Mem. S.A.It. Vol. 78, 172

Memorie della

BASS2000-Tarbes: current status and THEMIS data processing

N. Meunier, M. Lafon, P. Maeght, F. Grimaud, and Th. Roudier

Laboratoire d'Astrophysique de l'Observatoire Midi-Pyrénées, Université Toulouse III, 57 Avenue d'Azereix, BP 826, 65008 Tarbes Cedex, FRANCE
e-mail: meunier@ast.obs-mip.fr

Abstract. I will review the history and status of the data archive BASS2000 and will concentrate my presentation on the BASS2000-Tarbes data base, which contains a very large volume of THEMIS data, i.e. spectropolarimetric data. I will insist the implementation of the processing of MTR-THEMIS (multi-line spectropolarimetry) data by the BASS2000 team, which has been our main project in 2006. New data levels are Stokes profiles and clean spectra, maps of continuum intensity and line-center intensity, Dopplergrams, magnetograms and vector magnetic field maps. I will also present the tools and services that we are providing.

Key words. Sun: magnetic fields – Sun: sunspots – Sun: faculae, plages – Sun: atmosphere – Sun: photosphere – Sun: chromosphere – Sun: filaments – Sun: flares

1. Introduction

The purpose of the database BASS2000 (http://bass2000.bagn.obs-mip.fr/) is to archive most french solar data obtained with ground-based telescopes. The objectives are the following:

- It is necessary to standardize the data format in order to facilitate the data processing and the use of the large amount of data by a wide community.
- The data themselves must be preserved on the long-term and provided to the community on request.
- It is crucial to provide reduced data for non-specialists in order to reach as many users as possible.
- It is very important to provide tools (codes, informations) to help users to exploit the data.

Our teams have worked in these three directions over the past years. These activities correspond to two types of needs. First, it is very useful to have on-line observations in various spectral domains in order to prepare observations at a given observatory. For this purpose, the data must be put on-line very fast, typically daily. More specific observations, such as those provided by THEMIS, do not need to be available that fast but still need to reach a large community if one want to improve the scientific return. One should note that most observations of that type usually remain private, as only systematic observations are made public in most ground-based observatories. BASS2000 is therefore a unique database.

Send offprint requests to: N. Meunier

BASS2000 is constituted of two archives, one in Tarbes, which is the subject of this paper, and one in Meudon. The first archive contains mostly data from the THEMIS telescope, as well as all raw data from the Nançay Radio Heliograph, processed data from the Pic du Midi Coronagraph (on-going, already 5 years of data in the catalogue), and some data from the Lunette Jean Rösch, also at the Pic du Midi (mostly imagery). The second archive, in Meudon, provides an on-line selection of data, mostly full-disk images, which are very useful to prepare observing runs for example.

This paper mostly concerns the data from the THEMIS telescope, which is the subject of this conference and which also consitute most of the volume archived in our database. In Sect. 2, we describe in more detail the database itself and the solar physics related to the THEMIS data. Then in Sect. 3 we present our main on-going project, the implementation of the processing of multi-line spectropolarimetric data obtained at THEMIS. We conclude with our short-term and long-term priorities in Sect. 4 as well as on the calendar for next year (see also Meunier et al. 2005a, 2006).

2. The current data base

2.1. The content of the database

The Tarbes archive currently contains the following data:

- Raw THEMIS data in the MTR mode (multi-line spectropolarimetry) and MSDP mode (spectro-imagery with polarimetric capabilities): since 1999. The volume is about 7.1 Tb (see Sect. 2.2 for details).
- Processed THEMIS data in the MSDP mode: 2002-2004, for a volume of about 80 Gb.
- Processed THEMIS data in the MTR mode: part of 2004-2005 (see below, on-going work).
- Raw Nançay Radio Heliograph data (visibilities): since 1997. The volume is about 770 Gb.
- Calibrated data from the Pic du Midi Coronagraph (HACO) and catalogue of structures (jets, prominences) derived from this data set: 1997-2001 (1994-1996 and 2002-2006 : on-going). 80 Gb of compressed data are available.
- Imagery and spectro-imagery data obtained at the Lunette Jean Rösch (Pic du Midi), with a small field-of-view and with the new large-scale camera CALAS (Meunier et al. 2003, 2005b): on-going work.

