Machine Learning and Inversions of Mg II h&k spectra

A. Sainz Dalda[1,2],

J. de la Cruz[3], M. Gošic[1,2], & B. De Pontieu[1]

with the special guests

L. E. Fagernæs & M. Cheung

1. LMSAL 2. BAERI 3.University of Stockholm

Invited Talk

5. Opportunities and challenges

Machine learning and inversions of Mg II h/k spectra

A. Sainz Dalda[1], J. de la Cruz Rodríguez[2], M. Gošić[1], and B. De Pontieu [3]

[1] *Bay Area Environmental Research Institute - LMSAL*
[2] *University of Stockholm* [3] *Lockheed-Martin Solar and Astrophysics Laboratory*

Up to date, IRIS has obtained roughly 16500 datasets containing the Mg II h/k spectra. The variety of targets observed, the different set-ups used, and the large number of pixels involved in any of these datasets make their inversion a computationally expensive task. We present a new method to invert the Mg II h/k spectra observed by IRIS. This new approach is based on the STiC inversion of Representative Profiles (RPs). STiC takes into account non-LTE lines and continua including the effects of partial redistribution. The RPs are calculated using an easy-to-understand, easy-to-implement machine learning technique. Thus, a massive number of Mg II h/k profiles can be easily inverted and interpreted in a feasible, meaningful way. As a consequence of this new framework, we are able to recover in a few minutes the thermodynamics at the chromosphere from most of the IRIS Mg II h/k observations.

Machine Learning and Inversions of Mg II h&k spectra

A. Sainz Dalda[1,2],

J. de la Cruz[3], M. Gošic[1,2], & B. De Pontieu[1]

with the special guests

L. E. Fagernæs & M. Cheung

1. LMSAL 2. BAERI 3.University of Stockholm

Scientist's Happiness Plot

IRIS team *non-scientific* goal:

to make you feel (very) happy!

To provide the tools that allow you to earn valuable knowledge about the solar chromosphere in an affordable, easy way, and in a sort time

Knowledge

Time

Mg II h&k IRIS Observations

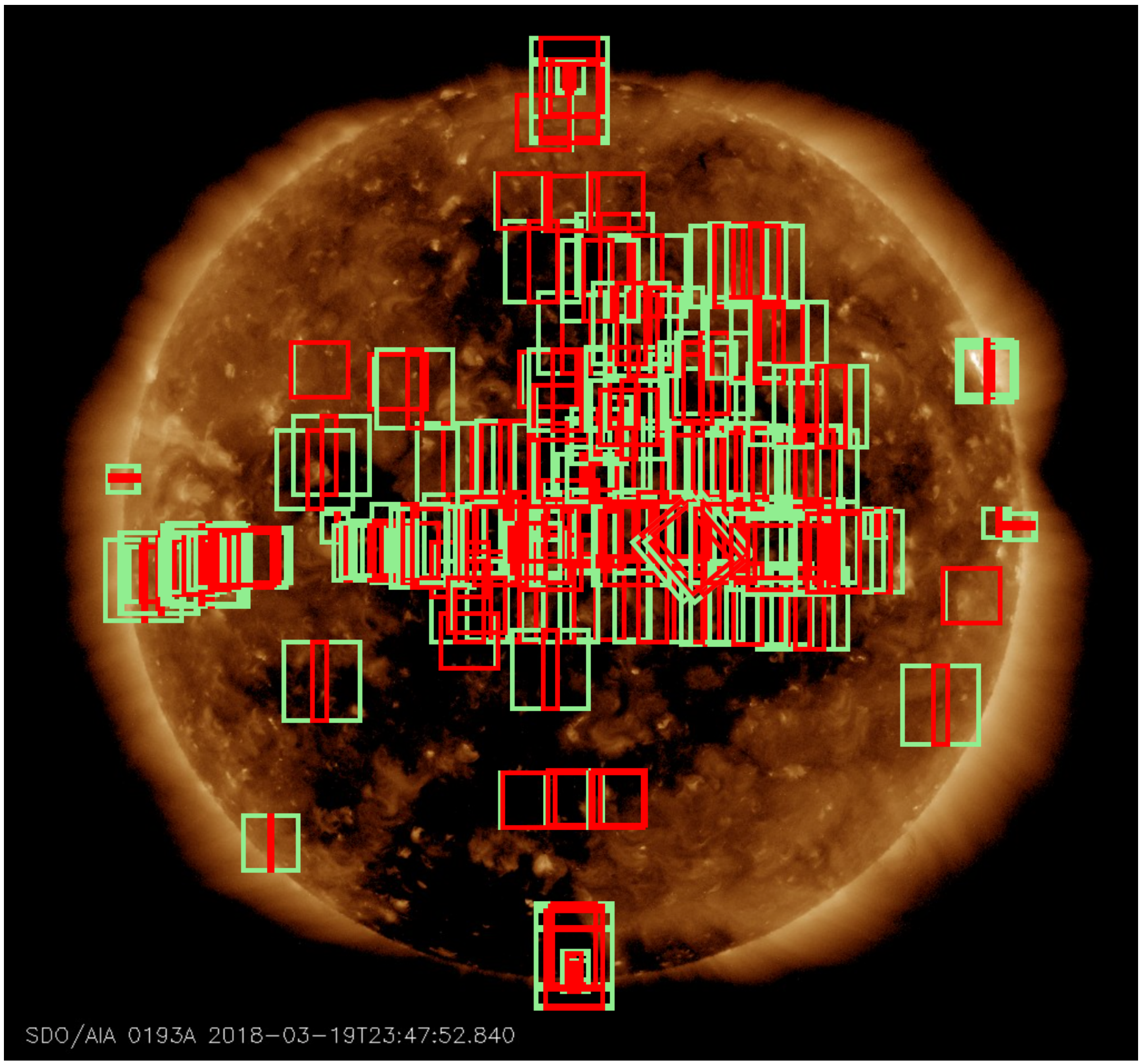
17192 data sets from 2013-06-27 to 2018-06-27

Sit & Stare

Small raster

Large Sparse 4-step raster

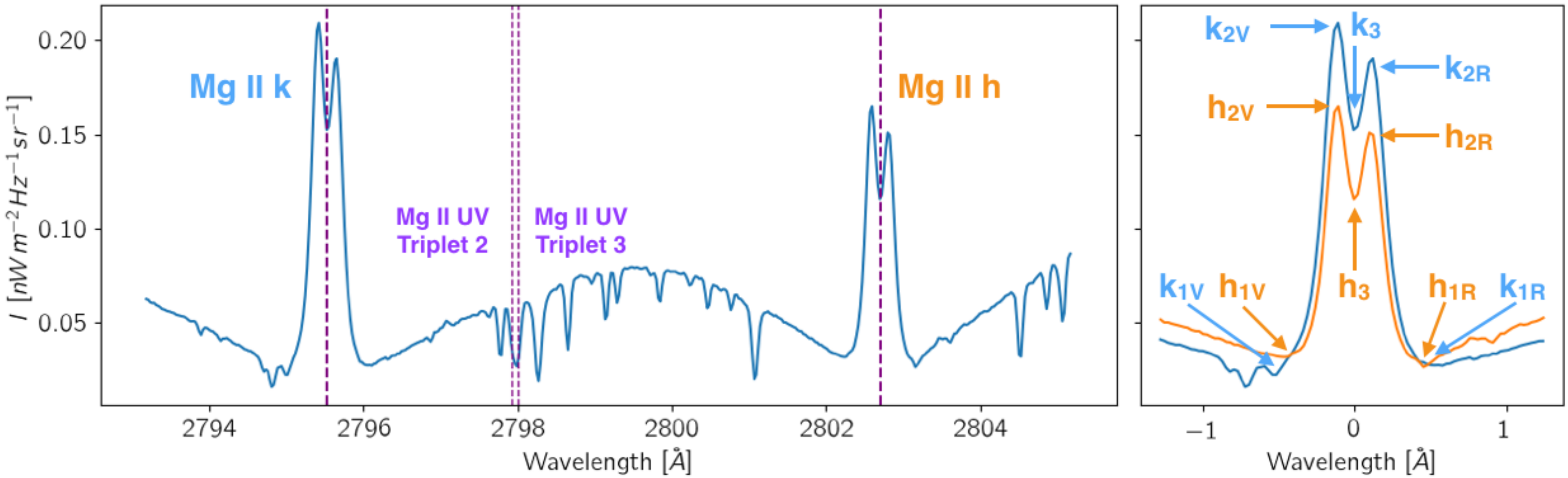
Medium dense 16/32-step raster



Very Large Sparse 64-step raster

Very large dense 320-step raster

Proxies based on Mg II h&k Profile Features



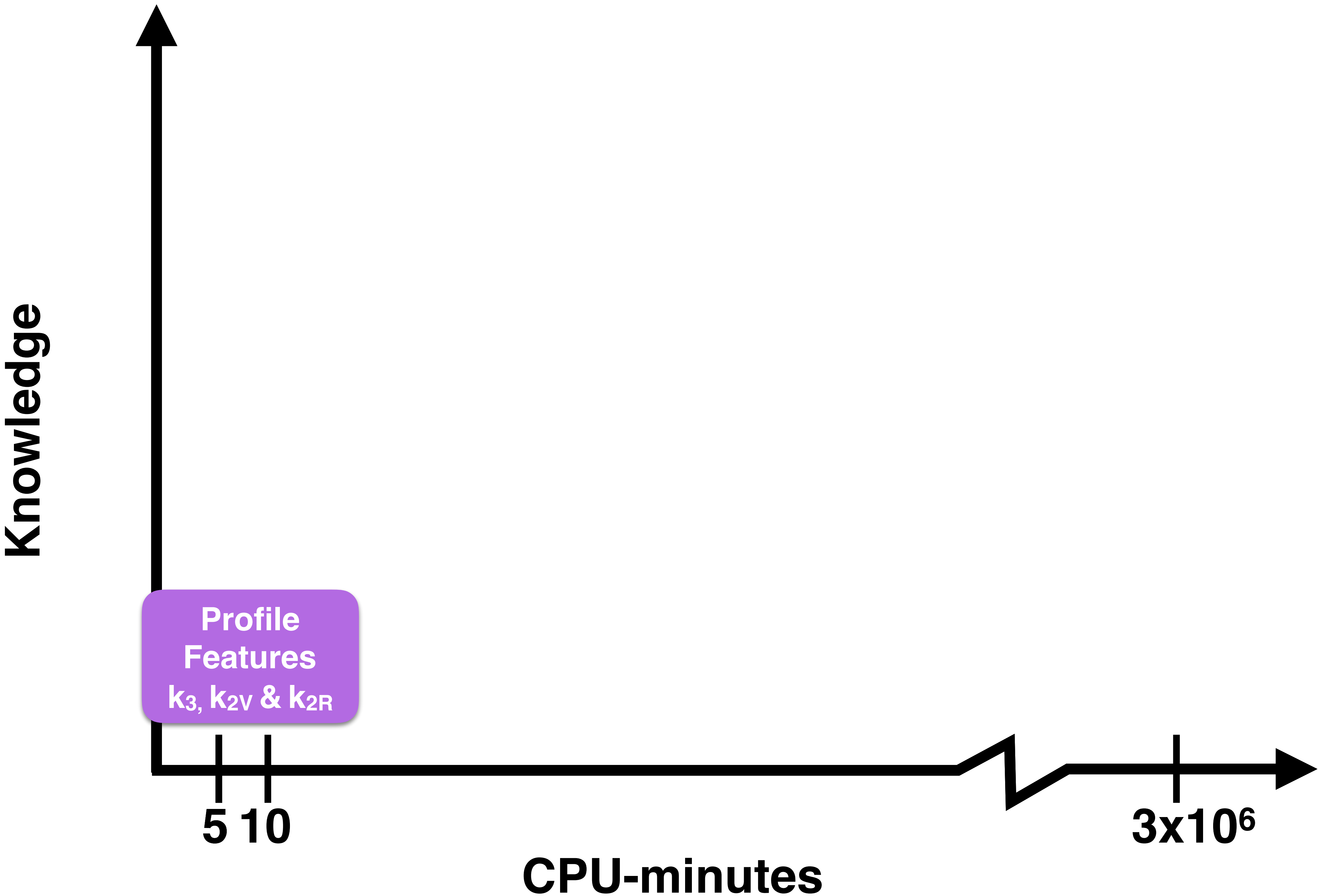
Correlation between Mg II Features and Atmospheric Properties

Spectral Observable	Atmospheric Property
Δv_{k3} or Δv_{h3}	Upper chromospheric velocity
Δv_{k2} or Δv_{h2}	Mid chromospheric velocity
$\Delta v_{\mathrm{k3}} - \Delta v_{\mathrm{h3}}$	Upper chromospheric velocity gradient
k or h peak separation	Mid chromospheric velocity gradient
k_2 or h_2 peak intensities	Chromospheric temperature
$(I_{\mathrm{k2}v} - I_{\mathrm{k2}r})/(I_{\mathrm{k2}v} + I_{\mathrm{k2}r})$	Sign of velocity above $z(\tau = 1)$ of k_2[a]

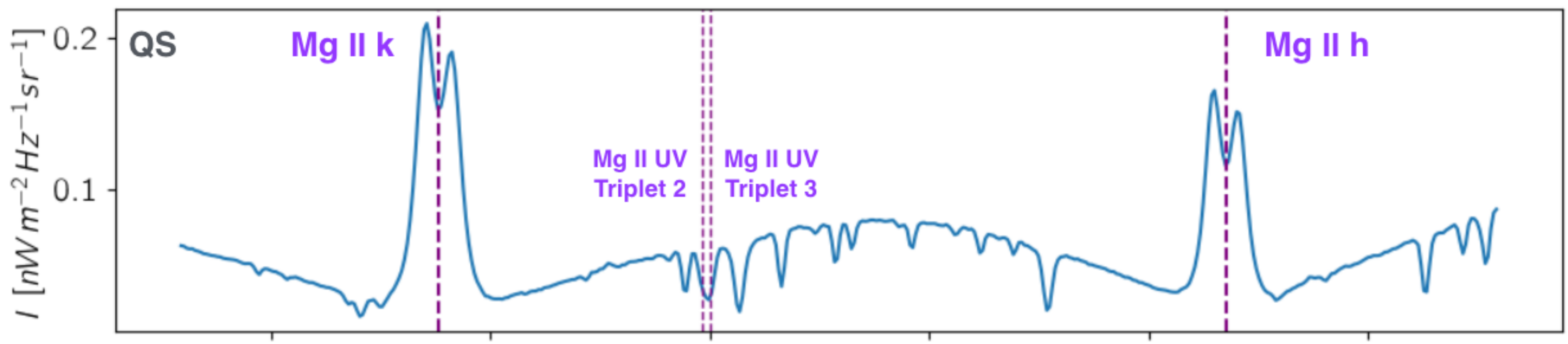
Notes. This is a simplified view and all correlations above have scatter.
[a] Likewise for the $h_2$peaks.

