Promoting access to and use of seismic data in a large scientific community

SpaceInn data handling and archiving

Eric Michel[1,⋆], *Kevin* Belkacem[1], *Reza* Samadi[1], *Raphael* de Assis Peralta[1], *Christian* Renié[1], *Mahfoudh* Abed[1], *Guangyuan* Lin[1], *Jørgen* Christensen-Dalsgaard[2], *Günter* Houdek[2], *Rasmus* Handberg[2], *Laurent* Gizon[3], *Raymond* Burston[3], *Kaori* Nagashima[3], *Pere* Pallé[4], *Ennio* Poretti[5], *Monica* Rainer[5], *Angelo* Mistò[5], *Maria Rosa* Panzera[5], and *Markus* Roth[6]

[1] *Observatoire de Paris – LESIA, UMR-8109, Université Pierre et Marie Curie, Université Paris-Diderot, PSL, Paris, France*
[2] *Stellar Astrophysics Centre, Department of Physics and Astronomy, Aarhus Univ., Aarhus, Denmark*
[3] *Max-Planck-Institut für Sonnensystemforschung, 37077 Göttingen, Germany*
[4] *Instituto de Astrofísica de Canarias/Department de Astrofísica, Universidad de La Laguna, La Laguna, Tenerife, Spain*
[5] *INAF-Osservatorio Astronomico di Brera, Merate, Italy*
[6] *Kiepenheuer-Institut für Sonnenphysik, Sconeckstr. 6, 79104 Freiburg, Germany*

Abstract. The growing amount of seismic data available from space missions (SOHO, CoRoT, Kepler, SDO,...) but also from ground-based facilities (GONG, BiSON, ground- based large programmes...), stellar modelling and numerical simulations, creates new scientific perspectives such as characterizing stellar populations in our Galaxy or planetary systems by providing model-independent global properties of stars such as mass, radius, and surface gravity within several percent accuracy, as well as constraints on the age. These applications address a broad scientific community beyond the solar and stellar one and require combining indices elaborated with data from different databases (e.g. seismic archives and ground- based spectroscopic surveys). It is thus a basic requirement to develop a simple and efficient access to these various data resources and dedicated tools. In the framework of the European project SpaceInn (FP7), several data sources have been developed or upgraded. The Seismic Plus Portal has been developed, where synthetic descriptions of the most relevant existing data sources can be found, as well as tools allowing to localize existing data for given objects or period and helping the data query. This project has been developed within the Virtual Observatory (VO) framework. In this paper, we give a review of the various facilities and tools developed within this programme. The SpaceInn project (Exploitation of Space Data for Innovative Helio- and Asteroseismology) has been initiated by the European Helio- and Asteroseismology Network (HELAS).

1 Introduction

With about 150,000 stars observed with CoRoT over periods going up to five months and about 200,000 stars observed with *Kepler* over periods up to 4 years, stellar seismology has known an unprecedented acceleration over the past decade. It has become common to use seismic data for hundreds or thousands of stars, and this information is fruitfully coupled with spectroscopic or astrometric informations available from various large surveys. This is true for studies addressing stellar structure and its evolution, but also for studies addressing galactic population formation and migration or for studies questioning planet formation. The stellar light-curves have also revealed signatures of other phenomena than just pulsation, like granulation, flares or activity, and the reference to the Sun is more than ever necessary to understand and interpret these phenomena which can be studied in great detail in the present wealth of solar data. Within the SpaceInn project, we have developed several facilities to make this information available for a large scientific community. Several data sources have been developed or upgraded and a portal has been created to help the access to this information.

2 The Seismic Plus Portal

The Seismic Plus Portal (see figure 1) is intended to be a hub fostering access to existing seismic data sources and to the most relevant complementary data sources.

