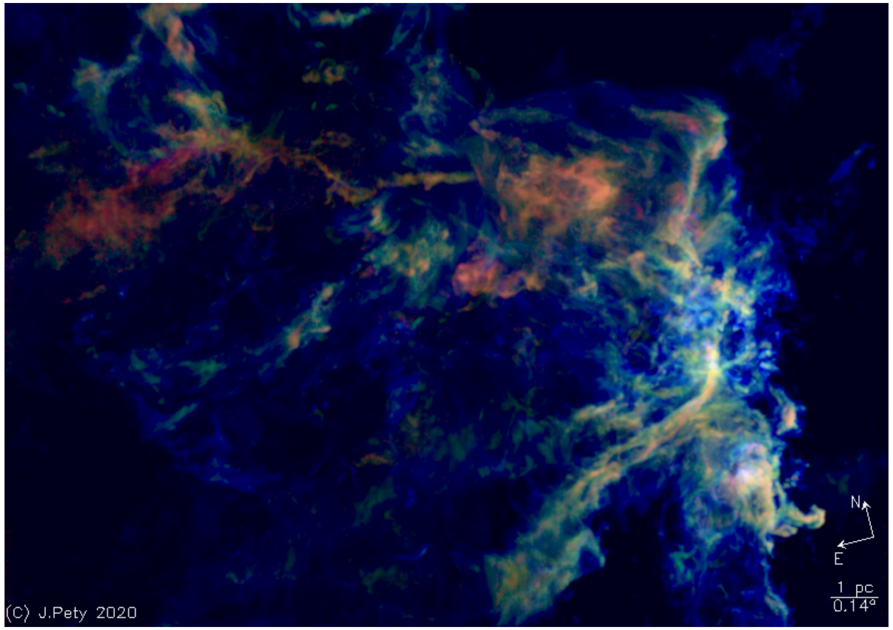
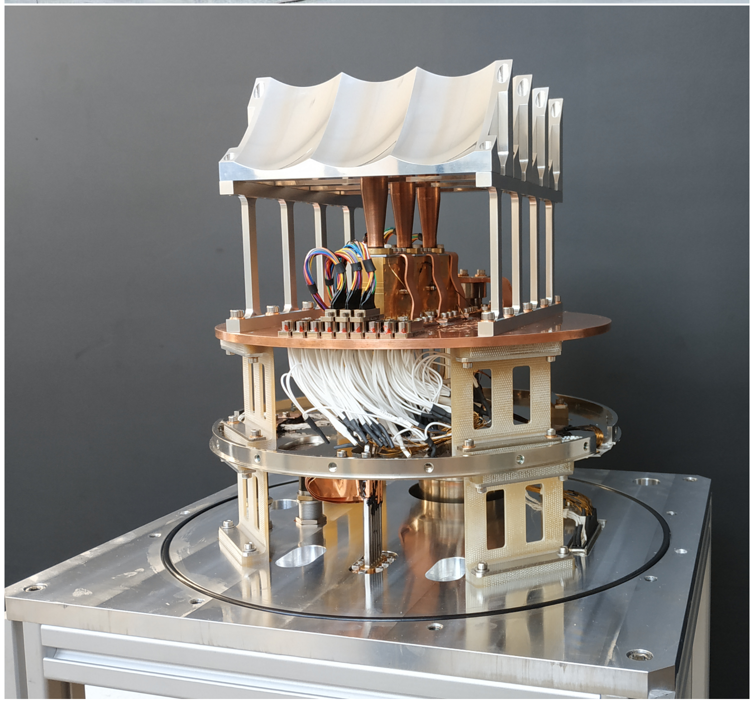
Heterodyne Data Processing

Jérôme Pety, Sébastien Bardeau, Emmanuel Reynier, Albrecht Sievers, Jean Tedros, Antonio Vazquéz, Kateryna Fal, Mathieu Houdeau





May 12th-14th 2025 - Paris

ALHAMBRA multi-beam project

Two bands

- 5×5 beams at 3 mm.
- 7×7 beams at 1 mm.

Derotation foreseen.

Mixer characteristics

- 2 polars x 2 sidebands x 8 GHz.
- Frequency range: $72 - 116.5$ GHz and $200 - 276$ GHz.
- Receiver temperatures: Better than 50 K at 3 mm and 83 K at 1 mm.
- At least 13 dB rejection.

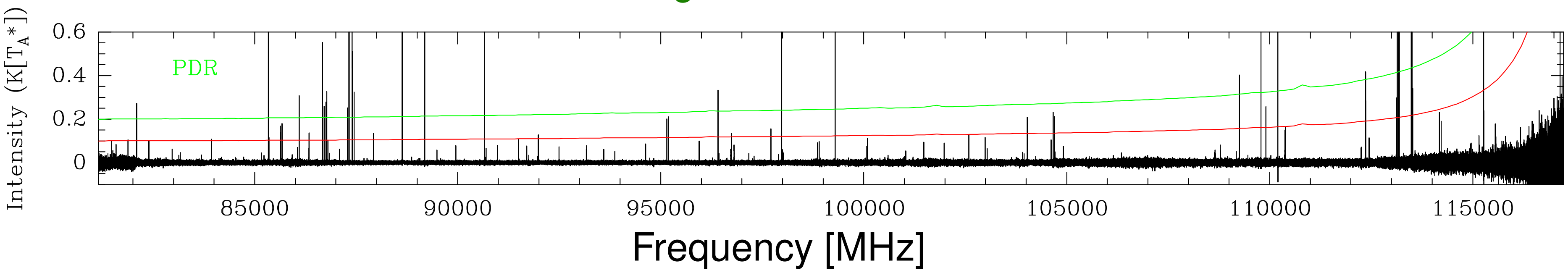
Spectrometer characteristics

- Full IF band at 64 kHz $\Rightarrow$ 13 and 25 Mchannels at 3 mm and 1 mm.
- Dump rate: 10 Hz.

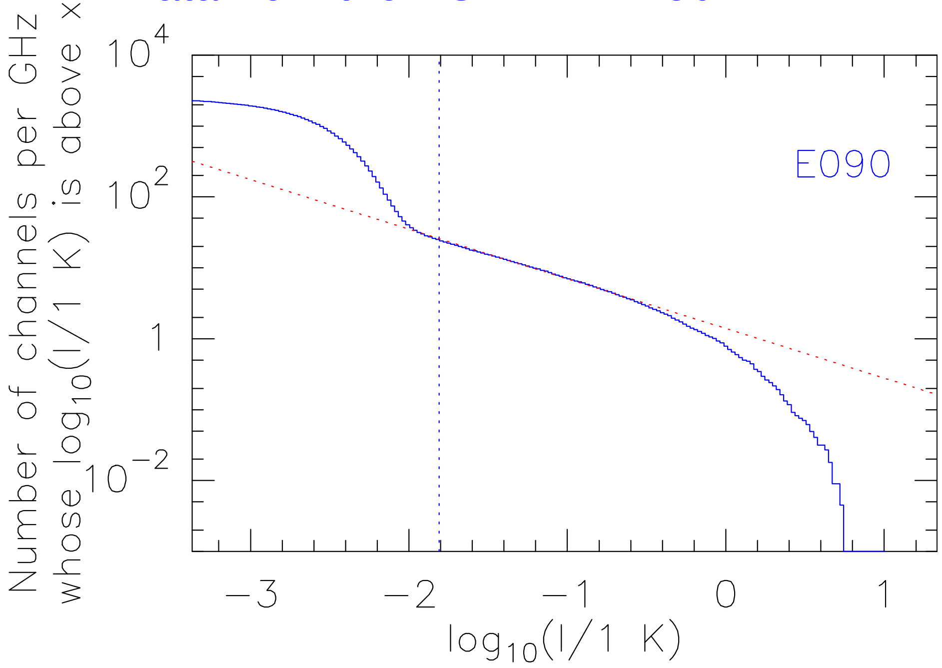
Timescale First light for at least one band in 2029.

Galactic spectra are full of faint lines!

Horsehead WHISPER line survey:
red curve $\Rightarrow 20\sigma$ and green curve $\Rightarrow 40\sigma$ with $\sigma \sim 5$ mK.



Data from the ASAI IRAM-30m LP



Percentage of channels above

1000 mK	0.03%
0100 mK	0.14%
0010 mK	0.68%
0001 mK	3.4%

$\Rightarrow$ More complicated baselining!

Data RateS

Sustainable peak data rate 700 MiB/s for A090 and 1400 MiB/S for A230 (including reduced data products).

Sustainable average data rate 4.5 PiB/yr for either A090 or A230 (including reduced data products).

No free lunch $\Rightarrow$ costs

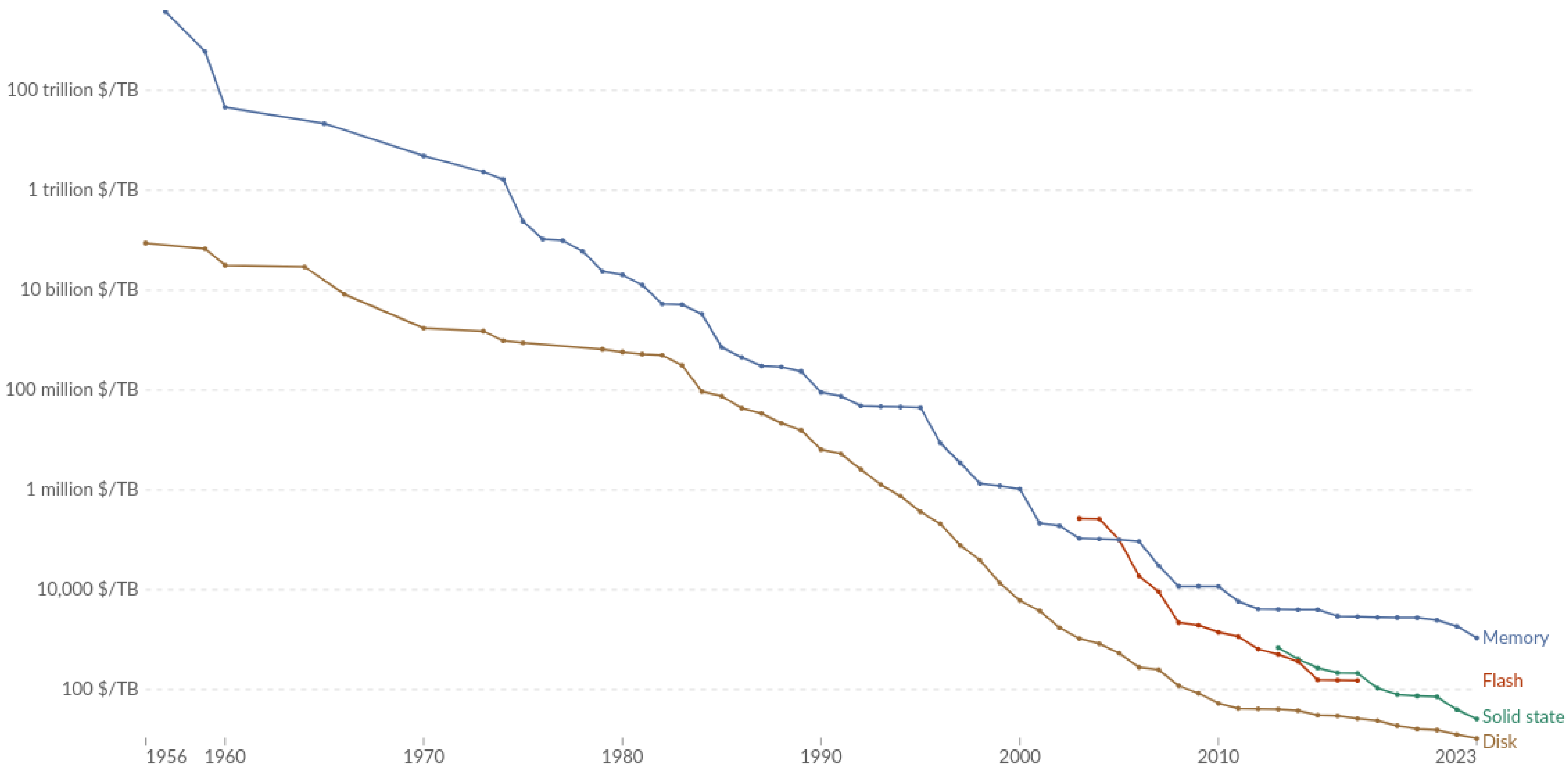
Hardware investment $\simeq 1.5$ Meuros for two copies for 5 years of operation.