Pereira et al., 2013, ApJ, 778,143

Recovering physical from IRIS Mg II h&k lines

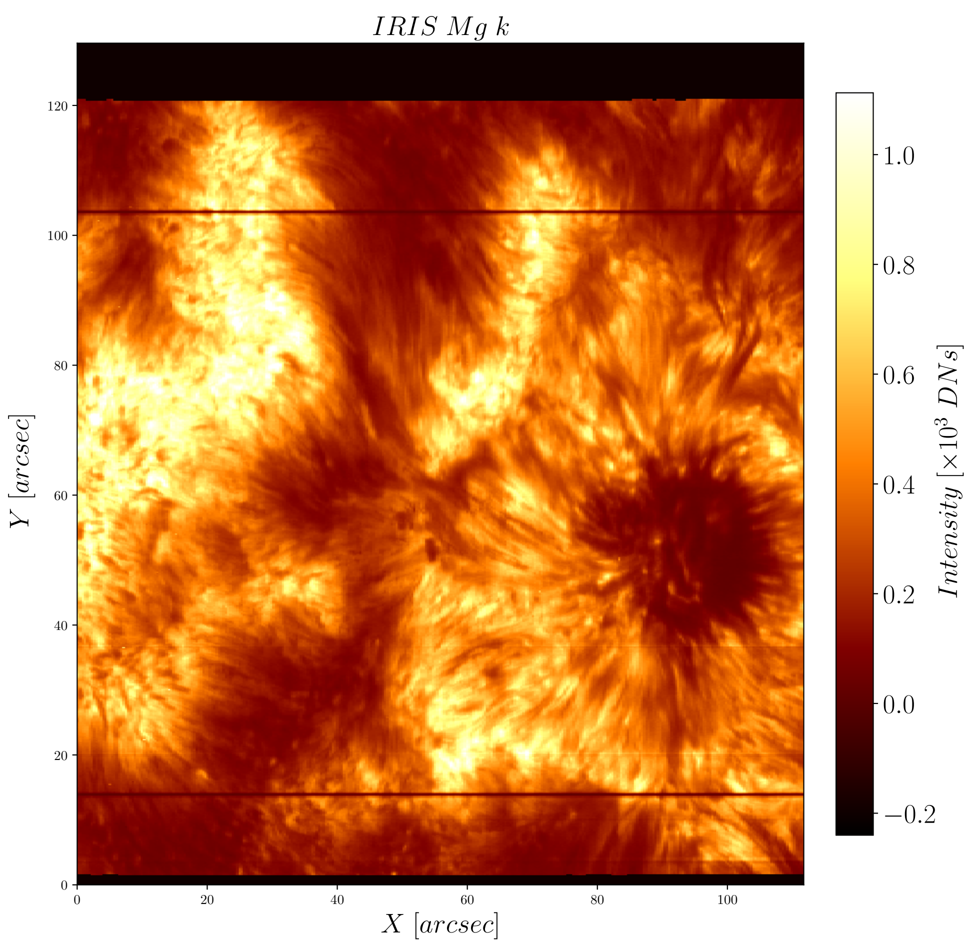


Proxies based on Mg II h&k Profile Features



IRIS Large Dense 320-Step Raster

2016-01-14 23:04:09



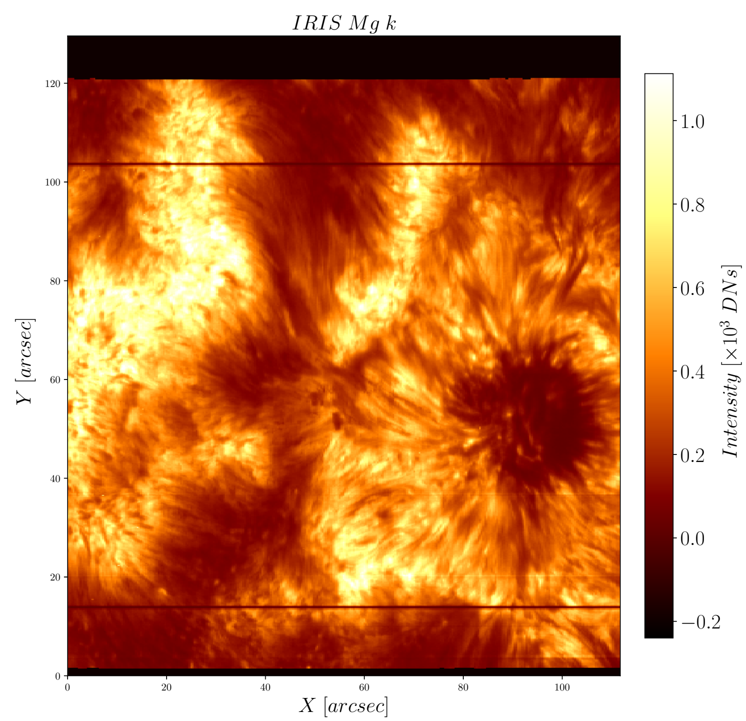
Data Information

- Location: (X,Y) = (559”, 121”), μ = 0.81
- Spatial scale in X, Y: 0”.35, 0”.16
- Spectral scale: 0.025 mÅ
- Analyzed spectral range: Mg II k&h
- Exposure Time: 8s
- FoV size: 112” x 119”
- Pixels in the FoV: **228941**

Inversion of the Full data by STiC

de la Cruz Rodríguez et al. 2016, ApJ, 830, 30D

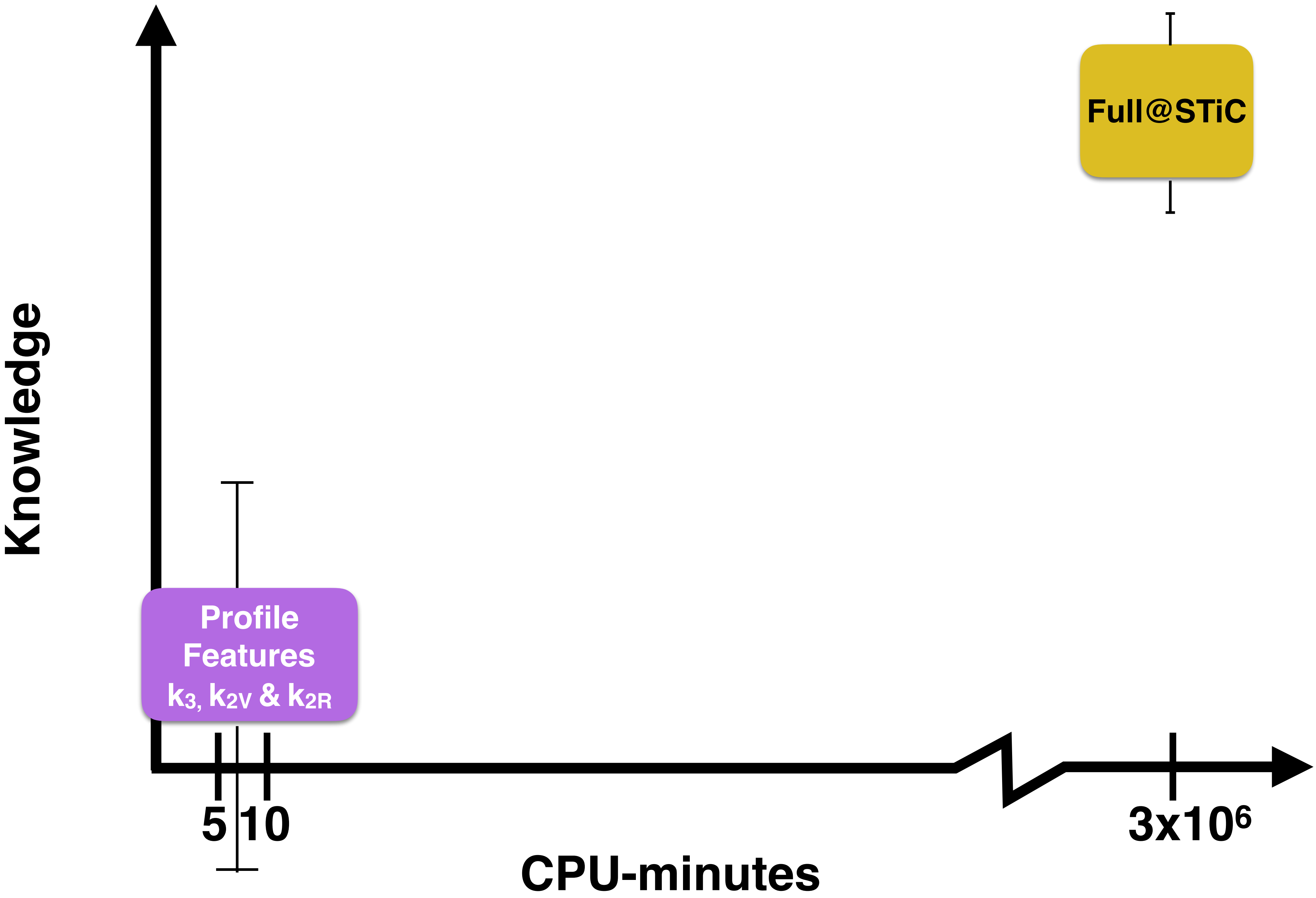
Inversion performance: **1 px / 2 CPU-hour** (non-LTE PRD)
Setup:inversions of only-Intensity profiles with 473 spectral points, model atmosphere in **39 optical depth heights**, 2 cycles, 7 nodes in T, and 4 nodes both in v_{los} and v_{turb}.



Estimated time:
228941px / (1px/2 CPU-hours) ~ 460000 CPU-hours

Estimated time ALL IRIS data 17192 x 50000 px/(1px/2 CPU-hours) ~ 1.7 10^9 CPU-hours

Recovering physical from IRIS Mg II h&k lines



How can we reduce the representation of our data without losing their significance?

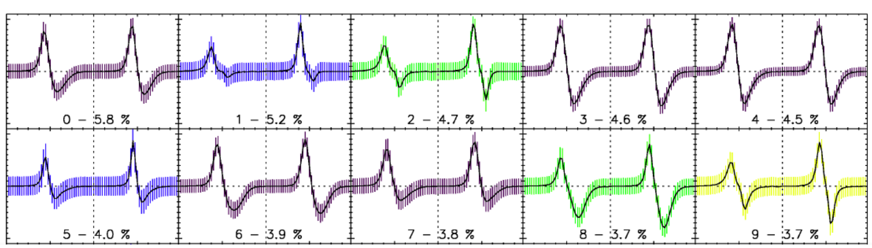
We want to reduce the number of profiles that we need to analyze while we keep most of the information of our data

k-mean Clustering Technique

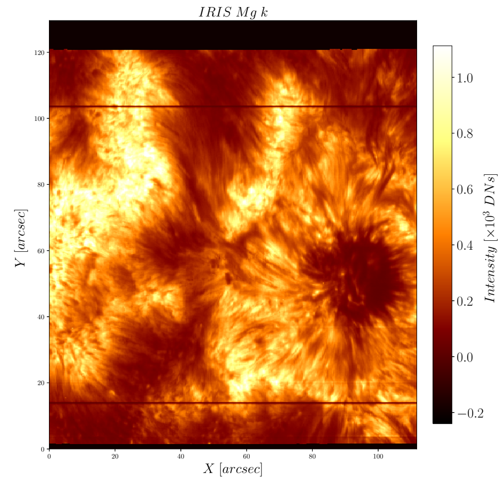
- Technique introduced by MacQueen,1967; Steinhaus, 1957; Lloyd, 1982

- In Solar Physics has been used by:
  Pietarila et al. 2007, ApJ, 663 (Ca II IR 8542)
  Viticchié and Sánchez Almeida 2011, A&A, 530 (Fe I 6302)
  Xu et al. 2016, ApJ, 819 (Mg II k)

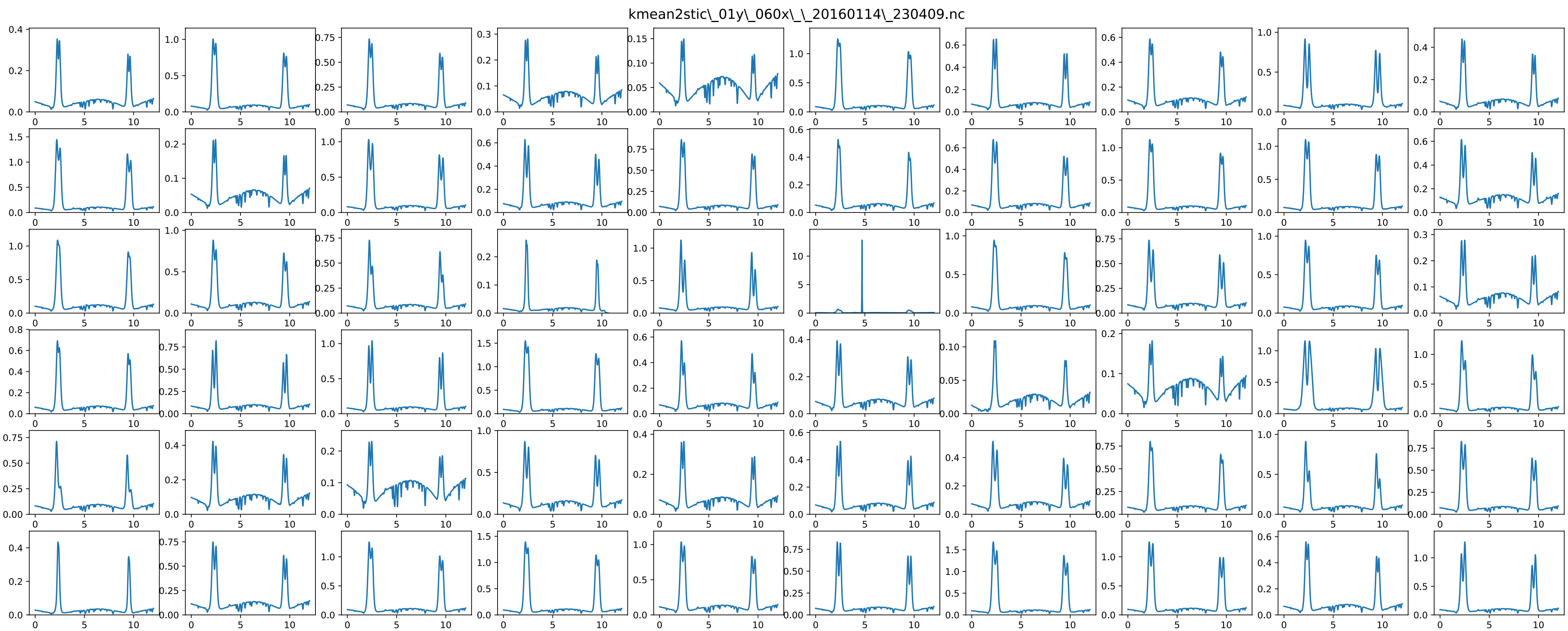
Viticchié & Sánchez Almeida: SOT/SP Stokes V profiles in the quiet Sun

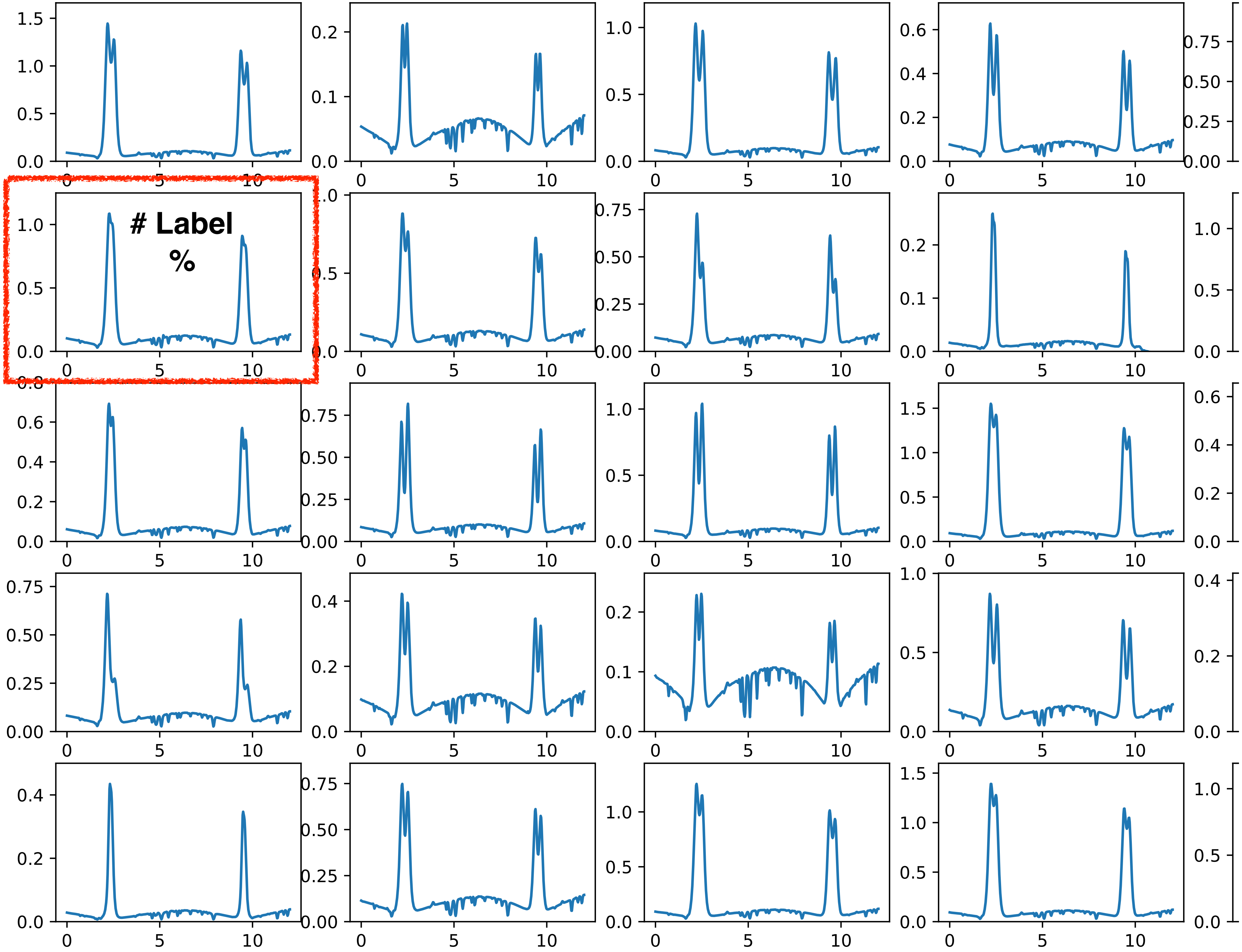


From 22894 profiles…



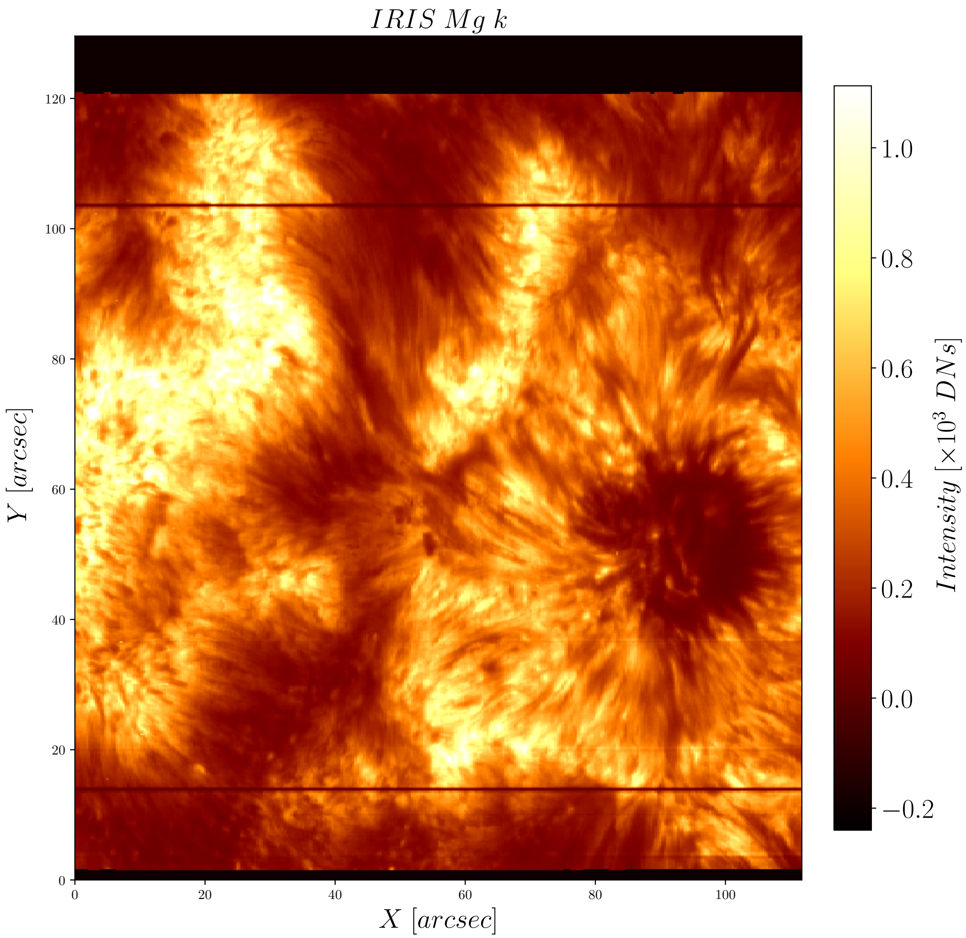
…to 60 Mg II h&k Representative Profiles (RPs)





