



# The Stellar Seismic Indices (SSI) data base: User's Guide and implementation guidelines

Réza Samadi, Raphaël Peralta, Mahfoudh Abed, Christian Renié, Eric Michel, Kévin Belkacem

LESIA - Observatoire de Paris

November 4, 2016 - Version 1.1

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Mainlines of the MLEUP method</b>	<b>3</b>
<b>3</b>	<b>The data base content and its query interface</b>	<b>4</b>
<b>4</b>	<b>The pipeline</b>	<b>5</b>
4.1	General description . . . . .	5
4.2	Data extraction . . . . .	6
4.3	Pre-processing . . . . .	6
4.4	Generation of the stellar indices . . . . .	8
4.5	Post-processing . . . . .	8
4.6	Implementation details . . . . .	8
<b>5</b>	<b>Structure of the data base</b>	<b>8</b>
5.1	Star table . . . . .	9
5.2	Indice table . . . . .	10



# 1 Introduction

Stars are cavity where waves can propagate and form standing waves (modes). The frequencies of these oscillation modes depend on the internal structure of the star. Their detection and measure enable us to probe the inner layers of the stars. Oscillations were detected in the Sun for the first time more than 50 years ago. Similar oscillations (thus named solar-like oscillations) were detected in other stars at the end of the nineties. Thanks to CoRoT (CNES and European partners) and *Kepler* (NASA) seismic observations, it is now possible to detect solar-like oscillations in a very high number of stars. The observed oscillation spectra can be quite complicate to analyse and interpret. Nevertheless, these spectra show characteristic pattern, which can easily be characterized with that we call *stellar seismic indices*. More specifically, the stellar seismic indices correspond to some characteristic global seismic numbers, which are extracted from the oscillation spectra of solar-like pulsating stars. Two indices are now frequently used:

- the peak frequency,  $\nu_{\max}$ : this is in general the frequency where the oscillation spectrum peaks in the Power Spectral Density (PSD) ;
- the mean large separation,  $\Delta\nu$ : this quantity corresponds to the mean frequency spacing between two consecutive p-modes (with same angular degree) ;

These seismic indices obey characteristic scaling relations that depend directly on the radius, mass and effective temperature of the star. From the knowledge of the effective temperature and the measure of these two seismic indices, it is then possible to estimate the mass and radius of the star, and subsequently the surface gravity ( $\log g$ ) of the star.

Because the seismic indices are relatively easy to measure thanks to CoRoT and *Kepler*, it has already been possible to measure seismic indices in about 15,000 red giants and sub-giants, and subsequently to derive their mass, radius and evolutionary status. These seismic indices are more and more used in the stellar physics but also for the study of Galactic population (see the reviews by [Belkacem 2012](#); [Belkacem et al. 2013](#)). These seismic indices have opened the way toward that we name ensemble asteroseismology.

Apart from these indices, it is also possible to extract from these spectra some characteristic parameters of the stellar granulation, namely the e-folding time ( $\tau_{\text{eff}}$ ) and the mean-square brightness fluctuations ( $\sigma_{\text{gran}}^2$ ) associated with the granulation background. The first parameter informs about the granulation lifetime while the second corresponds to the total integrated energy within the granulation background. Both parameters also obey characteristic scaling relations that can provide informations about the stars in addition to the seismic indices (see e.g. [Mathur et al. 2011](#); [Samadi et al. 2013a,b](#); [Kallinger et al. 2014](#)).

The Stellar Seismic Indices (SSI, <http://ssi.lesia.obspm.fr>) data base provides such sets of stellar parameters (both seismic indices and granulation parameters) extracted in *automatic* and *homogeneous* way for a large set of CoRoT and *Kepler* light-curves.

These stellar parameters were generated on the basis of the MLEUP method ([de Assis Peralta et al. 2016](#)). This new method relies on the Universal Pattern (UP [Mosser et al. 2011](#)) and takes advantage of the Maximum Likelihood Estimator (MLE) algorithm. It provides simultaneously and in a consistent way both seismic indices and stellar granulation parameters. For more details, please refer to [de Assis Peralta et al. \(2016\)](#). The method was applied on almost all *Kepler* long-cadence light-curves and almost all the targets observed from the CoRoT faint stars fields. A total of about 320 000 targets have been analyzed, among which seismic indices and granulation parameters have been extracted for about 18 000 of them.