Running costs Non-negligible running costs.

No free lunch ⇒ costs

Hardware investment $\simeq 1.5$ Meuros for two copies for 5 years of operation.

Running costs Non-negligible running costs.



Source: US bureau of Labor Statistics (2024)

Goals

End-to-end pipeline(s)

- Position-Position-Frequency cubes for the full bandwidth.
- Or at least around a few key lines.

Delivered products through DMS

- Diagnostic plots and reports.
- Raw data (IMBFITS or 30m files).
- Signal and noise PPV cubes. Mask of detected signal and associated moment maps.
- VO-compatible FAIR system.

Simplifying observation setups for simple and complicated projects.

Processing steps

Set of pipelines Automatically triggered when new data appear to avoid human intervention.

Minimum of hardcoded heuristics The data format have to be as much self-descriptive as possible.

Automatic diagnostic plots to observers and PIs Web pages and/or DMS.

Automatic filling of the calibration database Pointing. Focus. Receiver and system temperatures.

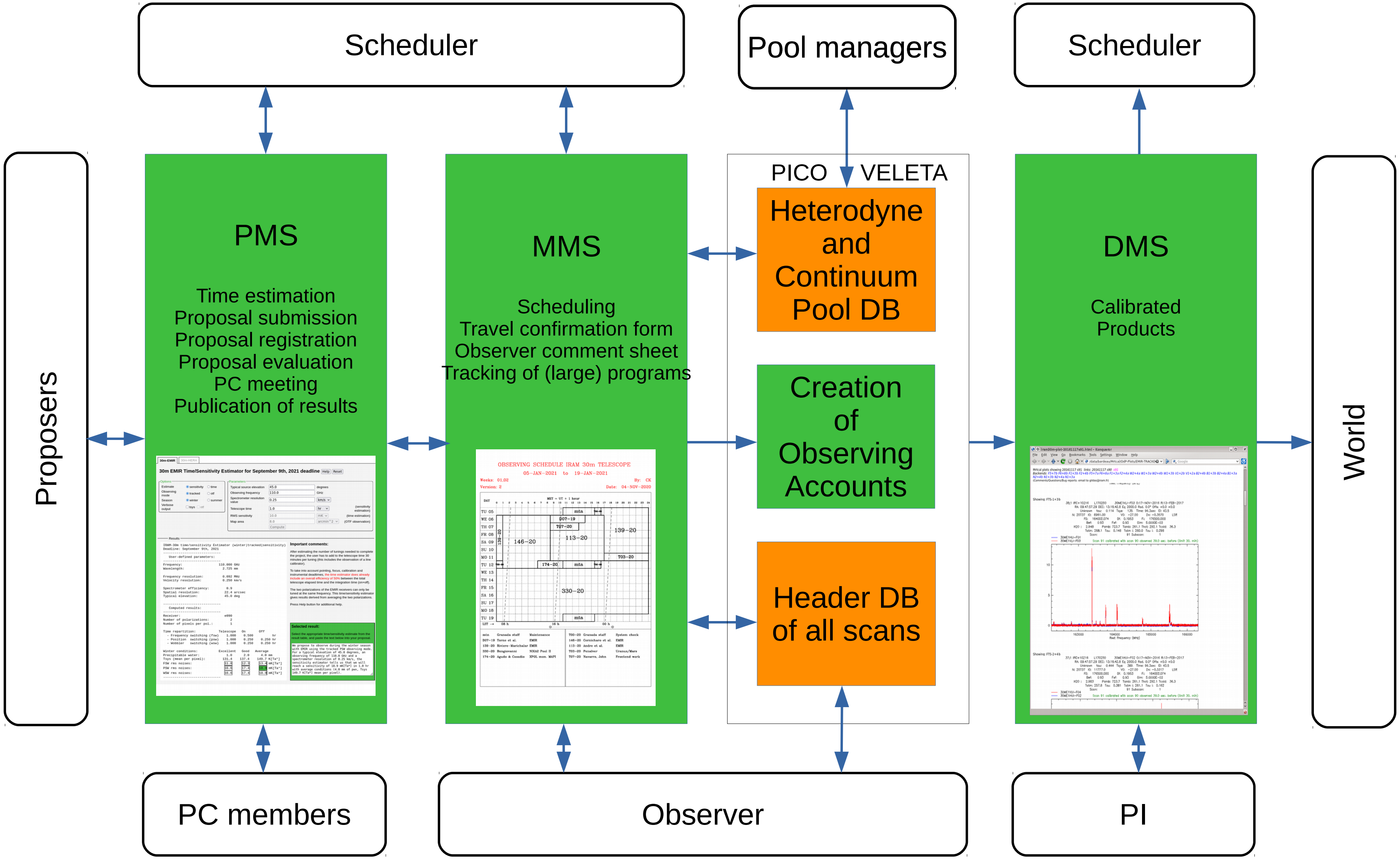
Automatic reports to IRAM staff once a day

Processing Step	Data formatting	Calibration	Editing & Processing per spectrum	Imaging & Simulation
Program	make-IMBFITS	MRTCAL	CLASS	CUBE
Operations	• Merging of antenna and spectrometer streams	• Determination of the beam Nasmyth offsets • Skydips • Pointing • Focus • Gain calibration • ON-OFF	• Data selection • Simple baselining • Filtering of RFI and receiver instabilities • Resampling	• Gridding • Advanced baselining • Deconvolution • Simulation • Characterization of data: noise images and spectra, detection of signal computation of moments