Label
%

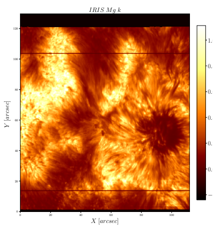
Spatial Distribution of the Representative Profiles



- **The RPs are *clustered* in a similar way that the solar features are in the solar atmosphere.**

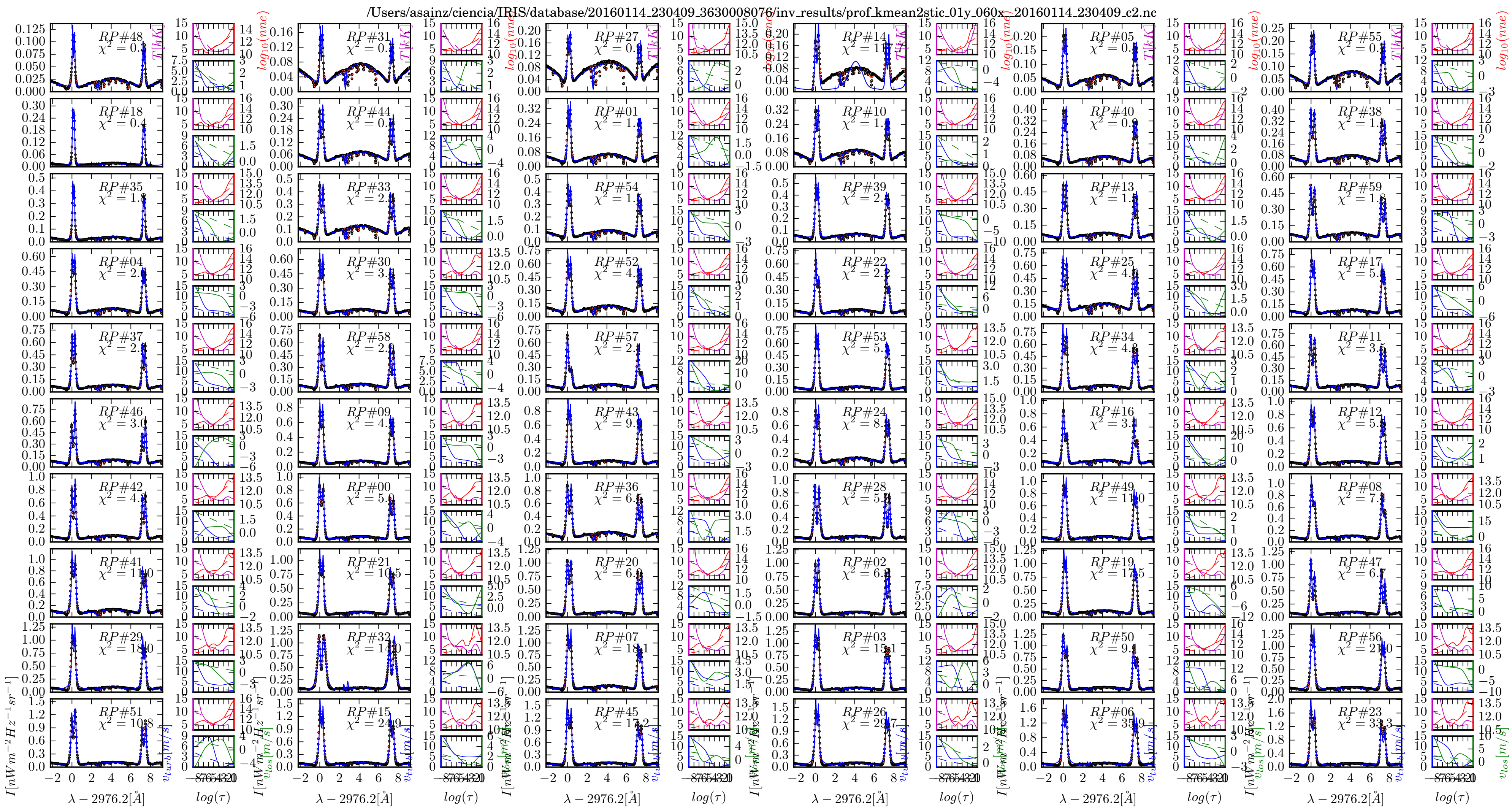
- **The RPs are the signature of different, spatially coherent conditions in the solar atmosphere, i.e., of Representative Model Atmospheres (RMAs)**

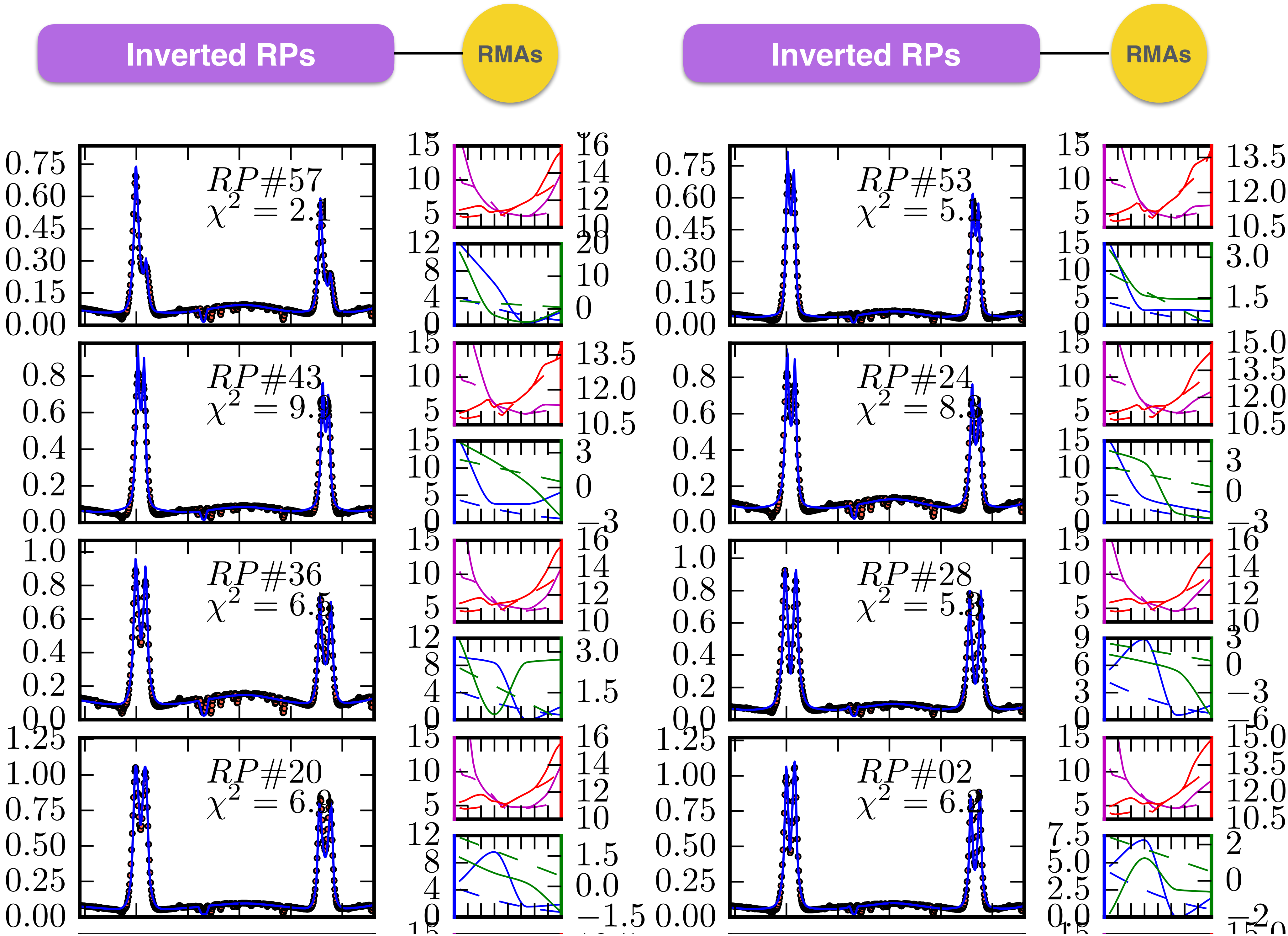
RPs Inverted by STiC



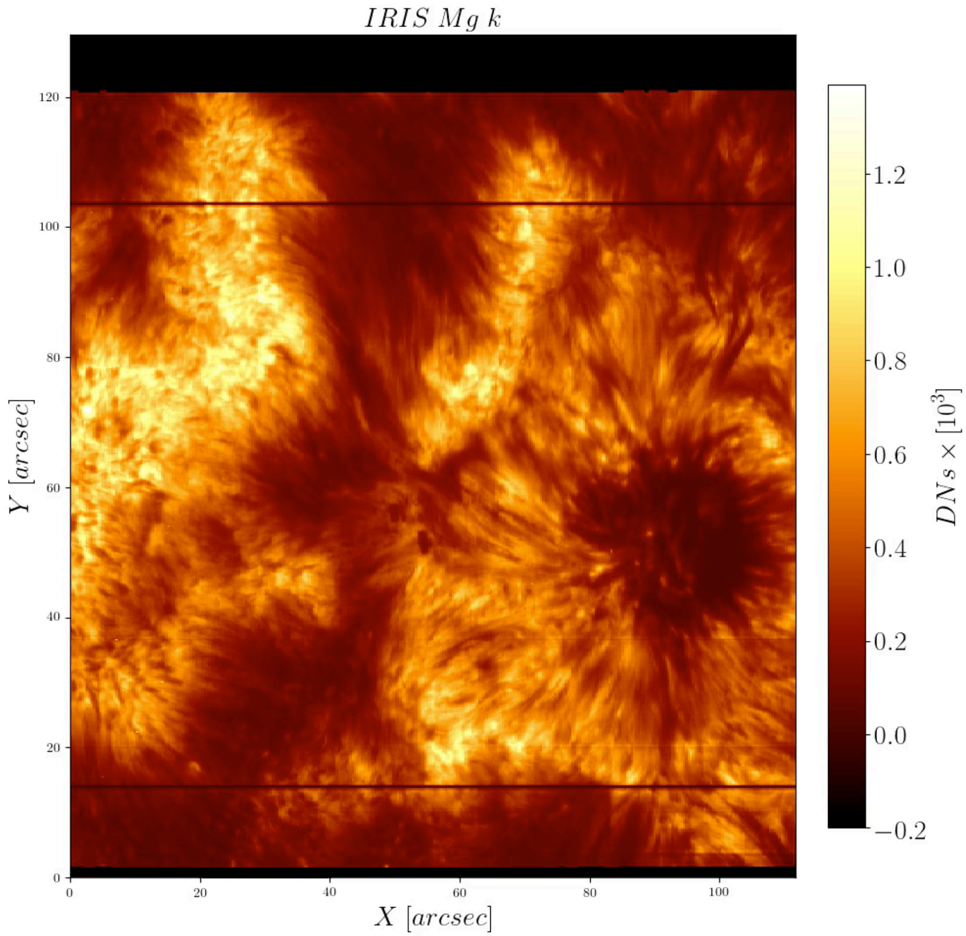
60 RPs → STiC → 60 RMAs

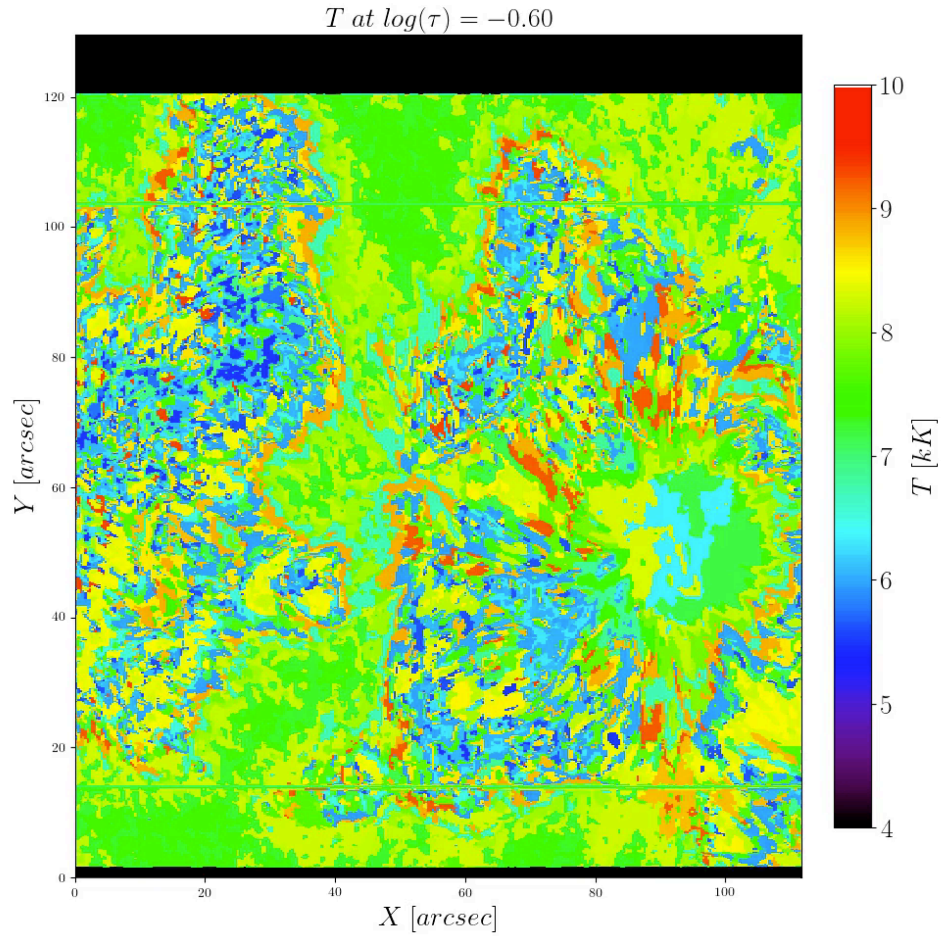
7200 CPU-minutes



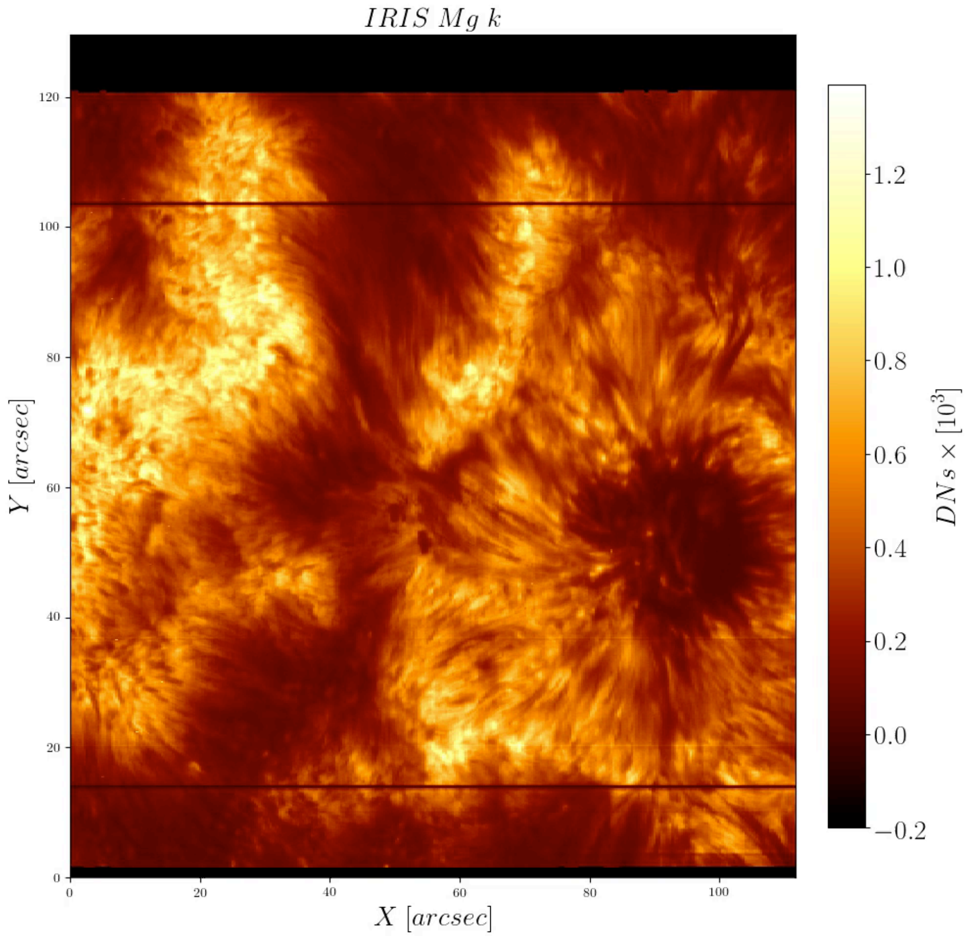

Inverted RPs
RMAs
Inverted RPs
RMAs
RP#57
RP#53
RP#43
RP#24
RP#36
RP#28
RP#20
RP#02

