Blindtests

1. Internal SHINE-BTs reports
   https://www.overleaf.com/project/6279671b884a06cf1d61cc9d

2. Data Challenge - Detection
   Cantalloube et al. (2021)
   https://arxiv.org/pdf/2101.05080.pdf

3. Data Challenge - Characterization
   Cantalloube et al. (2022)
   https://arxiv.org/pdf/2209.08120.pdf

COBREX

1. Internal SHINE BTs reports

Motivation:

- Not happy with the statistical robustness of the algorithms we were using for SHINE (cADI, TLOCI). Important number of False Positives (high false detection rates) mainly related to a non correct determination of our detection limits and SNR maps/thresholds.

- Alternative approaches: STIMmap, RSM, ANDROMEDA, PACO, TRAP…

- Important work/testing done for PACO for the detection robustness using SHINE observations. Decision to create a blindtests more focused on the characterization part.

- Data Challenge - Detection also organized at about the same time.

1. Internal SHINE BTs reports

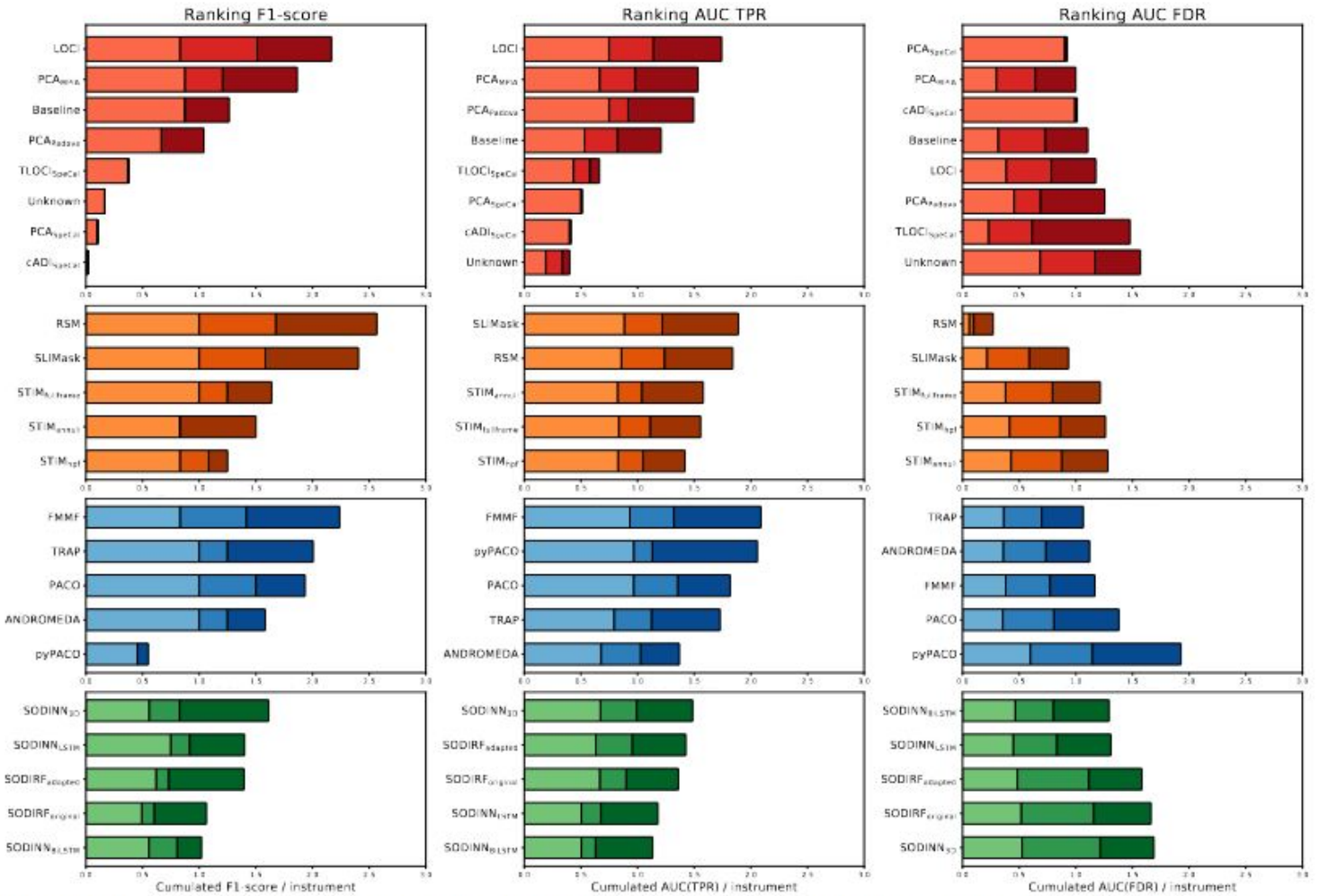


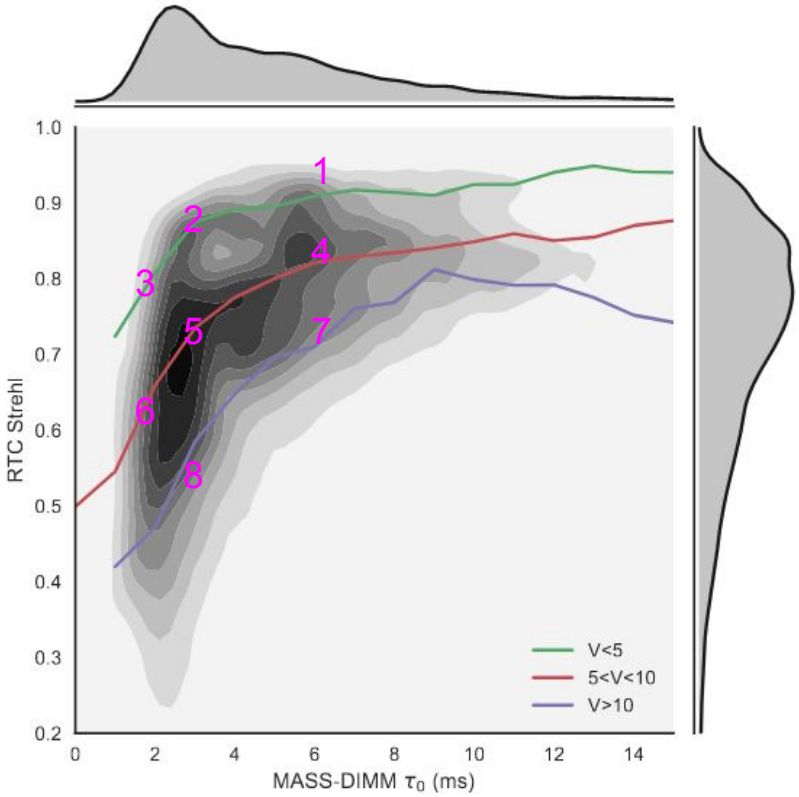