This document briefly presents mainlines of the MLEUP method (Sect. 2), the data base content and its web interface (Sect. 3) as well as the architecture of the associated pipeline (Sect. 4) and the structure of the data base (Sect. 5).

## 2 Mainlines of the MLEUP method

The model used to fit the power density spectrum is based on two Lorentzian-like functions for the granulation and activity component, a red giant parametric oscillations pattern based on the Universal Pattern [Mosser et al. \(2011\)](#) for the oscillations and a constant for the white noise (see Fig. 1). The oscillations are characterised by three seismic indices: The mean large separation,  $\Delta\nu_{\text{up}}$ , corresponding to the mean frequency spacing between two consecutive p-modes with same angular degree;  $H_{\text{env}}$  which is the maximum height of the oscillation envelope

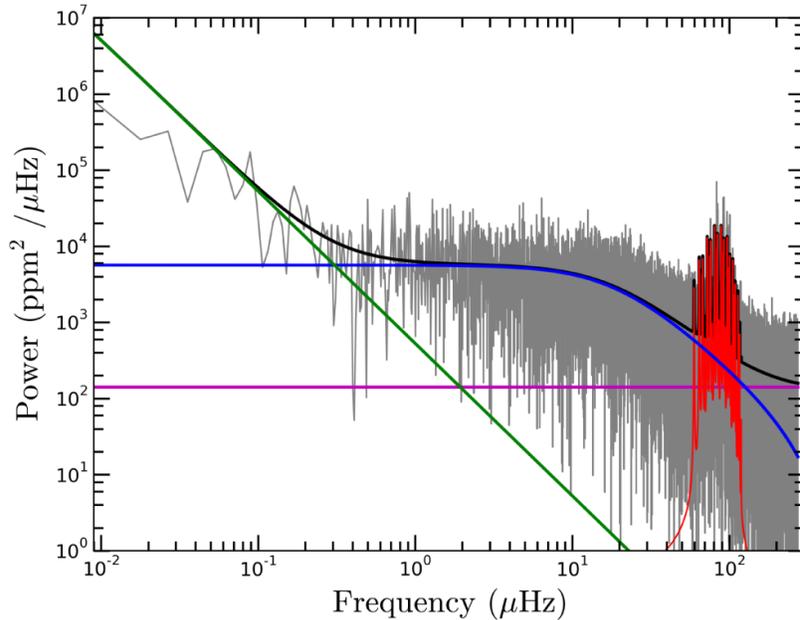


Figure 1: Results of the final adjustment of the background and oscillations of MLEUP (black line). In grey, the raw PSD of KIC 5527304. The green line and the blue one correspond respectively to the activity and granulation component. The magenta line represents the white noise component and the red one is the Universal Pattern (the oscillations component).

and  $\nu_{\max, \text{up}}$ , the corresponding frequency. Concerning the granulation, it is characterised by two parameters: the effective timescale  $\tau_{\text{eff}}$  (or the *e-folding time*), which measures the temporal coherence of the granulation in the time domain and the characteristic amplitude  $\sigma_{\text{gran}}^2$  which corresponds to the variance brightness fluctuation of the granulation. The method has been tested in order to characterize its bias and dispersion using Monte Carlo simulations. These simulations revealed that MLEUP presents low dispersions, especially for  $\Delta\nu_{\text{up}}$  and  $\nu_{\max, \text{up}}$ , and that the internal errors are reasonably representative of the real dispersions

### 3 The data base content and its query interface

Table 1: The stellar parameters generated by the MLEUP method

Name	Symbol	Short description	Unit
Deltanu_up	$\Delta\nu_{\text{up}}$	mean large separation $\Delta\nu$	$\mu\text{Hz}$
numax_up	$\nu_{\max, \text{up}}$	frequency of the maximum mode height $\nu_{\max}$	$\mu\text{Hz}$
Amax	$A_{\max}$	maximum of the auto-correlation function	-
H_env	$H_{\text{env}}$	height of the oscillations envelop	$\text{ppm}^2/\mu\text{Hz}$
taue_gran	$\tau_e$	e-folding time of the granulation background	s
tau_gran	$\tau_{\text{gran}}$	characteristic timescale of the granulation component	s
sigma2_gran	$\sigma_{\text{gran}}^2$	variance brightness fluctuation of the granulation component	$\text{ppm}^2$
P_gran	$P_{\text{gran}}$	height of the granulation component	$\text{ppm}^2/\mu\text{Hz}$
alpha_gran	$\alpha_{\text{gran}}$	slope of the granulation component	-

The SSI website is located at <http://ssi.lesia.obspm.fr>. The stellar parameters extracted with the MLEUP method and stored into the data base are listed in Table 1. Except for the parameter  $A_{\max}$ , the formal 1- $\sigma$  errors are systematically provided.