It offers three main functionalities:

- An extended census and an homogeneous and synthetic description of the most relevant data and data sources (see Sect. 2.1).
- localizing and quick-look of the data available for a given list of stars or for a given time period (see Sect. 2.2).
- Tools for data handling, including tools transmitting coordinated download queries to various data sources and

⋆e-mail: Eric.Michelatobspm.fr

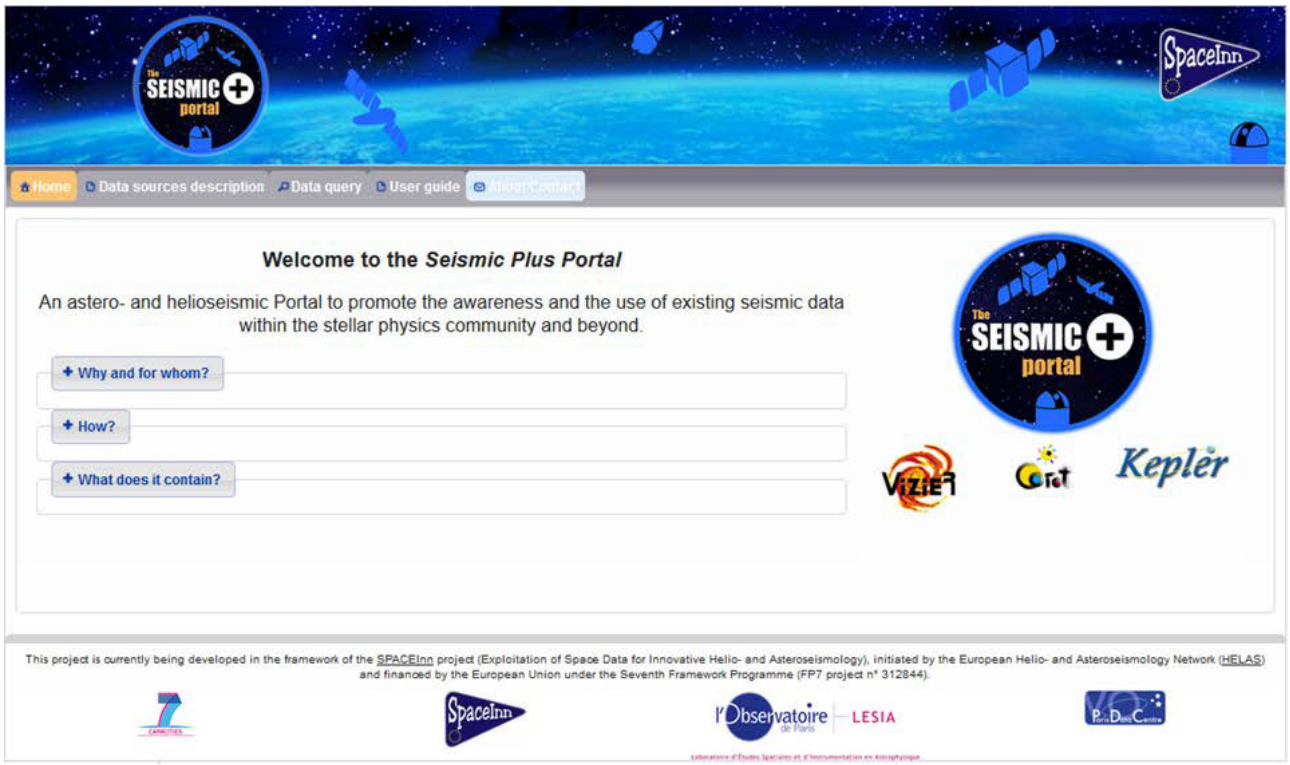

Figure 1: Home page of the Seismic Plus Portal, at http://voparis-spaceinn.obspm.fr/seismic-plus/

tools providing higher-level outputs combining information from various data sources (see Sect. 2.3).

2.1 The data and data sources

Beyond space and ground-based seismic data and data sources, the portal also addresses most relevant related sources of complementary data (e.g. astrometry, spectroscopy, ...) commonly used to complement seismic studies of stars.

