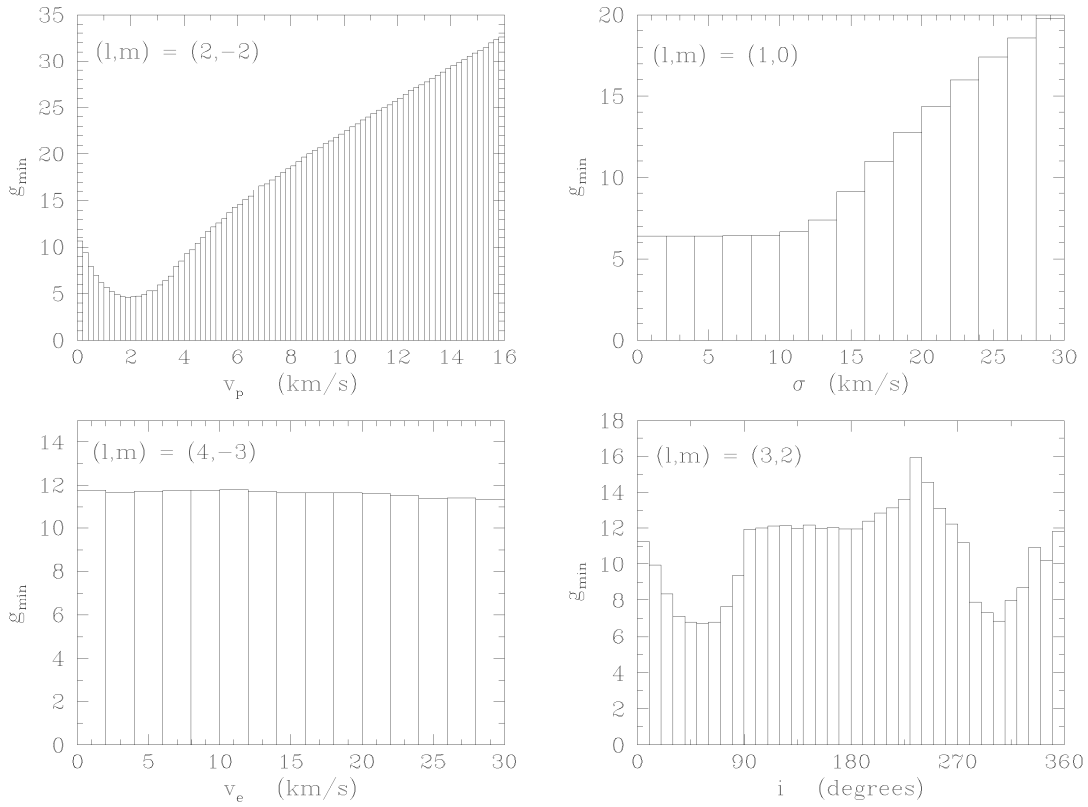


Figure 3: Representative examples of the minimal goodness-of-fit value for each sample interval of a parameter β_i, for the star HD181558. We remark that although the size of the intervals for the parameter v_p is fixed, the relevant range of v_p depends on the mode numbers (ℓ, m).

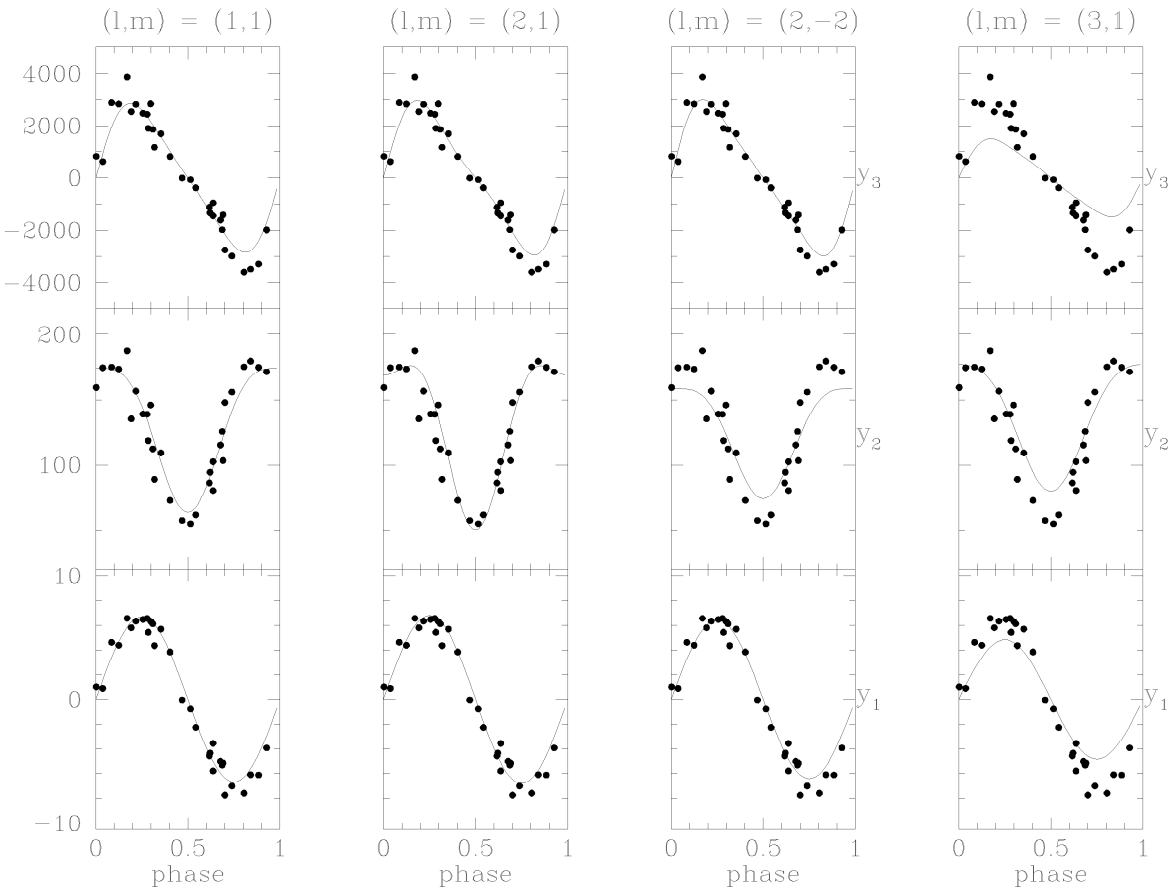


Figure 4: Theoretical models (solid lines) of the observed moments (bullets) for the three best fitting modes $(\ell, m) = (1, 1)$, $(\ell, m) = (2, 1)$, and $(\ell, m) = (2, -2)$, plus the less good fitting mode $(\ell, m) = (3, 1)$, with G^2 as a goodness-of-fit function. The theoretical models were obtained with the model parameters obtained with the GEE method. The first, the second and the third row are for the first moment y_1 (km/s), the second moment y_2 (km^2/s^2) and the third moment y_3 (km^3/s^3) respectively. The moments are shown as a function of the phase which is the fraction of the oscillation cycle. Note that the models of the different promising modes differ mainly for the second moment.