Pre-reduction

Getting the best data to the upstream algos

Rationale

Systematics in the data can kill even the most sophisticated post-processing.

→ Pre-processing should remove them, but some issues persist

IRDIS and IFS issues

Reduction is mostly OK, but :

-Frame centering can be improved → Jules

-Frame by frame photometric calibration might be possible → Maud

- Background subtraction K-band

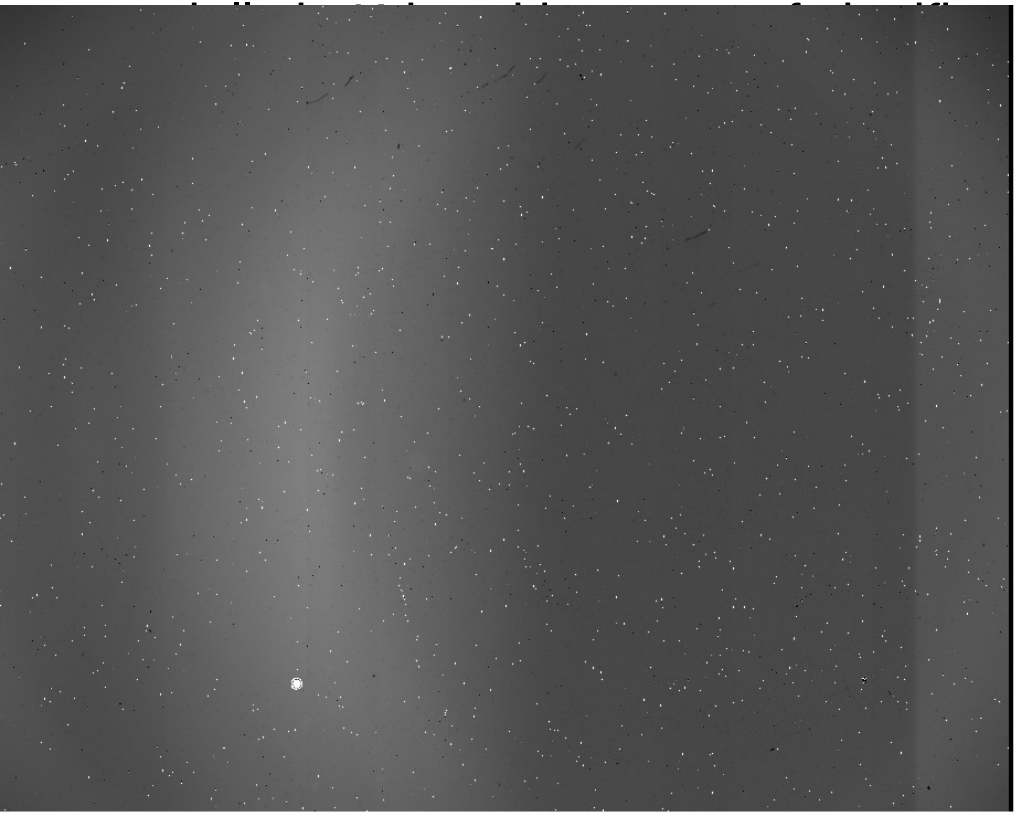
IRDIS and IFS issues

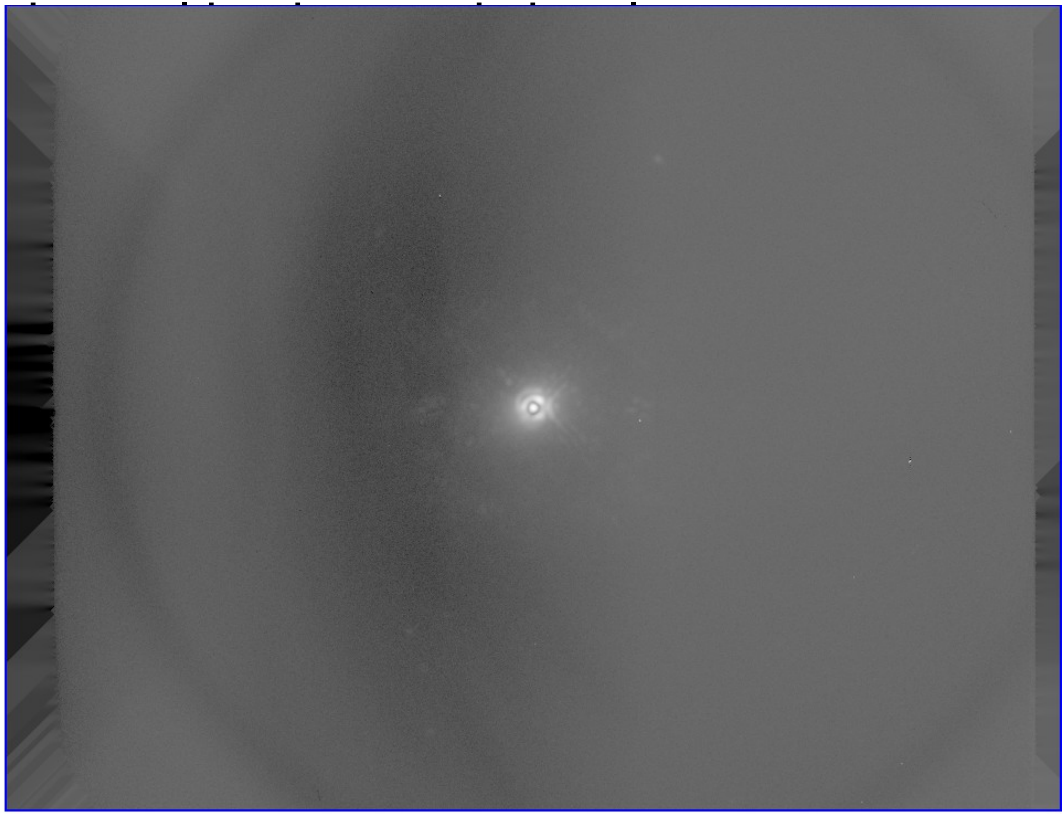
-Frame centering can be improved → Jules

-Frame by frame photometric calibration might be possible → Maud

IRDIS

Reduction is mostly OK, but some issue remain for background subtraction





IRDIS

Ideas to try :

make a master sky using many skies and fit it to the background in the science image ? How to mitigate target star flux impact in the flux ?

Make a PCA of a cube of skies and decompose the background dominated part of the science on the first components of this basis ?

IFS

Reduction is not good.