For convenience, data sources and data products are distibuted under the following categories:

- Time series: they constitute the main observational material for seismology. It can be composed of photometric measurements, radial velocities, magnetograms, ... It can be 1D time series for integrated-light measurements but also 2D or 3D for time series of spectra or magnetograms. The total duration, the time sampling, and the duty cycle are the specific characteristic parameters which, along with the more classical S/N, characterize these data and show how they can be used.
- Spectroscopy/individual spectra: seismic studies often rely on stellar modeling which benefits from additional constraints brought by spectroscopic analysis (e.g. for [Fe/H] and $T_{\rm eff}$ estimates). On the other hand, it has become common, for red giant stars at least, to improve spectroscopic detailed analysis by using so-called seismic log g estimates as additional constraints. It is thus important to bridge the gap between these types of data in the context of seismology space missions and spectroscopic large programs.
- Stellar parameters: With space missions like CoRoT and Kepler, a new concept appeared, namely 'ensemble asteroseismology'. It deals with large samples of stars and allows stellar population studies. The characterization of these large samples relies on observational indices. In addition to the classical photometric color indices or parallaxes, new indices have emerged which characterize the stellar pulsation and the stellar structure to first order (large separation, frequency of maximum amplitude, ...).
- others: The use and application of seismic data is growing and evolving quickly. The list of data sources and data products considered at the Seismic Portal has to remain open to changes in the observational landscape and in current scientific issues.

The various data are distributed under these categories in a synthetic table (see figure 2). A brief complementary information is also available locally, and links toward the reference site or publication are also given.

2.2 Localizing and inspecting data available for a list of stars or a time period

Two modes of query exist at the moment. The first one, for a given list of stars, returns a localization table stressing which kind of data are available and where (see figure 3). This information is available for each star of the list and quick-look tools allow a first inspection of the data in order to determine to which extent they are appropriate for the anticipated application.

Figure 2: A few entries in the list of data sources referenced at the Seismic Plus.

A second mode of query exists for solar data. For a given time period, it returns a table stressing which kind of time series exists and where. It also provides information about occurence of solar events (essentially coronal mass ejections (CME) and flares so far) over the same period. The intention here is to help establishing connexion between solar events and their signatures in a time series, thus helping the interpretation of such signatures in the light-curve of other stars.

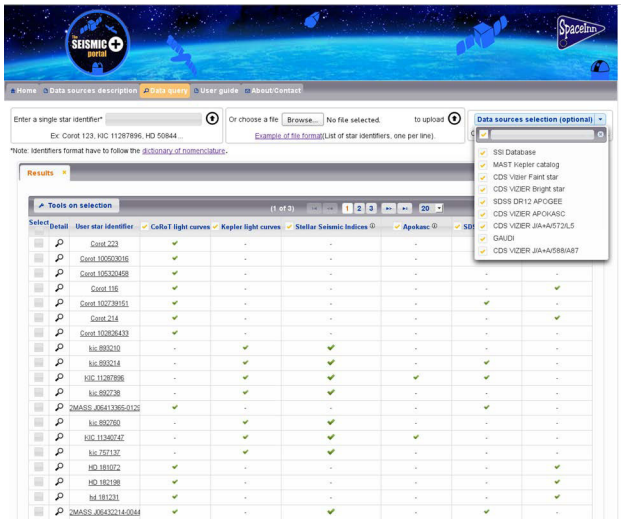

Figure 3: Result of a data query for a list of stars. Each line corresponds to a given star and green signs indicate existing data in the data source associated with the corresponding column.

2.3 Tools for coordinated download of the data and higher-level outputs

Once the result of the query is obtained, as in figure 3, it is possible to select a set of stars. The portal then offers the possibility to transmit a query to various data sources in order to download the corresponding data. Various formats are proposed. The portal is also intended to be the place where higher-level products can be built, combining several types of data from different sources. A first tool allows for instance to derive seismic estimates of the mass, radius and log g, for a selected list of stars and for various sources of seismic indices and effective temperatures. The procedures are based on scaling laws which are documented and the publication references are indicated.

3 Data sources developed or upgraded within the SpaceInn WP3 activities

Among the data sources referenced at the Seismic Plus Portal, several have been developed or upgraded within the SpaceInn activities. They are accessible via the Seismic Plus Portal but also via their dedicated web interface.