Figure 5. Ranking of the different families of submitted algorithms for the EIDC ADI subchallenge. Left column: Ranking based on the F1-score. Middle column: Ranking based on the AUC of the TPR. Right column: Ranking based on the AUC of the FDR. From top to bottom, the 5 different families of algorithms: classical speckle subtraction providing residual maps (red), advanced speckle subtraction building detection maps (orange), inverse problems (blue) and supervised machine learning (green). The light, medium and dark colors correspond to the three VLT/SPHERE-IRDIS, Keck/NIRC2, and LBT/LMIRCam data sets respectively.

1. Internal SHINE BTs reports

The scene: dataset selection

- In the line of Data Challenge: Detection
  But designed for characterization, and checking spectral fidelity during extraction (Cantalloube et al. 2021, SPIE)

- 16 datasets (IRDIS + IFS) for 8 targets covering different conditions of:
  - Brightness
  - Atmospheric conditions
    (Milli et al. 2017, SPIE)



Team: P. Delorme, M. Langlois, M. Janson, R. Galiche
M. Samland, F. Cantalloube, G. Chauvin

Deadline: Oct 29th, 2021

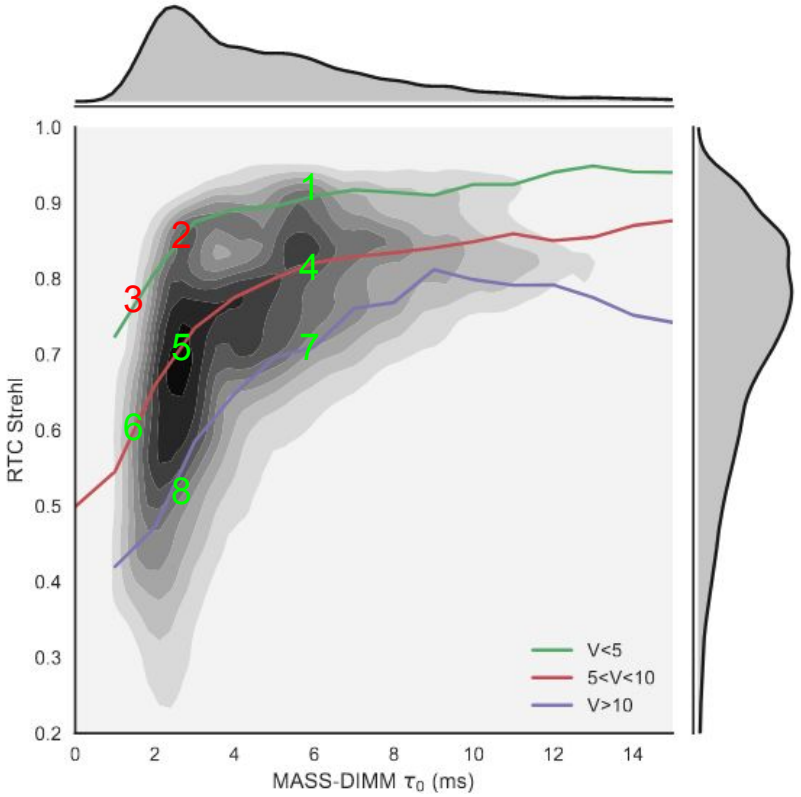
1. Internal SHINE BTs reports

The scene: dataset selection

- In the line of Data Challenge: Detection
  But designed for characterization, and checking spectral fidelity during extractio (Cantalloube et al. 2021, SPIE)