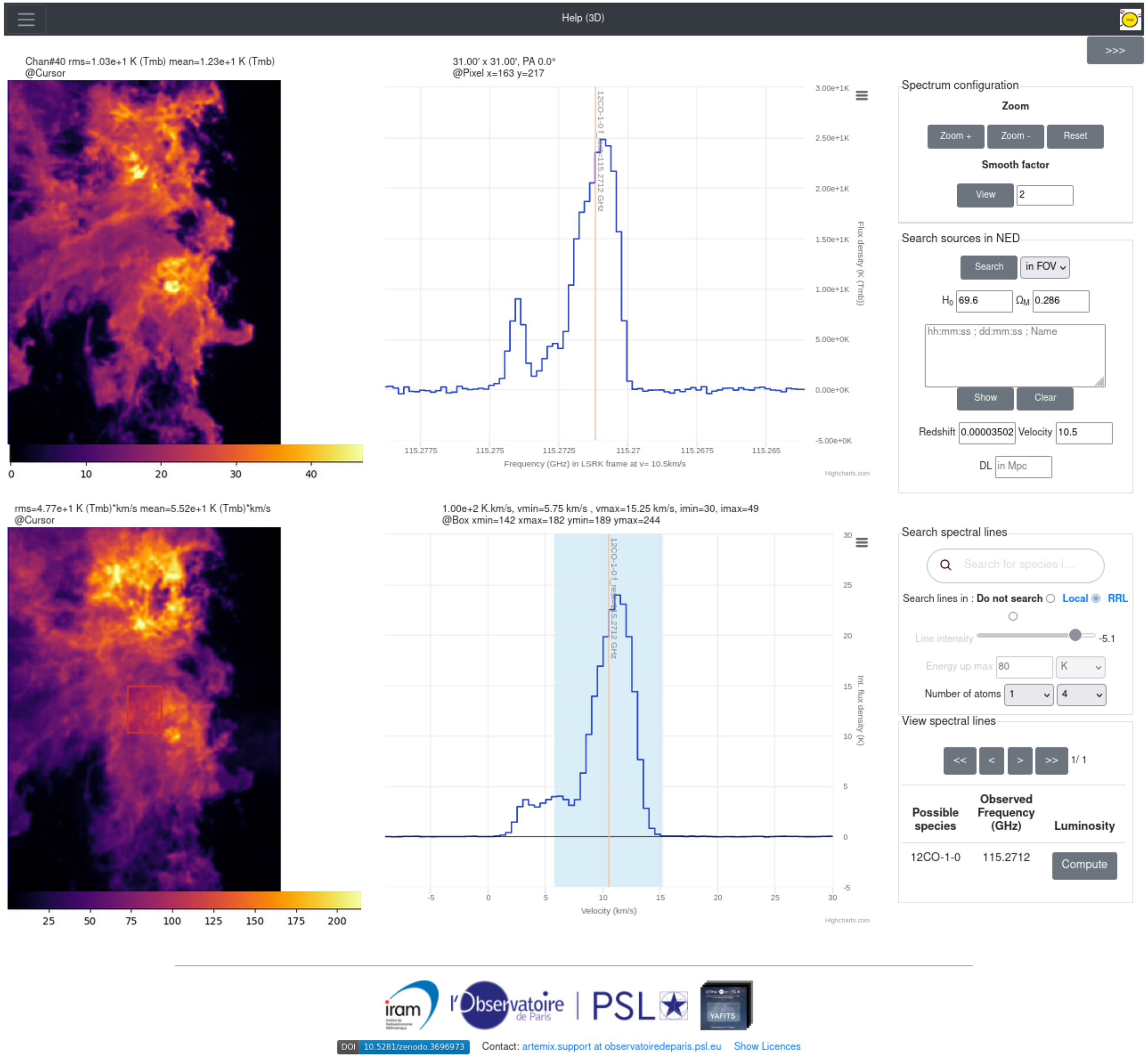
Observation Management System

https://oms.iram.fr

Goal for the 30m

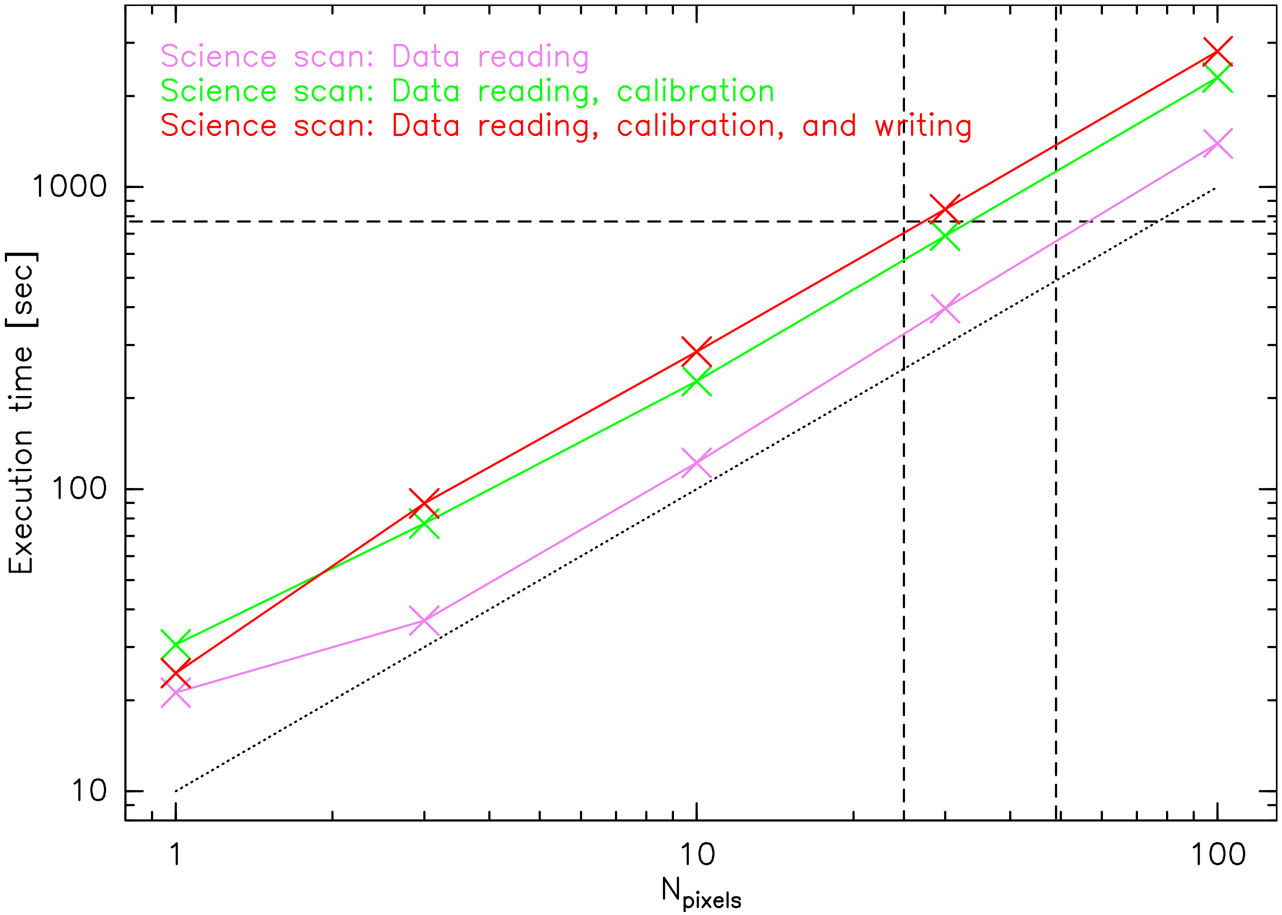


Interactive pre-visualization ⇒ YAFITS from Obs. de Paris (N.Moreau & P.Salomé)



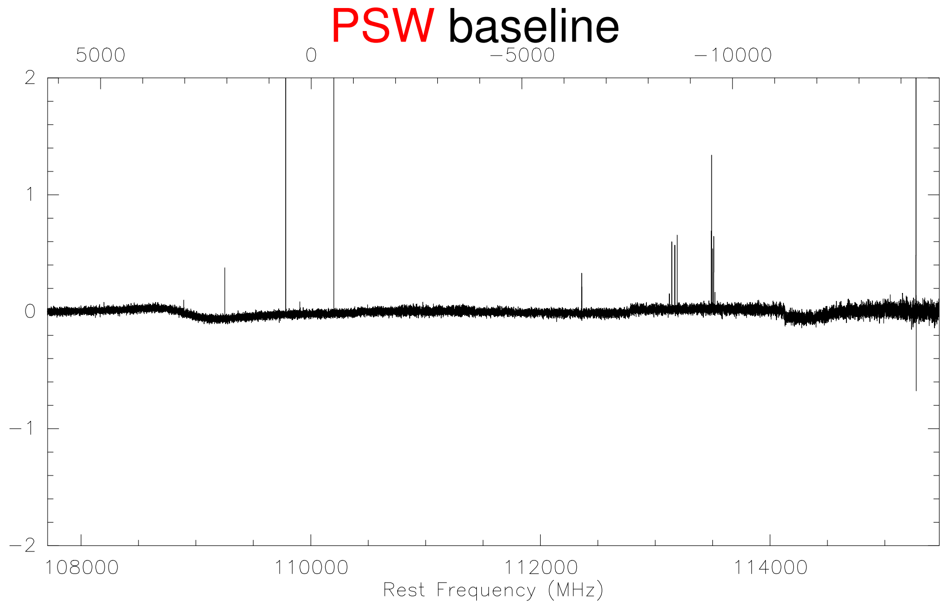
Linear CPU scaling as a function of the number of pixels

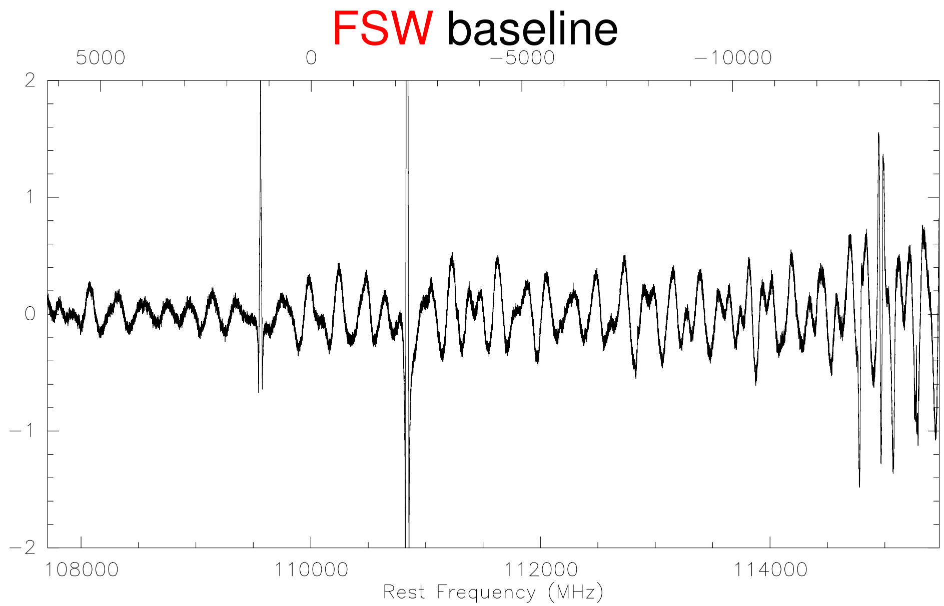
Processing one 13-min.-OTF scan on a 2010 PC (profiling made in 2017).



Calibration algorithms are READ/WRITE limited.

Observing strategy: I. PSW vs FSW baselines



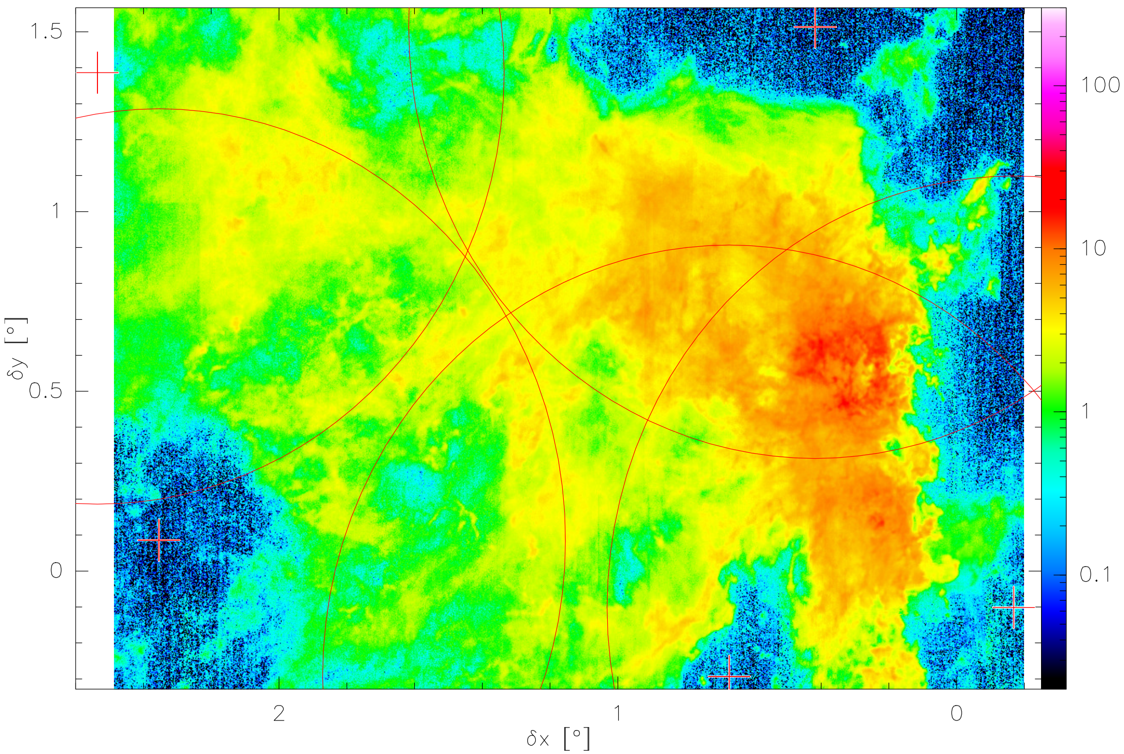


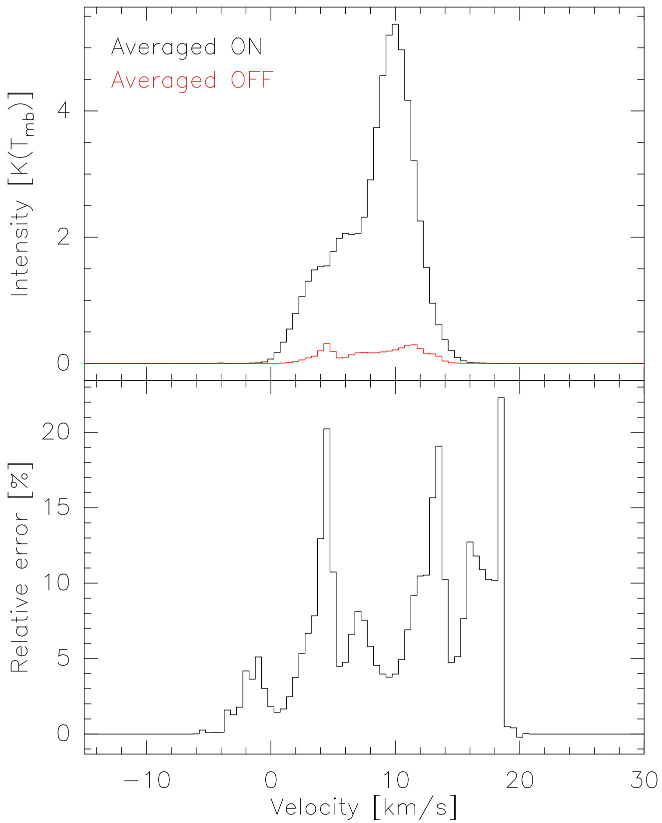
Observing strategy
II. Finding the best reference positions

Not too far for best baselines.