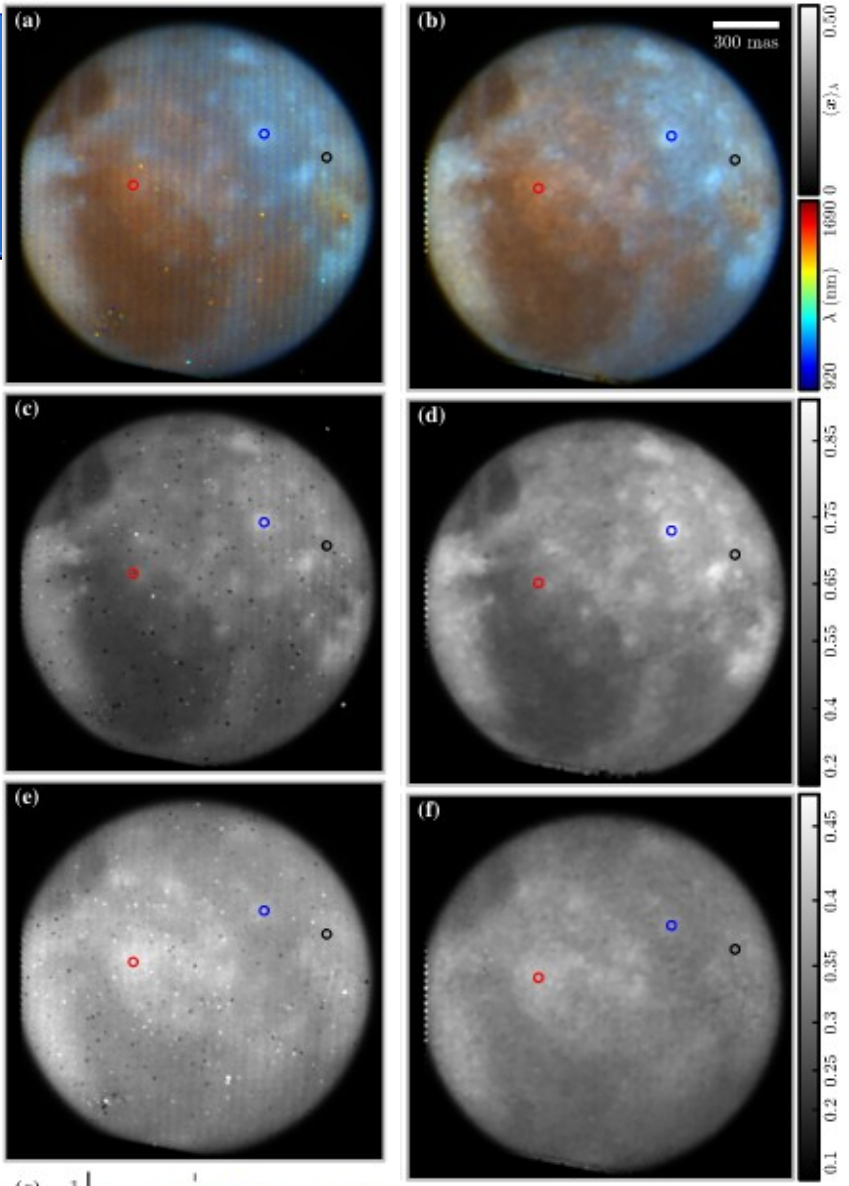
At best it is OK-ish, but it regularly crashes down because of wavelength calibration issue and/or spectra identification

