



**Laboratorio de Astrofísica Espacial y
Física Fundamental**
INSTITUTO NACIONAL DE TÉCNICA
AEROESPACIAL



“Evolution of circumstellar disks in pre-main sequence stars”

A proposal for a VO science case

Authors: E. Solano, R. Gutiérrez, B. Montesinos, B. Merín, L.M. Sarro


Date: May 7, 2004

Pages: 11

Version: 2

Reference: LAEFFSCICASE

Keywords: Virtual Observatory, Scientific Case, Data Mining tools

LABORATORIO DE ASTROFÍSICA ESPACIAL Y FÍSICA FUNDAMENTAL			
Evolution of circumstellar disks in Pre-Main Sequence stars			
Proposal for a VO Science Case			
GRUPO DE OPERACIONES Y ARCHIVOS	2004-04-07	Page 2 of 11	

Abstract

This document describes a science case to be developed in the framework of the Virtual Observatory aiming to characterize the circumstellar disks of pre-main sequence stars by fitting theoretical models to the observed spectral energy distributions. The document also describes the VO-compliant tool designed to make such an analysis. The project has been developed by the Group of Operations and Archives at LAEFF and it is proposed as candidate for the next VO demonstration to be held in Jan 2005. The project is driven by the concept that, in the VO context, the astronomical data archives should not limit themselves to act as data providers, but they should incorporate data mining tools in their systems to optimize the scientific return of the available datasets. The success of this objective requires maintaining a close feedback loop between the scientific community, in particular the potential users, and the VO tools developers and data providers. Both at individual and Institute level, LAEFF possesses the technical know-how required to carry out the design, development and implementation of the analysis tools described in this document as well as the scientific expertise to guarantee its optimum exploitation.

1.- Scientific Rationale


1.1.- Introduction

The study of protoplanetary disks is currently undergoing an exciting stage partly propelled by the discovery of extrasolar planetary systems following the detection of 51 PegB by Mayor & Queloz (1995, Na 378, 355). The possible discovery of telluric planets in the near future, in addition to the jovian-like planets already detected, will pose interesting questions on the formation of extrasolar planetary systems. Knowledge of the properties of protoplanetary disks and how they evolve to debris disks around main-sequence (MS) stars is, therefore, a crucial step in such a process.

According to current theories of planet formation, a dusty disk surrounds the star in the first few million years. After some time, the dust particles will start to aggregate in the disk to form planetesimals or planetary cores. These will eventually drag material from the disk to become planets. The existence of these dusty disks where planets are born has been determined from the excess emission at IR wavelengths of the very young stars (the so-called pre-main sequence stars) harbouring such disks and also recently by direct IR and millimetre imaging.

The IR excess characteristic of circumstellar dust has also been observed around another type of stars, still young but older than the PMS stars. These are called Vega-type stars, are main sequence objects and their IR excesses are different from those detected in pre-main sequence stars. After theoretical considerations on the size of their dust grains and other observed properties, it is believed that these "debris" disks may represent one of the final steps before planet formation.

Theories of dust evolution have given the general outline on how the disks evolve from the optically thick phase around the youngest stars (Herbig AeBe – HAeBe hereafter – and T Tauri stars) to the debris disks possibly harbouring planets around the main sequence Vega-

LABORATORIO DE ASTROFÍSICA ESPACIAL Y FÍSICA FUNDAMENTAL			
Evolution of circumstellar disks in Pre-Main Sequence stars			
Proposal for a VO Science Case			
GRUPO DE OPERACIONES Y ARCHIVOS	2004-04-07	Page 3 of 11	

type stars: after the growth of particles in the disk has reached a certain critical value and the planetary cores have formed -and subsequently swallowed the gas around them-, the remaining disk may have enough mass to force the planet to migrate inwards toward the star. As the mass decreases by accretion onto the star, the time may arrive where the forming planet stays in its orbit and the inner disk is swallowed by the star creating an inner hole in the disk. Gas dissipates in the outer disk, while solid particles remain in the midplane, colliding, growing and eroding, creating a population of small grains that characterises the far infrared emission of the debris disks. Since the material in them may be non primary, both their physical and chemical properties could be different from those of the primary material in pre-main sequence disks.