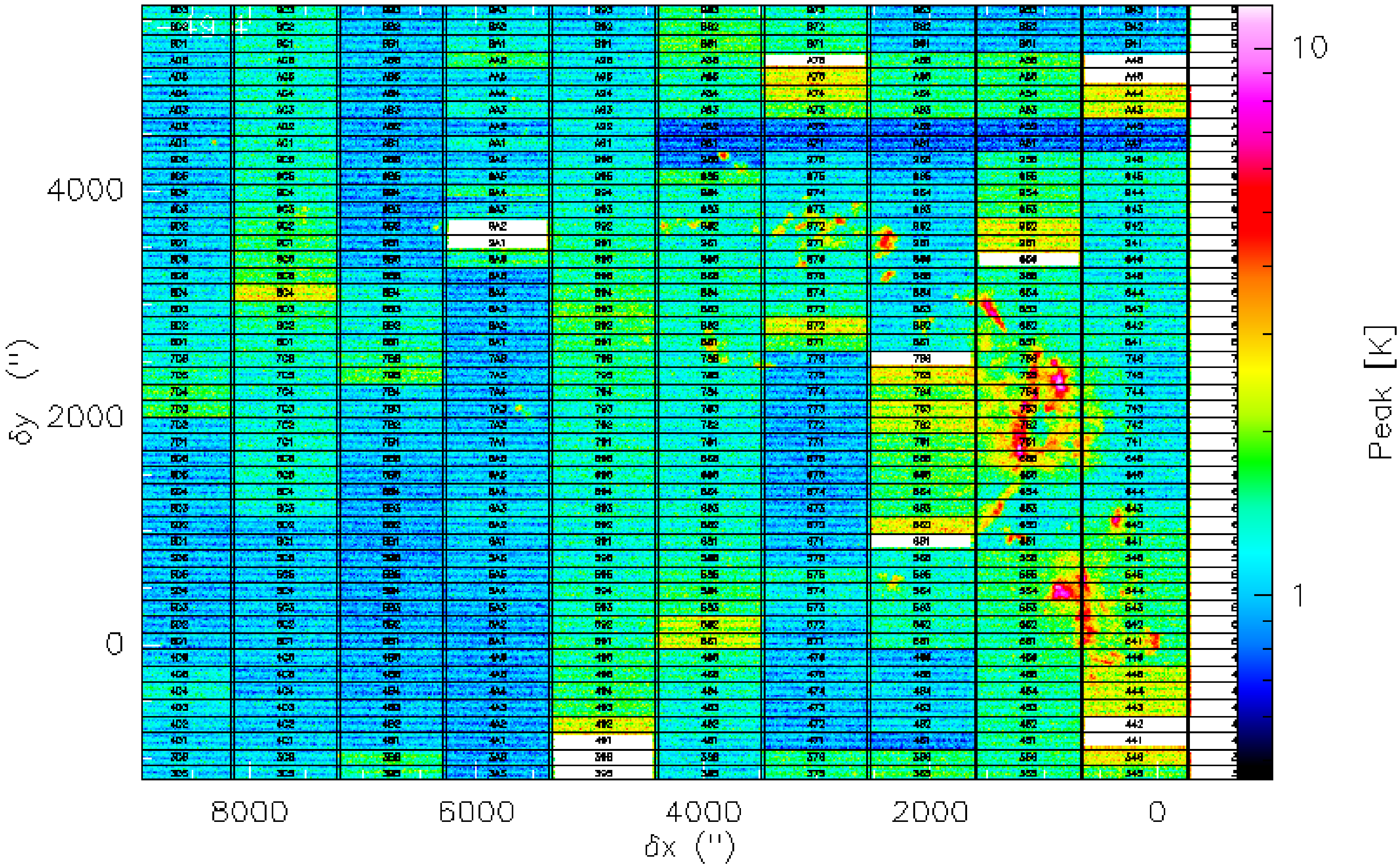
The most devoid of signal to avoid contaminating the result.

⇒ Dedicated observations of the reference position to check and correct if needed.





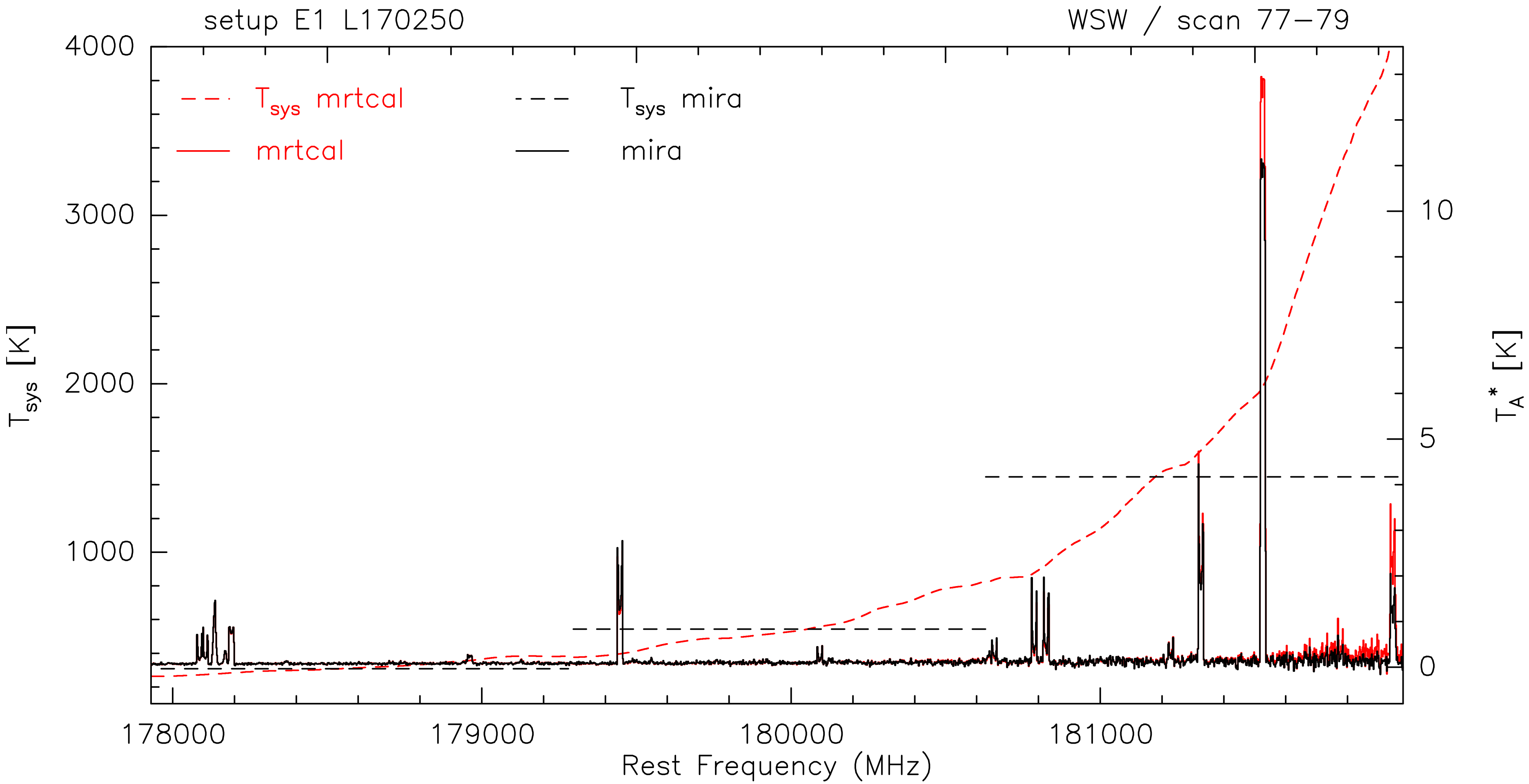
Observing strategy: III. Scanning



Wide Bandwidth (Marka et al. 2017)
⇒ Many parameters depend on frequency

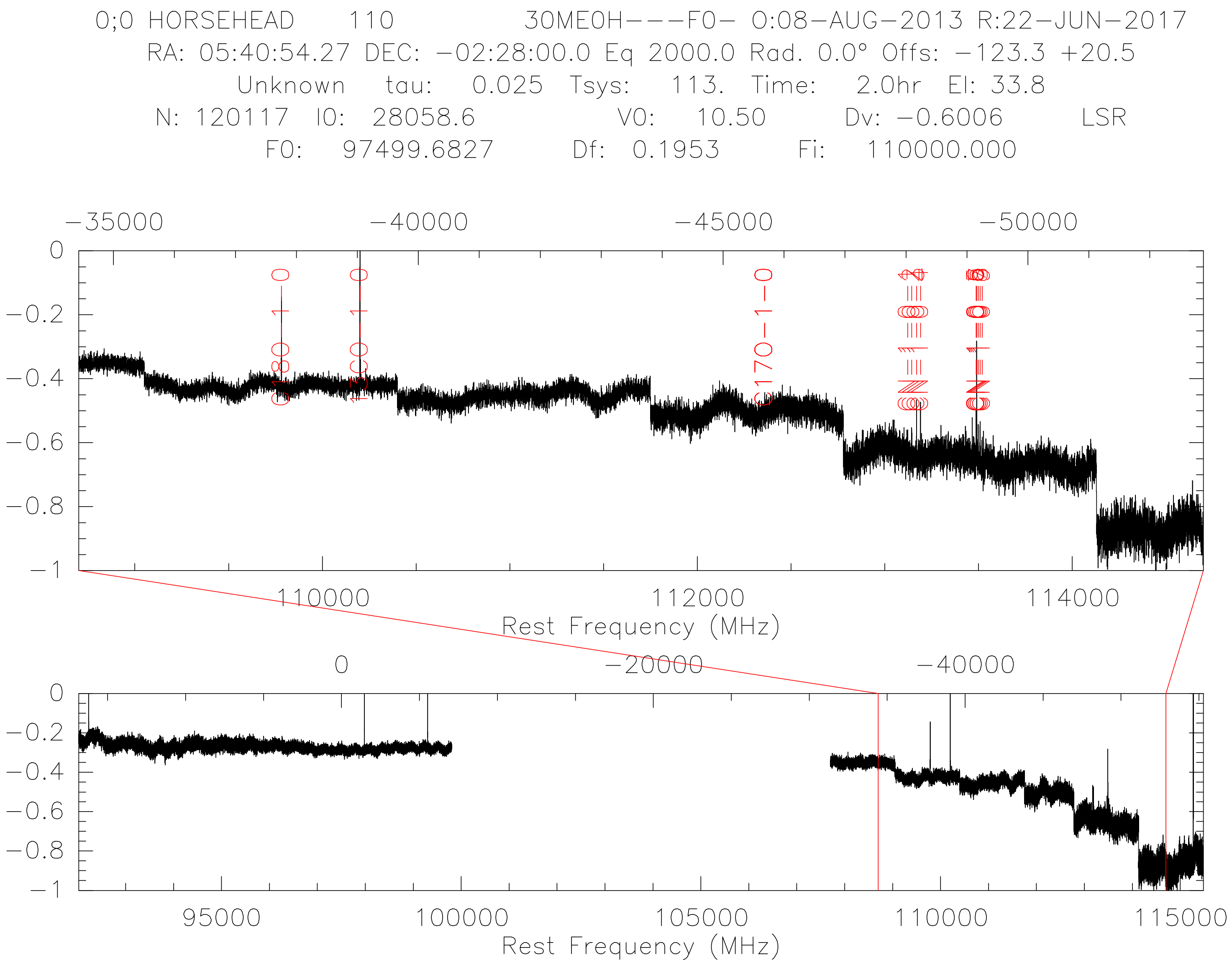
MIRA Calibration per baseband (1.35 GHz)

MRTCAL Tunable calibration bandwidth (default: 20 MHz) ⇒ Better traces change of atmospheric absorption close to water line at 183.3 GHz.