- 16 datasets (IRDIS + IFS) for 8 targets covering different conditions of:
  - Brightness
  - Atmospheric conditions

    (Milli et al. 2017, SPIE)



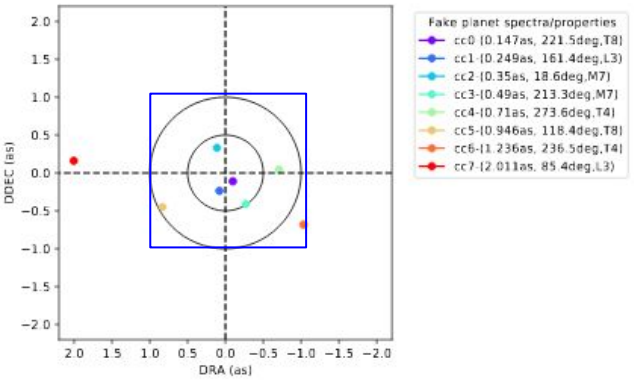
Results received:

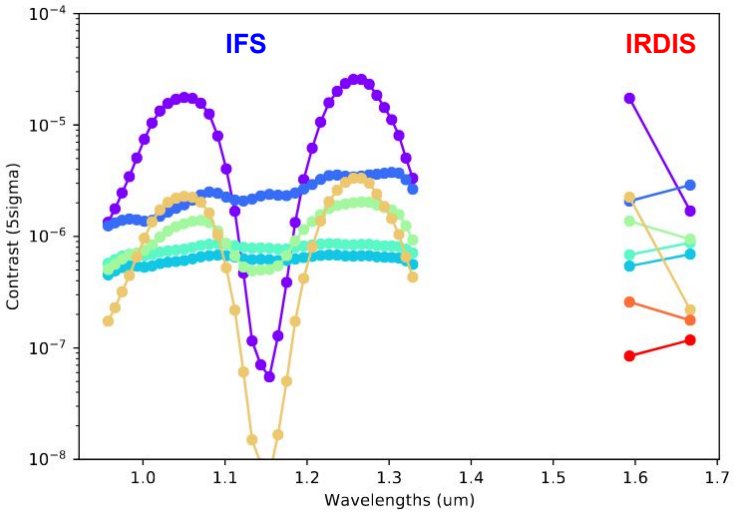
- Specal: IRDIS (TLOCI-ADI), IFS (PCAPad TLOC-ADI), eye
- TRAP: IRDIS(/IFS), automatic, <1.0as
- PACO: IRDIS (ADI/ASDI), IFS (ADI/ASDI), automatic/eye

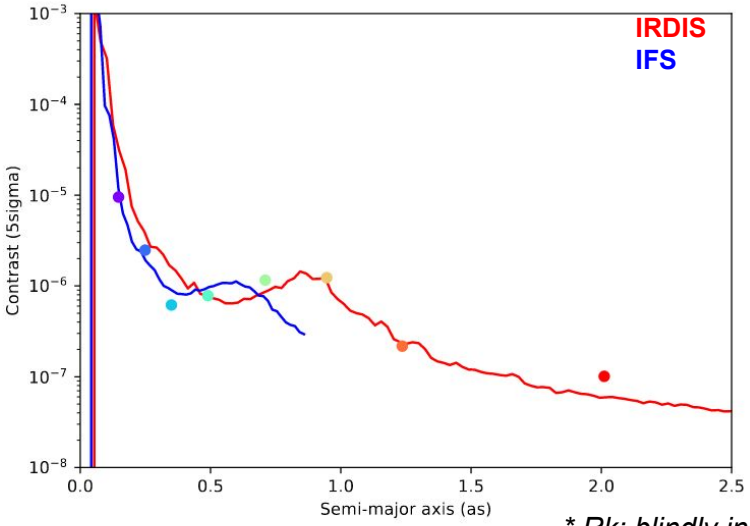
1. Internal SHINE BTs reports

The scene: fake planet(s) injection*

- 8 fake planets within 2as in IRDIFS mode (IFS-YJ, IRDIS-H2H3),
  - 8 in IRDIS FoV
  - 5 in IFS FoV
- Spectral types: late-M to late-T types,
- Close to the 5σ detection limit,
- Same astrophysical scene for both IRDIS and IFS