The data base also provides auxiliary information concerning the star (see Sect. 5.1), such as the magni-

StarID	Origin	alpha [deg]	delta [deg]	teff [K]	grav [cm/s <sup>2</sup> ]	mag_v	mag_r	specType	lumClass	numax_up [μHz]		Deltanu_up [μHz]		
										value	prec	value	prec	
100402467	corot	290.570746	1.644686	4329		-99	14.12	K	II	18.3	1	2.4912	0.012	8
100412751	corot	290.586053	1.634669	4343		-99	13.43	K	III	15.4118	0.586	1.9552	0.004	8
100422482	corot	290.600554	1.717	4549		-99	12.13	K	III	64.3284	0.963	6.0016	0.004	1
100434740	corot	290.618651	1.62033	5031		-99	12.87	G	V	47.502	0.39	4.1872	0.008	8
100440069	corot	290.626483	1.708309	4487		-99	12.87	K	II	51.6	1	5.605	0.005	8
100440565	corot	290.627243	1.647735	4507		-99	12.15	K	II	27.922	0.46	3.5088	0.004	1
100441421	corot	290.628492	1.636505	4753		-99	12.89	K	III	32.94	0.3	4.4695	0.005	1
100442460	corot	290.629967	1.714166	4549		-99	13.79	K	II	28.4526	0.958	3.5328	0.006	1
100448189	corot	290.638332	1.692805	4499		-99	13.53	K	II	16.3917	0.467	2.3356	0.004	1
100480626	corot	290.685877	1.647927	5673		-99	15.26	F	IV	19.1542	0.689	2.1554	0.002	2

Figure 2: Snapshot showing the query results

tude and the color in various photometry bands, the equatorial coordinates (ICRS, Epoch 2000), the effective temperature ( $T_{\text{eff}}$ ), the surface gravity ( $\log g$ ), luminosity class and spectral type. All these information have various origins and can be inaccurate. In particular the surface parameters ( $T_{\text{eff}}$  and  $\log g$ ), luminosity class and type are only indicative.

Access to the data is possible through the Search interface. The interface allows the users to perform a search using star ID, star properties (e.g.  $T_{\text{eff}}$ ,  $\log g$ , spectral type, ...), and finally stellar indices (seismic indices and granulation parameters). A list of star ID can be provided by downloading a file containing these ID. Access to the data is also possible through a VO compatible web service as well as through the Seismic Plus portal (<http://voparis-spaceinn.obspm.fr/seismic-plus/>).

The results of the search are displayed in the browser using a paginated data table (see Fig. 2). They can be filtered and sorted by any parameters. All parameters are by default not displayed but by clicking on “Show/hide columns”, it is possible to choose the parameters to be displayed. The results are displayed on the screen for a limited number of stars. The complete set can be downloaded on demand in a CSV file (the semicolon ‘;’ is used as delimiter).

The SSI pipeline was applied on about 114,000 stars observed by CoRoT in the exo-planet field and 210,000 stars observed during up to 4 years by *Kepler*. The corresponding light-curve were selected on the basis of their total duration (longer than 50 days) only, and accordingly without any prior concerning the evolutionary status of the targets. Stellar indices (seismic and granulation parameters) were extracted for total number of about 18,000 red giants (see the results presented in [de Assis Peralta et al. 2016](#)). Among the *Kepler* stars for which seismic indices were extracted, about 5,000 of them were not identified in the *Kepler* Input Catalog (KIC) as red giant. Some illustrative results are presented in Fig. 3.

The seismic indices  $\Delta\nu$  and  $\nu_{\text{max}}$  obey characteristic scaling relations that depend directly on the radius, mass and effective temperature of the star (for a review see e.g. [Belkacem 2012](#)). From the knowledge of the effective temperature and these two seismic indices, it is then possible to estimate the mass and radius of the star. Note that a tool to derive stellar masses and radii from the seismic indices and effective temperatures is implemented in the Seismic Plus portal (<http://voparis-spaceinn.obspm.fr/seismic-plus/>).

