Spectral survey analysis: the WEEDS package

P. Hily-Blant & S. Maret

Institute for Panetary science and Astrophysics of Grenoble (IPAG)
University Joseph Fourier

Collaborators: J. Pety, S. Bardeau, E. Reynier (IRAM)

October, 13th, 2011

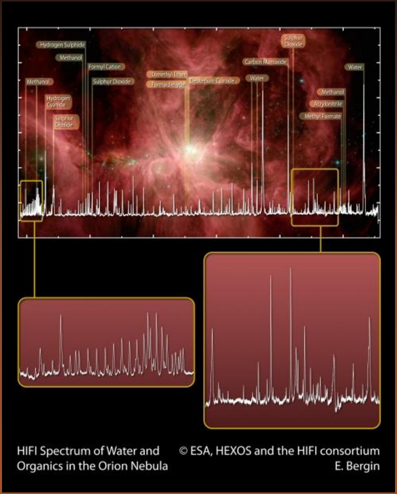
1. Introduction
2. Large Datasets
3. WEEDS
4. Issues

Introduction

Spectral survey: continuous scan in frequency over a certain range (*e.g.* an atmospheric window for ground-based telescopes)

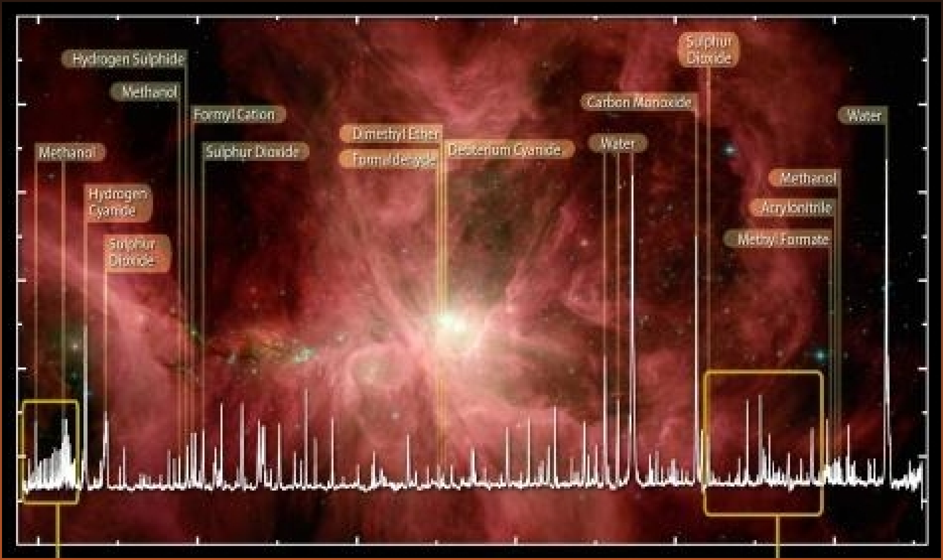
Unbiased spectral survey: a spectral survey with homogeneous sensitivity accross the full frequency range

Orion with Herschel/HIFI

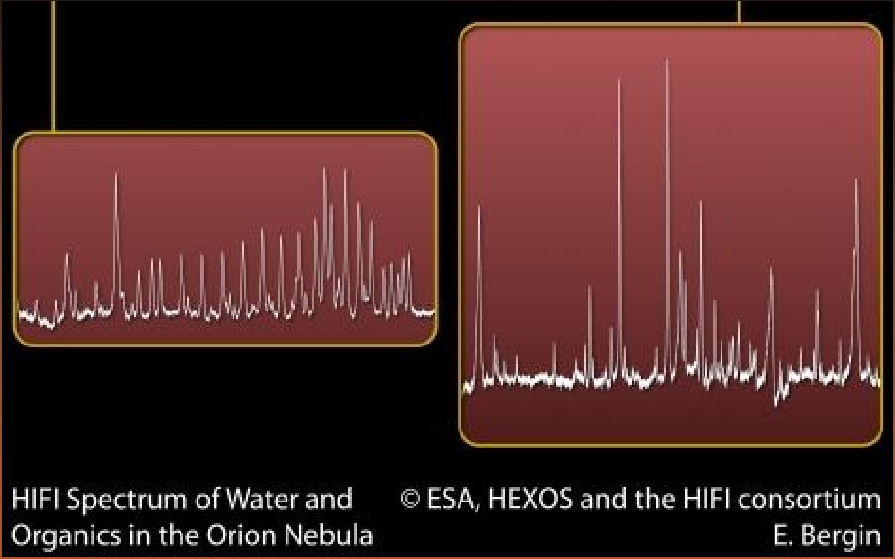


HEXOS key program (Bergin et al)