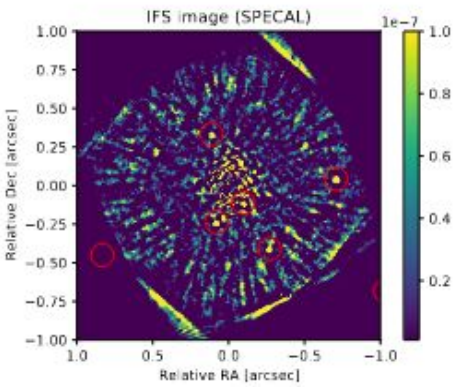
** Rk: blindly injected by Philippe through DC*

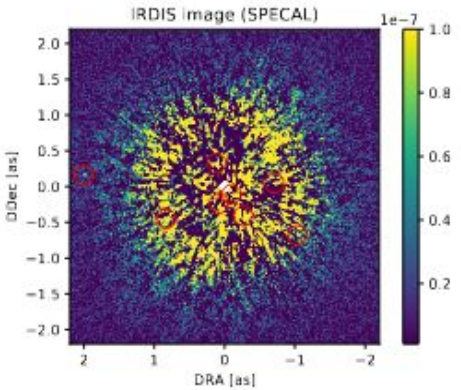
1. Internal SHINE BTs reports

Detection

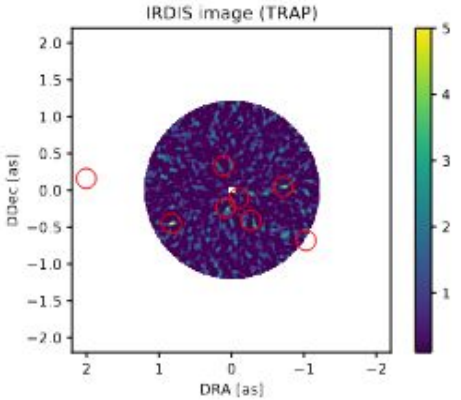
Case 1: Bright & Good
HIP12394 IRDIFS
Date: 2016/09/15