2.2. The THEMIS data

We focus now on the THEMIS data. THEMIS is a franco-italian telescope localized in Tenerife and operating since 1999. THEMIS includes three observing modes :

- The MTR mode : multiline spectropolarimetry. Three instrumental set-ups are possible : the 2x2' mode (field-of-view of 2', with the two optical beams on two different cameras) ; the 2x1' mode (field-of-view of 1', with the two optical beams on the same camera) ; the grid mode (small interlaced field-of-view on the same camera).
- The MSDP mode : double pass spectrograph, allowing spectro-imagery with polarimetric capabilities for a a few selected lines.
- The IPM mode : high spectral resolution imagery allowing a scan in wavelength across a few selected lines.

The data obtained with the two first instruments provided by France are archived at BASS2000 and are publicly available 1 year after the observation. The total volume of raw data is about 7.1 Tb. Because of the large volume involved with THEMIS data, the FITS files are not on-line but can be requested through a user-friendly interface allowing access to the on-line catalogue. Many JPEG files associated with other informations (scintillations, observation logbooks, etc., subset of data) help the users to prepare their requests. More than 40% of the volume of raw data has been provided by BASS2000 to users, mostly to the observers themselves (via magnetic tapes, DVD, ftp or on external disks). The reason is that in many cases the observers do not bring back all their observations but only a subset of data.

At this date, ~40 publications with referee are using THEMIS data which are publicly available in our database (access through requests on our website). The observations and publications cover a wide range of solar features, from active regions, sunspots and plages to the magnetic network and very quiet regions, as well as prominences and filaments. Note that these spectra do not only include many lines where the Zeeman effect is present, but many runs are also using the Hanle effect close to the limb or in prominences to determine small-scale turbulent magnetic fields.

3. Implementation of THEMIS data processing

The final objective of this work is to improve the scientific return of the THEMIS observations and BASS2000 work. In order to achieve this goal, we are proposing and will propose ready-to-use data, which should interest a wide community. Some of them will be used by scientist who are specialized in other types of data such as radio or UV: They will be mostly interested by maps, for example the vector magnetic field maps. Other data levels will focus on spectro-polarimetric specialists, that will be mostly the case of Stokes profiles or clean spectra. The processed data could be used in several ways :

- By themselves.
- Along with other data which have been obtained in a coordinated campaign including THEMIS (for example a Joint Observing Programs with SOHO).
- Along with other data in "a posteriori" coordination : this will mostly concern simultaneity with systematic observations, either groundbased or from space.

3.1. MSDP data processing

In the case of the MSDP data (spectro-imagery with polarimetric measurements), we have used the code developed by P. Mein (see for example Mein 2002). The BASS2000 team has then developed an interface in order to process a large amount of data at a time (typically one day of data), and to provide on-line useful JPEG files which should help the users to select the interesting data. This interface has also been made publicly available in order to help the community who is also processing these kind of data. So far, 3 years of MSDP data have been processed and are now available in the database, for the runs corresponding to standard observations. We propose intensity maps, Dopplergrams and magnetograms, as well as Stokes I and V profiles at a medium spectral resolution. The processing of 2005-2006 data as well as of older data should be pursued in 2007.

3.2. MTR data processing

In the case of the MTR data processing, the situation is quite different as the observations can be very different from one another: more variety in the observed objects and in the instrumental configuration, including three modes, and more than 100 observed spectral domains covering most of the visible spectrum. This means that the codes must be very robust in order to allow the easy processing of the large volume (several Tb). This also implies that to provide a good visibility of the processed data, we must avoid to include in the data base some data which would have been processed following different assumptions by various groups : in addition to be more difficult to deal with on the operationnal point of view (different input/output formats etc.), it would then be more difficult for the users to understand how the data has been processed. This is critical as the raw data are already quite complex. Furthermore, there is not a single code to process this type of data, as many groups have developped their own software. It was therefore necessary to select specific codes.