The understanding of the formation and evolution of planetary systems was one of the primary objectives of the EXPORT consortium (EXoPlanetary Observational Research Team, Eiroa et al. 2000, ASP Conf. Series, vol. 219, p. 3). The team obtained the entire International Time allocated in the Canary Islands Observatories in 1998 (5% of the total time available) and was able to use five large telescopes simultaneously during four observational runs of four days each. Photometric observations from optical to the near infrared, and optical medium and high resolution spectra were obtained for a sample of 70 stars (both PMS and MS stars). One of the driving goals of this effort was the study of the evolution and properties of protoplanetary disks by analysing the SEDs of the young stars, taking advantage of the fact that the EXPORT optical and near-IR photometry were obtained simultaneously. This observational approach is appropriate since T Tauri and HAeBe stars vary markedly in these spectral regimes and a significant part of the total luminosity of the object is radiated by the PMS stellar photosphere at these wavelengths.

Observed spectral energy distributions (SEDs) of PMS stars are widely used to study the properties of protoplanetary disks and to classify T Tauri and HAeBe stars into an evolutionary scheme (e.g. Hillenbrand et al. 1992, ApJ 397, 613). Although the interpretation of a given SED is, to some extent, model dependent, the theoretical modelling of SEDs constitutes an invaluable tool for understanding the structure and properties of protoplanetary disks (e.g. Chiang & Goldreich (1997, ApJ 490, 368; 1999, ApJ 519, 279), D'Alessio et al. (1998, ApJ, 500, 411; 1999, ApJ, 527, 893; 2001 ApJ, 553, 321), Dullemond et al. (2001, ApJ, 560, 957)). One of us (B. Merín) has had the opportunity to build during his Ph.D. thesis the multiwavelength SEDs for the stars in the EXPORT sample. He made use of simultaneous photometry from *U* to *K* bands plus any other photometric measurement published from UV to millimetre wavelengths or available in data archives (*IUE*, *IRAS* and *ISO*). In order to compare the observed SEDs with theoretical disk models, a collaboration with Drs. Paola D'Alessio (UNAM, Mexico) and Nuria Calvet (Smithsonian Astrophysical Observatory, USA) started in 2002. Using the D'Alessio's models it has been possible to fit some detailed SEDs of pre-main sequence stars, thus obtaining a great deal of information on the physical state of the protoplanetary disk (Merín et al. 2004, A&A, 419, 301; figures 1,2).

As stated above, the EXPORT consortium has compiled very complete SEDs for around 30 PMS and Vega-type stars with "debris" disks observed simultaneously from optical to the near infrared. The goal of this project is to repeat the time-consuming analysis made in Merín et al. (2004) for two stars in a much more efficient and uniform way covering the full EXPORT sample.