3.1 The SISMA spectral time series database

Thousands of high-resolution spectra were taken with the spectrograph HARPS mounted at the 3.6m ESO telescope at La Silla, Chile ([4]) during the ESO Large Programmes LP 182.D-0356 and LP 185.D-0056 as ground-based counterpart of the CoRoT ([1]) photometric measurements ([6]). The SISMA database (http://sisma.brera.inaf.it/ see figure 4) gathers 7103 fully reduced and normalized HARPS spectra, obtained for 261 stars, among which are 71 CoRoT bright targets observed simultaneously with the satellite. These spectroscopic time series allow to better exploit the photometric data and characterize the variability classes of CoRoT targets.

In addition to the spectra, SISMA offers mean line profiles (computed with the LSD software [3]), radial velocities, $V \sin i$, activity indices, differential rotation indicators, binarity and emmision flags. It also features the physical parameters of each target ($T_{\rm eff}$, log g, [Fe/H]) and, in the case of CoRoT targets, the relevant satellite light curves (see [5]).

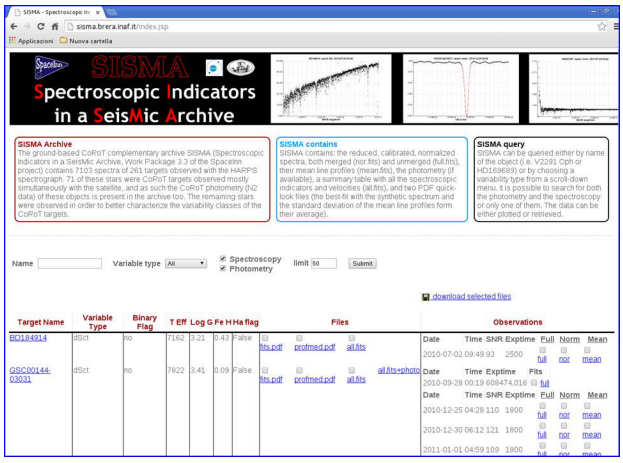

Figure 4: Home page of the SISMA database.

3.2 The KASOC and SpaceInn

The *Kepler* Asteroseismic Science Operations Center (http://kasoc.phys.au.dk/) provides asteroseismological data from the NASA *Kepler* mission to astronomers who are members of the *Kepler* Asteroseismic Science Consortium (KASC). In the framework of the SpaceInn project, efforts have been made to enhance reliability and efficiency of the access to *Kepler* data. A VO interface has been developed allowing for external access to KASOC database via the Virtual Observatory protocol. Various type of complementary information have been included in the database, like ground-based follow-up data and stellar models related to published studies of specific objects (figure 5).

In addition, a solution for long-term preservation of the data has been studied in collaboration with the Royal Library of Denmark.

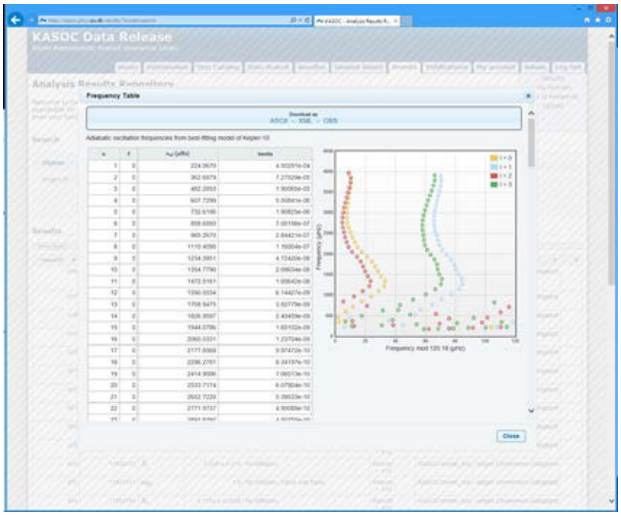

Figure 5: Interface allowing theoretical-frequency handling and visualisation at KASOC.