SPECAL

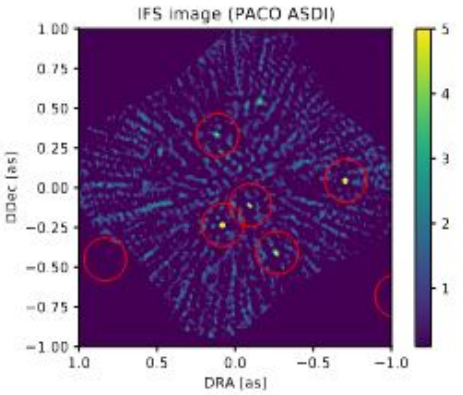


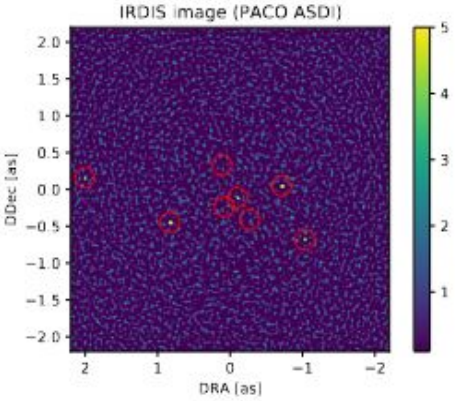


TRAP



PACO





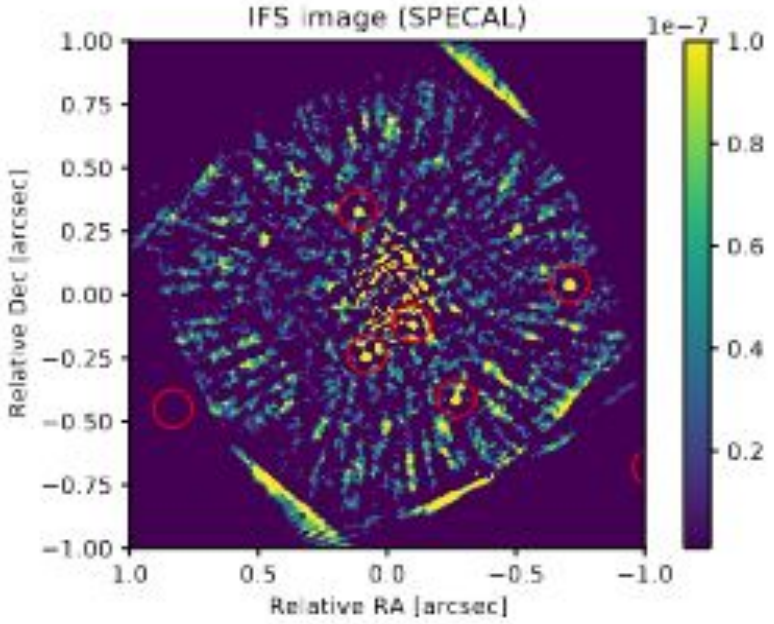
1. Internal SHINE BTs reports

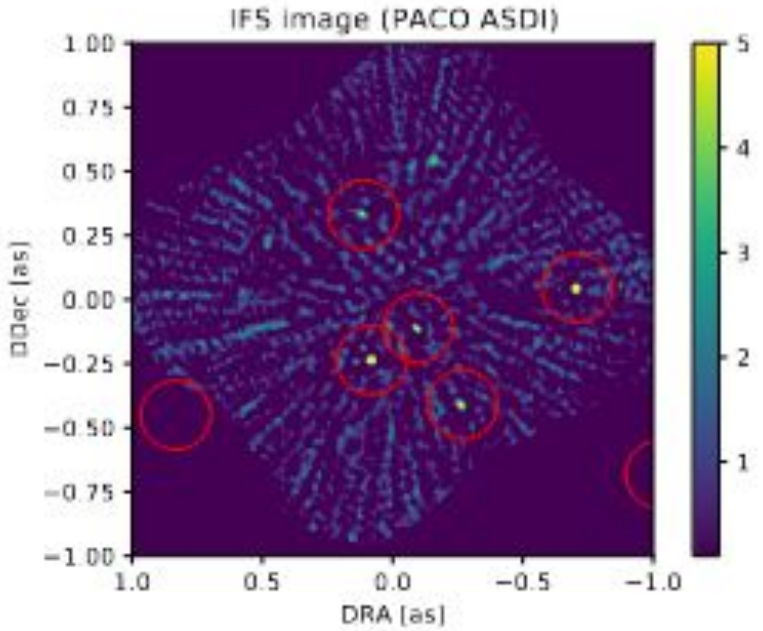
Detection

Case 1: Bright & Good
HIP12394 IRDIFS
Date: 2016/09/15

Table 1: Fake planet detection. (*) probably biased by real true "astrophysical" positive.

Algorithm	Instrument	True Positive (detection)	False Negative (no detection right position)	False Positive (detection false position)
Specal	IFS	5	0	0
	IRDIS	2	6	3*
TRAP	IFS	-	-	-
	IRDIS	3	5	0
PACO	IFS	5	0	0
	IRDIS	6	2	3*





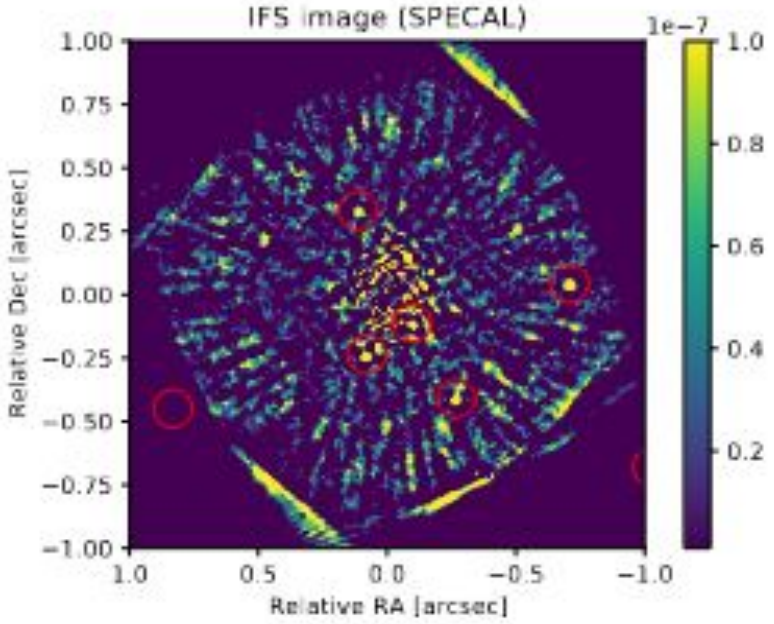
1. Internal SHINE BTs reports

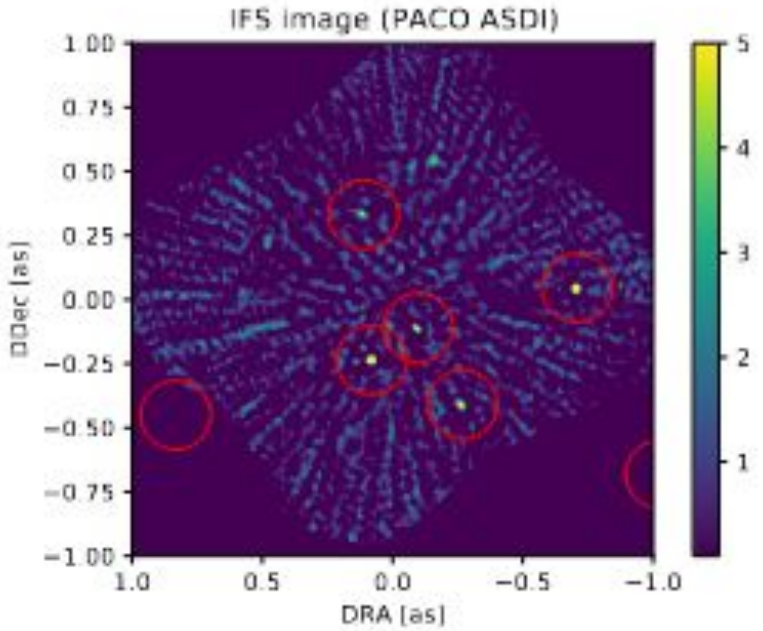
Detection

Case 1: Bright & Good
HIP12394 IRDIFS
Date: 2016/09/15

Table 1: Fake planet detection. (*) probably biased by real true "astrophysical" positive.