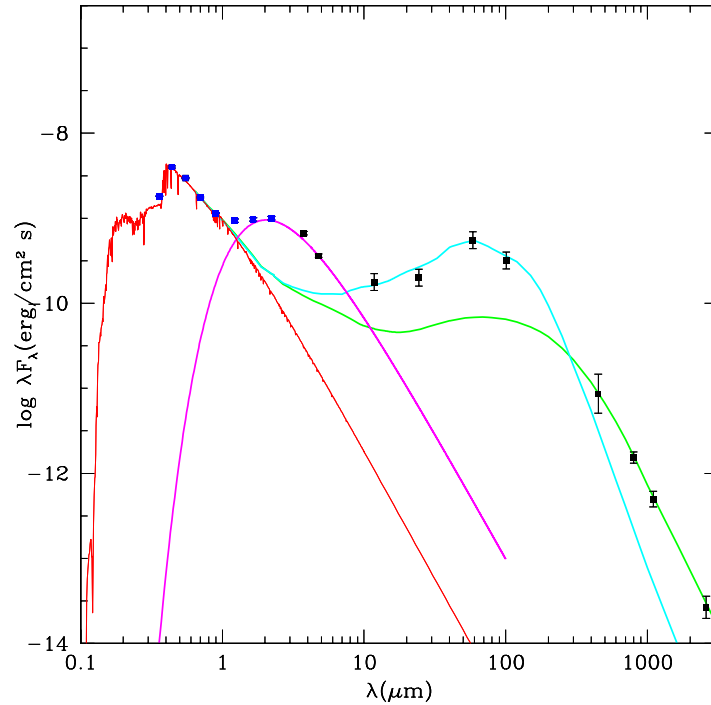


Figure 1. Spectral Energy Distribution of the pre-main sequence star HD 34282 with two disk models from D'Alessio et al. plus a black body to fit the near IR excess. The blue dots are EXPORT simultaneous photometry from *U* to *K* bands, black dots are photometry taken from the literature and IRAS and *ISO* archival data. The red line is the Kurucz model for the central star ($T_{\text{eff}} = 8720 \text{ K}$, $\log g = 4.2$). The purple line is a black body of 1400 K to fit to the $3 \mu\text{m}$ bump, the blue and green lines are the fluxes from two disk models calculated self consistently with the stellar parameters and with different size of the dust grains. The green line corresponds to a model with bigger grains ($a_{\text{max}} = 1 \text{ cm}$) than the blue line model ($a_{\text{max}} = 1 \mu\text{m}$). These two models were calculated with a disk radius $R_{\text{max}} = 800 \text{ AU}$, a mass accretion rate of 8.2×10^{-9} solar masses per year and an inclination angle of 60 degrees.

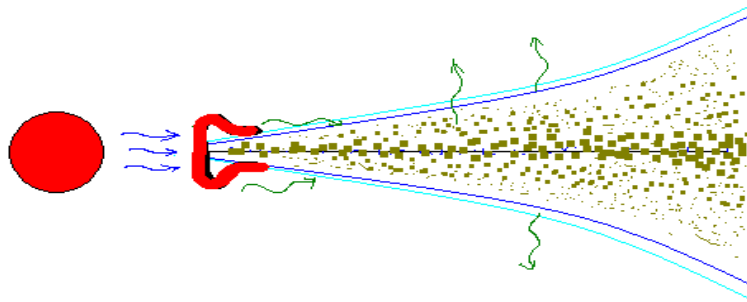



Figure 2. Sketch with the interpretation of the SED fitting. The black body emission at $3 \mu\text{m}$ comes from the inner edge of the disk, frontally illuminated by the star and not included in the disk models. The blue line would come from emission of thin dust particles in the surface of the disk, and the green line is the emission from the large particles in the mid-plane of the disk.

LABORATORIO DE ASTROFÍSICA ESPACIAL Y FÍSICA FUNDAMENTAL			
Evolution of circumstellar disks in Pre-Main Sequence stars			
Proposal for a VO Science Case			
GRUPO DE OPERACIONES Y ARCHIVOS	2004-04-07	Page 5 of 11	

1.2.- The on-line library of models

An on-line library of models covering the whole relevant physical parameter space for the central star and the disk is being built since December 2002 and will be available soon to the astronomical community. This library will be extremely useful to all scientists working in the field of star and planet formation and may become one of the best tools to analyse and interpret the multiwavelength emission from the protoplanetary disks around young stars.

The central stars considered for the library range from K7 to B9.5 with ages of one and ten million years. For each of them we have calculated models of protoplanetary disks with three different mass accretion rates, two values of the radius of the disk, two values for the viscosity of the disk, two kinds of dust size distributions and six different maximum grain sizes from $1\mu\text{m}$ to 10 cm. We have also calculated synthetic SEDs for two different inclination angles i . In this way we assure that all relevant physical parameter space is covered with the grid.


1.3.- Future applications

1.3.1.-New data

Despite the wealth of information provided by the *IRAS* and *ISO* satellites, the sensitivity of these missions did not allow to answer many key questions concerning the evolution of the disks and the possible formation of planets. For that purpose, future missions are being planned by both ESA and NASA to finally get the observational basis for a complete picture of the evolution of circumstellar disks around young stars and the possible mechanisms of planetary formation.

The first mission of this series is the Spitzer Space Telescope from NASA. It is already providing an unprecedented improvement in sensitivity allowing major advances in this area of research. To make use of Spitzer data, we have calculated synthetic magnitudes for the four IRAC (InfraRed Array Camera) photometers, operating at 3.6, 4.5, 5.8 and $8.0\mu\text{m}$, by convolving the model fluxes with the respective filter passbands. Thus we are able to predict the magnitudes and colours of the young stars harbouring planet-forming disks and interpret their physical state according to the IRAC colours. During its lifetime, Spitzer will observe many young stars with ages spanning a wide range of evolutionary stages, from objects still embedded in their parental clouds to stars with ages around 10 Myrs, the so-called Vega-type stars. This collection of observations will represent a big step towards an unbiased statistical study of the early evolution of stars and disks. The VO-tool described in this document will allow comparing between one of the best grids of circumstellar disks models with one of the best samples of observations of circumstellar disks in a very efficient way and will provide key results concerning the evolution of protoplanetary disks.