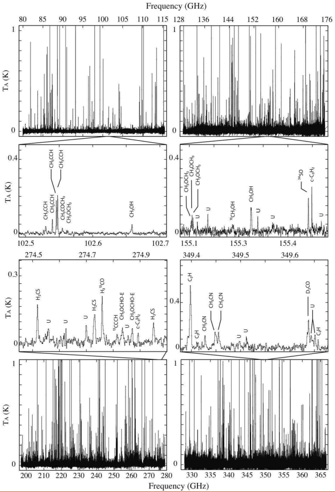
Orion with Herschel/HIFI



Orion with Herschel/HIFI

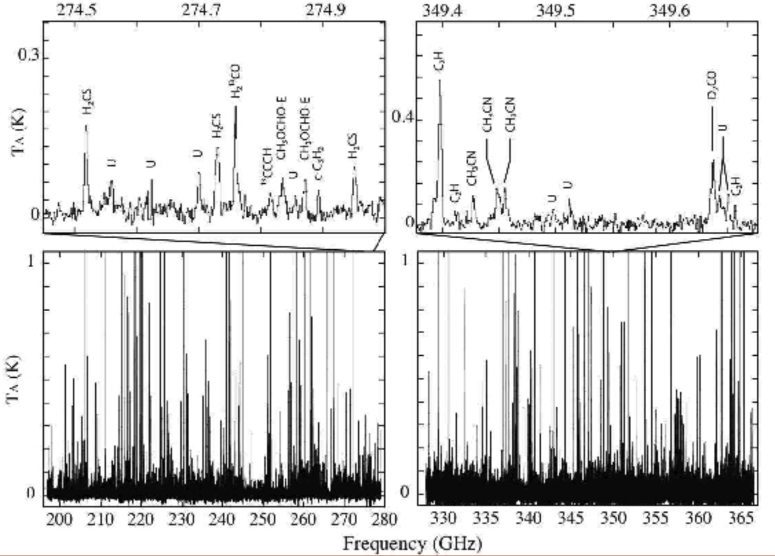


Iras16293-2422 with IRAM-30m

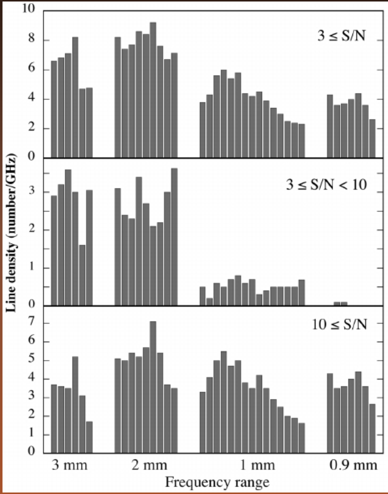


Caux et al 2010

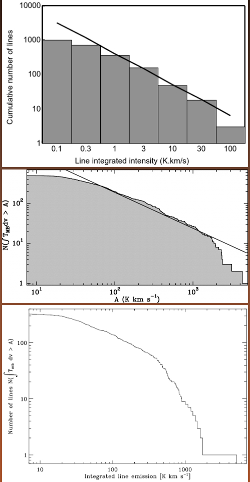
Iras16293-2422 with IRAM-30m



Line density

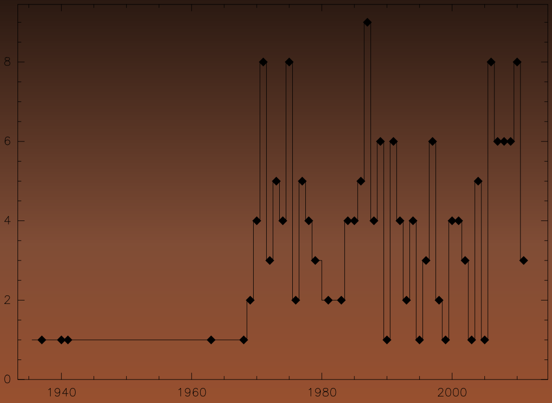


Caux et al 2011



Caux et al 2011, Comito et al 2005, White et al 2003

Line identification



- 174 molecules detected in the Interstellar Medium
- $\approx 10 - 15\%$ of U-lines in (ground-based) spectral surveys
- Spectral surveys from Herschel/HIFI analysis under way...

