The new HARPS-N pipeline

Rosario Cosentino

Global Architecture of Planetary Systems – GAPS

Progress Meeting #21

*Osservatorio Astronomico di Brera, **27-29 November 2024***

The HARPS-N DRS

1. **The DRS-32 Pipeline → 2012-now**
   - **Installed and operative at TNG**
   - **Installed and operative in Trieste (+ YABI)**
   - **Dismissed in Genève**

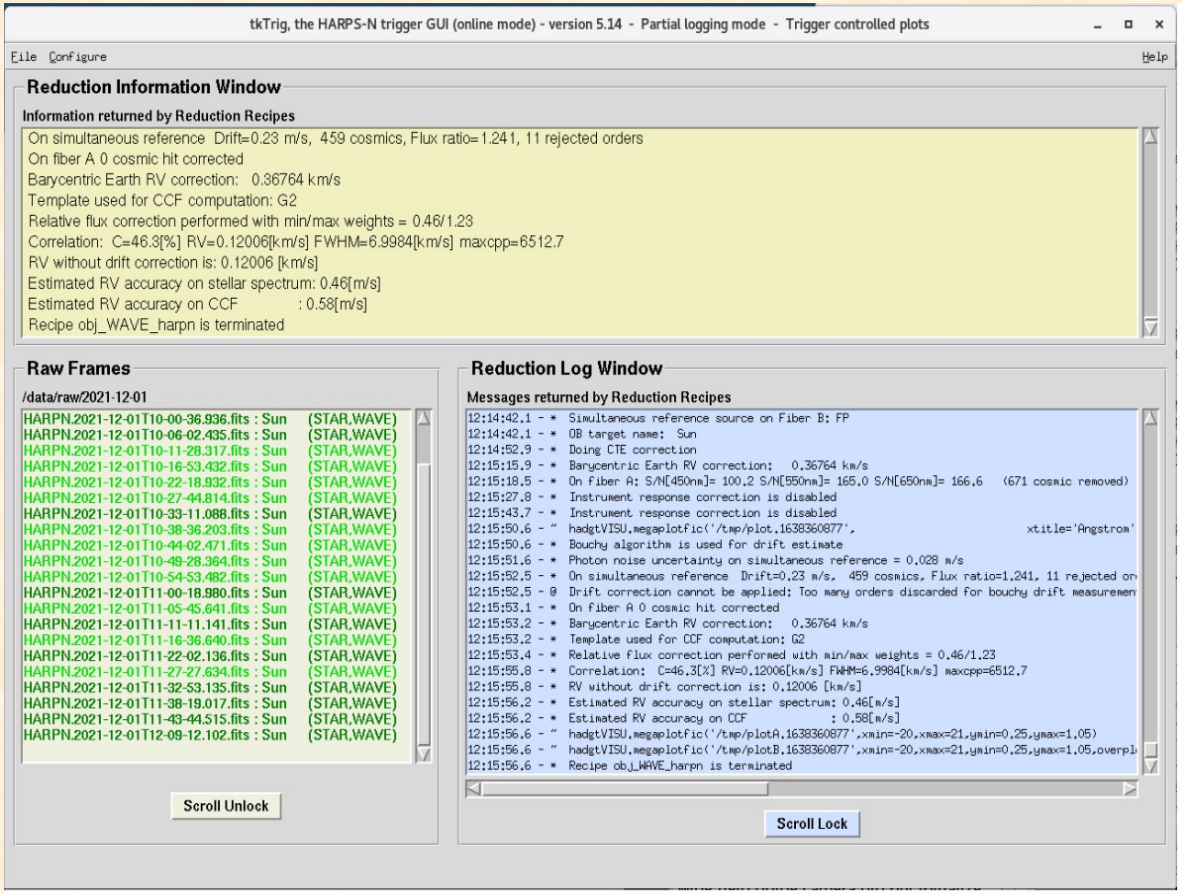
2. **The Espresso DRS → Oct 2021-now**
   - **Installed and operative at TNG**
   - **Installed and operative in Genève**
   - **Installation on going in Trieste (???)**

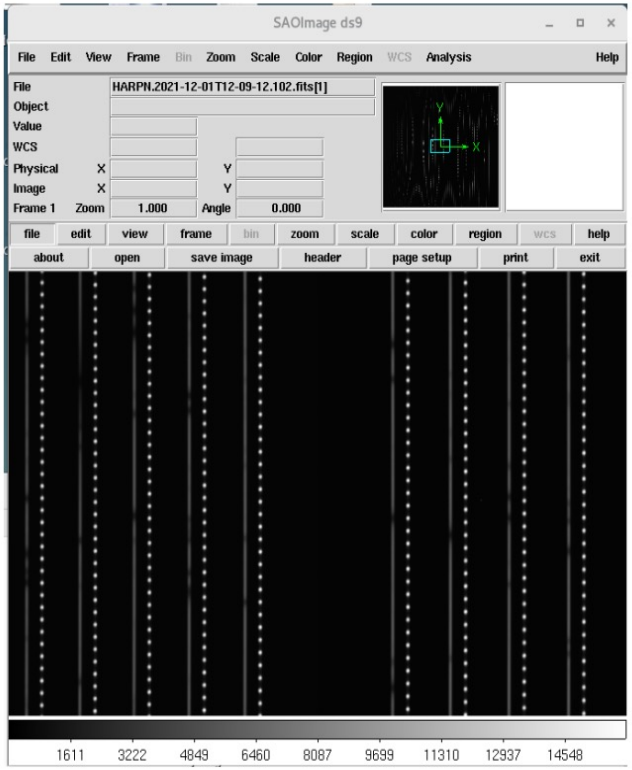
HARPS-N DRS User Manual (TNG-MAN-HARPN-0006_r1i3)
https://www.tng.iac.es/instruments/harps/data/tng-man-harpn-0006_harps-n_drs_manual_i1r3.pdf