Our goal is to provide ready-to-use Stokes profiles (for each pixel of the map derived from the scans), intensity maps, Dopplergrams and magnetograms, and as many vector magnetic field maps as possible.

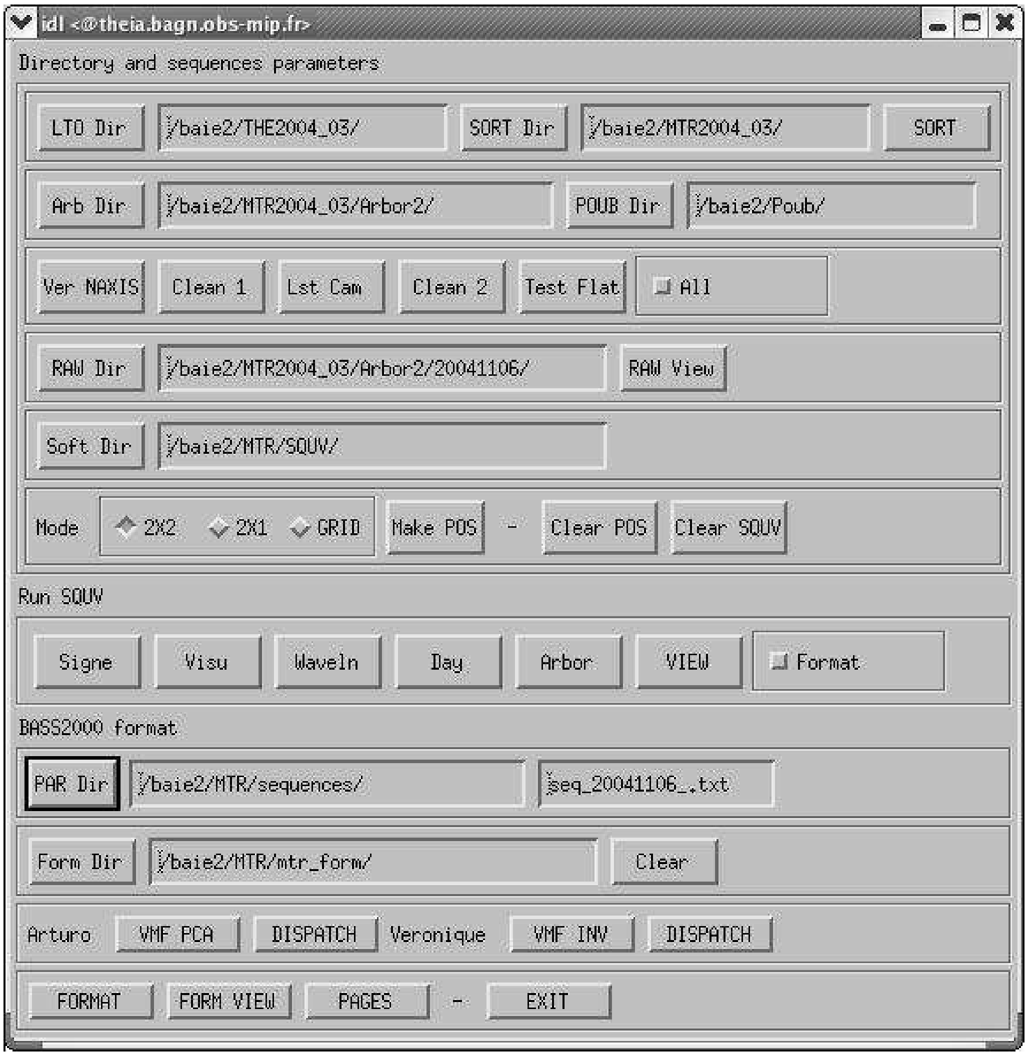

Fig. 1. Interface developped to process automatically a large amount (1 run of several days at a time) of MTR data.