The possibility to apply the expertise gained with these studies to other missions (e.g. PACS (Photodetector Array Camera and Spectrometer) onboard Herschel and operating between 60 and $210\mu\text{m}$ or MIRI, the Mid InfraRed Instrument planned for the James Webb Space Telescope, operating between 5 and $28\mu\text{m}$) will allow to achieve a good theoretical framework for a better scientific exploitation of the new data.

LABORATORIO DE ASTROFÍSICA ESPACIAL Y FÍSICA FUNDAMENTAL			
Evolution of circumstellar disks in Pre-Main Sequence stars			
Proposal for a VO Science Case			
GRUPO DE OPERACIONES Y ARCHIVOS	2004-04-07	Page 6 of 11	

1.3.2.-New models

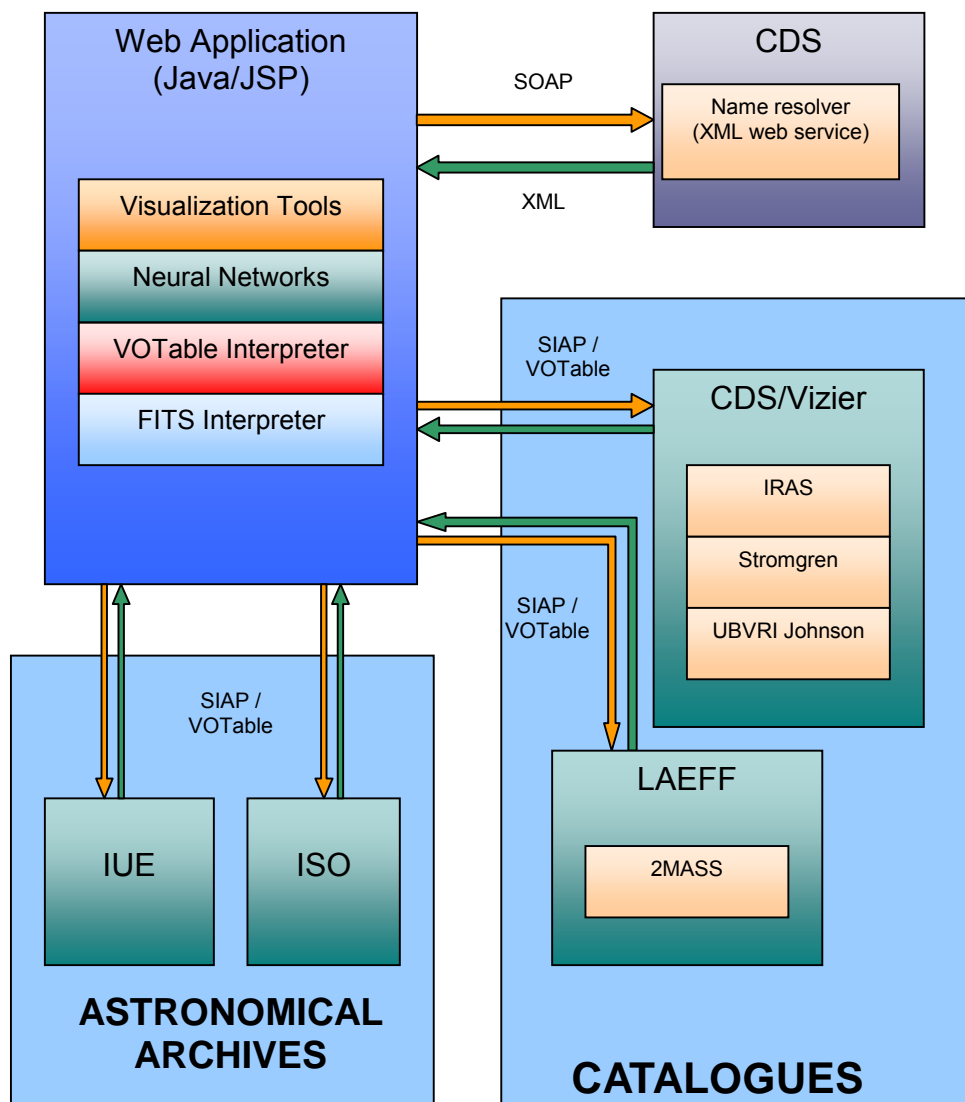
We plan to modify the numerical codes from D'Alessio et al. to simulate the conditions on the Vega-type stars and the so-called "debris" disks around main-sequence stars. The application of the same models to these new kind of disks would allow to answer very interesting questions about the transformation from the early gas-rich disks around the pre-main sequence stars to the tenuous gas-poor "debris" disks around main sequence Vega-type stars.