## 4 The pipeline

### 4.1 General description

The main goal of the pipeline is to extract a set of seismic indices from stellar light-curves acquired with CoRoT and *Kepler*. The main inputs of the pipeline are light-curves (one or several per target) from the CoRoT legacy data release (<http://idoc-corot.ias.u-psud.fr>) and the *Kepler* NASA data base (<https://archive.stsci.edu/kepler/>). A pre-processing of the light-curves is performed (see Section 4.3), which provides as outputs the light-curves corrected for different observational artifacts and long-term effects, as well

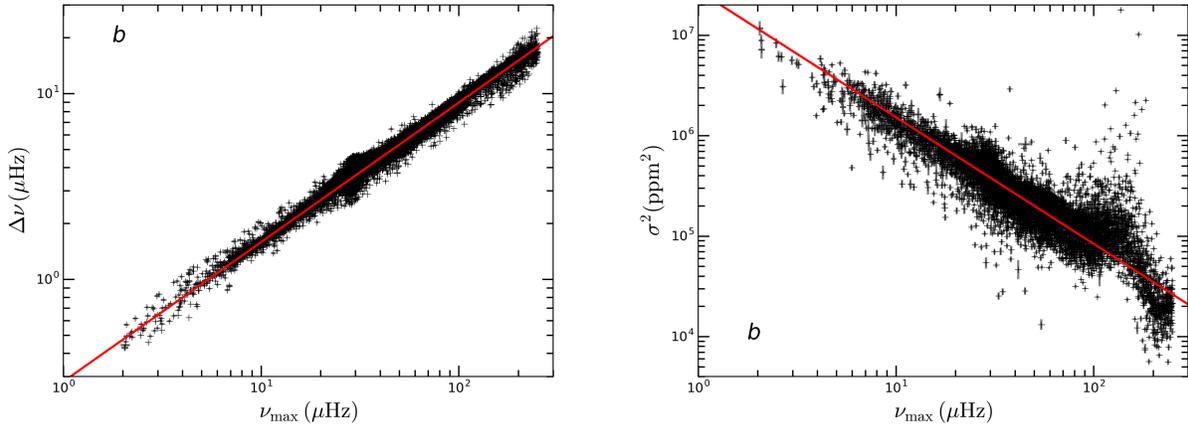


Figure 3: **Left panel:**  $\Delta\nu_{up}$  as a function of  $\nu_{\max,up}$  for about 13,000 *Kepler* red giants. **Right panel:**  $\sigma_{\text{gran}}^2$  as a function of  $\nu_{\max,up}$ .

as the associated Power Spectral Density (PSD). The pre-processing has also in charge the concatenation of the light-curves acquired at different periods for the same target. The seismic indices are then extracted from the pre-processed light-curves on the basis of the MLEUP method. The figure 4 shows the main data flows and the different data handled by the pipeline.

## 4.2 Data extraction

The inputs of the pipeline are either CoRoT or *Kepler* light-curves. Three different types of CoRoT light-curves are handled:

- AN2\_STAR: light-curve from the asteroseismology channel (sampling: 32s) ;
- EN2\_STAR\_MON: monochromatic light-curve from the Exo-planet channel (sampling 32s, 512s, or a mixed of both cadences) ;
- EN2\_STAR\_CHR: chromatic light-curve from the Explonat channel (sampling 32s, 512s, or a mixed of both cadences).

Note that we consider only the white flux of the chromatic light-curves (*i.e.* the sum of the three colors). The formats of the CoRoT legacy data are described in the document available at [http://idoc-corot.ias.u-psud.fr/sitools/common/html/doc/cII\\_4\\_data.pdf](http://idoc-corot.ias.u-psud.fr/sitools/common/html/doc/cII_4_data.pdf) (see also Chaintreuil et al. 2016).

The *Kepler* light-curves contain two type of fluxes: SAP\_FLUX and PDCSAP\_FLUX. The former corresponds to the flux acquired with the optimal aperture pixels while the latter corresponds to the flux corrected for instrumental perturbations (see the *Kepler* Archive Manual and the *Kepler* Data Processing Handbook). We consider the corrected flux. The cadence is either  $\sim 1$  mn (short cadence) or  $\sim 30$  mn (long cadence). Both cadences are not mixed in the same file.

## 4.3 Pre-processing

The pre-processing of the input light-curves consists in handling the perturbations listed below. Note that Some of these corrections are however already applied on the CoRoT light-curves (see Ollivier et al. 2016) while those performed at KASOC on the *Kepler* data are documented in the *Kepler* Data Processing Handbook.