Several softwares have been provided to us in order to perform some important steps of the processing :

- The code allowing to obtain the Stokes profiles (or clean spectra when not enough Stokes parameters were observed) has been provided by A. Sainz Dalda (THEMIS). This code, SQUV, starts with a set of files containing observations, dark-current and flat-field for a given spectral domain. It is very robust and suitable to process a large amount of data, as is the case here. The processing is optimized for a given line in the spectra, which is used to align the spectra with respect to each other: when several interesting lines are present, we repeat the processing for as many line as necessary. More details can be found in Sainz Dalda & López Ariste (2006).
- The code extracting the vector magnetic fields from the Stokes spectra in the 630

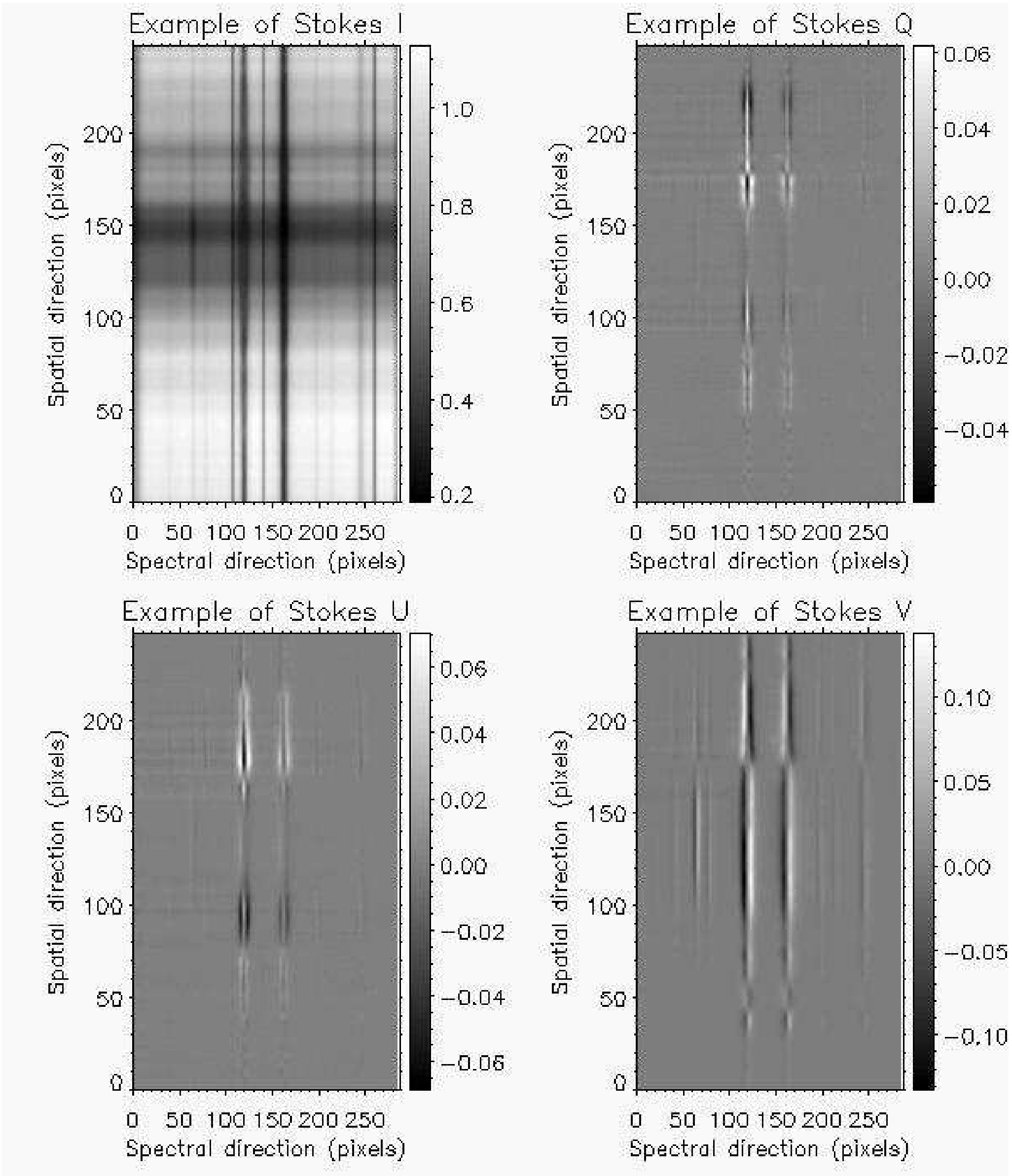

Fig. 2. Example of Stokes spectra for the 630 nm spectral domain (first step of a scan). Observations by G. Molodij (October 6th, 2004).

nm domain, using a PCA algorithm, has been provided by A. López Ariste (INVFE code), see e.g. Rees et al. (2000) for more details. Additionnal parameters are provided (thermodynamical parameters in particular) as well as uncertainties on the results.