Present situation

- Large instantaneous bandwidths of receivers
- Concomitant increase of the backend capabilities (FFTS, correlators)

⇒ Virtually any spectrum is (what was considered) a spectral survey (20hr to cover 80-115 GHz with few mK/(km/s))

Telescope	Band (GHz)	Bandwidth (GHz)
GBT	1–100	3.2
APEX	230–1000	4
IRAM	80–360	32

What does “large bandwidth” means and implies ?

- Bolometers: $\Delta\nu/\nu_0 \approx 0.2 - 0.5$
- Coherent receivers (e.g. 2SB): $\Delta\nu/\nu_0 \approx 0.1 - 0.3$
- Consequence: Resolution power $R = \nu_0/\delta\nu \approx 10^6$, $\delta\nu \approx 100$ kHz, hence
  #(channels) $= \Delta\nu/\delta\nu \approx 0.1 - 0.3 \times 10^6 \approx 10^5$
  $\Rightarrow$Need Tools to explore Large Spectra

Wishes

- Need frequent queries to spectral line catalogs (e.g. JPL, CDMS, Splatalogue)
- Need to "navigate" in a spectrum of several GHz
- Need modelling tools to identify lines

Data reduction: basic strategy

- Bandpass effects: 0th order baseline

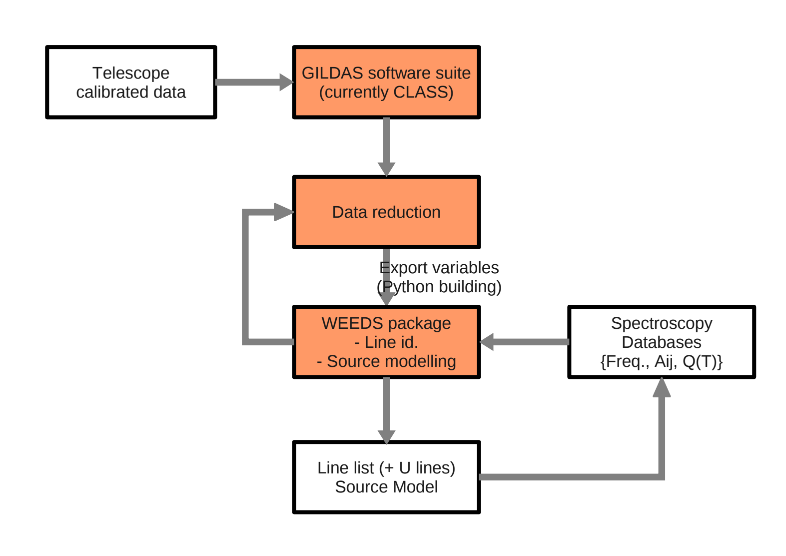
⇒ Problematic because not always are there free-of-signal channels

Data analysis: basic strategy

1. Identify *usual* species including isotopologues
2. Fit a model of the emission of these species to the full range spectrum
3. Eye-checking best fit
4. Subtract to the spectrum

The WEEDS package

- A CLASS extension to analyze spectral surveys written by Sébastien Maret and Pierre Hily-Blant (IPAG) with support from the IRAM scientific software team (J. Pety, S. Bardeau, E. Reyiner)
- Publically available as part of GILDAS (Linux, Mac, Windows)
- S. Maret, P. Hily-Blant, J. Pety, S. Bardeau, E. Reyiner al. A&A 2011
- Named after the so-called "weeds" by spectroscopists – "rogue" species with hundreds of ro-vibrational transitions that one needs to identify before picking up the "flowers".
- Maintenance: as part of CLASS
- There is a manual
- Python code, uses the GILDAS Python bindings


Telescope
calibrated data
GILDAS software suite
(currently CLASS)
Data reduction
Export variables
(Python building)
WEEDS package
- Line id.
- Source modelling
Spectroscopy
Databases
{Freq., Aij, Q(T)}
Line list (+ U lines)
Source Model