This adaptation will need important changes in the numerical code. First, the vertical height of the disk, which is calculated iteratively and self-consistently with the irradiation and diffuse fields, will have to be overridden and all dust will be set in the midplane in the new models. Also, since little or almost no mass accretion towards the central star is observed in these disks, the main heating agent will be the irradiation from the central star. A key question for the fitting of these new models to the observed SEDs will be the chemical composition and the abundances remaining in the dusty disk, since these will affect the opacity of the dust to the stellar irradiation and therefore the local temperature in the disk.


2.- The system: conceptual design

2.1.- System description

One of the main goals to be achieved is to maximize the accessibility to the system. The platform independence and the minimization of the software requirements are also desirable. That is why the proposed system is based in a web application using the HTTP protocol. In this way, the only requirements on the client side are an internet connection and a web browser. The proposed architecture of the system is represented here.



The user's interface is a HTML page where the user enters the object identification or coordinates and chooses the archives where the search has to be made. It is also possible to enter his/her own data. After this, the system will ask CDS to resolve the name of the object. This request is made via the SOAP protocol to one of the web services currently running at CDS. The service will return the object information (including its coordinates) in XML format. Having the coordinates of the object, a request is made to each of the data centres (INES,

LABORATORIO DE ASTROFÍSICA ESPACIAL Y FÍSICA FUNDAMENTAL			
Evolution of circumstellar disks in Pre-Main Sequence stars			
Proposal for a VO Science Case			
GRUPO DE OPERACIONES Y ARCHIVOS	2004-04-07	Page 8 of 11	

ISO, VIZIER,...) using the VO SIAP protocol. Each one of the data centres returns the data in VOTable format. The data will be interpreted by the web application using its VOTable interpreter module. Usually the VOTable output product only contains metadata and a link to the file containing the data itself (usually in FITS format). That is why a FITS interpreter module is necessary in the web application. Once the data retrieved by each of the data centres have been interpreted correctly, it can be combined, processed, or passed to the neural networks (see section 2.4) to perform the model fitting. After this, the user is able to visualize the result using the visualization tools implemented in the web application.

2.2.- Data sources

2.2.1.- Astronomical Archives

- INES (<http://sdc.laeff.esa.es/ines/>) Spectrophotometry in the range 1150 – 3350 Å
- ISO (<http://iso.vilspa.esa.es>) Photometric and spectroscopic information in the range 2.45 – 240 µm

2.2.2.- VizieR

- IRAS Point Sources Catalogue(<http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=II/125>) Spectrophotometry in four bands: 12, 25, 60 and 100 µm.
- Catalogue of uvby-β Data (Hauck B., Mermilliod M. 1998A&AS..129..431) This catalogue contains data for more than 63,300 stars in the Galaxy and Magellanic Clouds. Available at VizieR at <http://vizier.u-strasbg.fr/cgi-bin/VizieR?-source=II/215>
- “The Photoelectric Photometric Catalogue in the Johnson UBVRI system”. Lanzh T. (1986, A&AS, 65, 195)


2.2.3.- Locally available catalogues

- “2MASS All-Sky Catalog of Point Sources” (Cutri et al. 2003). JHK photometry for more than 470 million sources.
- “The Geneva-Copenhagen survey of the Solar neighbourhood” (Nordström et al. 2004 A&A, 418, 989). It contains Strömgren photometry for a sample of 16682 nearby F and G dwarf stars.

The Strömgren magnitudes will be converted into fluxes using the absolute calibration given in Gray (1998, AJ, 116, 482). 2MASS photometry will use the absolute flux calibration given in Cohen et al. (2003, AJ, 125, 2645) whereas the UBVRI magnitudes will be transformed into fluxes with the calibration given in Cohen et al. (2003, AJ, 126, 1090).

2.2.4.- User's data

The system will offer the possibility of including additional data that can either supersede the values available in the archives and catalogues or cover other regions of the electromagnetic spectrum.