– An inversion, also providing the vector magnetic field, is performed using the UNNOFIT algorithm and the code has

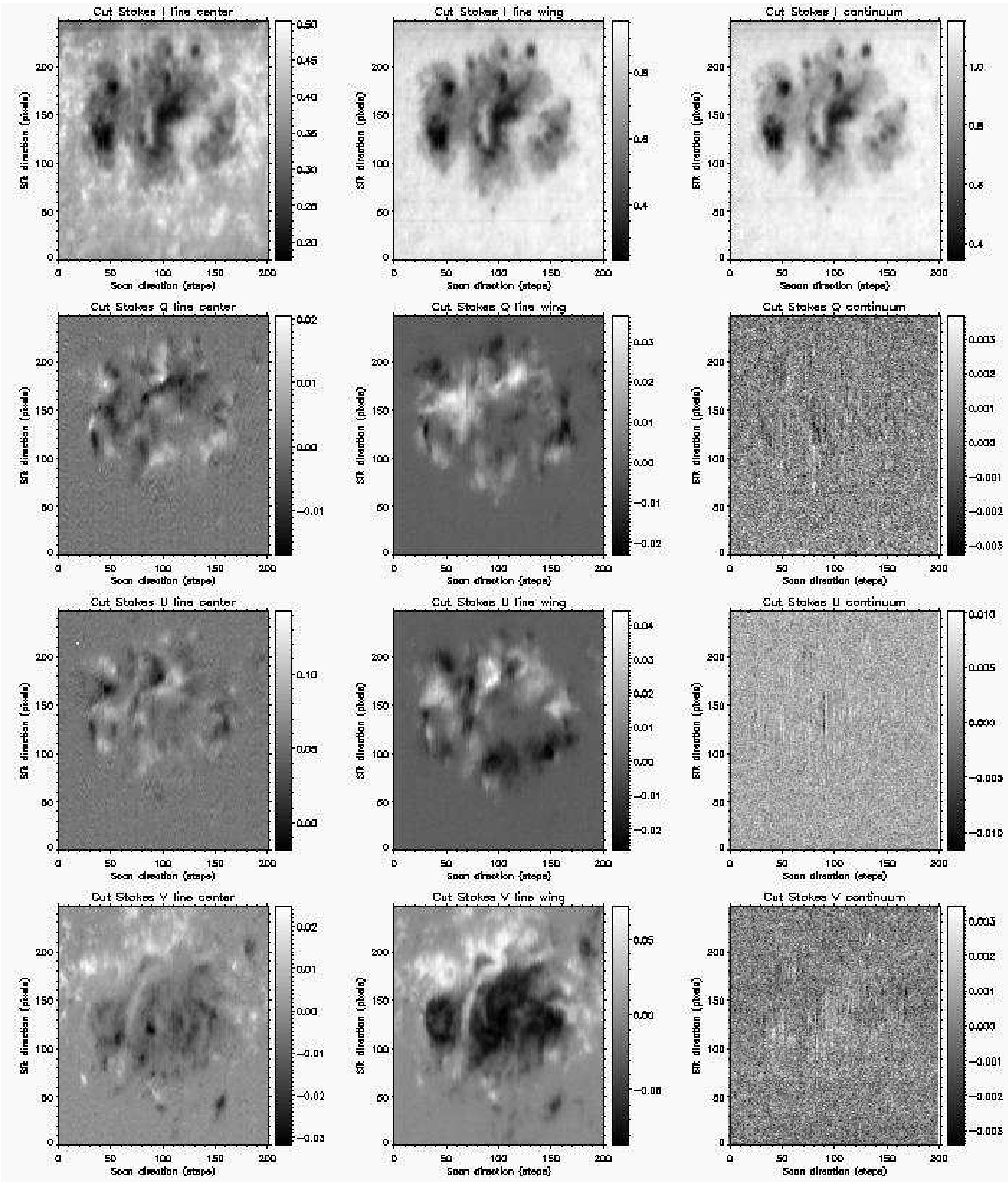

Fig. 3. Example of cuts through Stokes spectra for the 6301.5 Å line, at various wavelengths : at line center (left panels), in the line wing (center panels) and in the continuum (right panels). Observations by G. Molodij (October 6th, 2004).