The HARPS-N ESPRESSO DRS

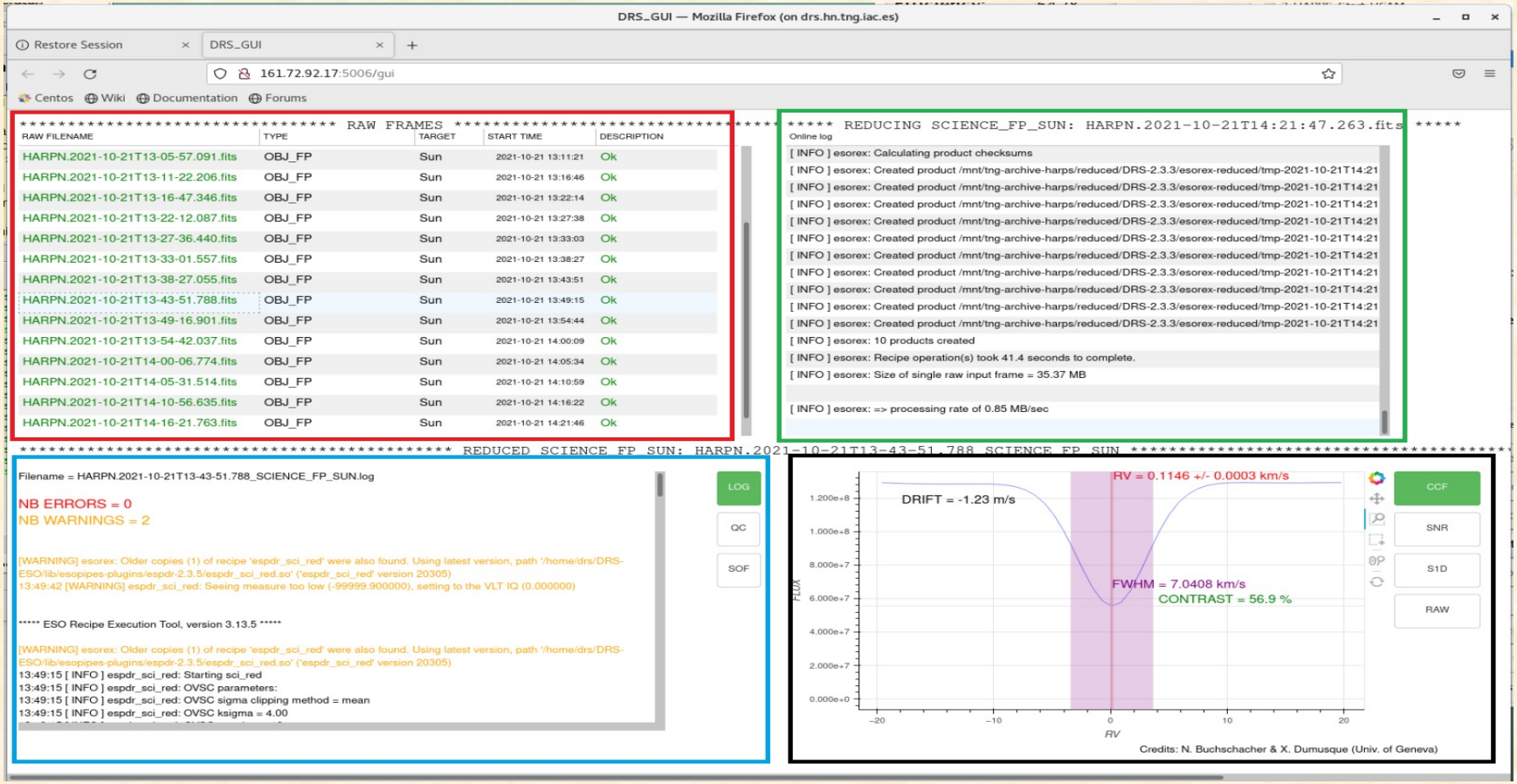
1. DRS-2.3.3 → Current version working @ TNG
2. DRS 3.0.1 → Virtual machine @ FGG (test)
3. DRS 3.2.0 → Public release
   - Installed in a Virtual machine @ FGG
   - Tested
   - Script for activity indexes extraction available
   - All data reduced
     - RV updated
     - Centroid updated
     - Test with modified mask successfully
     - Test of Activity index extraction successfully
     - Verification on-going
4. DRS 3.3.0 → Next release

The HARPS-N DRS-32 (original)