3.3 The MARK-I database

This historical archive (1976 till 2012) has been developed with the solar observations gathered at the Observatorio del Teide (IAC) with the pioneer Mark-I resonant-scattering spectrophotometer (http://svo2.cab.inta-csic.es/vocats/marki/). Observations extend from 1976 to 2012 with only summer campaigns from 1976 to 1983.

The data products are radial velocity measurements of the Sun-as-a-star at the Potassium KI 7699A absorption solar line.

3.4 The German Data Center for SDO and SpaceInn

The German Data Center for SDO (GDC-SDO) hosts the European depository for HMI data. The data processed and made available from the Joint Science Operations Center (JSOC) headquartered at Stanford University USA are immediatly transfered to GDC (several TBs per day). This constitutes the largest source of data available today for helioseismology (online storage capacity over 1 PB). The data products are line-of-sight Dopplergrams, line-of-sight magnetograms, continuum intensities, subset of Milne-Eddington inversions. They are available at various cadences.

Within the SpaceInn Project, a dedicated GDC-SDO Help Desk has been developed for SpaceInn (http://www2.mps.mpg.de/projects/seismo/GDC-SDO/spaceinn_access.htm). It contains information for setting up NetDRMS and receiving relevant data and users support.

3.5 The Stellar Seismic Indices database

The Stellar Seismic Indices (SSI) database (http://ssi.lesia.obspm.fr/ see figure 6) has been developed within the SpaceInn programme (see also [7]). It contains stellar seismic indices of solar-like oscillating stars as well as characteristic parameters of stellar granulation. These seismic indices obey characteristic scaling relations that depend directly on the radius, mass and effective temperature of the star. From the knowledge of these quantities it is possible to estimate the mass and radius of a star, and subsequently the surface gravity of the star (log g). The indices populating the current database were extracted from about 13,000 Kepler and 5,000 CoRoT red-giant stars using an automatic analysis pipeline [2].

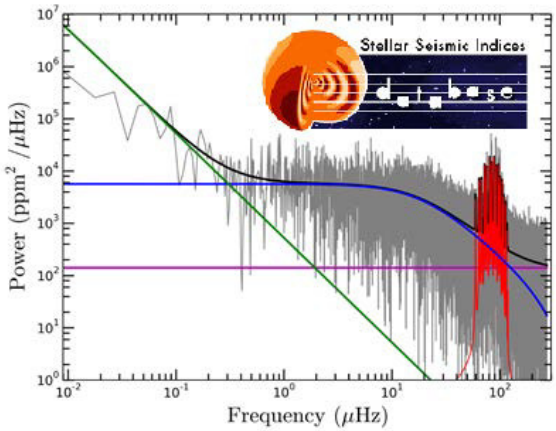

Figure 6: The SSI database provides seismic indices and granulation parameters characterizing solar-like pulsators.

4 Perspective

In the perspective of future projects like TESS (launch 2017) and PLATO (launch 2025) which will enrich considerably the seismic material available to the community and in the perspective of the Gaia wealth of information, it seems very relevant to foster as efficiently as possible the access to these data and techniques necessary to exploit them with the best outcome.

Acknowledgements

This project has been developed within the European FP7 SpaceInn project (SpaceInn.eu). We acknowledge the great help from Spanish VO (E. Solano), Paris Astronomical Data Centre (P. Le Sidaner), Copenhagen Royal Library, INAF-Osservatorio Astronomico di Treste (M. Molinaro and F. Cepparo).

References

[1] Baglin, A. et al., AIPC **895**, 201 (2007)
[2] de Assis Peralta, R. et al. submitted
[3] Donati, J.-F. et al., MNRAS **291**, 658 (1997)
[4] Mayor, M. et al., The Messenger **114**, 20 (2003)
[5] Rainer, M. et al., Astronomical Journal **152**, 207 (2016)
[6] Rainer, M. et al., AN **333**, 1061 (2012)
[7] Samadi, R. et al., these proceedings