Algorithm	Instrument	True Positive (detection)	False Negative (no detection right position)	False Positive (detection false position)
Specal	IFS	5	0	0
	IRDIS	2	6	3*
TRAP	IFS	-	-	-
	IRDIS	3	5	0
PACO	IFS	5	0	0
	IRDIS	6	2	3*

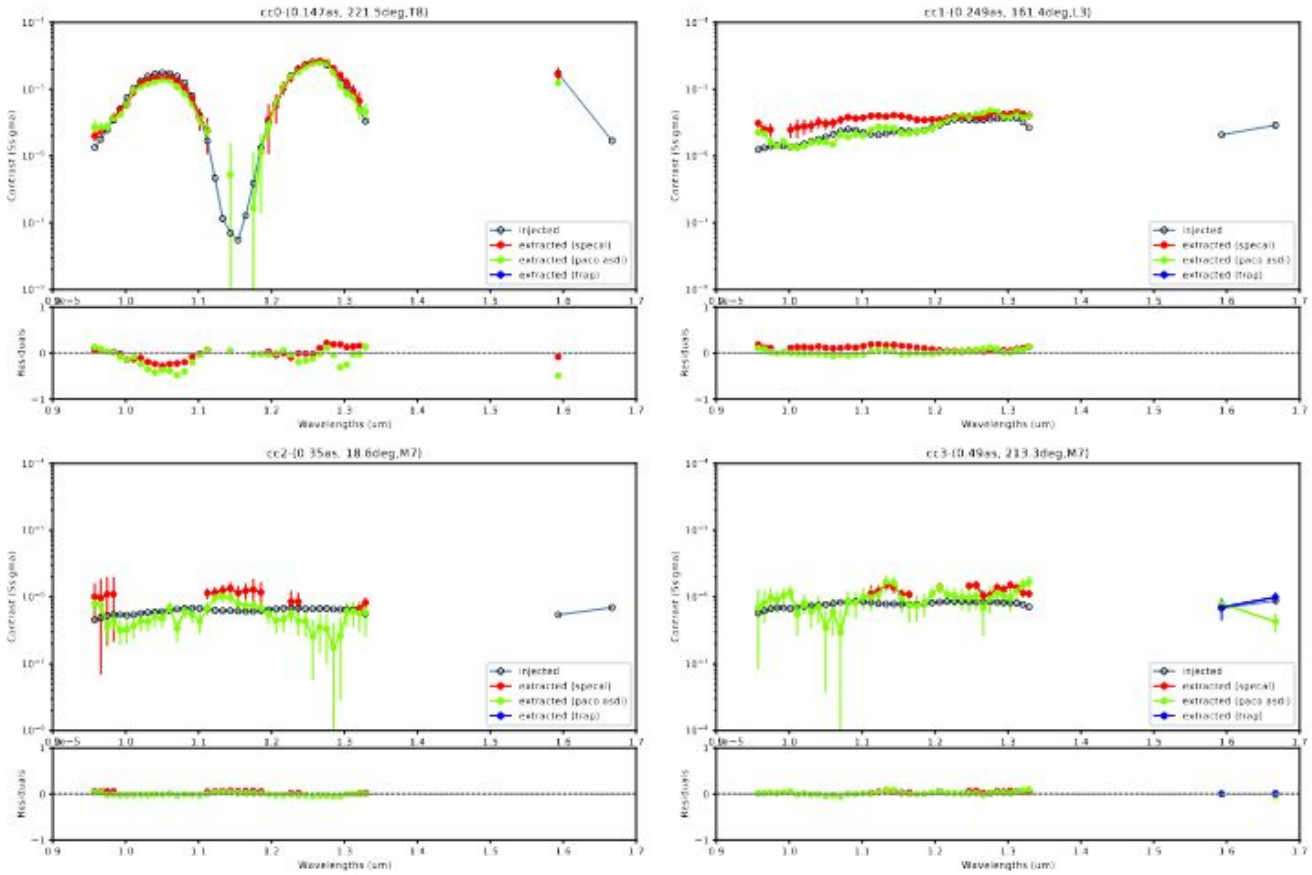




1. Internal SHINE BTs reports

Spectral fidelity

Case 1: Bright & Good
HIP12394 IRDIFS
Date: 2016/09/15



1. Internal SHINE BTs reports

Metrics to rank the different algorithms

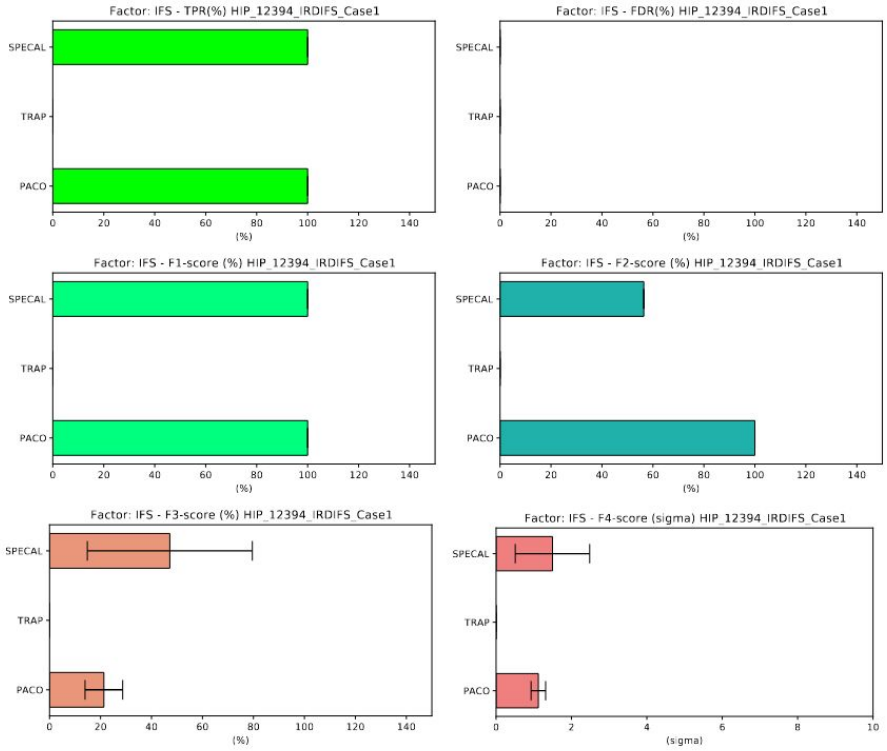
1. Separate IFS from IRDIS in the metrics

2. Add detection scores: TPR, FDR and F1-score
   a. True Positive Rate: TPR = TP/(TP + FN),
   b. False Discovery Rate: FDR = FP/(FP + TP)
   c. F1-score (harmonic mean of precision and sensitivity): F1-score = 2 ·TP/(2 ·TP + FP + FN).