Bad pixels propagated though interpolations

Wrong pixtable because wrong wavelength

Biased extracted spectra

PIC Berdeu et al .2020

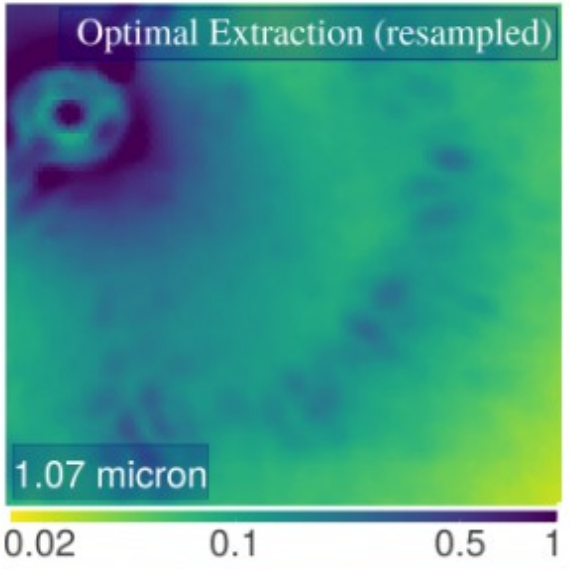


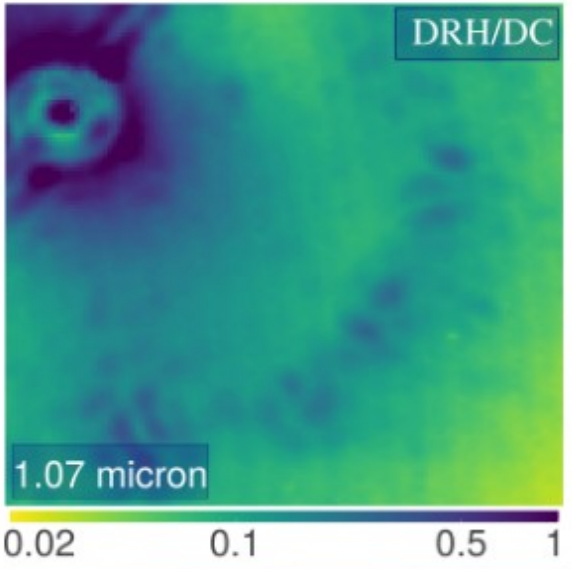
IFS

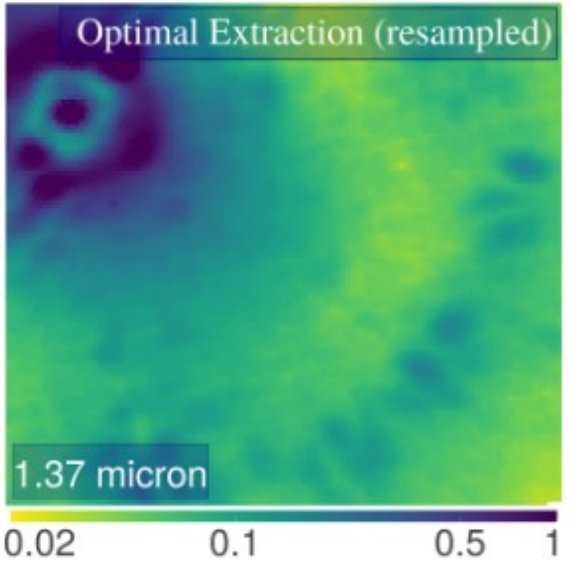
Issues most visible in poor dataset and on extended sources

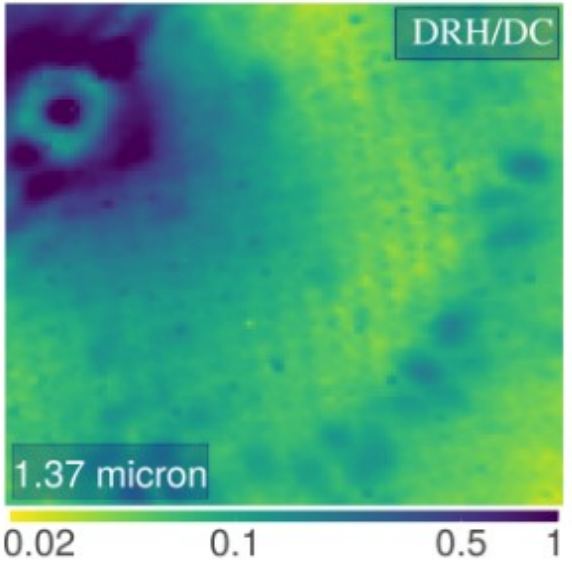
But the issues are still there on good dataset and point sources !