Catalog queries

- Database are accessed on-line using a VO-compliant protocol (SLAP)
- SLAP isn't widely adopted yet $\Rightarrow$WEEDS can also access database using their own specific protocol
- Can access JPL, CDMS and Splatalogue (thanks to Brian Kent and Tony Remijan for their help!)
- Can also make a copy (cache) of the database on one's computer (to work "offline")

Line Identification

- Strong lines:
  - likely to be a *usual* species
  - likely to have $E_u \sim kT$
  - and/or large A_{ul}
- Weaker lines:
- Strong case for line identification: identify several lines of a given species

$\Rightarrow$ Need *filters*

  - species
  - sub-catalogs (*e.g.* Splatalogue, CDMS)
  - A_{ul}, E_u

```
file in toto.30m   ! open the data file
find               ! read the file
get f              ! load the first observation
dbselect jpl       ! Select a database
lid                ! Interactive search in the current band
lid /i             ! Interactive search in the image band
lid /s co /f       ! search for CO accross the full band
```

Line id: methodology

- Make a model for a given species
- Search for all predicted lines in the survey
- Ensure that all lines are emitted from the same region (follow-up interferometric observations)

Line modelling

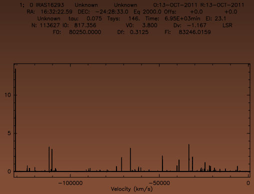
Input parameters *for each species*

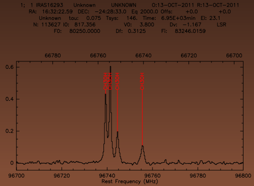
- source size, telescope diameter
- excitation temperature
- column density, linewidth
- centre line velocity wrt systemic source velocity
- continuum background (default is CMB @ 2.73 K)
- emission / absorption

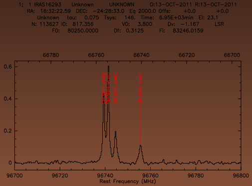
Spectroscopic inputs

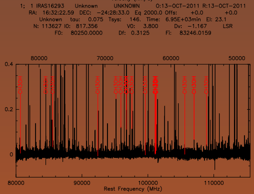
- Rest frequencies
- Einstein coefficient
- Partition functions ($\mathcal{Q}(T_{\rm ex})$)

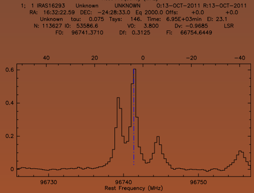
Demo

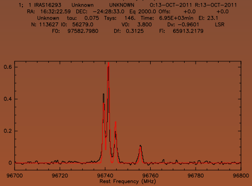

Velocity (km/s)


Rest Frequency (MHz)


Rest Frequency (MHz)


Rest Frequency (MHz)


Rest Frequency (MHz)


Rest Frequency (MHz)

Issues

- varying spatial resolution accross the spectrum
- inhomogeneous thermal noise accross the spectrum
- Current surveys: mostly with single dish telescopes (IRAM, CSO, HIFI...) Large spectra (up to 1.5 THz) but on single pixel. Analyse takes a lot of time, but still doable
- Future surveys: large datacubes (thousands of pixels per direction):
  - OTF map on single dish telescopes with wide band receivers (e.g. EMIR)
  - Interferometers (ALMA: 8 GHz, NOEMA)

Spectral images

- Dealing with datacubes of several GHz bandwidth or more is a HUGE challenge: how do we analyse this?
- Thousands of lines $\times$ millions of pixels:
  - requires some automatic fitting routines
  - continuum subtraction (2D): use spatial (and/or time) correlations, spectral correlations
  - Large # of free parameters (N, $T_{\rm ex}$, source size, FWHM, #(species), non-unique solution
- Probably requires some high level tools (with GUI) built on top of the low-level utilities provided by WEEDS
- Database are ESSENTIAL: they need to be maintained on a regular basis (addition of new species). They should provide partition functions to allow for LTE modelling.

Conclusions

- Modelling: LTE first, non-LTE afterwards
- Speed up modelling: takes long to model one species over $\sim$ THz
- Consider 2D fitting: one species over large ν range and full map
- Consider help from “amateurs”
- Build sub-catalogs/templates (*e.g.* splatalogue, CDMS): “cold gas”, “hot core”
- We have to change our minds: ISM through new glasses