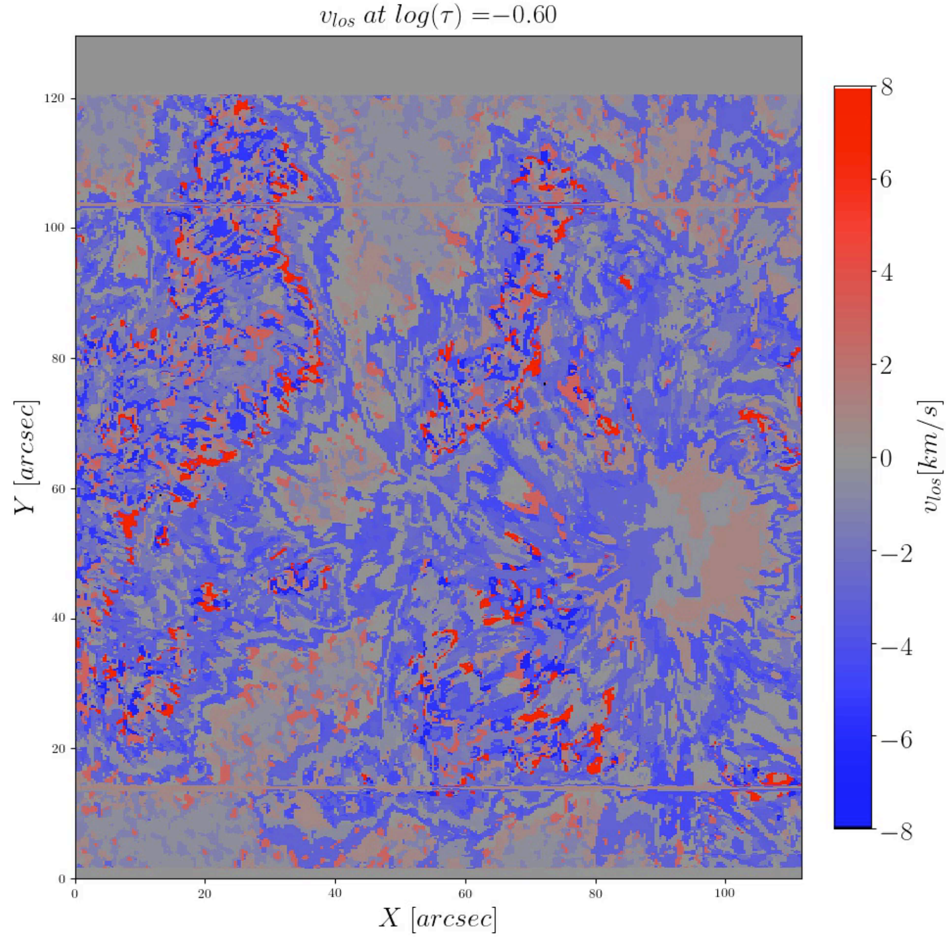
Temperature recovered from STiC inversion on 60 Mg II h&k RPs



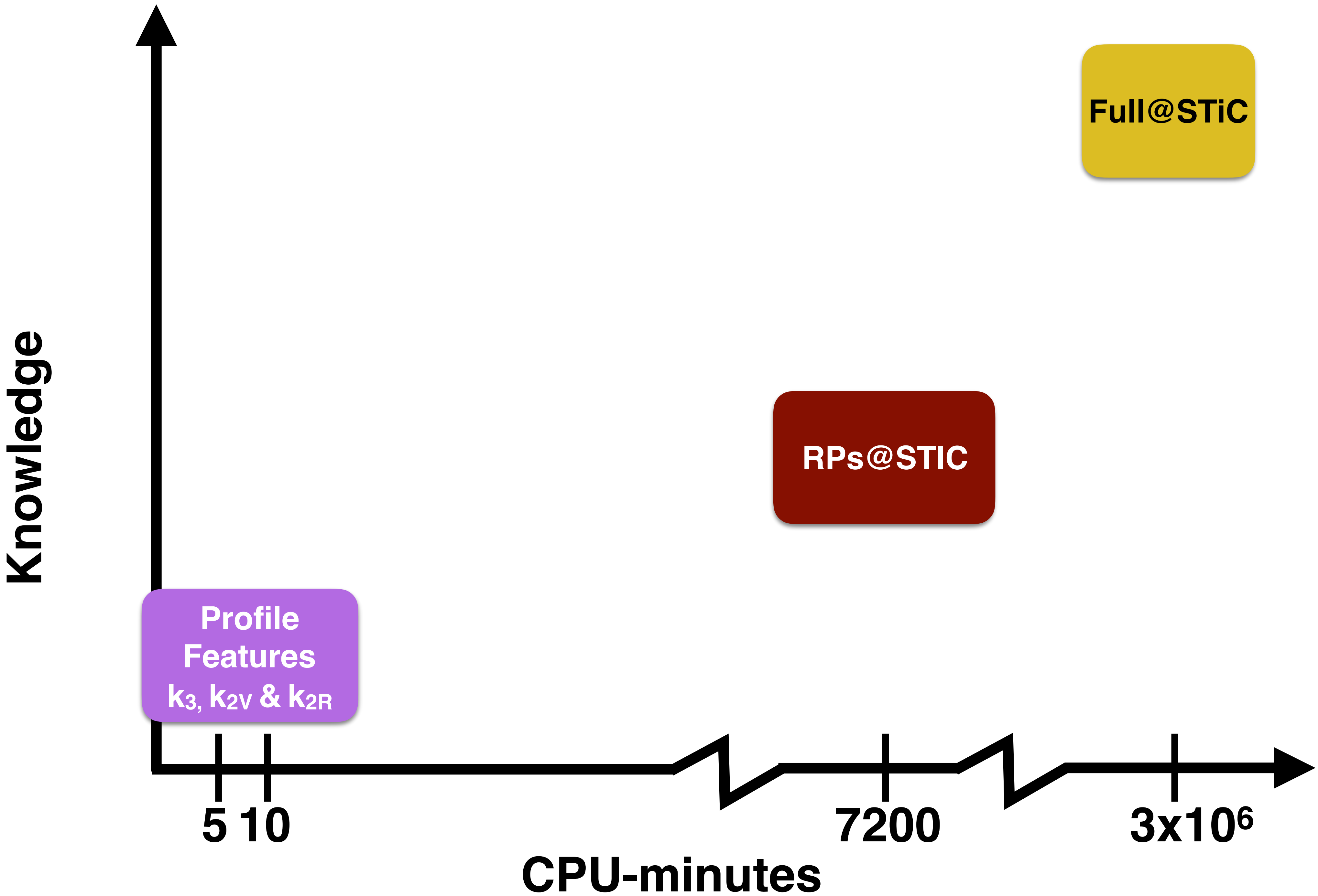


V_{los} recovered from STiC inversion on 60 Mg II h&k RPs

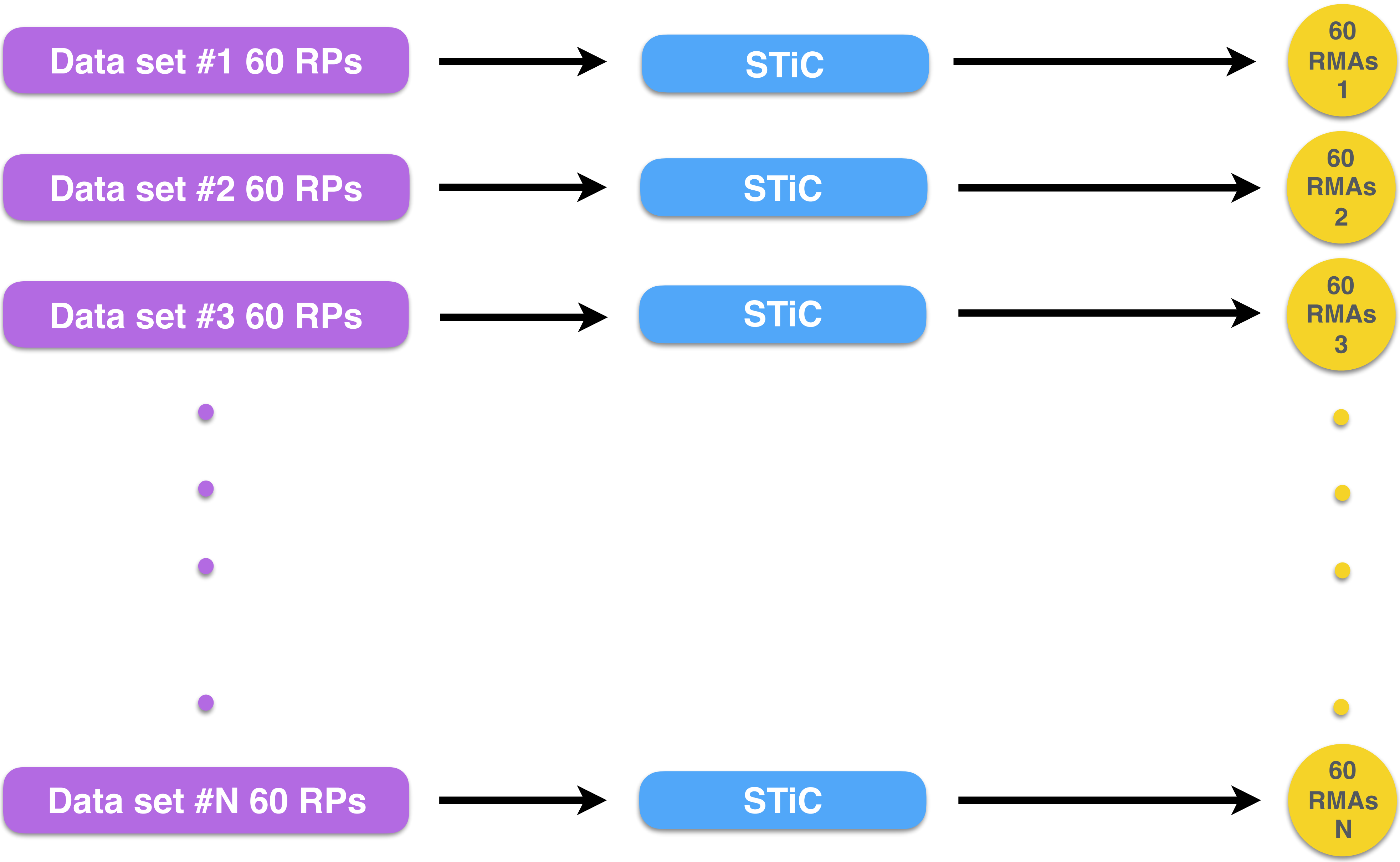




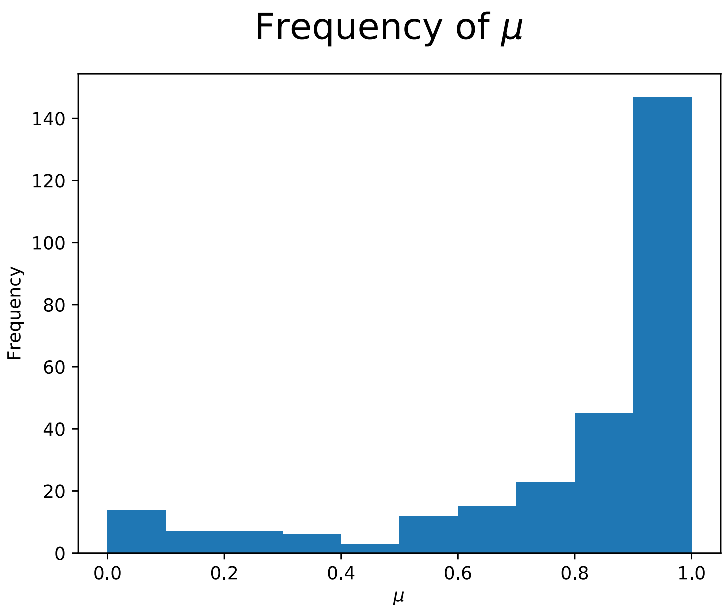
Recovering physical from IRIS Mg II h&k lines

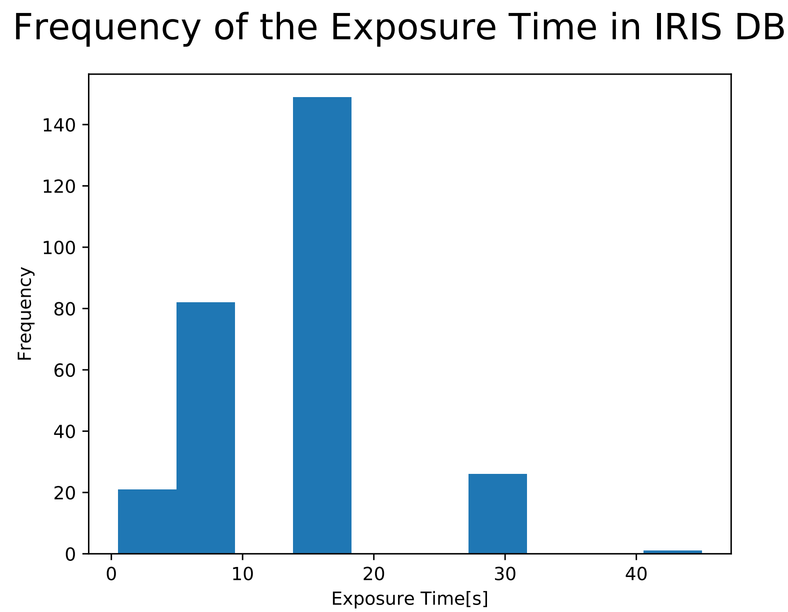


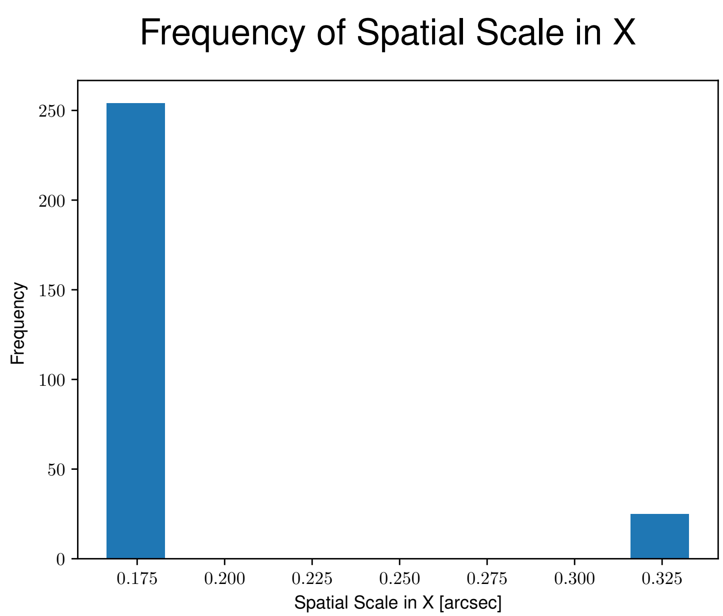
IRIS Inversion based on the RPs Inverted by STiC