Samland et al.2022 Accepted



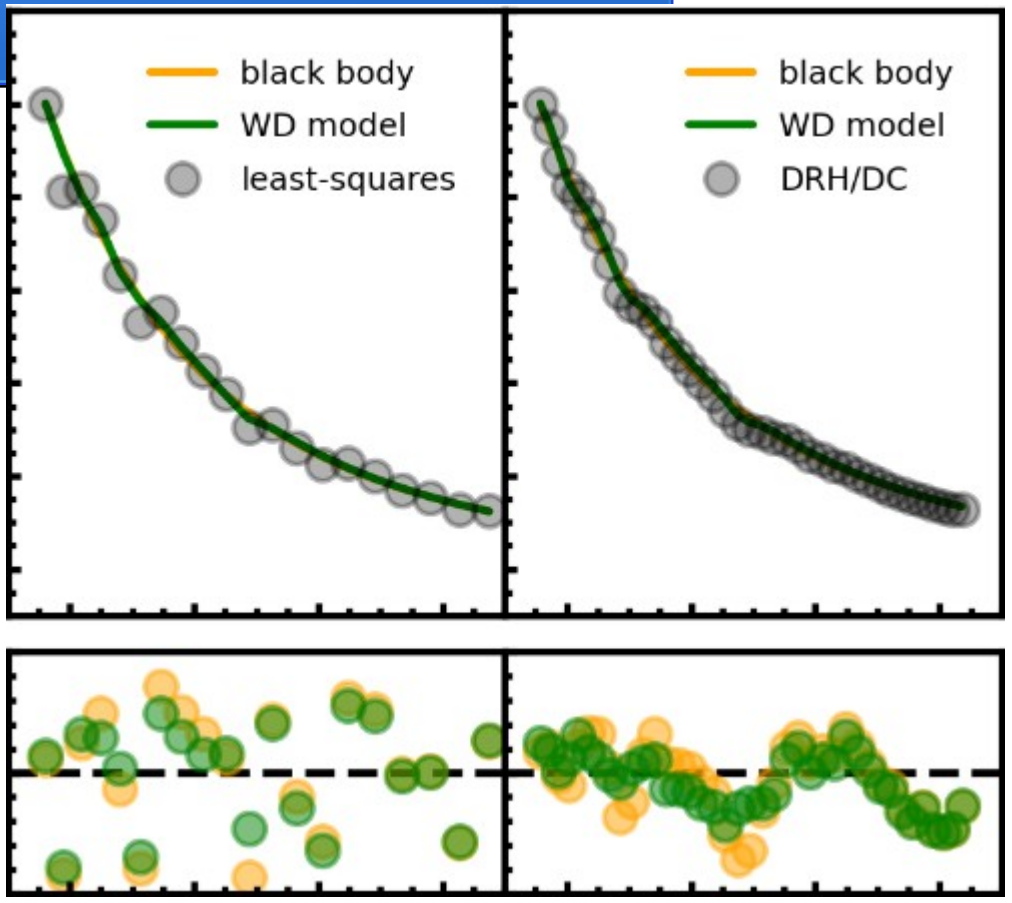






IFS

Biase spectra and introduce non gaussian, correlated noise



Samland et al.2022 Accepted

IFS

Integrate Samland et al pipeline in DC ?

Inverse model approach could replace all the reduction chain up to « convert »

COBREX -Data Center

Philippe Delorme for the COBREX team

COBREX kick-off Meeting

Existing infra :SPHERE-DC

- Provide homogeneous reduction for SPHERE data
  - Extensive knowledge of SPHERE reduction allows to avoid most problems
  - Specific patches developped to correct problems not solved by public software
  - Keep memory of reductions (routines, versions, parameters, date, products) through dedicated database
- Reduce PI/GTO data on request using improved pipeline
  - Provide private reduced data to registered users. Password protected.
- Provide access to reduced public SPHERE data
  - All public data to be released, reduction rate 1 year of data per trimester

Existing infra : Some Numbers

Reduced more than 100 ESO programs including SHINE
> 340000 raw SPHERE files in data base
> 1200000 reduced SPHERE files
> 165000 existing processes

84To

All these files and processes, are lines of a database : easy queries/selections using associated metadata.

Reducing a full semester of SPHERE data public data takes typically 2-3 weeks, most of it computer time, which could be decreased to ~ 1 week by running more workflows in parallel and having someone on it at 50% ETP for the week.

Once a workflow is set, human time required is low : data import, data check, and manual resolution of some specific issues

COBREX -Data Center

Philippe Delorme for the COBREX team

COBREX kick-off Meeting

Existing infra :SPHERE-DC

- Provide homogeneous reduction for SPHERE data
  - Extensive knowledge of SPHERE reduction allows to avoid most problems
  - Specific patches developped to correct problems not solved by public software
  - Keep memory of reductions (routines, versions, parameters, date, products) through dedicated database
- Reduce PI/GTO data on request using improved pipeline
  - Provide private reduced data to registered users. Password protected.
- Provide access to reduced public SPHERE data
  - All public data to be released, reduction rate 1 year of data per trimester

Existing infra : Some Numbers

Reduced more than 100 ESO programs including SHINE
> 340000 raw SPHERE files in data base
> 1200000 reduced SPHERE files
> 165000 existing processes

84To

All these files and processes, are lines of a database : easy queries/selections using associated metadata.

Reducing a full semester of SPHERE data public data takes typically 2-3 weeks, most of it computer time, which could be decreased to ~ 1 week by running more workflows in parallel and having someone on it at 50% ETP for the week.

Once a workflow is set, human time required is low : data import, data check, and manual resolution of some specific issues