3. For the detected planets, number of channels if the planet is detected. F2: fraction of channel detected (average, stdev)

4. Check spectral fidelity using absolute deviation normalized by injected spectrum
   F3 = | Ext - Inj | / Inj (%)
   At the end take the median value for one fake planet, take the average and stdev for all FPs

5. Check spectral fidelity (with error bars) using absolute deviation normalized by error bar
   F4 = | Ext - Inj | / sig_Ext (sigma)
   At the end take the median value for one fake planet, take the average and stdev for all FPs



1. Internal SHINE BTs reports

Metrics to rank the different algorithms

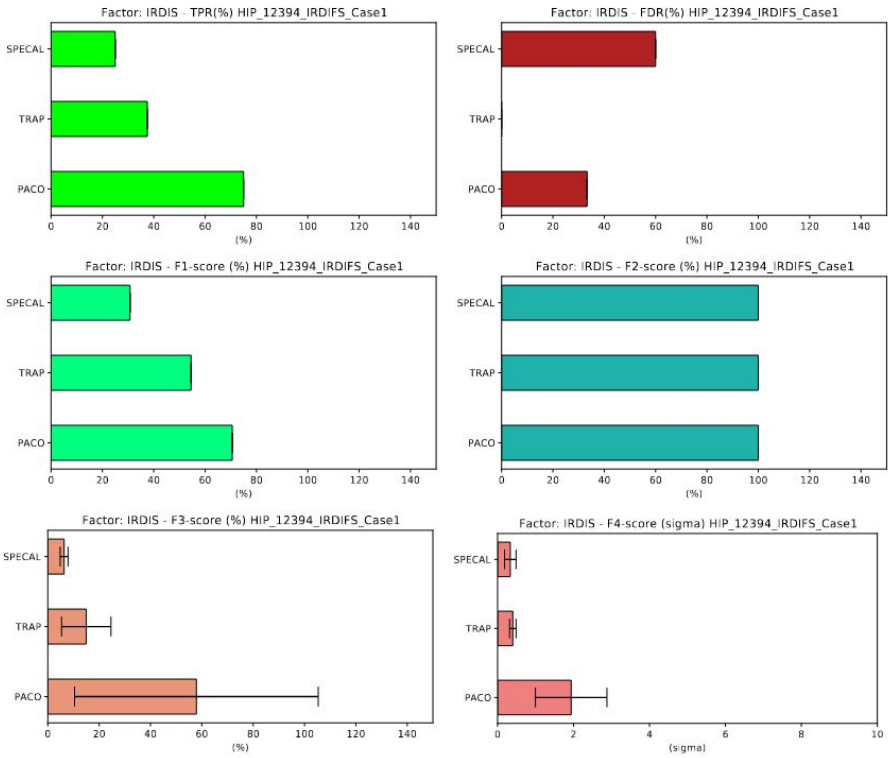
1. Separate IFS from IRDIS in the metrics