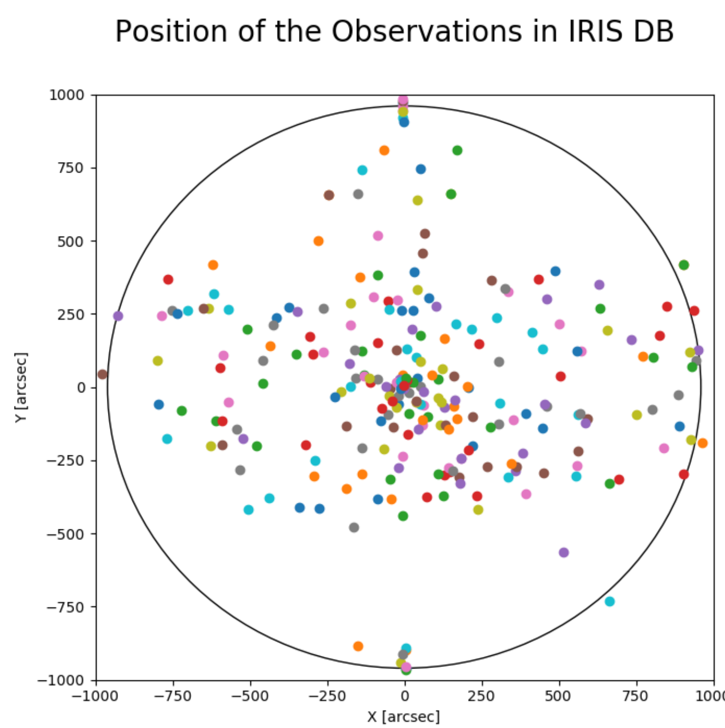


Building IRIS² database

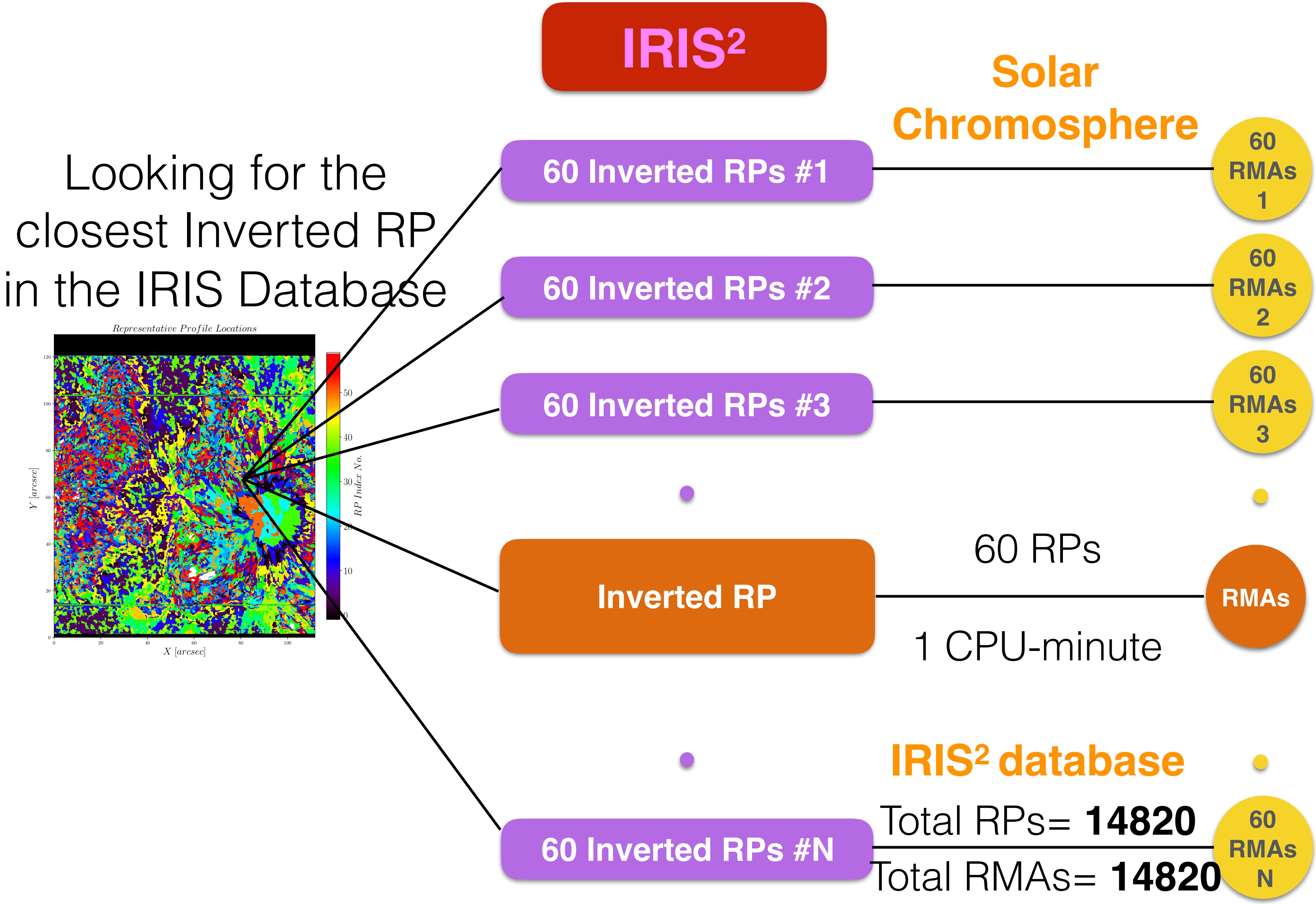








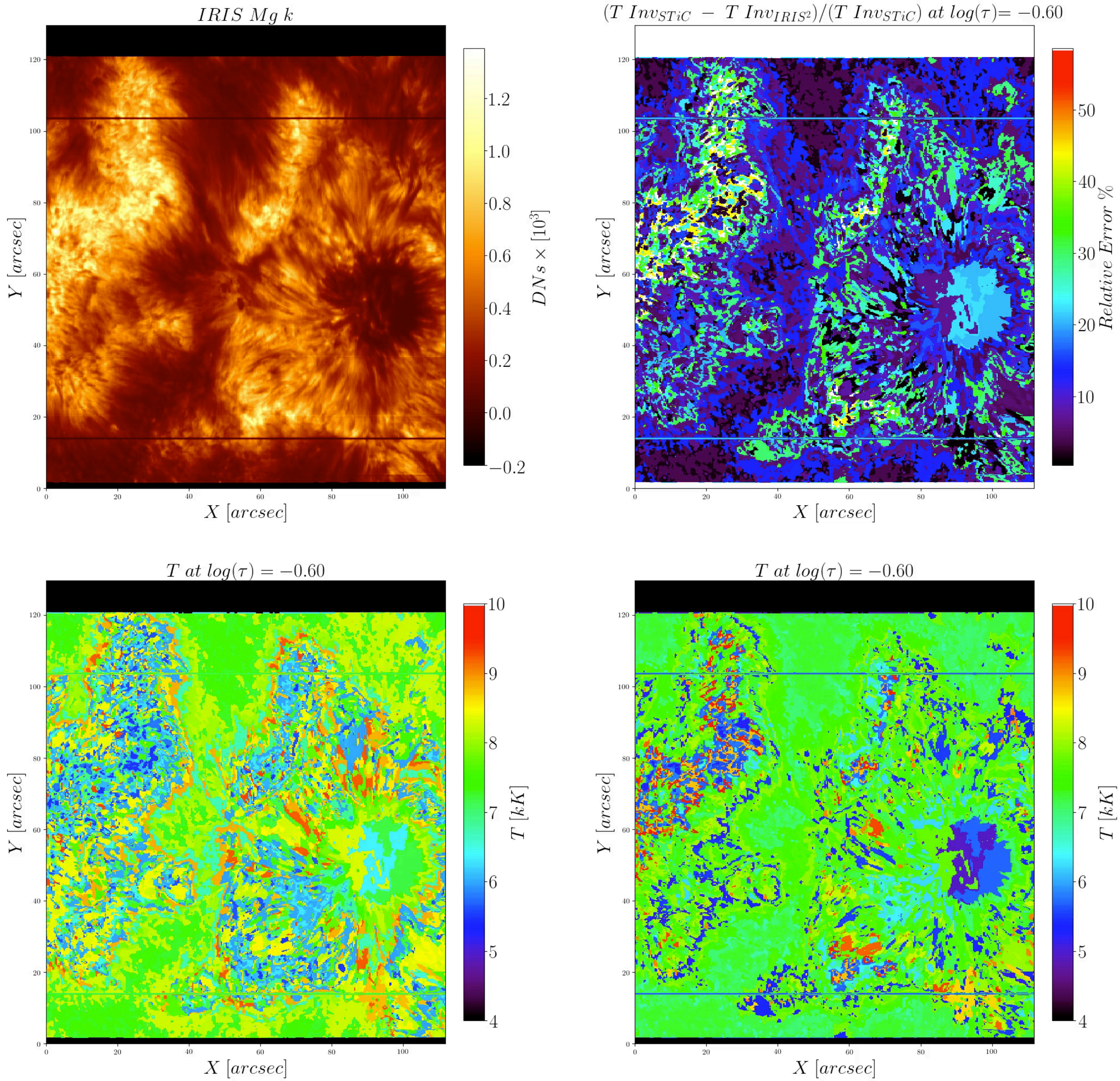
IRIS Inversion based on the RPs Inverted by STiC



INVERSION of RPs by IRIS²

IRIS
Mg II k&h
2016-01-14
23:04:09

Rel.
Error
[%]
T_{STiC}
vs
T_{IRIS^2}



T[kK]

from

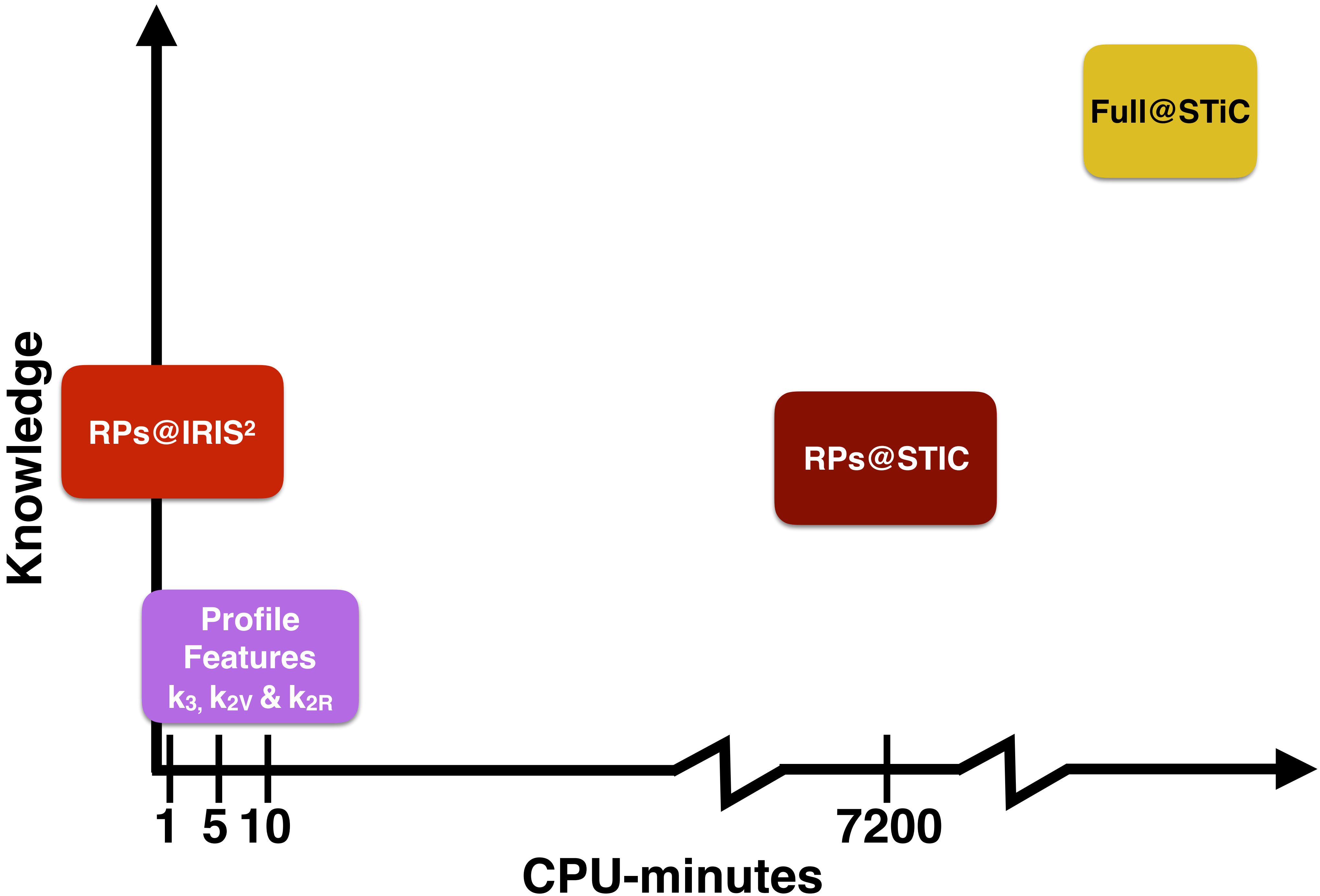
RPs
@
STiC

T[kK]

from

RPs
@
IRIS²

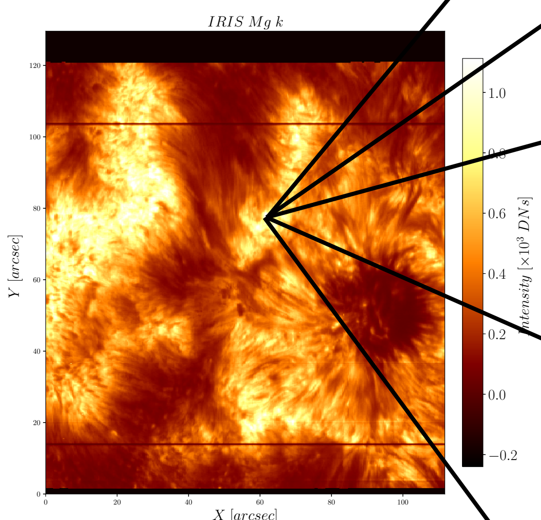
Recovering physical from IRIS Mg II h&k lines



IRIS Inversion based on the RPs Inverted by STiC

IRIS²

Looking for the closest Inverted RP in the IRIS Database



Solar Chromosphere

60 Inverted RPs #1 — 60 RMAs 1

60 Inverted RPs #2 — 60 RMAs 2

60 Inverted RPs #3 — 60 RMAs 3

Inverted Observed Profile — RMAs

228941 profiles

30 CPU-minutes

IRIS² database

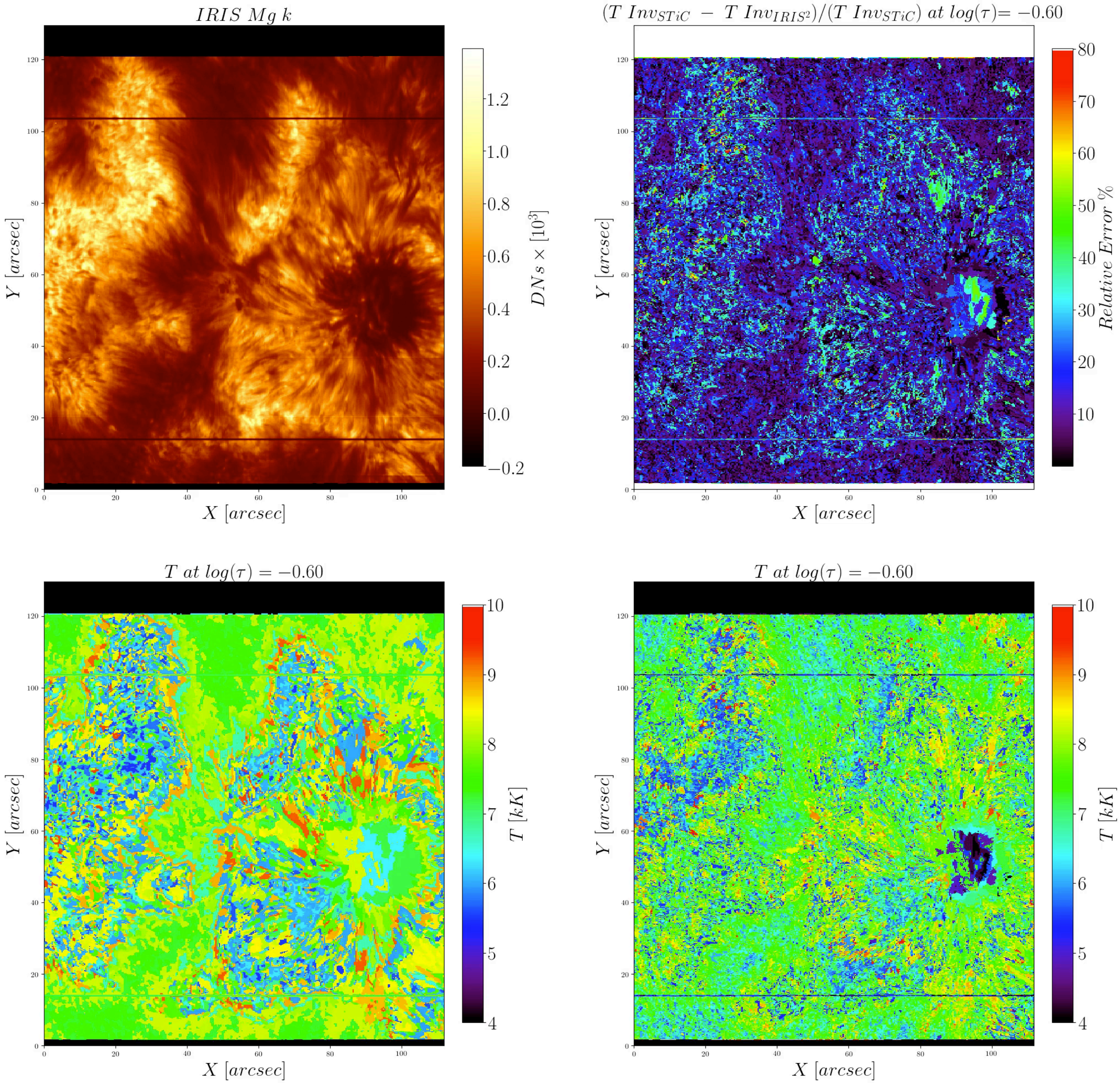
60 Inverted RPs #N — 60 RMAs N

Total RPs= **14820**

Total RMAs= **14820**

INVERSION FULL FoV "PIXEL by PIXEL" by IRIS²

IRIS
Mg II k&h
2016-01-14
23:04:09



Rel. Error [%] T_{STiC} vs T_{IRIS^2}

T[kK]