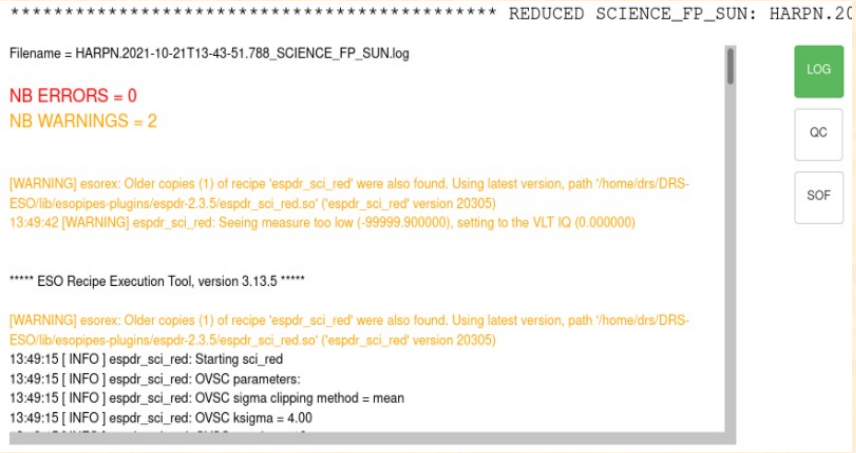


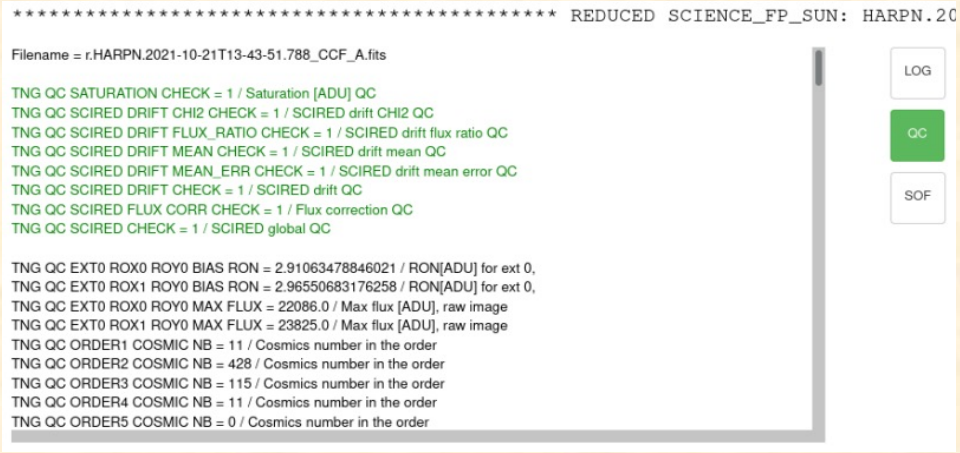


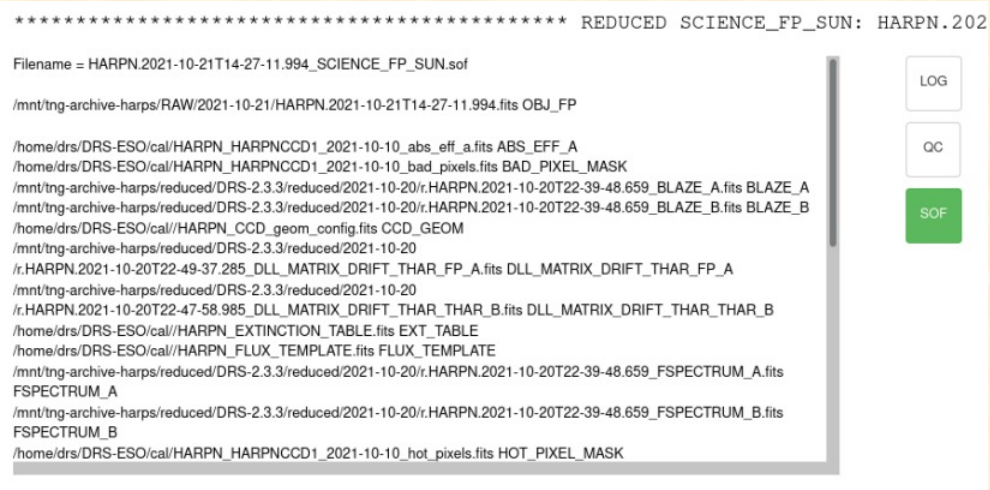
The HARPS-N Espresso DRS



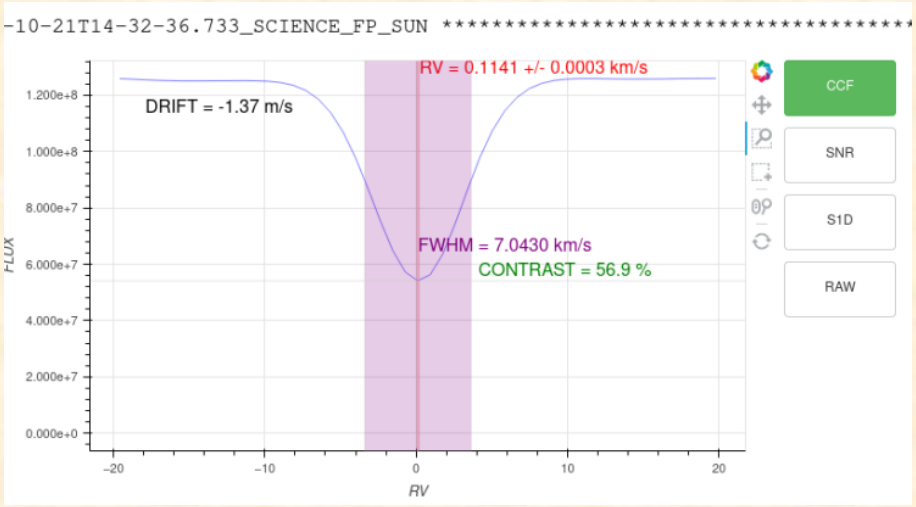
The HARPS-N Espresso DRS (left panel)