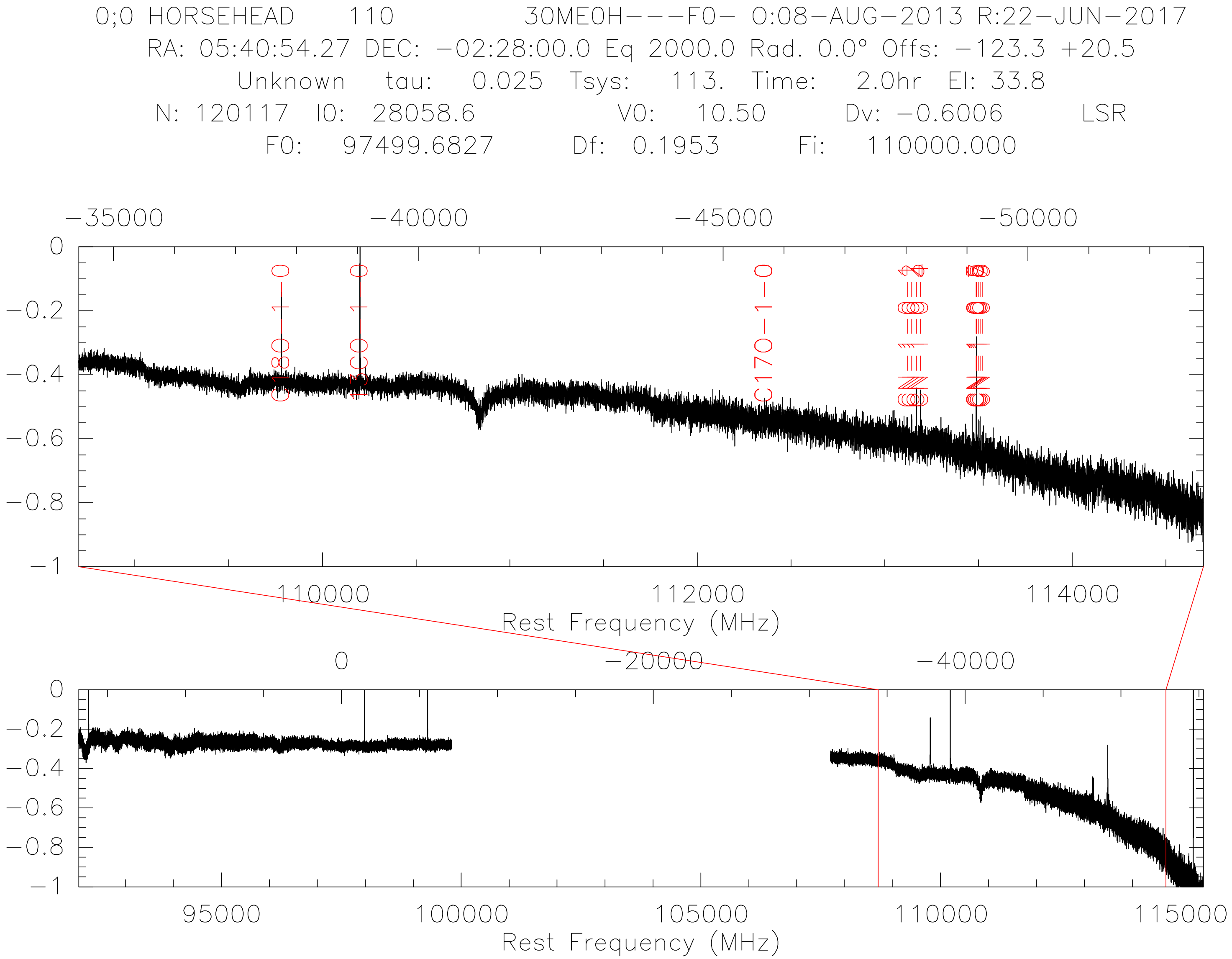
Wide Bandwidth (Marka et al. 2017)
⇒ Many parameters depend on frequency

MIRA solution One atmospheric calibration per 1.35 GHz, no interpolation.



Wide Bandwidth
⇒ Many parameters depend on frequency

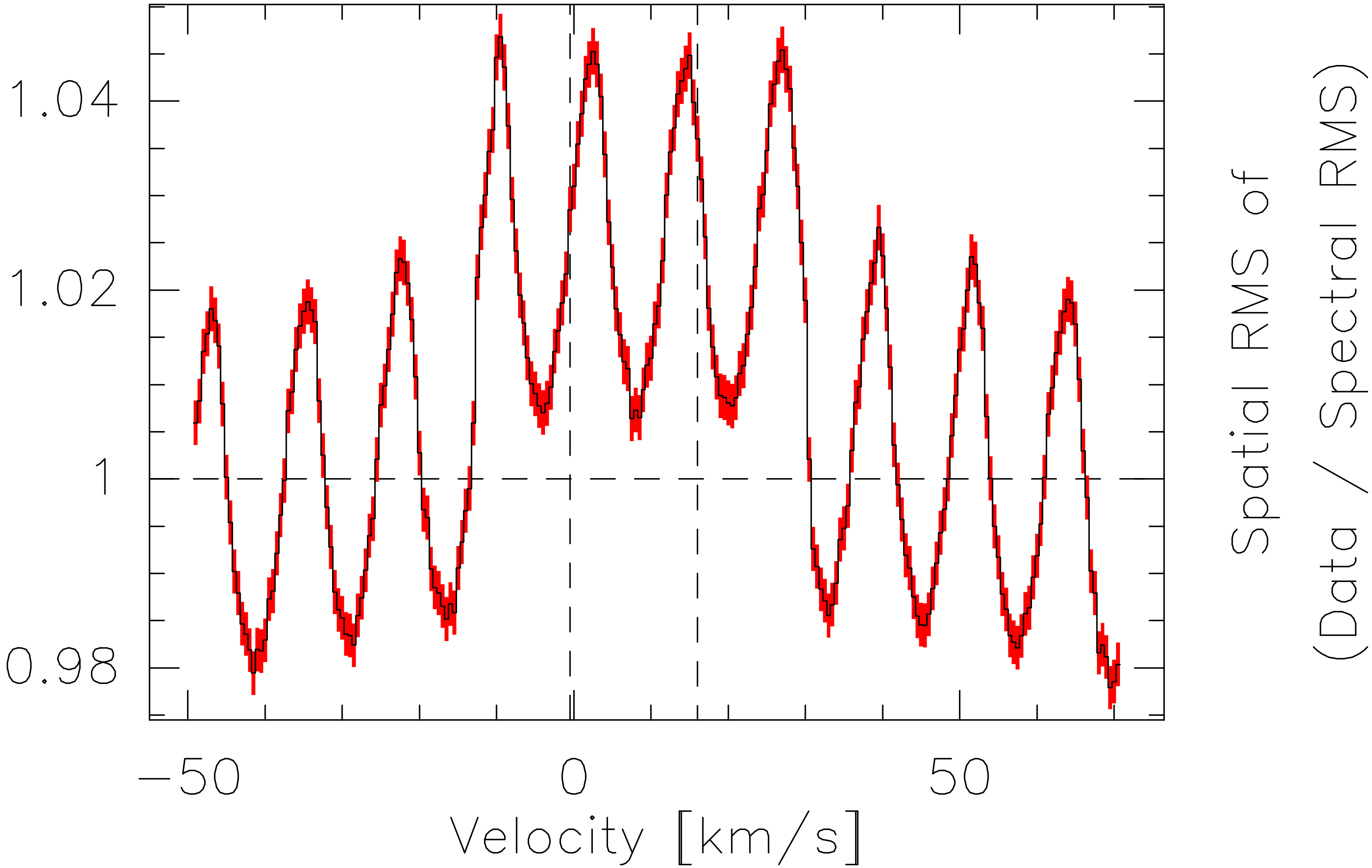
MRTCAL solution One atmospheric calibration per 20 MHz, linear + linear interpolation.



Wide Bandwidth (Einig et al. 2023)

⇒ Many parameters depend on frequency

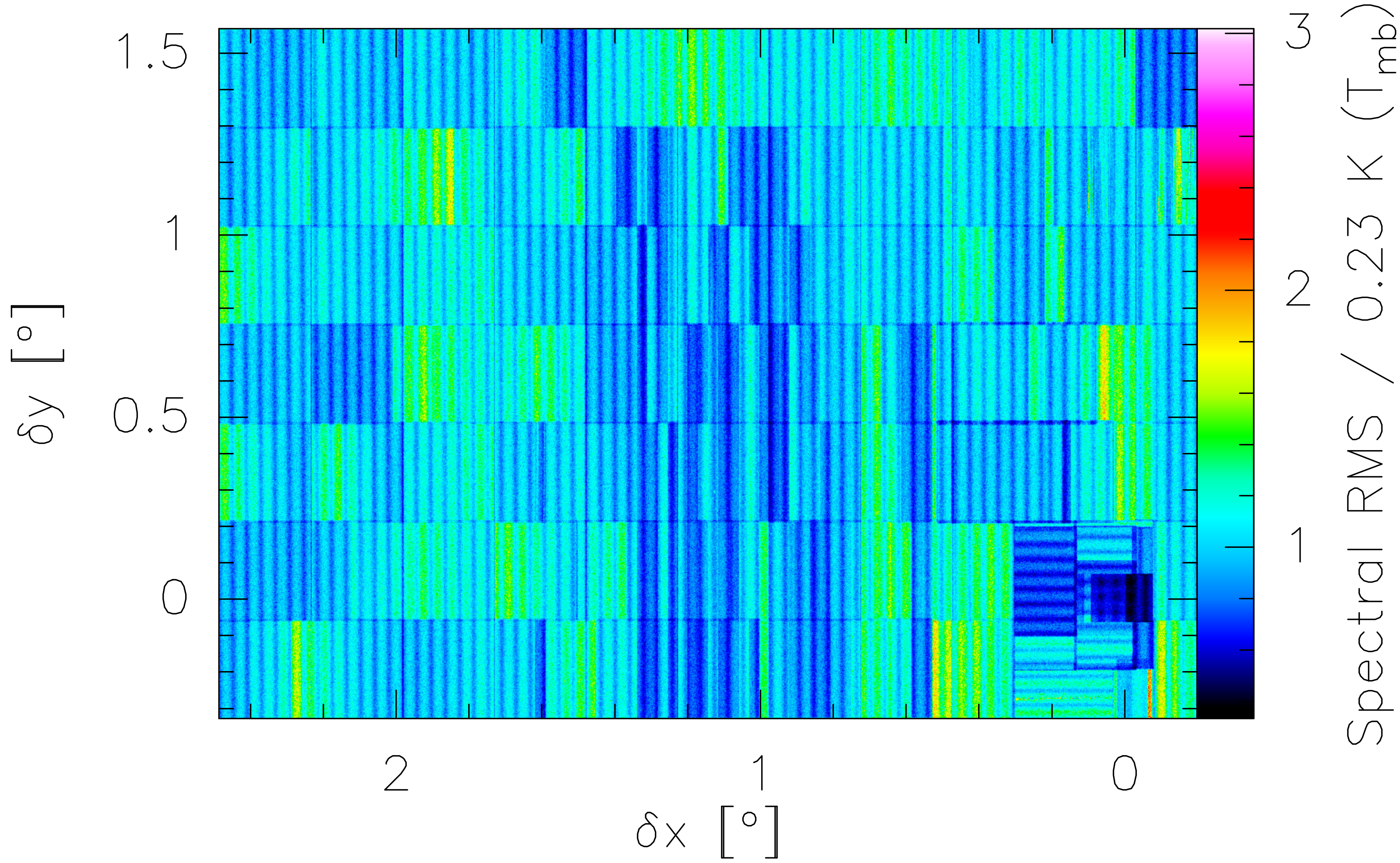
Effect of 1) baselining, and 2) spectral resampling