from

RPs @ STiC

T[kK]

from

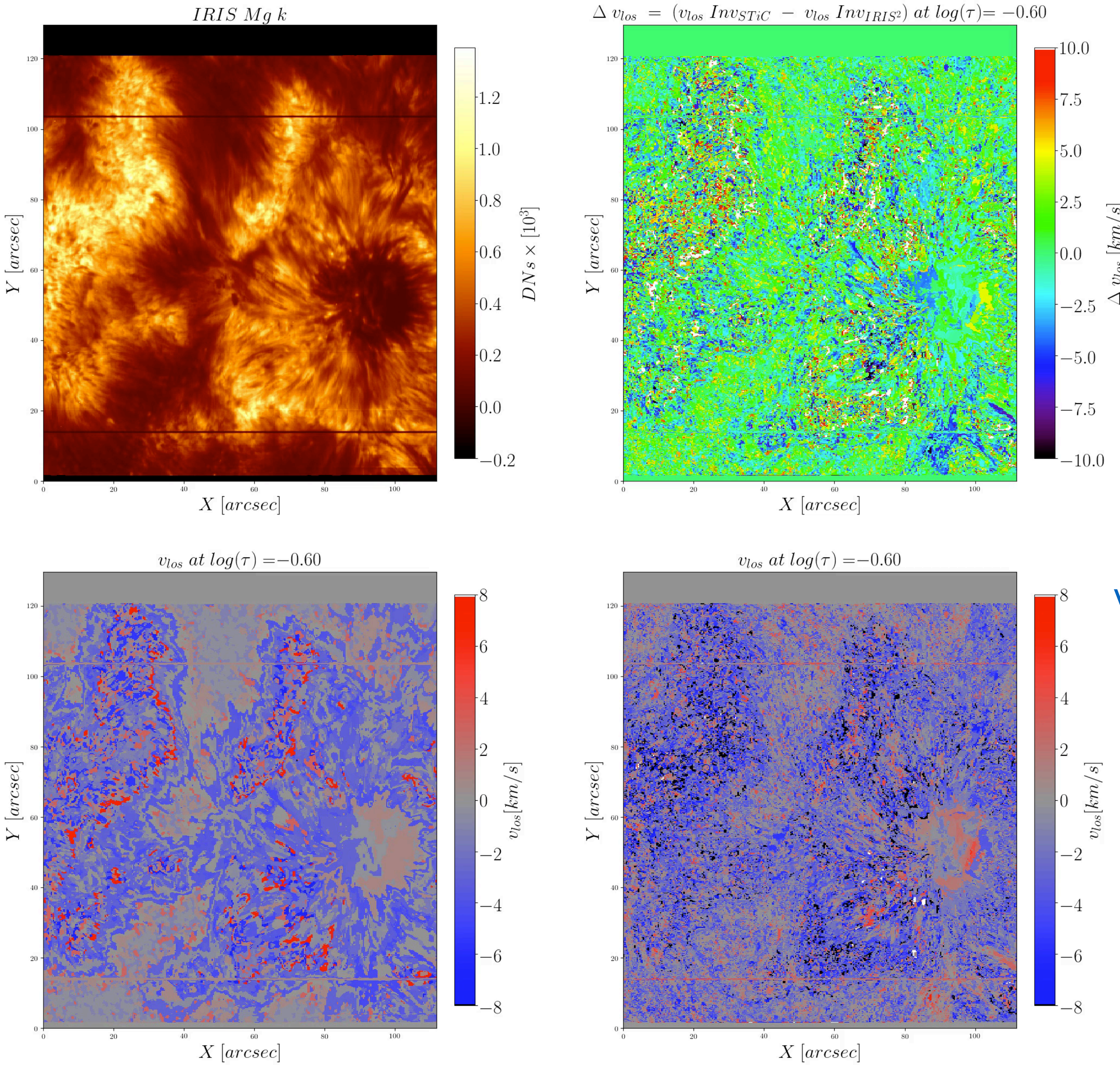
Px by Px @ IRIS²

INVERSION FULL FoV “PIXEL by PIXEL” by IRIS²

IRIS
Mg II k&h
2016-01-14
23:04:09

Diff. [km/s]

v_{los}STiC - v_{los}IRIS^2



v_{los}[km/s] from **RPs @ STiC**

v_{los}[km/s] from **Px by Px @ IRIS²**

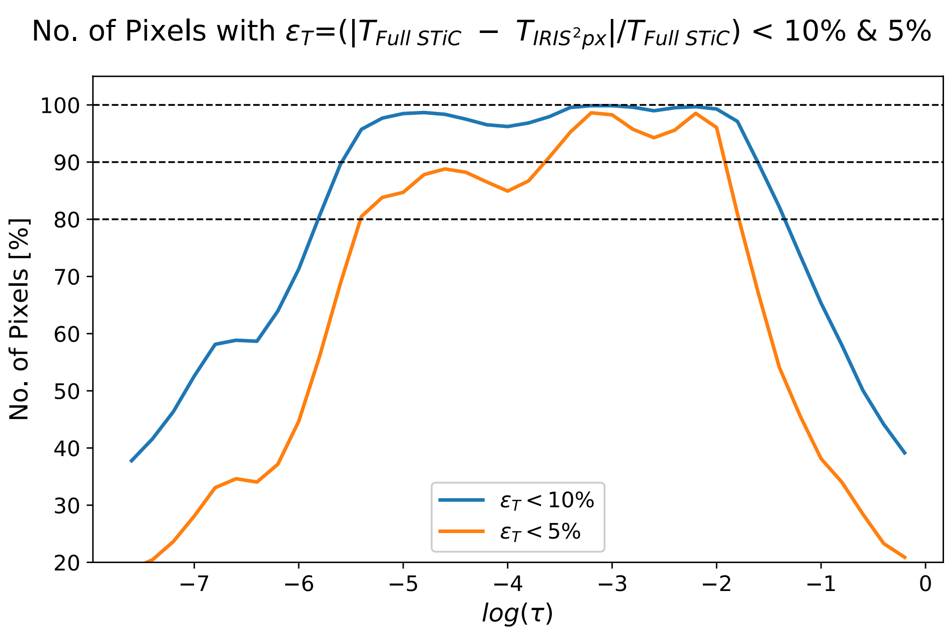
Performance of Full STiC (1px every 100px) vs IRIS²

2280 (original) px inverted by STiC and IRIS²

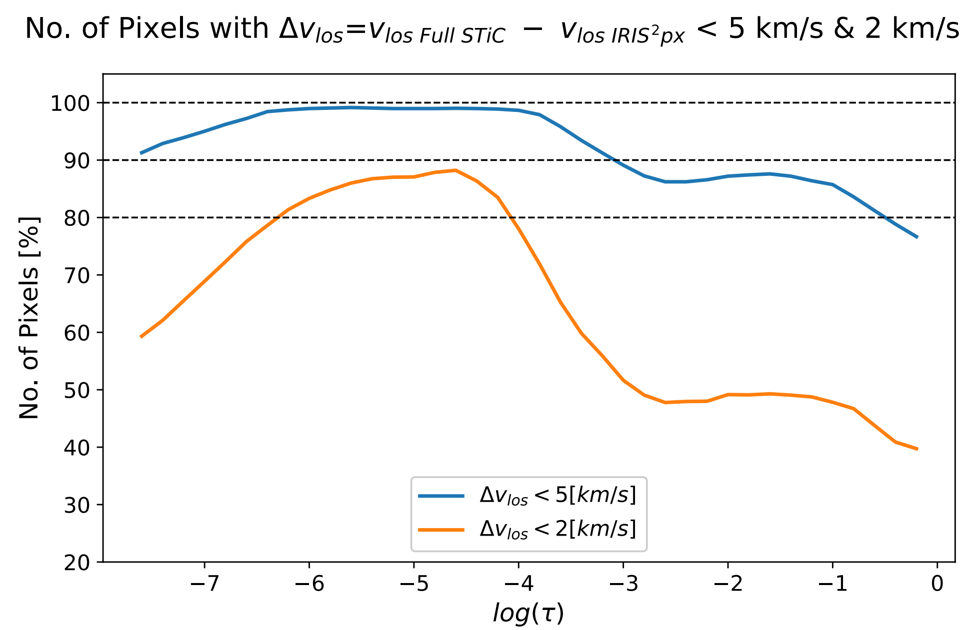
Relative Error with respect Full STiC, $\varepsilon_T = |T_{Full\ STiC} - T_{IRIS^2 px}| / T_{Full\ STiC}$

Difference in v_{LOS}, $\Delta v_{LOS} = v_{LOS\ Full\ STiC} - v_{LOS\ IRIS^2 px}$