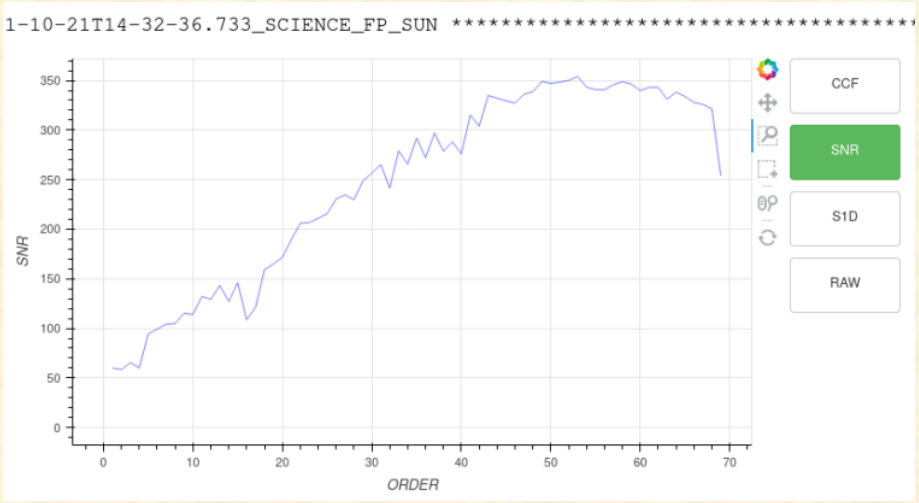


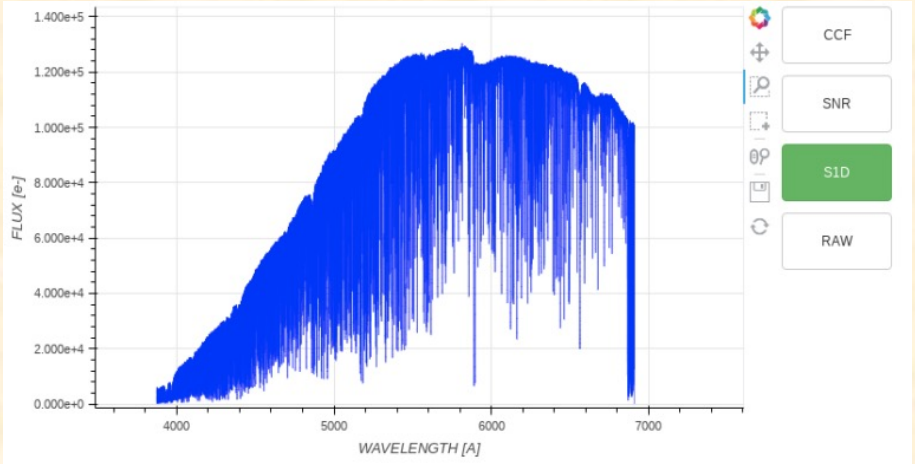


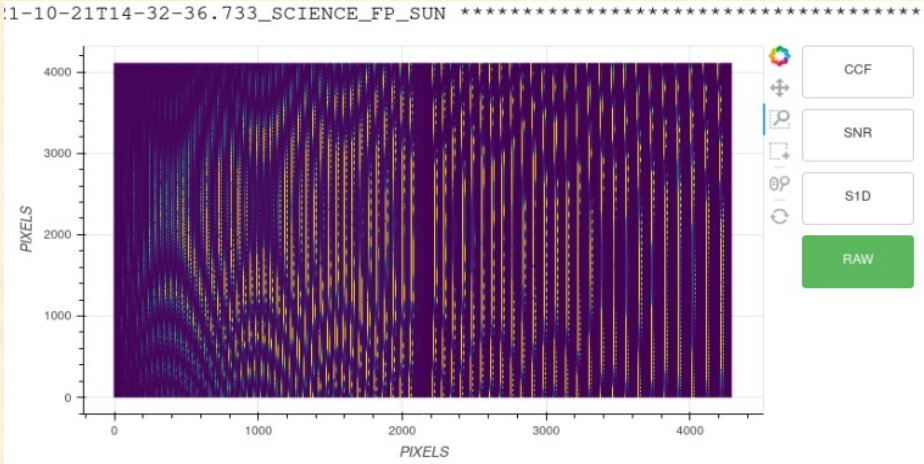


The HARPS-N Espresso DRS (right panel)









Comparison between the DRS

The transition from the original **DRS-32** to the **ESPRESSO DRS** brings improvements in efficiency and precision for Radial Velocity (RV) calculation. These changes affect both the reduced results and the input data requirements (RV, masks, fits keywords Etc.).

Improved RV Precision

The **mask set** used in the pipeline to calculate RV has been extended to include more spectral types, thus improving the precision of RV measurements.

- **Masks available** in ESPRESSO DRS: **F9**, **G2**, **G8**, **G9**, **K2**, **K6**, **M0**, **M2**, **M3**, **M4**, **M5**.
  (Note: The old DRS used only **G2**, **K0**, **K5**, **M2**, and **M4**.)