2. Add detection scores: TPR, FDR and F1-score
   a. True Positive Rate: TPR = TP/(TP + FN),
   b. False Discovery Rate: FDR = FP/(FP + TP)
   c. F1-score (harmonic mean of precision and sensitivity): F1-score = 2 ·TP/(2 ·TP + FP + FN).

3. For the detected planets, number of channels if the planet is detected. F2: fraction of channel detected (average, stdev)

4. Check spectral fidelity using absolute deviation normalized by injected spectrum
   F3 = | Ext - Inj | / Inj (%)
   At the end take the median value for one fake planet, take the average and stdev for all FPs

5. Check spectral fidelity (with error bars) using absolute deviation normalized by error bar
   F4 = | Ext - Inj | / sig_Ext (sigma)
   At the end take the median value for one fake planet, take the average and stdev for all FPs

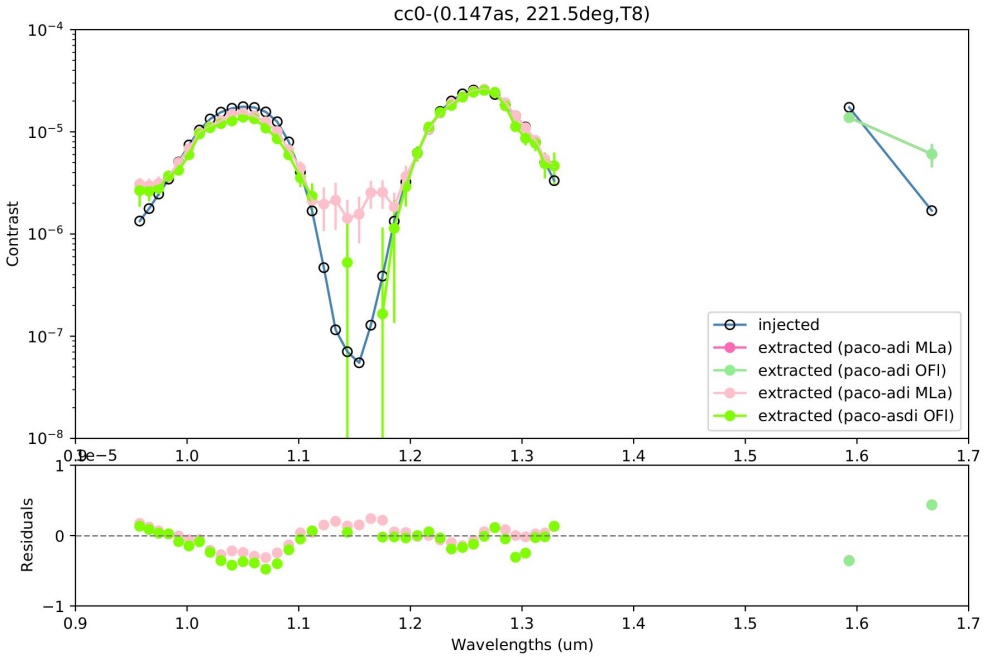


1. Internal SHINE BTs reports

Detection & Spectral fidelity

Action: Implementation of ASDI PACO results

- 8 x 2 x 6 = 96 datasets (astrometry, indexes check - very long and fastidious process)
- IFS: 80% complete, no huge differences, no impact on the previous conclusions. Plan to add the flag region proposed by Olivier on the graph.
- IRDIS: Problem to understand indexation (and rho/theta astrometric results) between IRDIS ASDI and IRDIS ADI. To be checked with Olivier.
- NEW: TRAP IFS results received, currently checking how to implement them in the final report
- Latest version of the BT-report shared in April



1. Internal SHINE BTs reports

Conclusions (April 2022)

- Proceed with final actions and validate preliminary conclusions,
  - F400 roadmap (validated by WP0):
    i. Run Specal for all pending F400 observations,
    ii. PACO as baseline for detection & characterization,
    iii. TRAP used as support for specific targets of interest,
    iv. ANDROMEDA (potentially same strategy than TRAP)?

- Proceed with pipeline update required and proceed with F400 reduction/analysis,
  - validate pipeline tools in Oct/Nov 2021 > December 2021
  - reprocess everything btw Dec 2021 to Dec 2022,
  - WP3 analysis in spring 2023,
  - first draft of complete F400 appears for Fall 2023