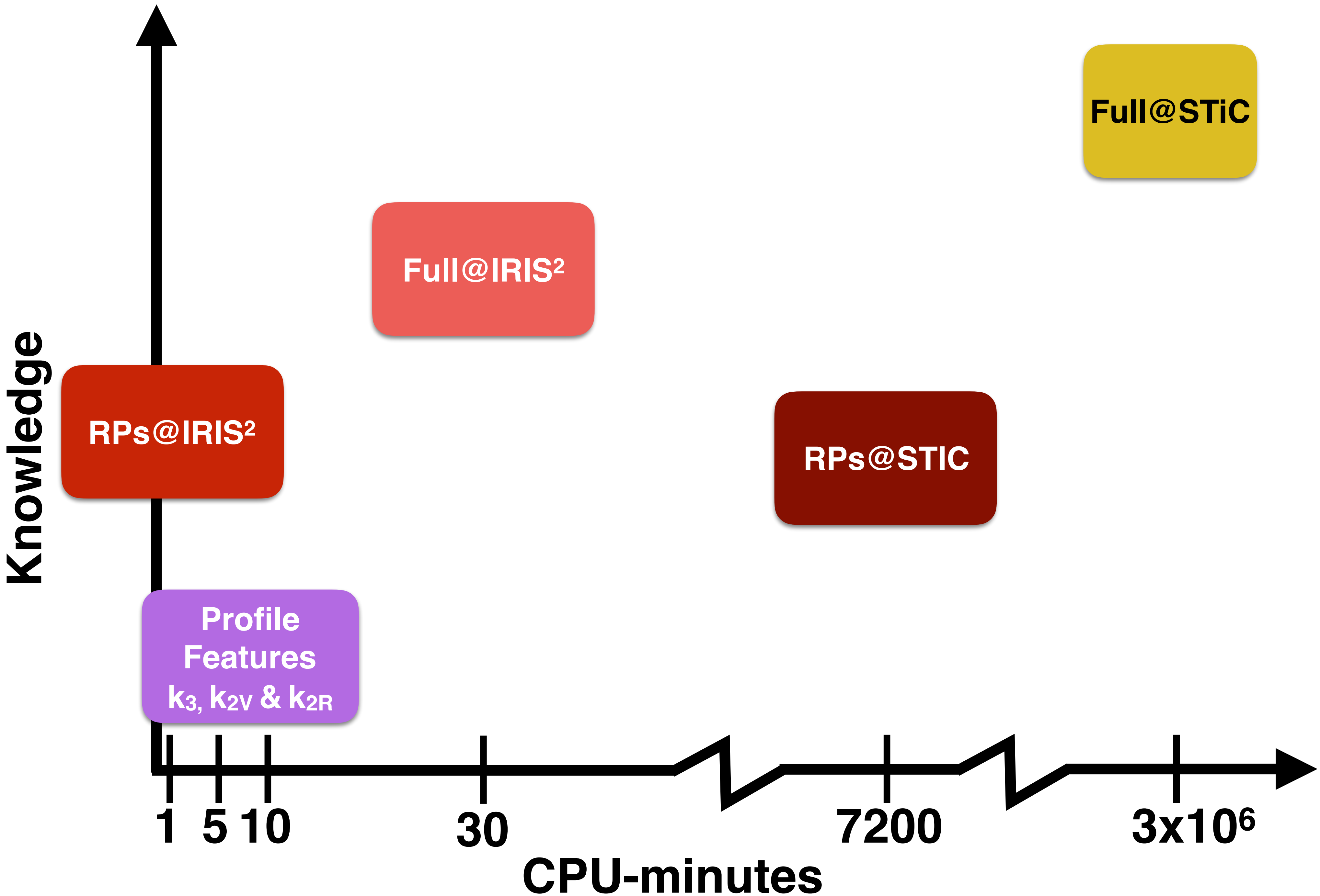
Performance in Temperature



Performance in Velocity

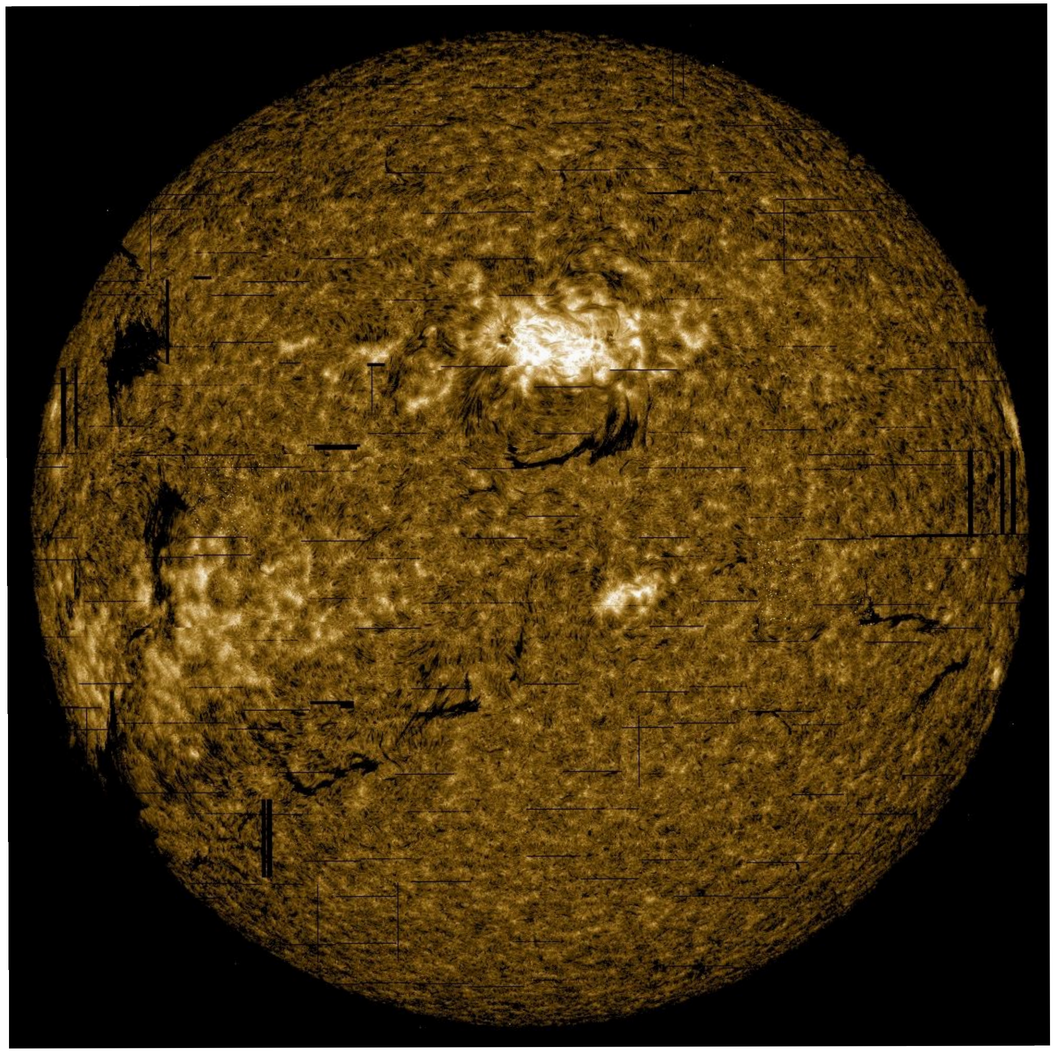


Recovering physical from IRIS Mg II h&k lines

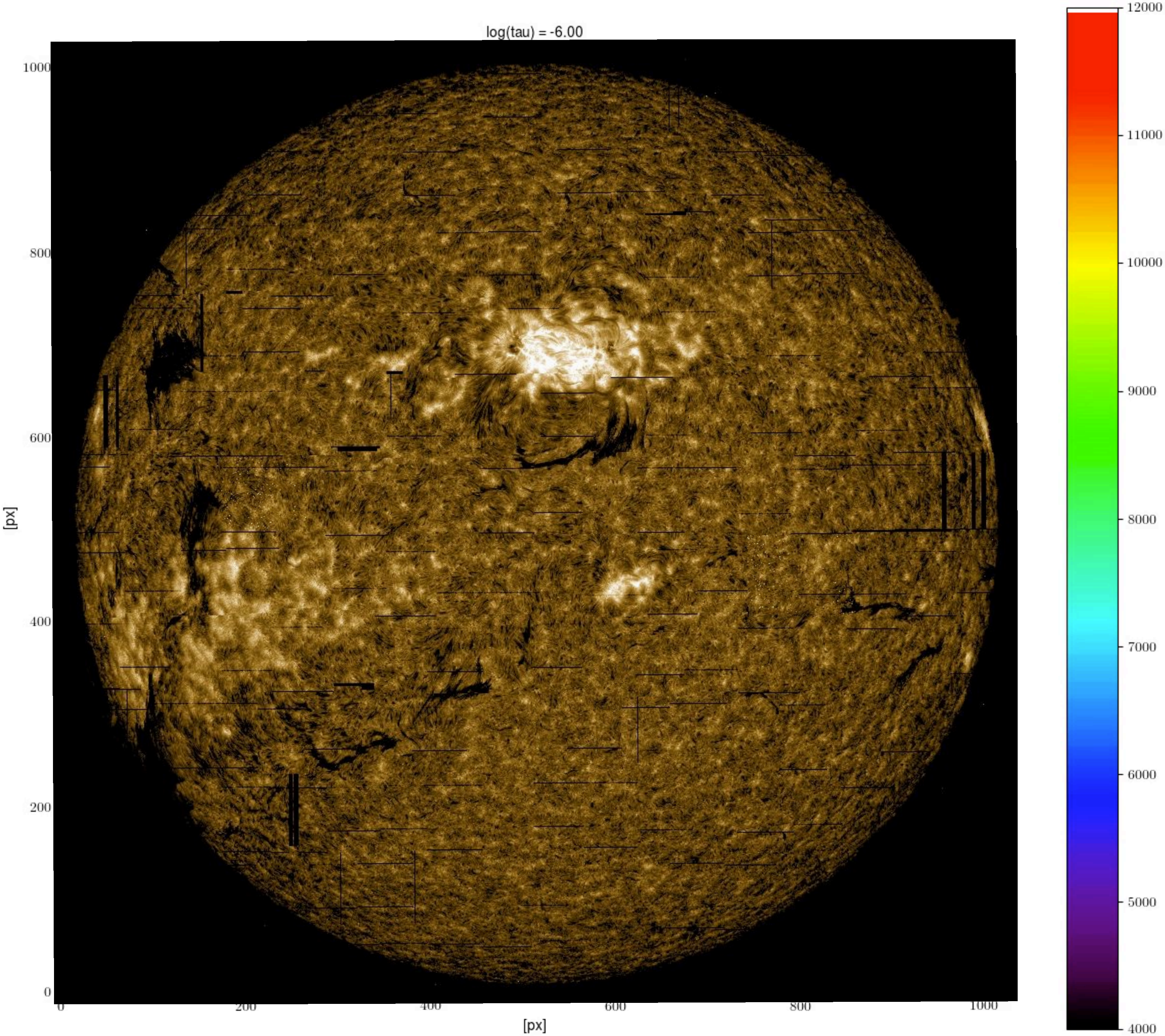


MOSAIC Mg k 2015-08-23

~4,000,000 pixels



~8,000,000 CPU-hours or…
…10 CPU-hours using IRIS^2

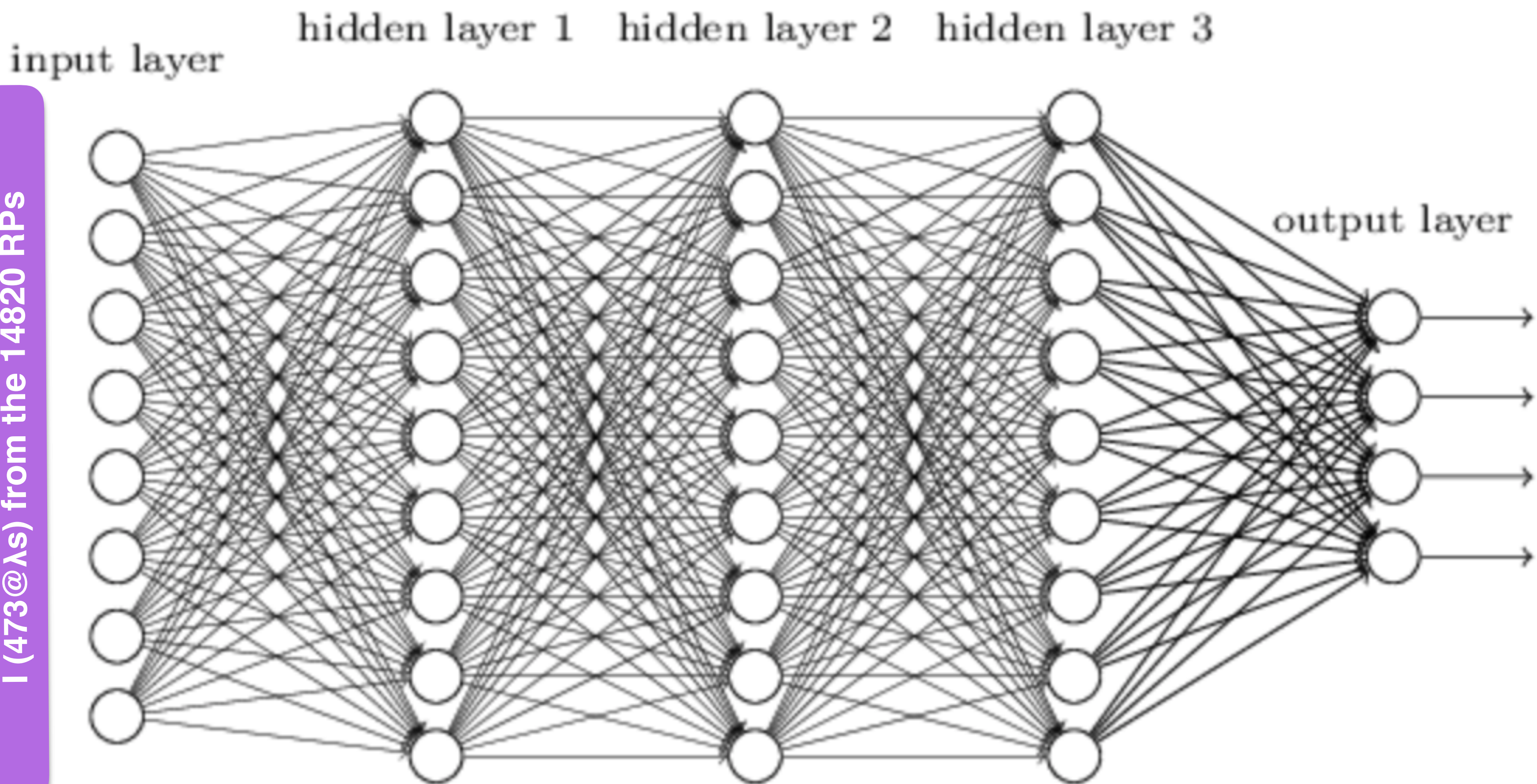


log(tau) = -6.00
[px]
[px]
0
200
400
600
800
1000
12000
11000
10000
9000
8000
7000
6000
5000
4000

Deep Learning on IRIS Inversion based on the RPs Inverted by STiC

deepIRIS²

I (473@λs) from the 14820 RPs



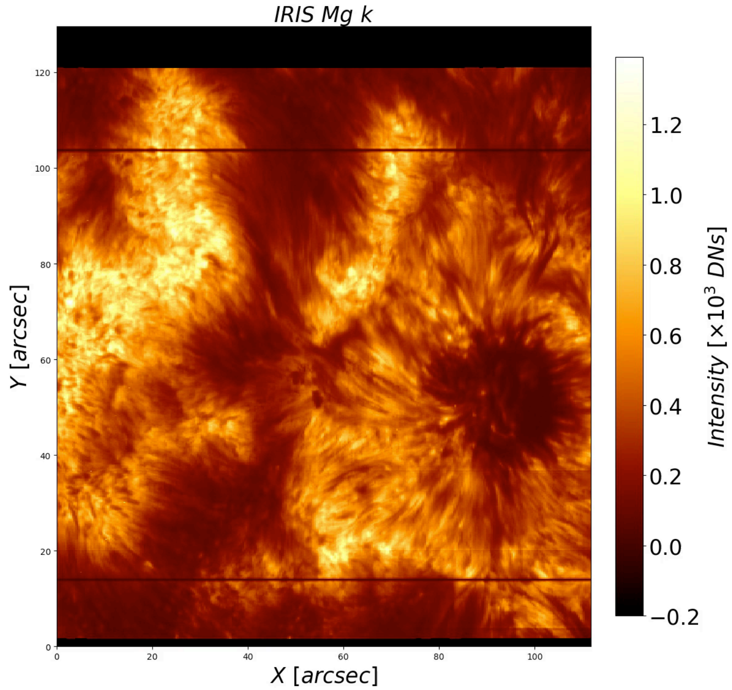
T, vlos, vturb, and ρ at 39 τ from 14820 RMAs

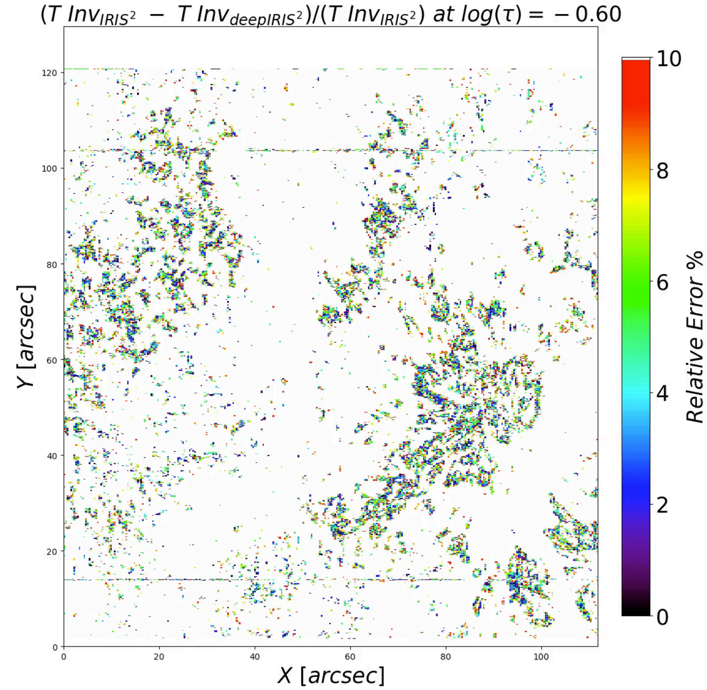
After few hours of training using 1 CPU…

full map is inverted in…

INVERSION FULL FoV "PIXEL by PIXEL" by IRIS² & deepIRIS²

IRIS
Mg II k&h
2016-01-1
23:04:09



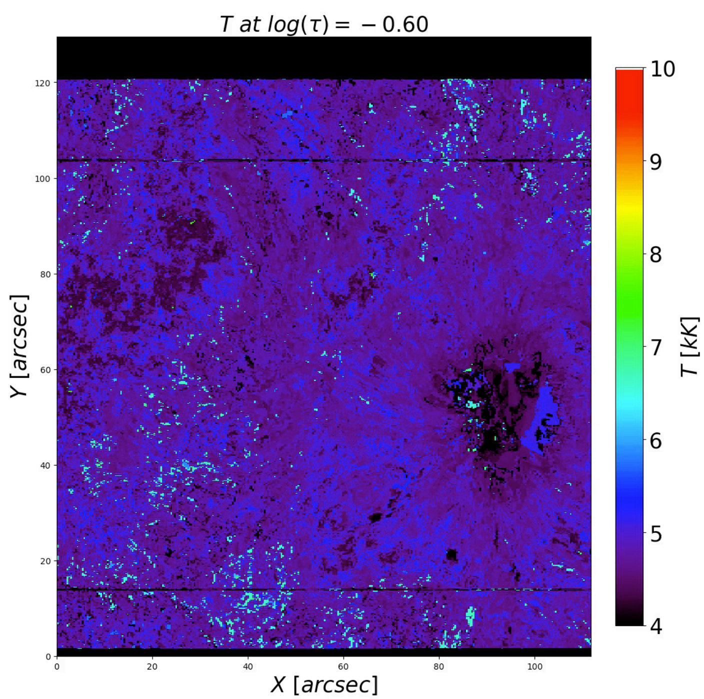


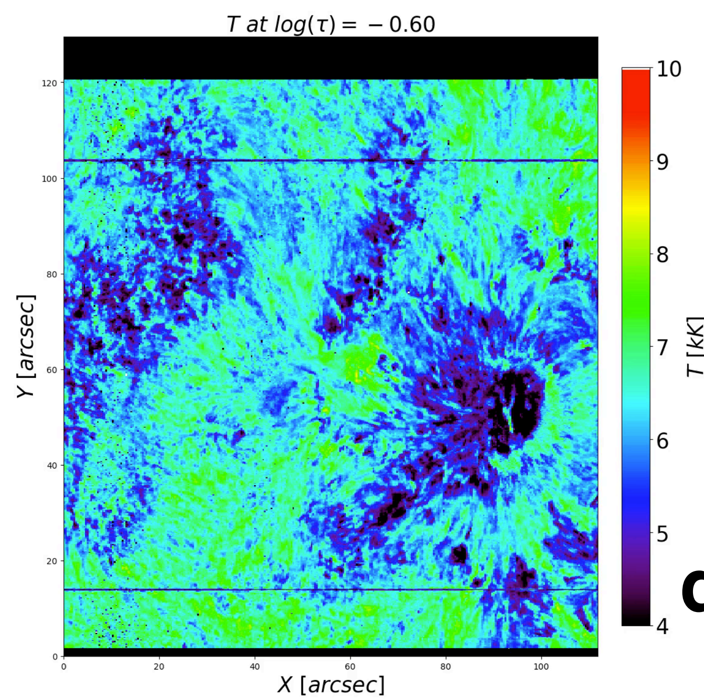
Rel.
Error
[%]
T_{IRIS}^2
vs
$deepIRIS^2$

T[kK]

from

Full
@
IRIS²





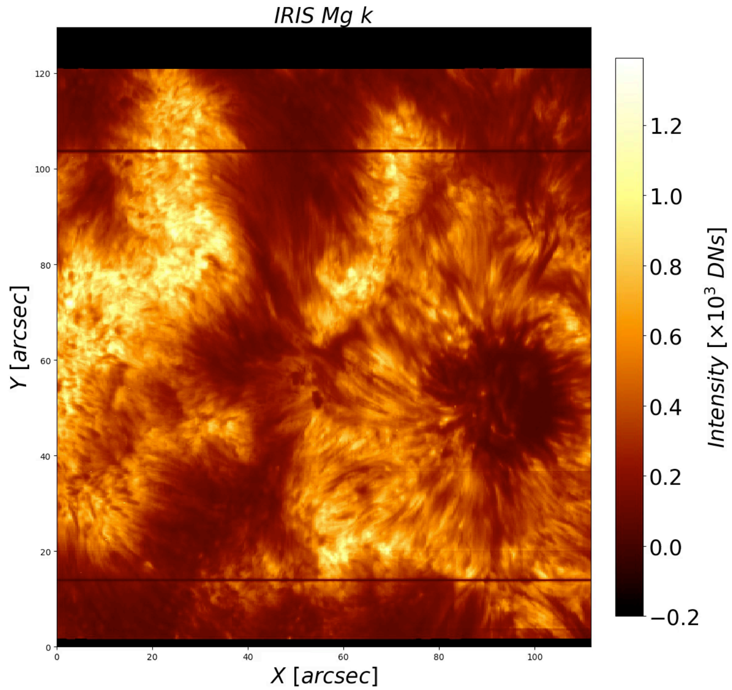
T[kK]