LABORATORIO DE ASTROFÍSICA ESPACIAL Y FÍSICA FUNDAMENTAL			
Evolution of circumstellar disks in Pre-Main Sequence stars			
Proposal for a VO Science Case			
GRUPO DE OPERACIONES Y ARCHIVOS	2004-04-07	Page 9 of 11	

2.3.- Determination of the physical parameters

In order to fit the best model to the observations it is necessary for the model atmospheres (Kurucz, 1993, <http://cfaku5.cfa.harvard.edu>) to have a first estimation of the stellar physical parameters (effective temperature, surface gravity and metallicity) as well as the interstellar reddening. In all cases, the parameter estimation will follow a pre-defined hierarchy in which the values provided by the user are on the top. Significant differences between the input values and the values obtained after the iterative process may indicate either a non-adequate set of input values (due, for instance, to variable reddening and lack of simultaneity between the Strömgren and 2MASS photometry) or the presence of peculiarities in the spectral energy distribution that might deserve further and more detailed analysis.

2.3.1.- Interstellar reddening

The following hierarchy has been adopted:

- Value provided by the user.
- Strömgren calibration:
 - For A-type stars ($2.72 < \beta < 2.89$) the relation given in Crawford (1979, AJ 84, 1858) will be used.
 - For F0-G2 ($2.58 < \beta < 2.72$), Olsen (1988, A&A 189, 173) will be adopted. Valid for luminosity class III to V and all populations except the extreme population II stars. This calibration gives similar values to Crawford's (1975, AJ 80, 955) calibration but cover a wider parameter space. For spectral types later than G5, the beta index cannot be used to determine E(b-y).
- Hipparcos distances: If the distance to the star is less than 100 pc, the reddening is neglected.
- Otherwise, the reddening is set to 0.0.


The extinction curve of Rieke & Lebofsky (1985, ApJ 288, 618), from 4400 Å to longer wavelengths and that of Steenman & Thé (1991, Ap&SS 184, 9) for shorter wavelengths are used.

Future implementations

- **Determination of E(B-V) from the graphite interstellar absorption band around 2200 Å.** If observations at that wavelength are available (e.g. INES archive), the process consists of correcting the spectra for reddening with different values of E(B-V) until the absorption feature disappears (Mora et al. 2001, A&A 378, 116). This method is only applicable when there is a substantial amount of flux in the continuum around 2200 Å, i.e., only suitable for stars hotter than F0.

2.3.2.-Effective temperatures

- Value provided by the user
- Use of 2MASS photometry: This has proven to be a powerful method in the range 4000 – 8000 K. (Ribas et al. 2003, A&A 411, L501)
- Use of Strömgren photometry: The Moon & Dworetzky calibration (1985, MNRAS 217, 305) gives excellent results in the range 6000 – 20000K as reported by Napiwotzki et al. (1993, A&A 268, 653).

LABORATORIO DE ASTROFÍSICA ESPACIAL Y FÍSICA FUNDAMENTAL			
Evolution of circumstellar disks in Pre-Main Sequence stars			
Proposal for a VO Science Case			
GRUPO DE OPERACIONES Y ARCHIVOS	2004-04-07	Page 10 of 11	

- Use of Johnson photometry: The (B-V)₀ calibration given in Napiwotzki et al. (1993) will be used.

In the last three cases, the reddening value calculated in the previous section will be used.

2.3.3.-Surface gravities

- Value provided by the user.
- Use of Strömngren photometry: Moon & Dworetzky (1985) will be used with the correction suggested by Napiwotzki et al. (1993) for stars with Teff > 9000K.

2.3.4.-Metallicities

- Value provided by the user.
- Use of Strömngren photometry and the following calibrations:
 - Smalley (1993, A&A 274, 391) for A2 – F0 stars.
 - Haywood (2002, MNRAS 337, 15) for F – G stars
 - Olsen (1988, A&AS, 57, 443) for G – K stars.

2.4.- Data mining tools


We plan to implement in the system a series of tools based on well established Artificial Intelligence developments in the field of Decision Theory and pattern recognition. In the following we consider the problem of, given a set of points in a Spectral Energy Distribution (SED), how to obtain the physical parameters of both the star and the disk. The main aim is to help astronomers to quantitatively analyse the data in terms of evidence favouring one or several sets of parameters, what other alternative models can compete with the most a posteriori probable one, and what are the most discriminant observations to discard alternatives. These functionalities will be linked with the capacity of the server to query other archives in search of values for the discriminant observables.