been provided by V. Bommier (LERMA-Observatoire de Paris). It is based on an analytical solution of Unno-Rachkowski, which provides the polarization emerging from a Milne-Eddington atmosphere for a given magnetic field. The fit is done using a Marquardt algorithm. We refer to Bommier et al. (2006) for more details. This approach is based on the work of Landolfi et al. (1984), to which an addi-

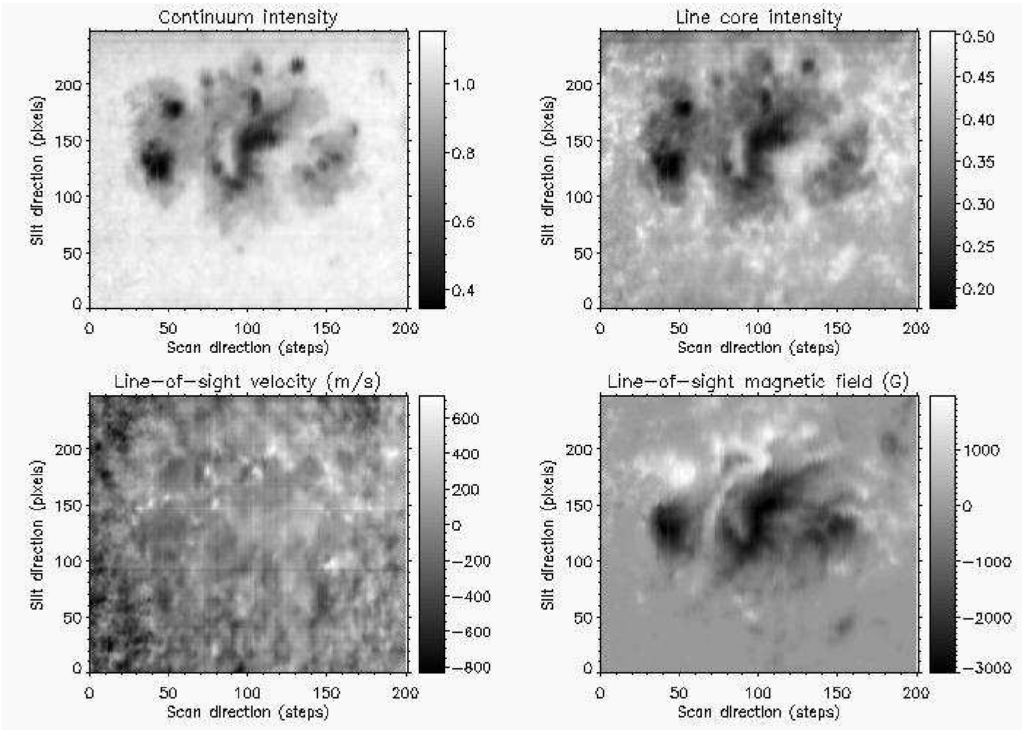

Fig. 4. Example of maps for the 6301.5 Å line: Intensity in the continuum and line core, Doppler map and line-of-sight magnetic field. Observations by G. Molodij (October 6th, 2004).

tionnal parameter is added (the magnetic filling factor). Parameters corresponding to the Milne-Eddington models are also provided. In principle, the code can be used for any lines. In practice, due to the CPU cost, we are performing the computations routinely only for the 6302.5 Å and 5250.2 Å lines. The possibility to offer to perform the inversion on specific sequences on request for many other lines will have to be examined. More lines may be processed if the speed of the code is improved: this is on-going work.