Mandatory RV Estimate for Calculation

To correctly extract the RV of a star, an **estimated RV value** (with a precision of about **10 km/s**) is now **mandatory**.

- The RV calculation requires the **estimated RV** of the star, either from the catalog (e.g., **radvel** value) or from the **TEL.TARG.RADVEL** setting in NSTS.
- The value **radvel = -99999** for unknown RV is **no longer allowed**.

Barycentric Julian Date (BJD) in FITS Headers

- In the **original DRS**, the BJD was stored in **Universal Time Coordinate (UTC)** under the keyword: **"HIERARCH TNG DRS BJD"**.
- In the **ESPRESSO DRS**, the BJD is now stored in **Barycentric Dynamical Time (TDB)** under the keyword: **"HIERARCH TNG QC BJD"**.

Calculation: Both BJD values are calculated at the **flux-weighted mid-exposure time**.

Differences in the S1D output file

- In the **original DRS**, the S1D file contains the flux vs wavelength in air.
- In the **ESPRESSO DRS** the S1D file contains two set of flux vs wavelength data (in air and in vacuum)

TARG RADVEL value VS RV Precision

Name	DEST	TARG RADVEL	Exptime	TNG QC CCF RV
HARPN.2021-12-01T01-50-22.278	TOI-4524	-10	600	-20.6574
HARPN.2021-12-01T02-00-48.293	TOI-4524	-20	600	-20.8945
HARPN.2021-12-01T02-11-14.692	TOI-4524	-30	600	-20.7946
HARPN.2021-12-01T02-21-40.707	TOI-4524	-40	600	-21.7437
HARPN.2021-12-01T01-50-22.278(1)	TOI-4524	-45	600	-66.5865
HARPN.2021-12-01T02-32-06.722	TOI-4524	-50	600	-71.5220
HARPN.2021-12-01T02-42-33.121	TOI-4524	-60	600	-71.5289
HARPN.2021-12-01T02-52-58.368	TOI-4524	-65	600	-71.5283
HARPN.2021-12-01T03-03-24.767	TOI-4524	-70	600	-71.5282
HARPN.2021-12-01T03-13-51.166	TOI-4524	-80	600	-71.5275
HARPN.2021-12-01T03-34-43.605	TOI-4524	-90	600	-71.4427
HARPN.2021-12-01T03-24-17.590	TOI-4524	-100	600	-83.7018
HARPN.2021-12-01T02-00-48.293(1)	TOI-4524	-105	600	-87.1394
HARPN.2021-12-01T03-45-09.620	TOI-4524	-110	600	-126.7585

Catalogue file entry and NSTS variables table

Catalog name	NSTS Name	Fits Header	Column Name	Value/Unit
alpha	**TEL.TARG.ALPHA**	**HIERARCH TNG TEL TARG ALPHA**	**right ascension**	**nn:nn:nn.nn**
delta	**TEL.TARG.DELTA**	**HIERARCH TNG TEL TARG DELTA**	**Declination**	**nn:nn:nn.nn**
mualpha	TEL.TARG.PMA	HIERARCH TNG TEL TARG PMA	proper motion alpha	arcsec/year
mudelta	TEL.TARG.PMD	HIERARCH TNG TEL TARG PMD	proper motion delta	arcsec/year
mv	**TEL.TARG.MAG**	**HIERARCH TNG TEL TARG MAG**	**magnitude V**	**magnitude V**
bv	Not implemented	Not implemented	Bv	Bv
TypSp	**Not used**	**Not present**	spectral type	spectral type
radvel	**TEL.TARG.RADVEL**	**HIERARCH TNG TEL TARG RADVEL**	**mean radial velocity**	**KM/sec**
snr	S/N (550nm):		Signal/Noise	Signal/Noise at 550nm
spectr	**TEL.TARG.SPTYPE**	HIERARCH TNG QC CCF MASK	spectral type for the mask	F9, G2, G8, G9, K2, K6, M0, M2, M3, M4, M5
remarks	comment		Remarks	Remarks
acquisition	TPL.NAME	HIERARCH TNG TPL NAME	Acquisition template	e.g.. HARPN_ech_acq_objA
equinox	TEL.TARG.EQUINOX	HIERARCH TNG TEL TARG EQUINOX	Equinox	Equinox
progid	**OBS.PROG.ID**	**HIERARCH TNG OBS PROG ID**	**program identificator**	**es. TAC_xx**
piname	**OBS.PI-COI.NAME**	**HIERARCH TNG OBS PI-COI NAME**	**PI name**	**char**
exptime	**DET.WIN1.UIT1**	**EXPTIME**	**Exposure time**	**seconds**
name	OBS.TARG.NAME	HIERARCH TNG OBS TARG NAME	**DEST name**	**Char**

Reduction products

Reduction product	Description
S2D_A	Echelle-order extracted spectrum blaze corrected for fibre A
S2D_B	Echelle-order extracted spectrum blaze corrected for fibre B
S2D_BLAZE_A	Echelle-order extracted spectrum Without blaze correction for fibre A
S2D_BLAZE_B	Echelle-order extracted spectrum Without blaze correction for fibre B
S1D_A	Order-merged spectrum for fibre A
S1D_B	Order-merged spectrum for fibre B
S1D_FLUXCAL_A	S1D flux calibrated spectrum for fibre A
CCF_A	CCF for fibre A
CCF_B	CCF for fibre B
CCF_RESIDUALS_A	residuals from CCF computation
S2D_SKYSUB_A	S2D spectrum sky-subtracted for fibre A
S1D_SKYSUB_A	S1D spectrum sky-subtracted for fibre A
CCF_SKYSUB_A	CCF sky-substracted for fiber A
S1D_SKYSUB_FLUXCAL_A	S1D spectrum sky-subtracted flux calibrated for fibre A