The main part of the system will be based in model parameter estimation using Bayesian methods. Given a set of priors (prior probability densities) for the different parameters (which are inclination angle, disk radius, mass accretion rate, viscosity, dust size distribution and maximum size for the disk and effective temperature, surface gravity and metallicity for the star) we may compute the posterior probabilities for the data obtaining in this way the most likely set of parameters both for the star and the disk.

Let $P(\{\theta\})$ be the prior probability distribution for the set of parameters describing a model from the grid ($\{\theta\}=\{\theta_1, \theta_2, \dots, \theta_7\}$); let $f(\{\lambda\})$ be the observed SED and let $P(\{f(\lambda)\} | \{\theta\})$ be the sampling model of the observations which can be derived from instrumental characteristics at each wavelength band $\{f(\lambda)\}$ (i.e. the likelihood or evidence of the observations given the set of parameters); finally, let $P(\theta | \{f(\lambda)\})$ be the sought posterior probability of the parameters given the observations. Then, the posterior probability can be computed, according to Bayes theorem, as

$$P(\{\theta\} | \{f(\lambda)\}) \propto P(\{f(\lambda)\} | \{\theta\}) P(\{\theta\})$$

where the proportionality constant is the inverse of the probability density for the observations

LABORATORIO DE ASTROFÍSICA ESPACIAL Y FÍSICA FUNDAMENTAL			
Evolution of circumstellar disks in Pre-Main Sequence stars			
Proposal for a VO Science Case			
GRUPO DE OPERACIONES Y ARCHIVOS	2004-04-07	Page 11 of 11	

$$P(\{f(\lambda)\}) = \int P(\{f(\lambda)\} | \{\theta\}) P(\{\theta\}) d\{\theta\}$$

The posterior probabilities can be marginalized in order to provide probability distributions for individual parameters:

$$P(\theta_i | \{f(\lambda)\}) = \int \dots \int P(\{\theta\} | \{f(\lambda)\}) d\theta_1 d\theta_2 \dots d\theta_{i-1} d\theta_{i+1} \dots d\theta_7$$

and given two competing sets of parameters with similar probabilities, the system can provide wavelength ranges where observations are most discriminant and propose the astronomer queries to archives able to provide the SED completion. This will be achieved by predicting the values of most discriminant observables sampling from posterior probabilities.

Finally, one of the most important inference tools to be implemented in the demo will be a bayesian assistant to quantitatively assess the best model degree to sufficiently explain the data in the philosophy of Occam's razor in the sense that a simpler theory with a more compact parameter space will have a larger evidence that a higher order (more complicated) model unless the latter performs significantly better at explaining the data. Let M_1 represent the set of model with a single dust component and M_2 the set of all possible models with two dust components (Figure 1, for instance). Then, using the formalism sketched above, the evidence supporting M_i can be estimated by integrating over the parameter space

$$P(\{f(\lambda) | M_i\}) = \int \dots \int P(\{f(\lambda)\} | \{\theta\}, M_i) P(\{\theta\} | M_i) d\{\theta\} = \int \dots \int P(\{\theta\} | \{f(\lambda)\}, M_i) d\{\theta\}$$

where the latter integral can be computed by Markov-Chain Monte Carlo Methods.

In the future, when new models for debris type disks become available the same formalism can be used to help astronomer decide the evolutionary nature of the disk best supported by the data.

A completely different tool based on a connectionist approach is already implemented as a Self-Organized Map (SOM) that provides a 2D representation of the parameter space $\{\theta\}$. At present, the SOM can be queried with a given observed SED and the neural network module responds with the coordinates of the most active unit, the set of theoretical models grouped around its codebook vector and the topological properties of the surrounding neighbourhood.

3.- Conclusions

A VO-science case based on the study of protoplanetary disks using spectral energy distributions is presented here. In short, it is based on the fitting between a library of synthetic spectral energy distributions computed using the physical stellar parameters, the Kurucz model atmospheres and the physical disk models by D'Alessio et al, and the observations of spectral energy distributions of PMS and MS stars with disks made by the EXPORT consortium. The comparison will be performed making use of data mining tools. The result of this analysis will be an excellent initial approach for a further, more detailed modelling, based on refined initial parameters, on improvements in any part of the analysis or on any other physical approach not included in the computation of the library of SEDs.