Wide Field

⇒ Many parameters depend on sky position

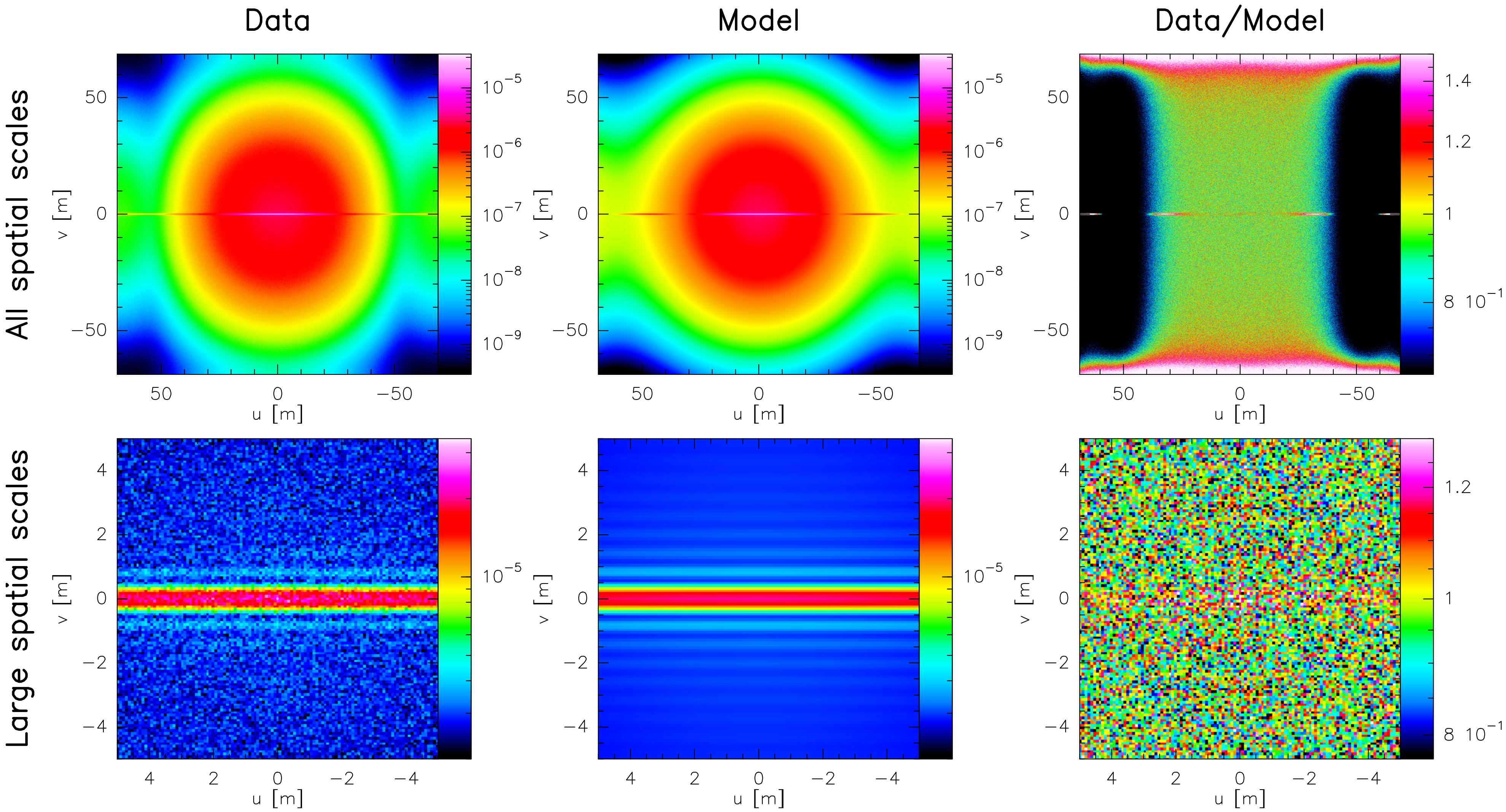
Noise level From weather variations and scanning strategy.



Wide Field (Einig et al. 2023)

$\Rightarrow$ Many parameters depend on sky position

A more subtle effect The noise is correlated along the scanning direction due to the switching strategy because the same reference position (OFF) is shared among many ON dumps.



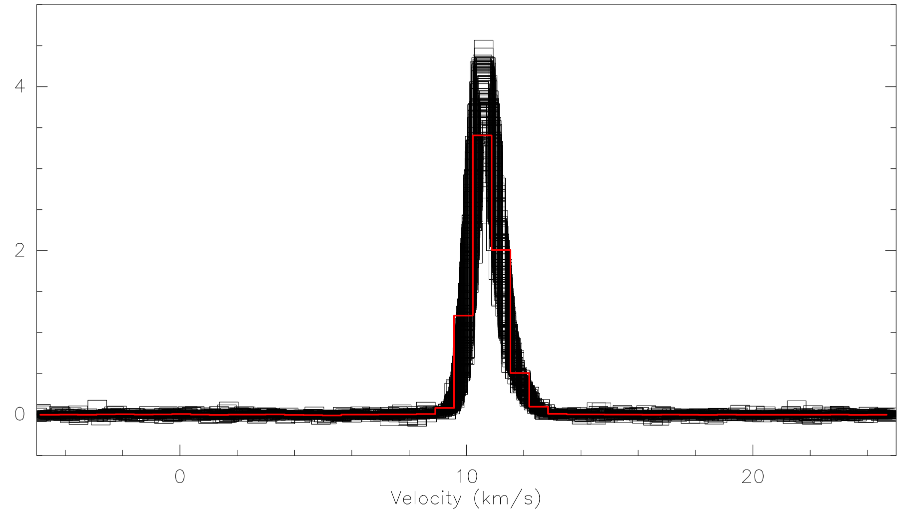
On the importance of the homogeneity of pixels

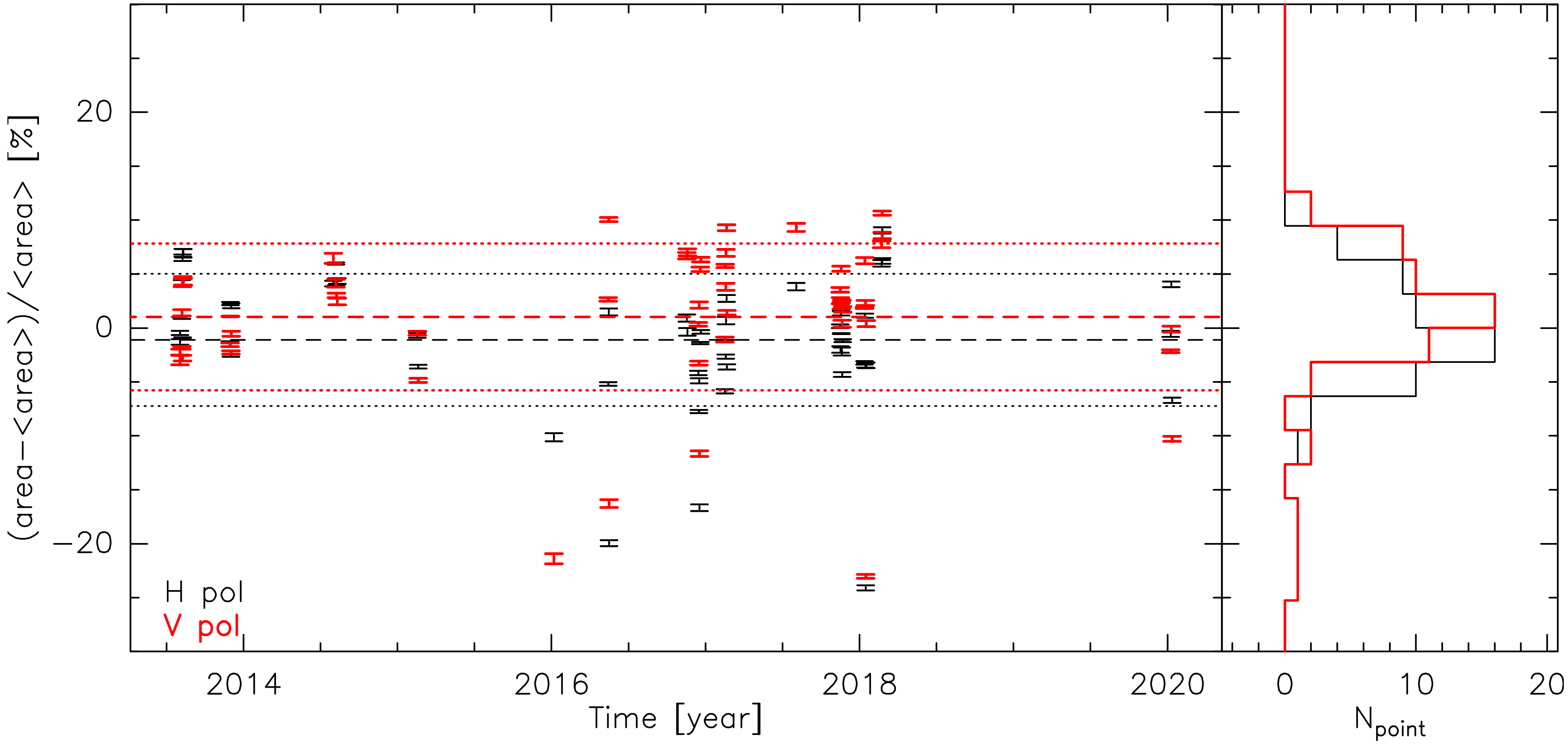
Same bright position Horsehead core #1.

Same observing strategy 6 minutes at beginning of each 8hrs observing session of the ORION-B project.

Same data reduction and analysis Baselining + Gaussian Fit.

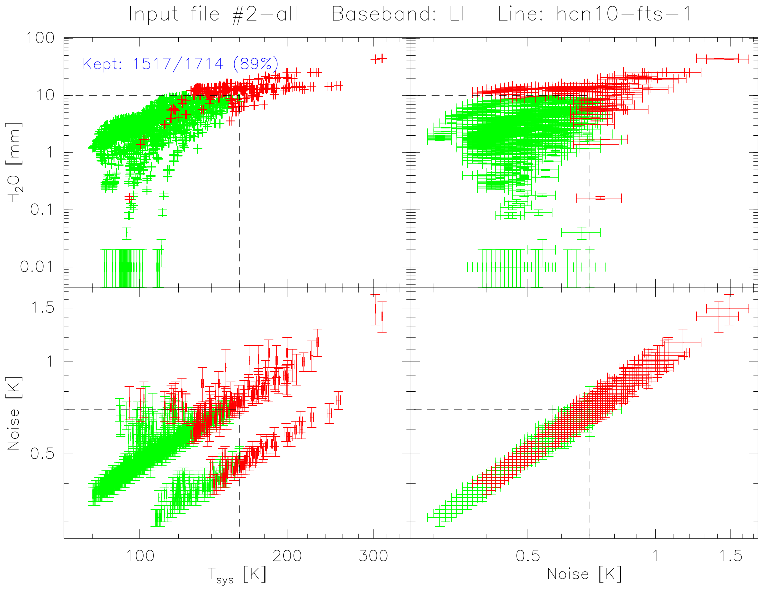
⇒ Systematic calibration offset between H and V polar.