The reduced files format is:
r.HARPN.YYYY-MM-DDThh-mm-ss.sss_ReductionProduct_FIBER.fits
where:

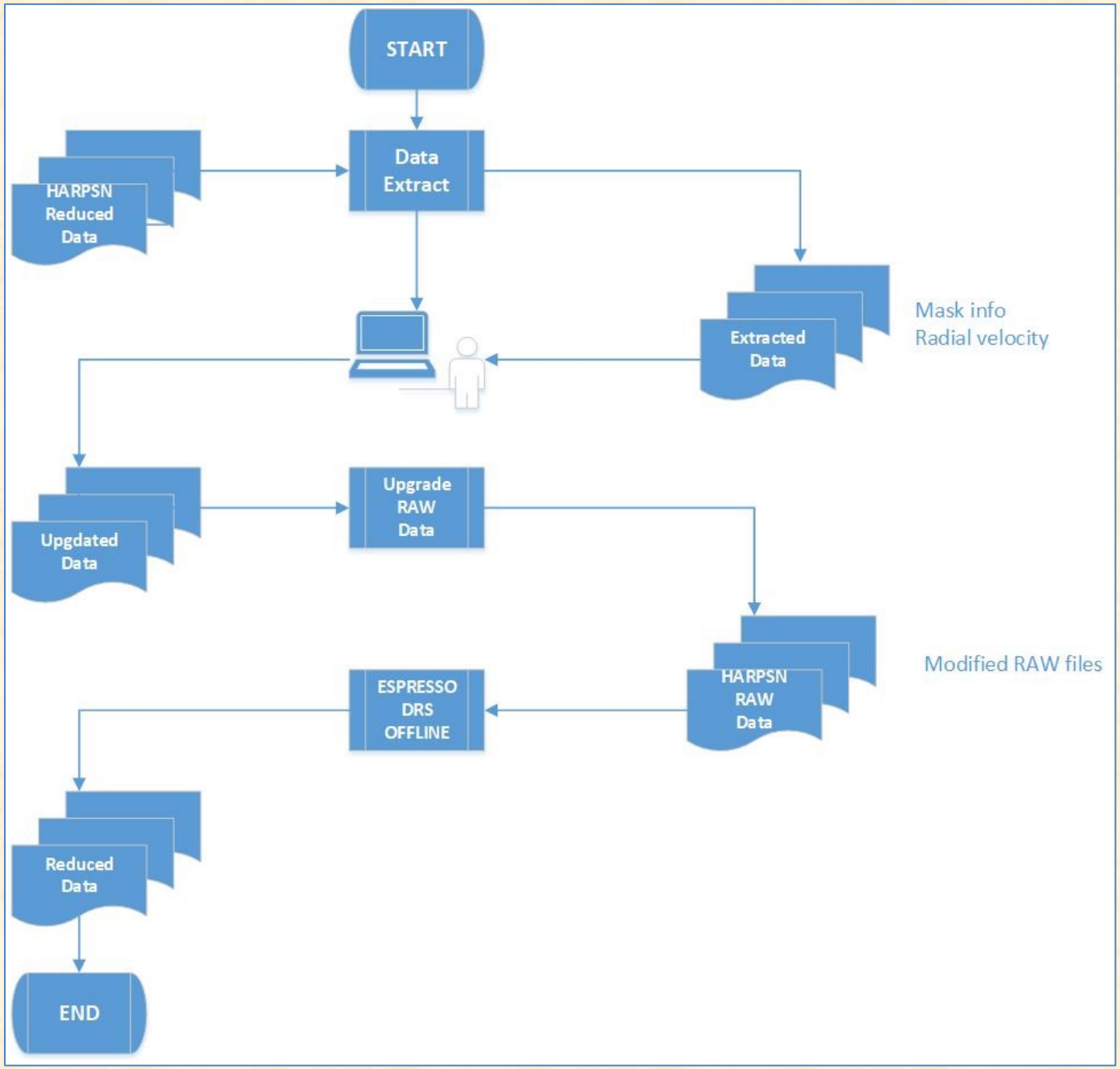
- ✓ yyyy-mm-dd = date of the observation (year, month, day)
- ✓ hh-mm-ss.sss = hour of the observation (hour, minuter, seconds)
- ✓ ReductionProduct is the name of the product (Table 4)
- ✓ Fiber can be A or B

Processing the original HARPSN data with the ESPRESSO DRS

In the ESPRESSO DRS, to correctly extract the **Radial Velocity (RV)** of a star, it is **mandatory** to provide an **estimated RV value** with a precision of at least **10 km/s**. The use of **radvel = -99999** (indicating an unknown RV) is **no longer allowed**.