- outliers (e.g. proton impacts): Proton impacts can occur while monitoring a target. This is particularly so for CoRoT when it crosses the South Atlantic Anomaly.

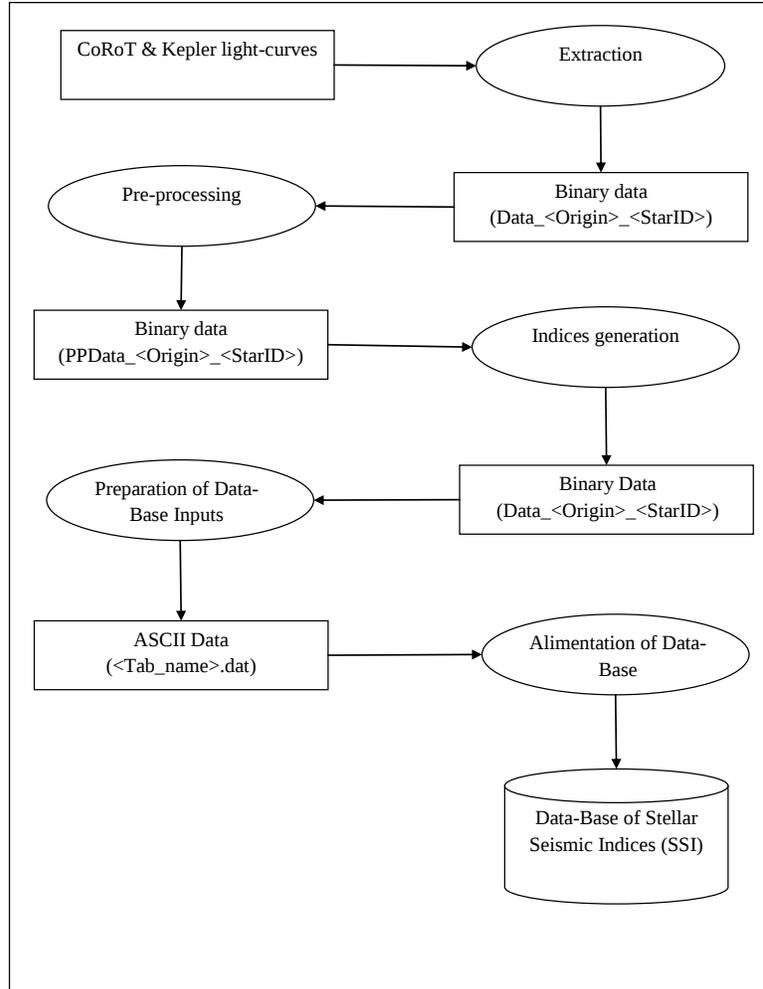


Figure 4: Data flows

- large jumps: big jumps occur some times in the light-curves. They are usually due to the occurrence of a bright pixel, which is typically generated by a particularly energetic proton impact. However, the origin of such jumps are not always clear.
- gaps: the interruptions occurring during an observation sequence introduce aliases in the spectrum, which can alter the determination of the seismic indices.
- flagged measurements: doubtful measurements are in general flagged. For instance for CoRoT, measurements considered as outliers (e.g. proton impacts) are flagged as doubtful measurements by the pipeline. The flagged data are often periodic (for instance the crossing of the South Atlantic Anomaly by CoRoT is periodic). Discarding these measurements then produces aliases in the Fourier spectrum. It is therefore required to replace the flagged measurements by some representative values.
- merging together several light-curves: light-curves acquired during different observing sequences are not comparable in intensity. Therefore, it is not possible to build a single light-curve just by joining together the individual light-curves. It is required to match the individual light-curves in an appropriate manner.
- long term effects: The global gain of an instrument is expected to vary with the temperature. In addition to the variation of the global gain with the temperature, we face long-term decrease of the intensity

because of the aging of the optics and of the CCD. All these variations occur at long time-scale (from weeks to several months). These variations are corrected by the recent version of the CoRoT pipeline.

Once the individual light-curves are processed and merged together, the pre-processing module computes the Power Spectral Density (PSD) associated with the merged light-curve. The time series are in general non-equispaced in time. Accordingly, the PSD is computed using the Lomb-Scargle periodogram, which contrary to the Fast Fourier Transform does not require equispaced measurements. Among the fast available algorithms, we consider the one proposed by Leroy (2012), which relies on the NFFT algorithm (<http://www-user.tu-chemnitz.de/~potts/nfft/>). A Python version of this algorithm named PyNFFTLs is available at <http://pypi.python.org/pypi/pynfftl/>.