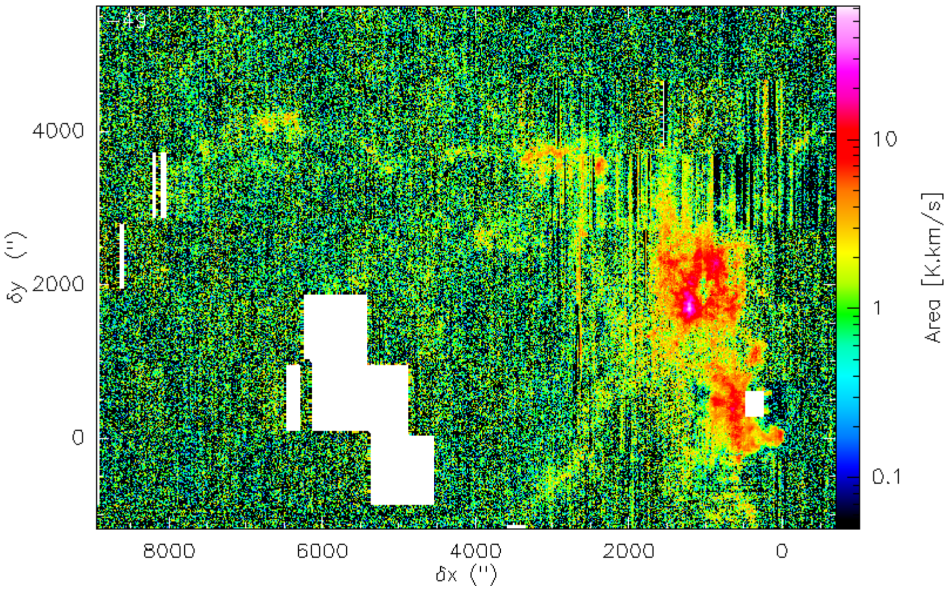




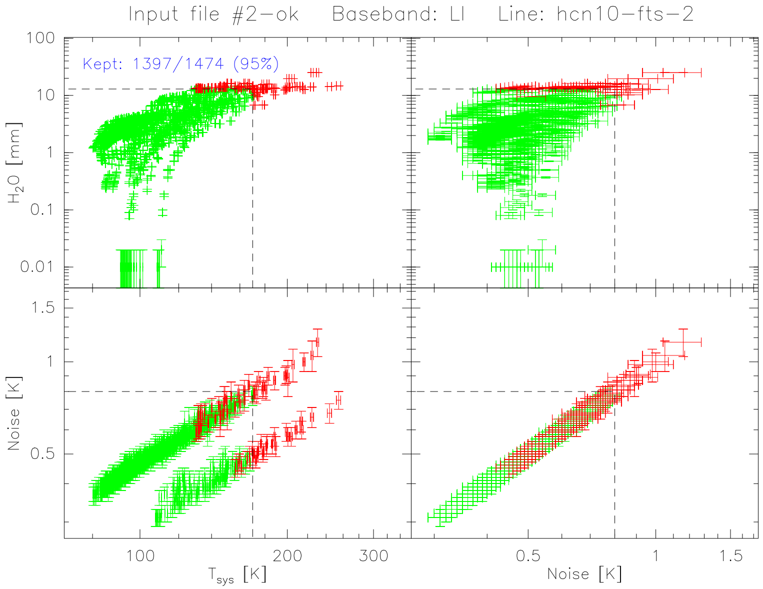
Data filtering: Use the empirical RMS noise

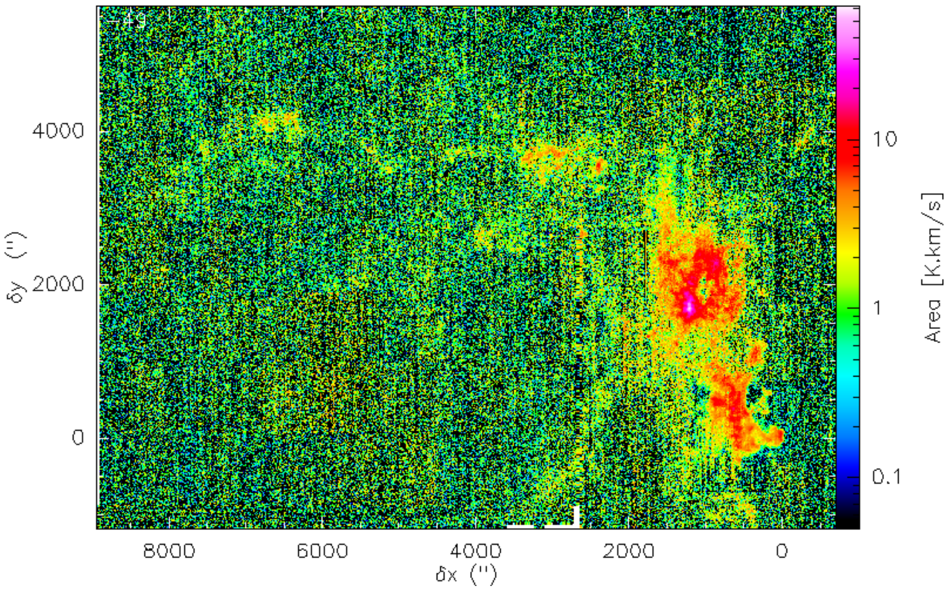
Bad weather data are filtered but data with receiver instabilities remain





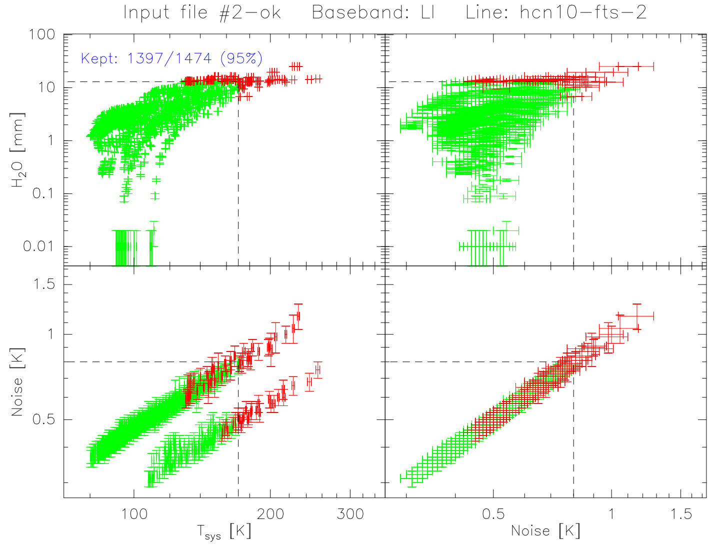
Bad weather data and data with receiver instabilities are filtered



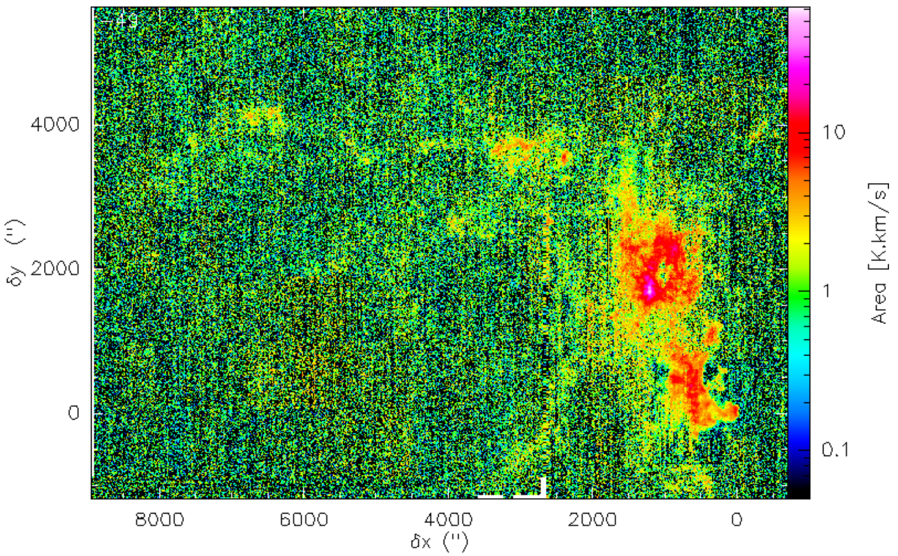


Data filtering:
Weather conditions matter, even at 3 mm

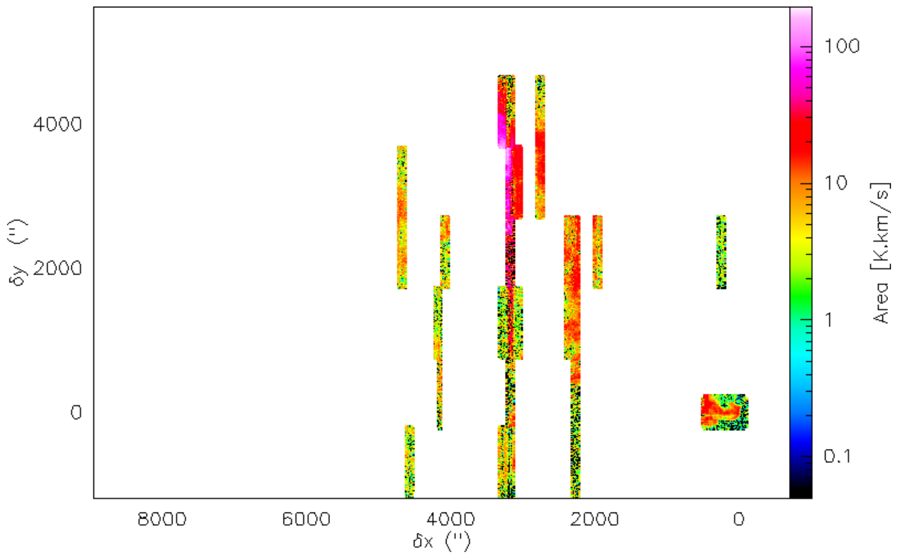
Bad weather data and data with receiver instabilities are filtered