1. Extracts the information of the processed file that contains a not usable value of "TARG RADVAL" keyword. (this step was completed and the output file is available)
2. Update the information of the output file
    - RV update (automatic)
    - Centroid correction (automatic)
    - Mask update (no correction or on-request correction)
3. Generate the file for the keyword update script
4. Upgrade the keywords in the RAW files and generate the list of file to be reduced
5. Process the data with the Espresso DRS (all data has been reduced)
6. Extraction of data from the processed data (DSS-32 and Espresso DRS) → done on a data sample
7. Comparison of the results → on-going on a data sample

Processing HARPS data flow chart



First step: Extract the information from the archive

This operation requires the data reduced by the original DRS and the rawextract_debug.py script.

The "**rawextract_debug**" script extract the needed information from the archive files that contain the wrong value of the HIERARCH TNG TEL TARG RADVEL keyword (-99999.0).

Field	Keyword	Use	Note
Datadir	NO	To locate the date	Date directory
Name	NO	To identify the file	File name
tname	HIERARCH TNG OBS TARG NAME	Info	DEST name
rvel	HIERARCH TNG TEL TARG RADVEL	To be updated	Catalog DEST RV (-99999.0)
rv	HIERARCH TNG DRS CCF RV	info	Baryc RV (no drift correction)
spytype	HIERARCH TNG TEL TARG SPTYPE	To be updated	Catalog DEST mask
ccfmask	HIERARCH TNG DRS CCF MASK	Info	DRS used mask
program	PROGRAM	info	Program name
piname	HIERARCH TNG OBS PI-COI NAME	info	PI name

Example of extracted information

Datadir	Name	DEST	Catalog RV	DRS RV (calculated)	spytype	ccfmask	program	piname
2018-01-02	HARPN.2018-01-02T19-14-14.947	HR10	-99999	-247.2343897	A2	G2	CAT17B_78	Montesinos
2018-01-02	HARPN.2018-01-02T19-28-00.452	HR10	-99999	-267.5990406	A2	G2	CAT17B_78	Montesinos
2018-01-02	HARPN.2018-01-02T20-12-32.344	EP246214735	-99999	-62.73625262	G5	G2	GTO	Pepe
2018-01-03	HARPN.2018-01-03T19-20-18.105	HIP103371	-99999	-305.2593605	O6	G2	GAPS	Sozzetti
2018-01-03	HARPN.2018-01-03T19-37-12.688	HD3765	-99999	-63.07340591	K2	K5	GAPS	Sozzetti
2018-01-03	HARPN.2018-01-03T20-03-23.468	EP246214735	-99999	-62.74248434	G5	G2	GTO	Pepe
2018-01-03	HARPN.2018-01-03T20-22-04.824	HD4628	-99999	-10.08185511	K2	K5	GTO	Pepe
2018-01-03	HARPN.2018-01-03T20-48-37.059	HD3651	-99999	-32.81713954	K0	K5	GTO	Pepe
2018-01-03	HARPN.2018-01-03T20-56-42.177	HD3651	-99999	-32.81732172	K0	K5	GTO	Pepe
2018-01-03	HARPN.2018-01-03T21-42-26.771	HD10476	-99999	-33.50605285	K1	K5	GTO	Pepe
2018-01-03	HARPN.2018-01-03T21-48-01.424	HD10476	-99999	-33.50633274	K1	K5	GTO	Pepe
2018-01-03	HARPN.2018-01-03T21-53-35.813	HD10476	-99999	-33.50664152	K1	K5	GTO	Pepe
2018-01-03	HARPN.2018-01-03T23-43-26.695	HD32147	-99999	21.7054167	K3	K5	GTO	Pepe

Second step: Update the information of the output file

The responsible of the program or the PI of the observations must upgrade the RVEL and SPYTYPE values.

By default, shall be used the **rv** value for the **rvel** keyword and the **ccfmask** for the **spytype** keyword (which are the DRS calculated/set values).

Third step: generate the file for the keyword update

After the step two has to be generated the file with the information needed for *the upgrade of the fits keywords in the raw files.*

*The file has t*o contain the following fields, each field separated by TAB:

Field	Keyword	Use	Note
Date	No	Used by the script	
filename	No	Used by the script	
Object name	No	info	
RV	HIERARCH TNG TEL TARG RADVEL	Used by the script	New value
Mask	HIERARCH TNG TEL TARG SPTYPE	Used by the script	New value

2021-11-30 HARPN.2021-12-01T02-52-58.368 TOI-4524 -71.53158797 G2
[…]
2021-11-30 HARPN.2021-12-01T03-58-55.342 TOI-0480 5.207028658 G2

Fourth step: Upgrade the keywords in the RAW files

This operation requires the raw data, the text file with the upgraded RV and mask information, and the **upg_keyword.py** script.

The **upg_keyword.py** script reads the text file, updates the FITS keywords, and generates a list of the modified raw files with their complete paths.

Output:

Listfile.txt - A file containing the full paths of the modified raw files. This file can be used for the offline reduction of the modified files using the **ESPRESSO DRS**.

Fifth step: Processing of the modified raw-files with the Espresso DRS

For the processing of the modified raw files is needed the list of the file to be reduced (listfile.txt) and the Espresso DRS (see attached file).