#### 4.4 Generation of the stellar indices

The pipeline implements the MLEUP algorithm as described in de Assis Peralta et al. (2016, see also the brief description made in Sect. 2). The program takes as input the PSD associated with a given pre-processed light-curve and returns a set of stellar indices for each star. The generated stellar indices are listed in Tab. 1.

#### 4.5 Post-processing

The post-processing consists in the analysis of the generated stellar indices and their insertion into the data base. The stellar indices are first corrected for biases and their validity checked. For a detailed description see de Assis Peralta et al. (2016). Granulation parameters for redgiants with  $\nu_{\max, \text{up}} > 100 \mu\text{Hz}$  are not stored in the data base. As explained in de Assis Peralta et al. (2016), this is because the MLEUP is unable to derive reliable granulation parameters above this threshold.

#### 4.6 Implementation details

The SSI pipeline is coded under Python language (version 2.7) and relies on the following libraries:

- Numpy: fundamental package for scientific computing with Python (<http://www.numpy.org/>) ;
- Scipy: scientific computing tools for Python (<http://www.scipy.org/>) ;
- FFTW3: library for computing discrete Fourier transforms (<http://www.fftw.org/>) ;
- NFFT: library for computing nonequispaced discrete Fourier transforms (<http://www-user.tu-chemnitz.de/~potts/nfft/>) ;
- PyNFFTLs: a fast algorithm for the Lomb-Scargle periodogram (<http://pypi.python.org/pypi/pynfftl/>).

The SSI pipeline has a multiprocessing capability (the stars are processed in parallel).

The code is open source released under the GNU General Public Licence (<http://www.gnu.org/licenses/>). It can be freely downloaded from the SSI website.

## 5 Structure of the data base

The data base contains five tables (see Figure 5) with the following meaning:

- **Star**: characteristics of the stars (e.g. name, origin, position, magnitude, spectral type ... ) ;
- **Indice**: the seismic indices (e.g. value, precision) ;
- **File**: characteristics of the set of files (light-curves) that have been merged together for a given target ;
- **Code**: versions of the different components of the pipeline ;
- **Use**: this table ensures the links between the indices and the files.

The data base is developed under `postgresql`, an open source data base system (<http://www.postgresql.org/about/>). We only describe below the tables named **Star** and **Indice**. A more complete description is given in the document available for download in the SSI website.

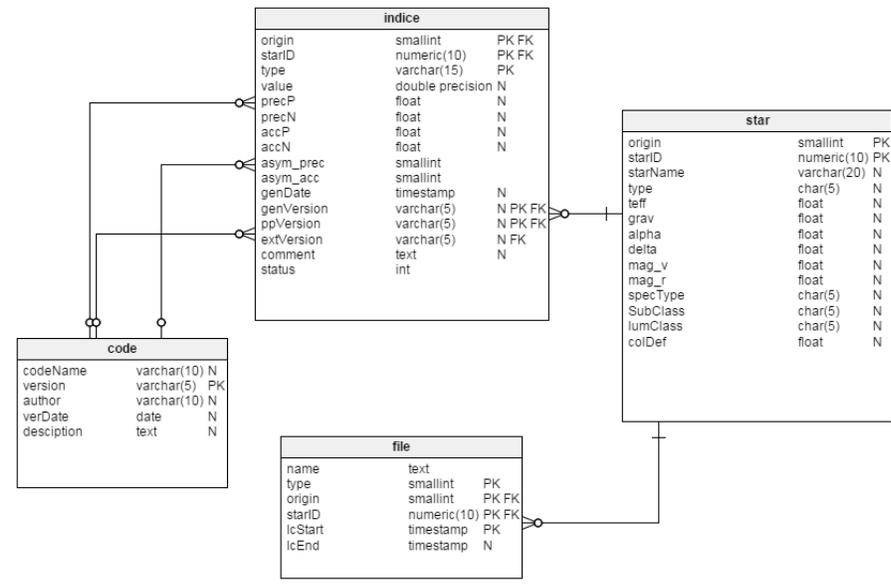


Figure 5: Structure of the date base

## 5.1 Star table

This table contains informations related with the star:

- **Origin:** origin of the data ;
  - = 0 : for an unknown origin
  - = 1 : for CoRoT data
  - = 2 : for Kepler data
  - = 3 : for OGGLE data (not yet implemented)
- **starID:** the star ID (CoRoTID in the case of a CoRoT target, and the KIC number in the case of a *Kepler* target) ;
- **type:** type of light-curve (depending of the instrument, the light-curves can be of different type. For instance for CoRoT: monochromatic or chromatic light-curves, or light-curves from the asteroseismology channel) ;
- **teff:** effective temperature of the star (in K) ;
- **grav:** surface gravity of the star in log decimal ;
- **alpha:** right ascension (ICRS, Epoch 2000) ;
- **delta:** declination (ICRS, Epoch 2000) ;
- **mag\_v:** magnitude in the V band (depends on the origin and type of the light-curve) ;
- **mag\_r:** magnitude in the R band (depends on the origin and type of the light-curve) ;
- **specType:** spectral type ;
- **subClass:** spectral type sub-class ;
- **lumClass:** luminosity class ;
- **colDef:** star color (B-R for *Kepler*, V-R for CoRoT).

## 5.2 Indice table

This table contains the seismic indices, the granulation parameters together with the following auxiliary information:

- **Origin:** as in the `Star` table ;
- **starID:** as in the `Star` table ;
- **type:** name of the seismic indice (see below) ;
- **value:** value of the seismic indice (Negative if unknown) ;
- **precP:** formal error (+1- $\sigma$  uncertainty) ;
- **precN:** negative component of the formal error (-1- $\sigma$  uncertainty) (only for those parameters with asymmetric formal errors) ;
- **status:** status of the indice, `status=0` for a valid indice otherwise a non zero value ;
- **genDate:** generation date ;
- **genVersion:** version of the generation pipeline ;
- **ppVersion:** version of the pre-processing ;
- **extVersion:** version of the extraction program ;
- **comment:** any relevant comment associated with the indice.

## 6 Acknowledgements

The Stellar Seismic Indices Data Base has been developed in the framework of the SPACEInn project (Exploitation of Space Data for Innovative Helio- and Asteroseismology), initiated by the European Helio- and Asteroseismology Network (HELAS) and financed by the European Union under the Seventh Framework Programme (FP7 project number 312844). We acknowledge the Paris Data Centre for their financial and technical support. We thank S. Cnudde (LESIA) for the conception of the SSI logo and F. Henry (LESIA) for technical supports. We acknowledge financial supports from Programme National de Physique Stellaire (PNPS).

The CoRoT space mission has been developed and operated by CNES, with contributions from Austria, Belgium, Brazil, ESA (RSSD and Science Program), Germany, and Spain. Funding for the *Kepler* Discovery mission were provided by NASA's Science Mission Directorate.

## References

- Belkacem, K. 2012, in SF2A-2012: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, ed. S. Boissier, P. de Laverny, N. Nardetto, R. Samadi, D. Valls-Gabaud, & H. Wozniak, 173–188, arXiv:1210.3505
- Belkacem, K., Samadi, R., Mosser, B., Goupil, M.-J., & Ludwig, H.-G. 2013, in Astronomical Society of the Pacific Conference Series, Vol. 479, Progress in Physics of the Sun and Stars: A New Era in Helio- and Asteroseismology, ed. H. Shibahashi & A. E. Lynas-Gray, 61
- Chaintreuil, S., Deru, A., Baudin, F., et al. 2016, II.4 The "ready to use" CoRoT data, ed. CoRot Team, 61
- de Assis Peralta, R. A., Samadi, R., & Michel, E. 2016, MNRAS, submitted
- Kallinger, T., De Ridder, J., Hekker, S., et al. 2014, A&A, 570, A41
- Leroy, B. 2012, A&A, 545, A50

- Mathur, S., Hekker, S., Trampedach, R., et al. 2011, ApJ, 741, 119
- Mosser, B., Belkacem, K., Goupil, M. J., et al. 2011, A&A, 525, L9
- Ollivier, M., Deru, A., Chaintreuil, S., et al. 2016, II.2 Description of processes and corrections from observation to delivery, ed. CoRot Team, 41
- Samadi, R., Belkacem, K., & Ludwig, H.-G. 2013a, A&A, 559, A39
- Samadi, R., Belkacem, K., Ludwig, H.-G., et al. 2013b, A&A, 559, A40