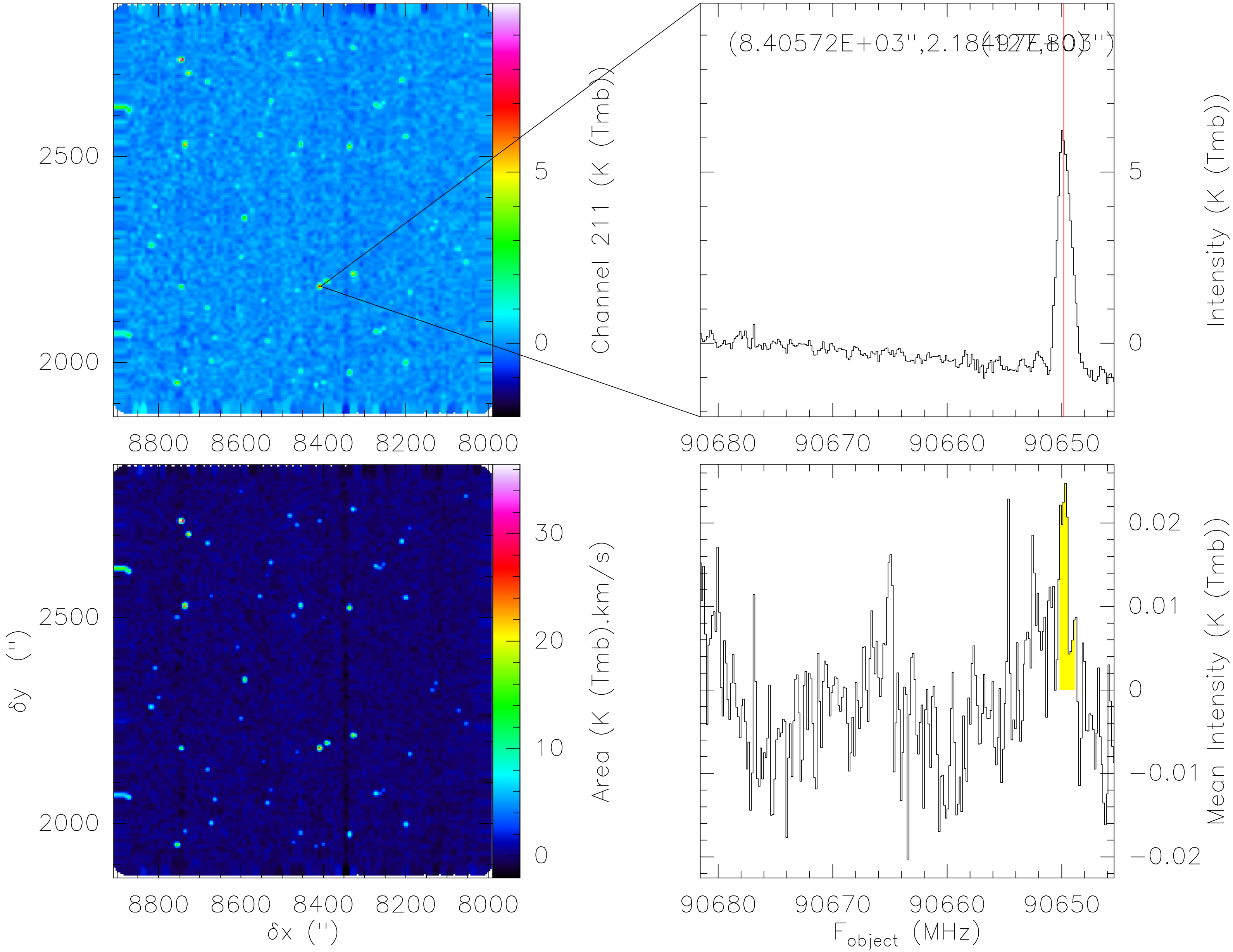
Good to average weather



Bad weather

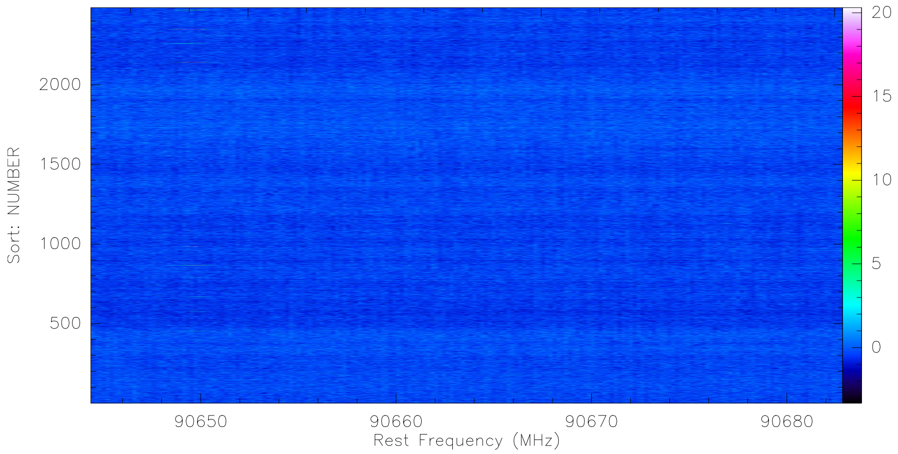


RFI from cellular phones, internet switches, car radars, etc

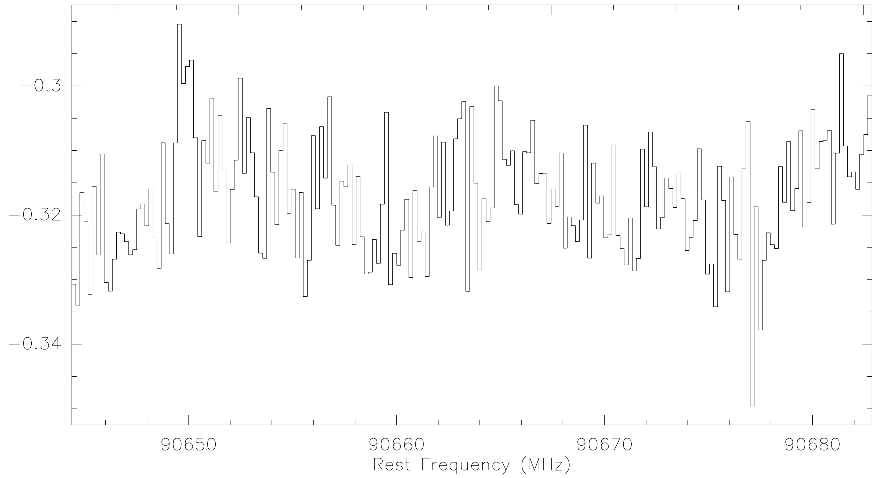


RFI from cellular phones, internet switches, car radars, etc

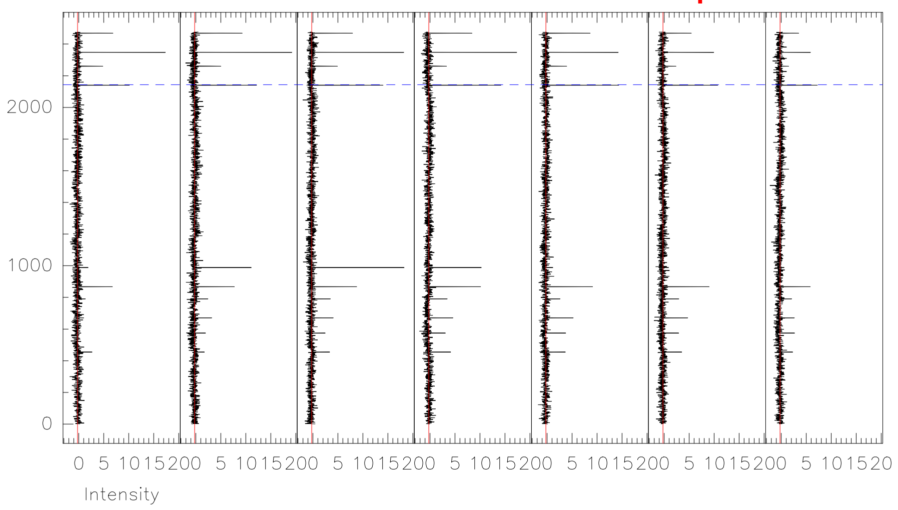
Spectro-Time drift ⇒ visible if you know where to search for



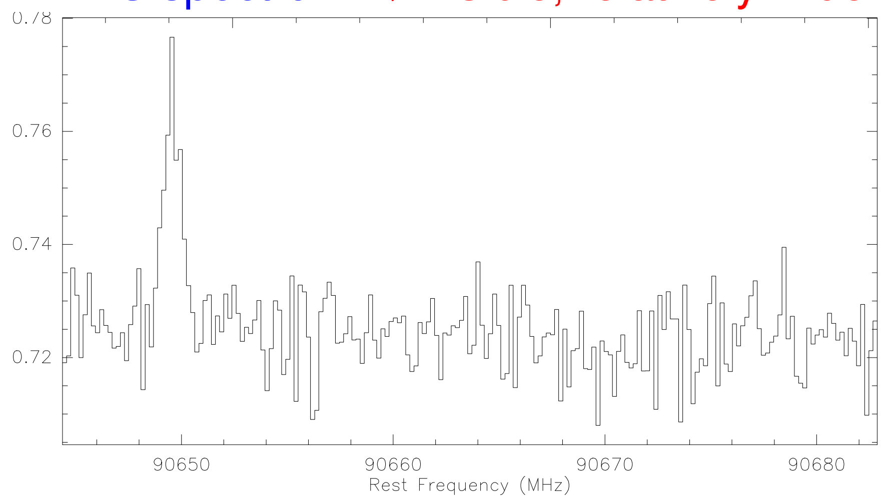
Average spectrum ⇒ invisible



Time drift ⇒ transient random pulses



RMS spectrum ⇒ visible, relatively wide



Take-home messages

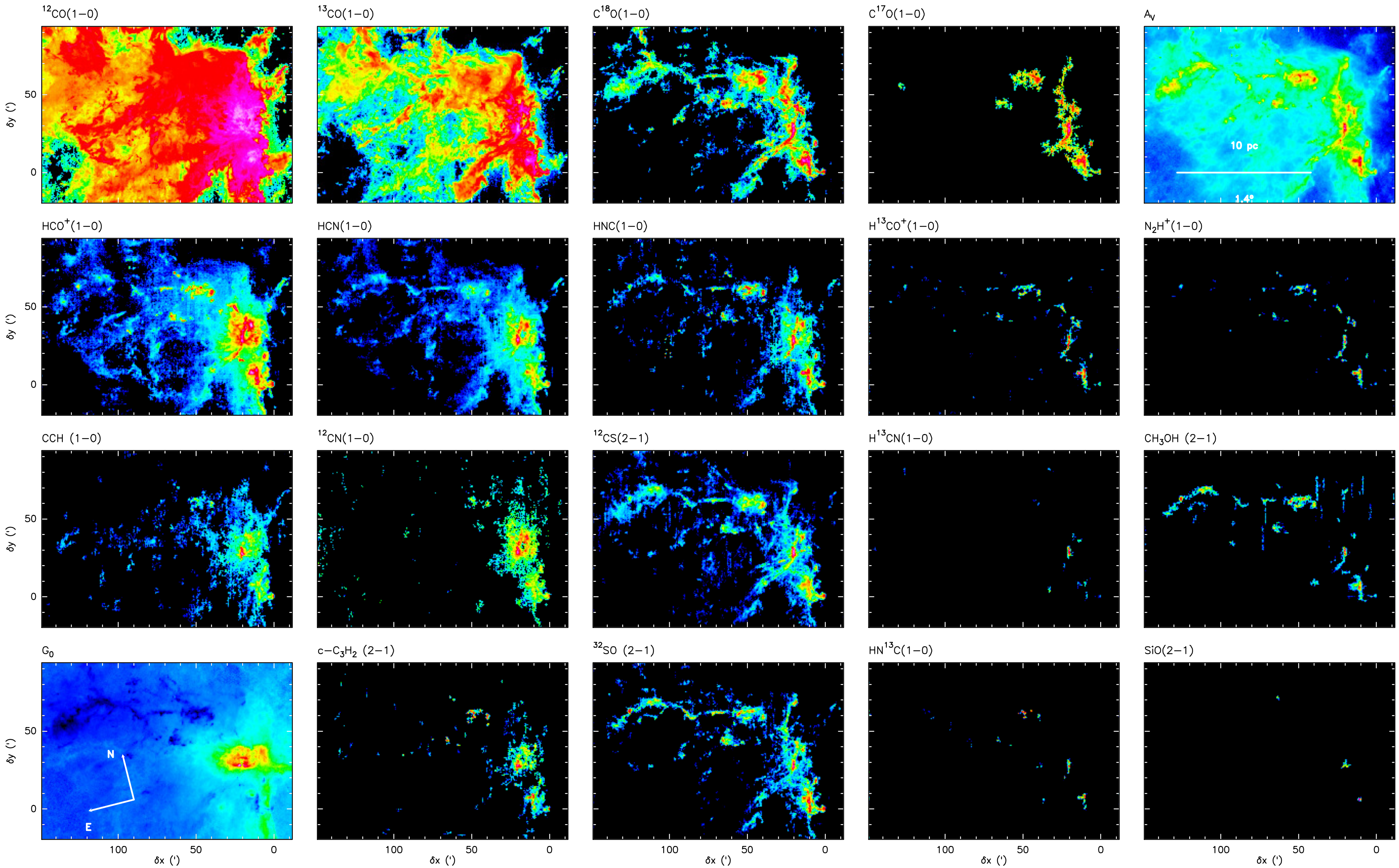
Hardware limitations Astronomers have to leave with them.
$\Rightarrow$ Many thanks to hardware engineers to keep improving the frontend/backends!
$\Rightarrow$ Observe at wavelengths where the telescope is good to excellent!

Datarates Think about the minimum datarate required by a science goal, not the maximum datarate that the instrument can deliver (our children will thank us!).

Atmospheric contribution Think about nearby signal-“free” reference positions (all the more difficult than the sensibility increases)!

Flat fielding is as important as having a good reference position.

It is worth!


^{12}CO(1−0)
^{13}CO(1−0)
C^{18}O(1−0)
C^{17}O(1−0)
A_V
10 pc
1.4°
HCO$^+$(1−0)
HCN(1−0)
HNC(1−0)
H^{13}CO$^+$(1−0)
N_2H^+(1−0)
CCH (1−0)
^{12}CN(1−0)
^{12}CS(2−1)
H^{13}CN(1−0)
CH_3OH (2−1)
G_0
N
E
c−C_3H_2 (2−1)
^{32}SO (2−1)
HN^{13}C(1−0)
SiO(2−1)
δy (')
δx (')
50
0
100
50
0