Use of the script (in the Espresso DRS machine):

1. Move in "/home/drs/Trigger" directory
2. Execute the script: python Trigger.py -i HARPN –f [list_of_frames]

Activity index extraction

$ **python Indexes_extract.py -h**
Usage: index_extract version 1.0

Options:

-h, --help	show this help message and exit
-f S1D, --file=S1D	s1d filename with the path
-r RV, --RV=RV	Radial velocity in meter por seconds
-b BV, --BV=BV	color index b-v
-d DEST, --DEST=DEST	destination files DEST_(Ca, Ha, Na, Rhk, Smw).txt

Example:
python Indexes_extract.py -f /mnt/tng-archive-harps/tmp/gaps/AMS17/r.HARPN.2024-09-23T19-50-26.782_S1D_A.fits -r -6.362 -b 0.677 -d ASM17

Output files:
DEST_Ha.txt
DEST_Na.txt
DEST_Rhk.txt
DEST_Smw.txt
DEST_Ca.txt

The script uses the original: SpectralActivityIndexesS1d.py

Data sample re-reduction

Obj(GAPS)	Obj	Simbad	RV	Spectral Type	Mask-DRS32	Mask DRS 3.2.0	B-V
AMS01	TOI-2411	G 271-53	-13.57	K5V	K5	K6	0.84
AMS17	TOI-1860	HD 134319	-6.362	G5V	G2	G2	0.677
ASM18	K2-233	BD-19 4086	-10.248	K3	K5	K2	0.939
AMS20	TOI-6078	BD+16 2904	-45.677	K0	K5	G9	0.854

1. Mask Upgrade in the RAW files
2. Re-reduction of the data set with the DRS 3.2.0
3. Activity indexes extraction
4. Data analysis

DRS 32 masks: G2, K0, K5, M2, M4
DRS 3.2.0 masks: F9, G2, G8, G9, K2, K6, M0, M2, M3, M4, M5

AMS20 - BD+16 2904

Spectral type: K0
Vmag: 10
Low activity: logR'hk = - 4.9
RV: -45.677 Km/sec

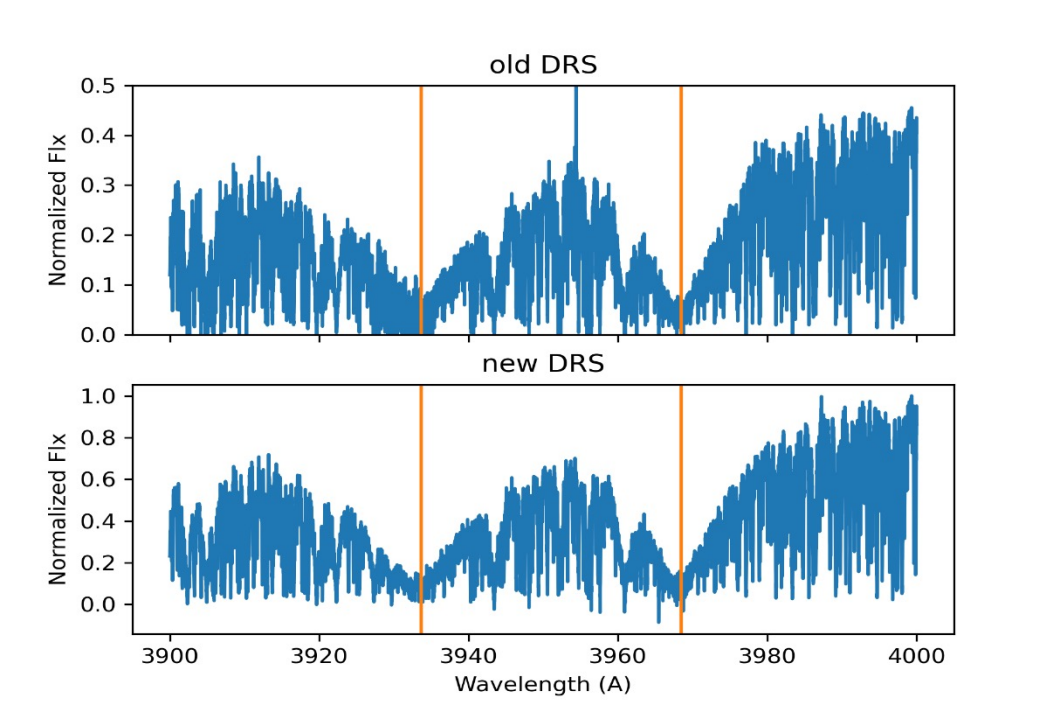
ESPRESSO DRS Comparison

RV rms (v2.3.3) = 3.3481027 m/s
RV rms (v3.2.0) = 3.5085695 m/s
DIFF rms = 0.16046680 m/s

RV err (v2.3.3) = 1.1609062
RV err (v3.2.0) = 1.1627056
DIFF err = 0.0017994694 m/s

Espresso DRS vs DRS32

No difference in activities index
No difference in H, K and Ca lines
Clearest periodogram with Espresso DRS



The improved quality of the periodograms derived from the data of the Espresso DRS enables a more precise discrimination of planetary signals

Thanks to Serena Benatti and Claudia Di Maio

CONCLUSION

Status of the re-processing

1. Extraction of the list of the file to be re-processed → done
2. Upgrade of RV and mask for a data sample → Done
3. Upgrade the keywords in the RAW fits file (all data) → done
4. Re-processing of the data with the Espresso DRS 3.2.0 → done
5. Re-processing with optimized mask (data sample)
6. Extraction of activity indexes (data sample)
7. Comparison with the DRS-32 outputs (data sample)

Next steps

1. The reduced data are available for Trieste archive
2. The script for the Activity indexes extraction can be installed (TBD)
3. Installation of the DRS 3.2.0 at Trieste archive (TBD)
4. Development of a YABI-like platform (TBD)