The work of the BASS2000 team has then been to develop and test a number of routines to perform automatically several steps of the complete analysis:

- Organize the files of an observing run as a function of the various modes (same spectral domain on a given camera for example).
- Eliminate bad data (intrinsically bad data or data which are not standard, ...) or very short scans.
- Prepare the inputs for the processing codes, which include a comparison of the raw spectrum with a reference spectrum in order to calibrate the spectra in term of wavelength, as well as to extract the position of the useful spectral lines for which we want to optimize the processing.
- Call the provided routines
- Compute the intensity maps (continuum and line center), Dopplergrams and magnetograms.
- Compute some additionnal parameters that will help the users to select the data.
- Finalize the formatting of the data and prepare the on-line JPEG files.

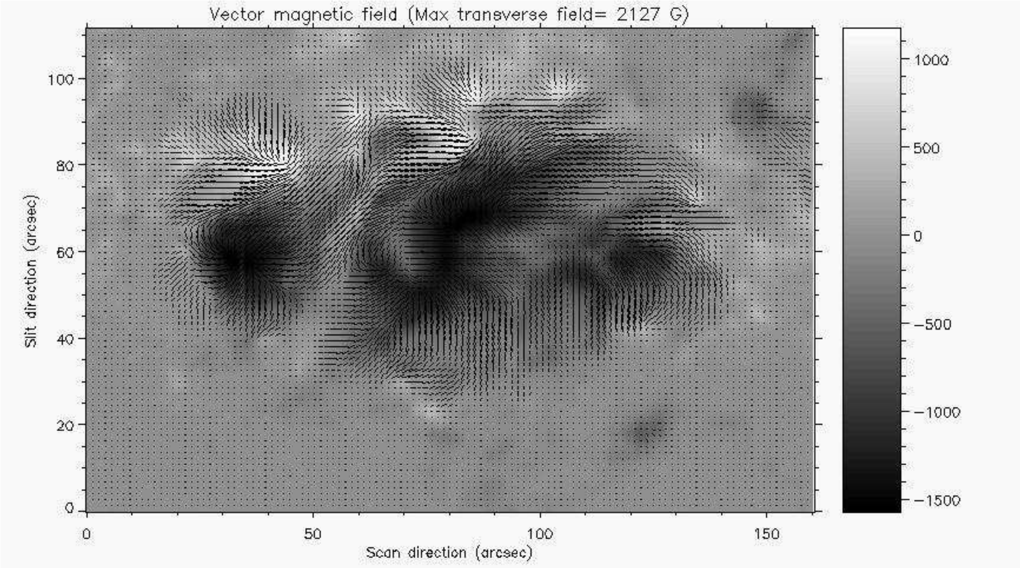

Fig. 5. Example of a vector magnetic field map for the 6302.5 Å line. Gray levels shows the line-of-sight magnetic field, while dashed lines show the transverse component (no arrow are shown as the 180° fundamental ambiguity has not been solved). Observations by G. Molodij (October 6th, 2004).

– Include the results in a test base for validation.
– Save the results and include them in the on-line data base. The data can then be requested through the request form at http://bass2000.bagn.obs-mip.fr/Query/index.php4.

These different steps have been tested for the various modes. Fig. 1 shows an example of the interface that has been developped. An example of Stokes spectra obtained with a standard processing is shown in Fig. 2. The cuts at various wavelengths are shown in Fig. 3. The corresponding maps of intensities, line-of-sight velocity fields and magnetic fields are also illustrated in Fig. 4. The vector magnetic field using the UNNOFIT inversion is shown in Fig. 5. The pixel sizes (spatially along the slit and spectral dispersion) are computed at BASS2000 as they are not provided in the headers. The former is done by considering that the height of the observed field-of-view, as defined by a threshold corresponding approximately to the inflexion point of cuts through the spectra, is known and fixed (for example 2 arcmin in the 2x2' mode). The former is computed by comparing the spectrum to a reference spectrum. This comparison also allows to identify the interesting lines. The pixel size of the map in the scan direction is given by the scan step in the header. It should be noted however that this step may be smaller than the slit size. This is not necessarily a problem since the seeing is usually larger than the slit width. Concerning the final products, the vector magnetic fields are given with respect to the line-of-sight and not in the solar reference. Finally, the fundamental 180° ambiguity on the azimuth is not corrected.