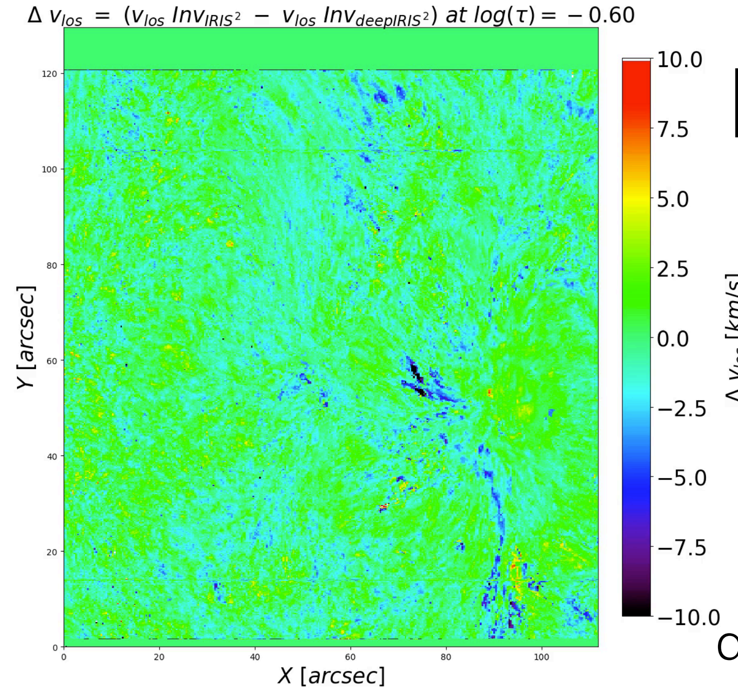
from

Full
@
deepIRIS²

INVERSION FULL FoV “PIXEL by PIXEL” by IRIS² & deepIRIS²

IRIS
Mg II k&h
2016-01-1
23:04:09





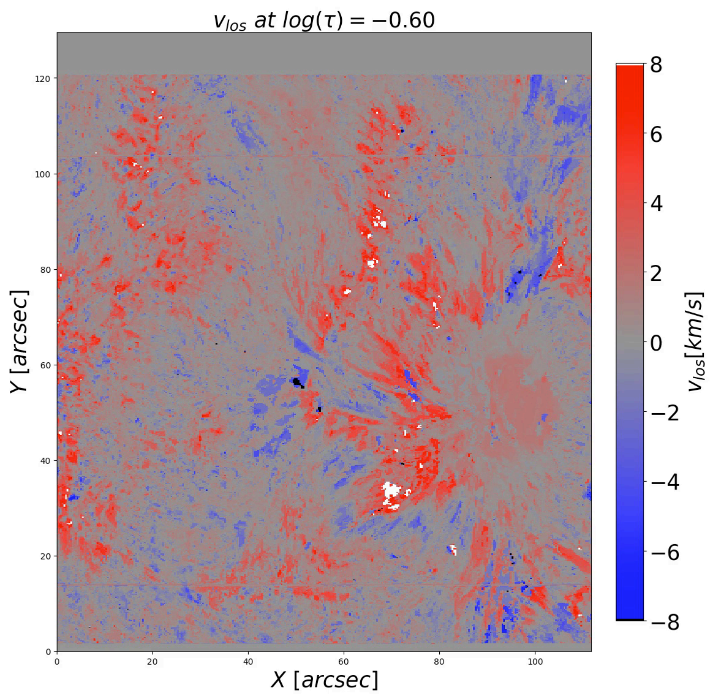
Diff. [km/s]

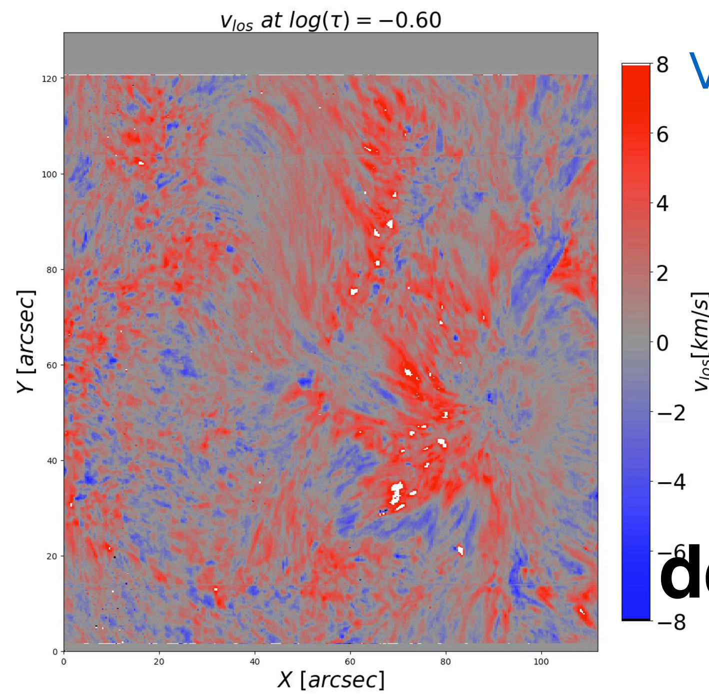
v_{los} $IRIS^2$ - v_{los} $deepIRIS^2$

v_{los}[km/s

from

Full @ IRIS²





v_{los}[km/s]

from

Full @ deepIRIS²

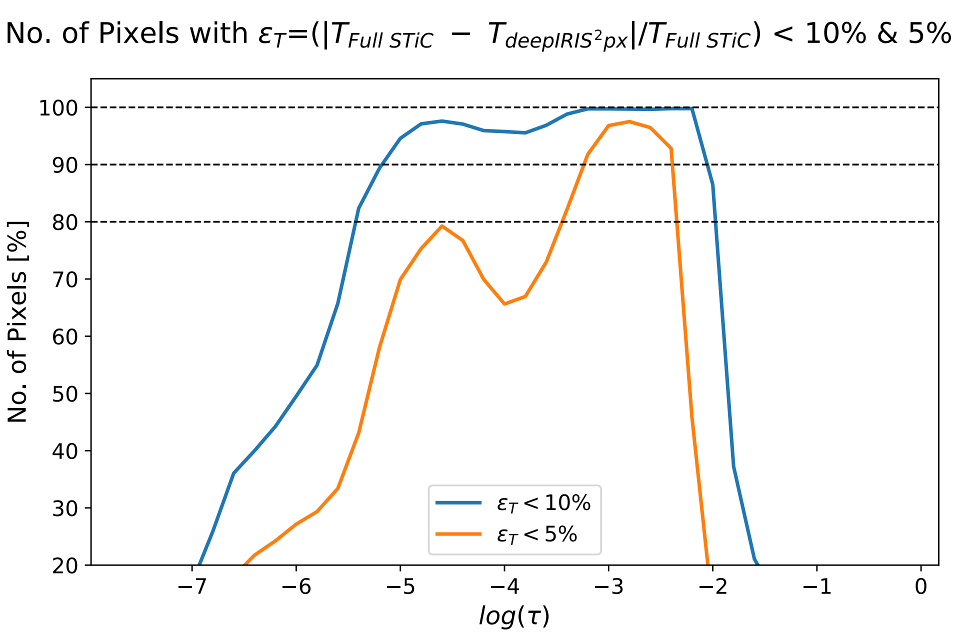
Performance of Full STiC (1px every 100px) vs deepIRIS²

2280 (original) px inverted bSTiC and IRIS²

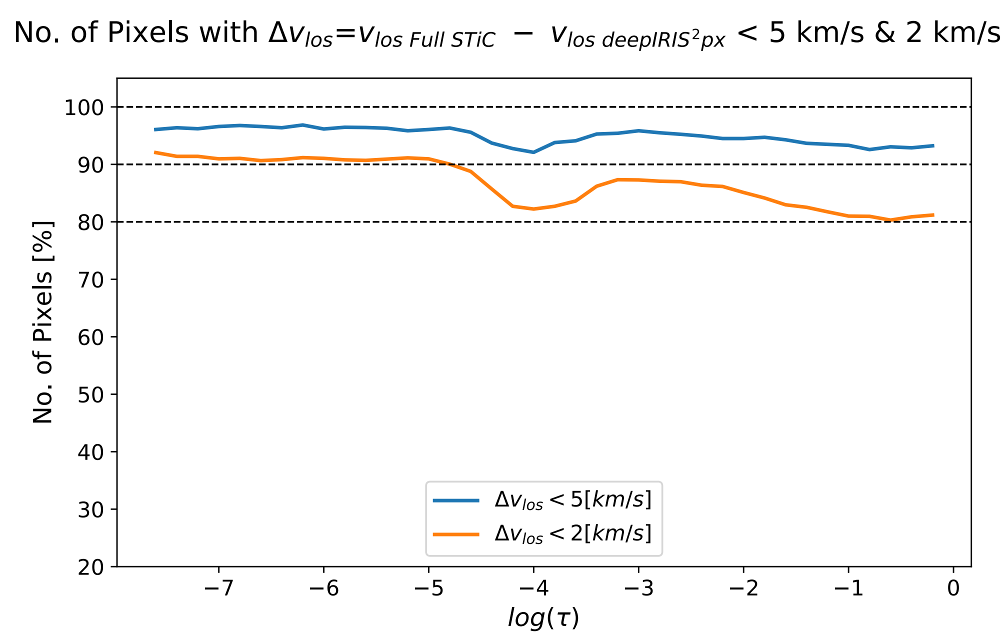
Relative Error with respect Full STiC, $\varepsilon_T = |T_{\text{Full STiC}} - T_{\text{deepIRIS2}}| / T_{\text{Full STiC}}$

Difference in v_{LOS}, $\Delta v_{LOS} = v_{LOS\ \text{Full STiC}} - v_{LOS\ \text{IRIS}^2\text{px}}$

Performance in Temperature



Performance in Velocity



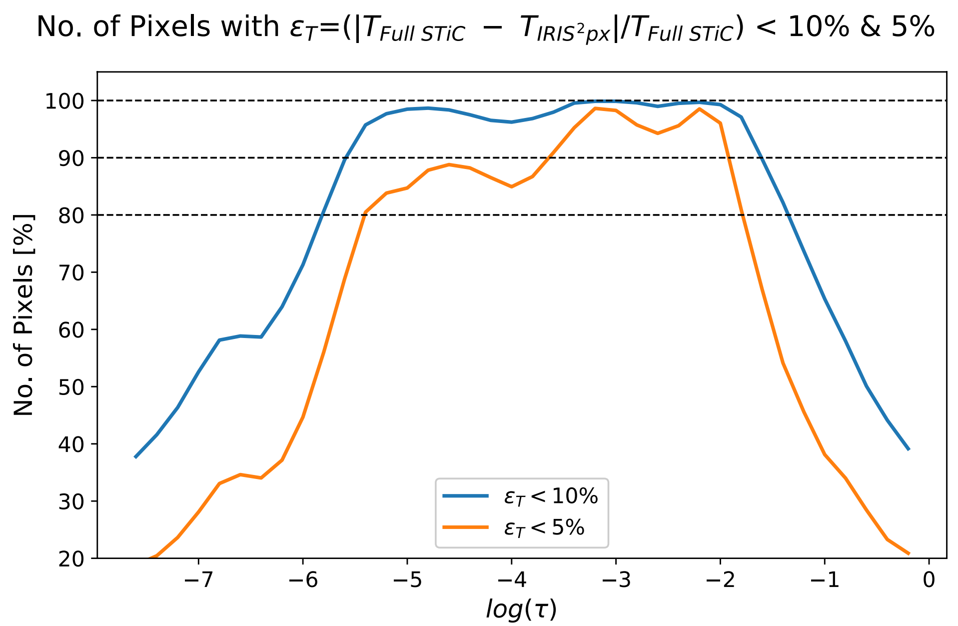
Performance of Full STiC (1px every 100px) vs IRIS²

2280 (original) px inverted by STiC and IRIS²

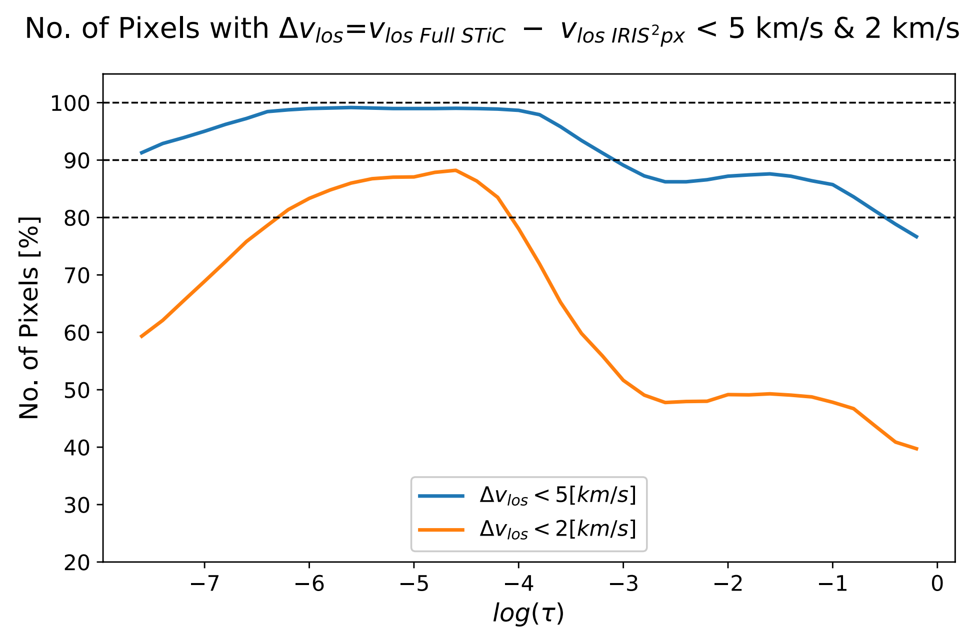
Relative Error with respect Full STiC, $\varepsilon_T = |T_{Full\ STiC} - T_{IRIS^2 px}| / T_{Full\ STiC}$

Difference in v_{LOS}, $\Delta v_{LOS} = v_{LOS\ Full\ STiC} - v_{LOS\ IRIS^2 px}$

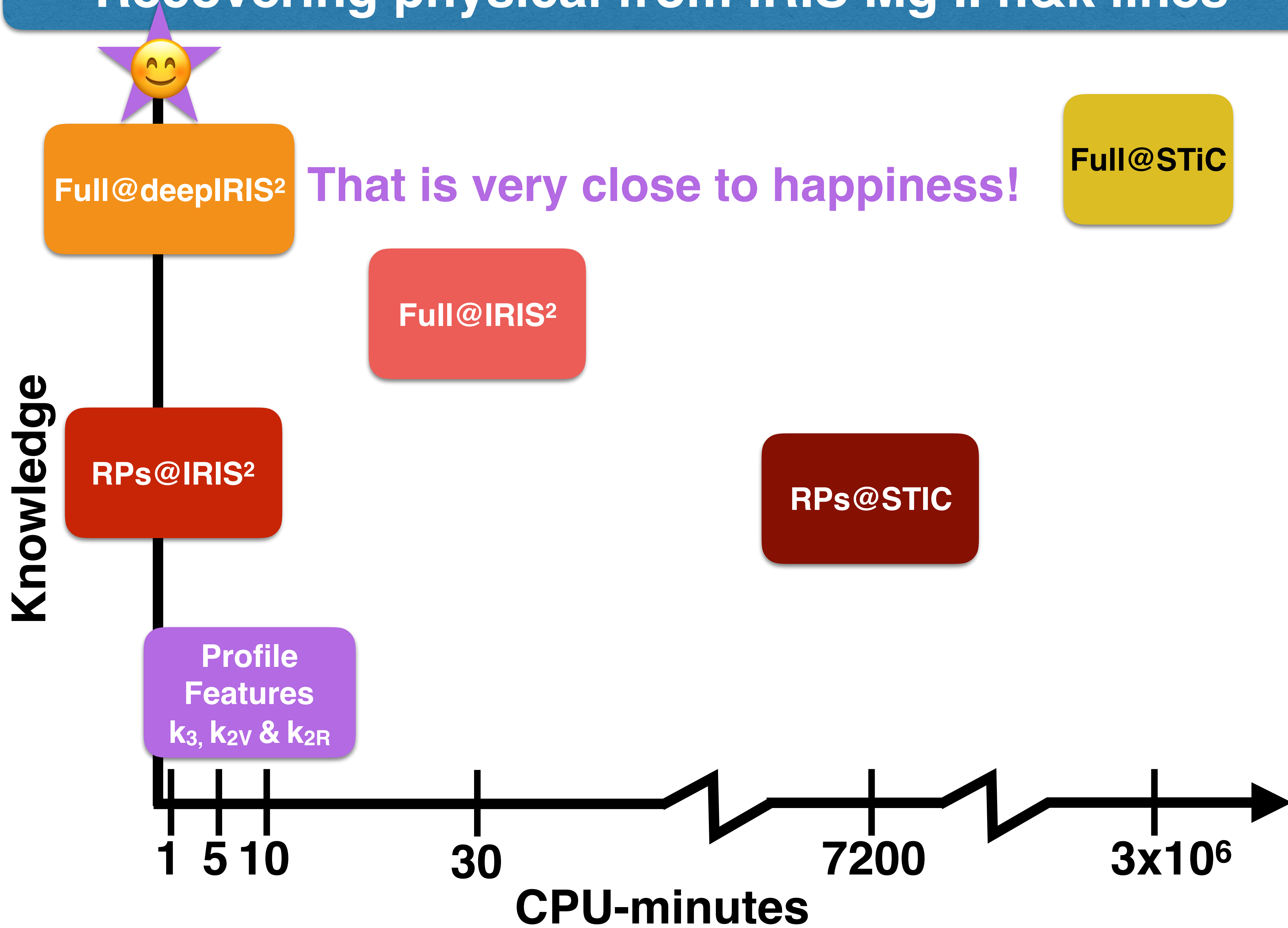
Performance in Temperature



Performance in Velocity



Recovering physical from IRIS Mg II h&k lines



Conclusions
Machine Learning and Inversions of Mg II h&k Spectra

- **IRIS²** allows to invert a man in a 10-CPU desktop machine in **3 min.**
- **deepIRIS²** allows to invert a map in a desktop machine in < **1 min.**
- **Inversion** of **ALL** IRIS Mg II h&k data sets with **deepIRIS² < 15 (+75) days**
- Currently, the bottleneck is reading (accessing a disk through the local network) and preparing (radiation calibration, cropping, interpolate (for only-Mg-II-k or only Mg-h-only) the data ~ 5min.
- **Open question: How many profiles does IRIS² need to represent the solar chromosphere?**
- **Is the IRIS² database good enough? Yes, it is. It can be improved, though.**

Thanks for your attention

asainz.solarphysics@gmail.com