We plan to process data obtained after 2004 (included) first for active regions and filaments, as well as quiet sun observations (outside Hanle effect at the limb). The observations of the solar second spectrum close to the limb will be processed in a second step, possibly with an updated version of the code. The prominence observations will probably be left out of the processing as they are not standard. It is an open question whether it will be possible to process older data, as some crucial informations are missing from the headers.

Furthermore, depending on the amount of requests following this implementation (all these data will be publicly available), it will be necessary to identify, in the near future, what kind of additionnal data level would be useful to the solar community.

4. Conclusion

On the short term, we will focus our efforts on the processing of a large amount of spectropolarimetric data obtained at the THEMIS telescope. These processed data will be publicly available. Data from other instruments will also be included or continue to be included in the database. Data from the Lunette Jean Rösch will be included, including granulation images which have been numerized (covering more than a solar cycle), more recent granulation images obtained with a CCD, and future observations at high spatial resolution and on a large field-of-view obtained with CALAS (Meunier et al. 2003, 2005b). Rondi et al. (2006) shows some first images of CALAS. We have started to add some data from the Pic du Midi Coronagraph (HACO), and this work must continue. A new instrument, CLIMSO, will replace it in HACO, and the data will also be available at BASS2000 in 2007. Of course current and future THEMIS and NRH observations will continue to be included in the catalogue.

On a longer term, it is very important that the solar databases BASS2000 and MEDOC get involved in a Virtual Observatory in relation with the plasma community (Earth and planets of the Solar System). In order to follow the implementation for the data processing and possible new data levels elaborated from the vector magnetic fields, it will be necessary that the solar community gets more involved in BASS2000, and all these new projects will be developed only if that it the case.

Acknowledgements. BASS2000 in Tarbes is funded by the Programme National Soleil-Terre, the Observatoire Midi-Pyrénées and the Laboratoire d'Astrophysique de l'Observatoire Midi-Pyrénées. The fortran code to process MSDP data has been provided by P. Mein. The IDL code to process raw MTR data into Stokes profiles has been provided by A. Sainz Dalda, and the fortran codes to compute the vector magnetic fields have been provided by A. López Ariste (PCA algorithm) and V. Bommier (UNNOFIT).

References

Bommier, V., Landi Degl'Innocenti, E., Landolfi, M., & Molodij G., 2006, to appear in A&A

Landolfi, M., Landi Degl'Innocenti, E., & Arena, P. 1984, Sol. Phys. 93, 269

Mein, P. 2002, A&A 381, 271

Meunier, N., Rieutord, M., & Beigbeder, F. 2003, in SF2A-2003: Semaine de l'Astrophysique Francaise, ed. F. Combes, D. Barret, T. Contini, & L. Pagani, 93

Meunier, N., Lafon, M., Maeght, P., Grimaud, F., & Roudier, Th. 2005a, SF2A-2005: Semaine de l'Astrophysique Francaise, F. Casoli, T. Contini, J.M. Hameury and L. Pagani eds., EdP-Sciences, Conference Series, 169

Meunier, N., Rondi, S., Tkaczuk, R., Rieutord, M., & Beigbeder, F. 2005b, in ASP Conf. Ser. 346: Large-scale Structures and their Role in Solar Activity, ed. K. Sankarasubramanian, M. Penn, & A. Pevtsov, 53

Meunier, N., Lafon, M., Maeght, P., Grimaud, F., & Roudier, Th. 2006, SF2A-2006: Semaine de l'Astrophysique Francaise, F. Casoli, T. Contini, J.M. Hameury and L. Pagani eds., EdP-Sciences, Conference Series, in press

Rees, D. E., López Ariste, A., Thatcher, J., & Semel, M. 2000, A&A 355, 759

Rondi, S., Roudier, Th., Molodij, G., Bommier, V., Malherbe, J.-M., Schmieder, B., Meunier, N., Rieutord, M., & Beigbeder, F. 2006, this volume

Sainz Dalda, A., & López Ariste, A. 2006, submitted to A&A