

# OU-LE3 Status and Open Issues



Agenzia Spaziale Italiana



**Alfonso Veropalumbo (Univ. Roma Tre)**  
**Ben Granett (INAF – OA-Brera)**  
**Emiliano Munari (INAF – OA-TS)**

5<sup>th</sup> National Euclid Meeting 2022

# OU-LE3 P1/2 PFs status

WP	PF Priority 1 - 2	ML status	Test data	Pipeline status	Run status
Internal Data	VMSP-ID	ML3A	Proposed	Complete	PPO
	SEL-ID	ML3A	Proposed	Complete	PPO
External Data	GALEXT-ED	ML3A	On EAS	Complete	PPO
Galaxy Clustering	2PCF-GC	ML3A +30%	On EAS	Complete	PPO
	PK-GC	ML3A +64%	On EAS	Complete	PPO
	3PCF-GC	ML3A +20%	On EAS	Complete	PPO
	BK-GC	ML3A +62%	On EAS	Complete	PPO
	CM-2PCF-GC	ML3A	On EAS	Complete	PPO
	CM-PK-GC	ML2B +85%	On EAS	Complete	PPO
Clusters	DET-CL	ML2B	On EAS	Complete	PPO

## Outstanding Issues

- Science requirements update, codes upgrade and validation tests
- PF integration in the data analysis pipeline: in progress.
- Data flow: Interface with IST
- Scientific validation: LE3 involvement and impact

# OU-LE3 Galaxy Clustering

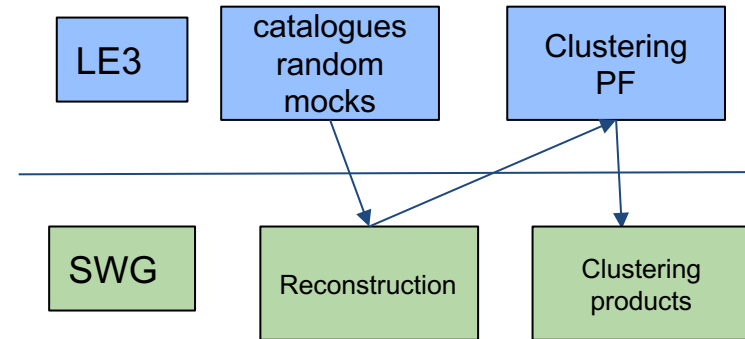
## Validation activities

- **Maturity Level 3B** (requested for the production phase) requires 100% validation test coverage. This is our **current main effort**.
- Most of these tests **use specific datasets** (not generated by OU-SIM). Use / maintenance / production of these datasets **require manpower**.
- **Additional data products** will likely be requested by SWGs (i.e. cross correlation statistics, anisotropy estimators etc). We tried to **anticipate** them by adding new options in the codes. That eventually will need to be validated.
- The need to **validate the whole galaxy clustering pipeline** has emerged at the last MKP. We will contribute to this effort, which implies defining additional tests and writing the codes to perform them.
- We also contribute to **scientific validation tests** (under SWG responsibility and within the SPV2 framework) on top of the SGS ones.

# Reconstruction

## Scientific background:

- Techniques to reduce the non-linearity of the galaxy density field.
- Used to sharpen the BAO peak, to improve cosmological constraints.
- **Responsibility of implementation and tests by SWG.**



## OU LE3 perspective:

- PF **ready** to ingest reconstructed data and random samples, and to produce outputs compatible with EAS.
- **Open issue:** implementation of estimator for Zel'dovich reconstruction.

# OU-LE3 Galaxy Clustering

## Interface with ISTs and SWGs

- **Interface with SWGs** mainly involves definition of **science requirements**, definition and execution of validation tests, contribution to S/W production.
- Over the last year the need of a **more effective interface with ISTs** has also emerged. First to guarantee that **ISTs are aware** of the existing OU-LE3 **data products** and their associated documents. Then to **identify potential missing products** and finally **agree on the responsibility for their production**. Example: window function for Fourier-space based clustering analyses.

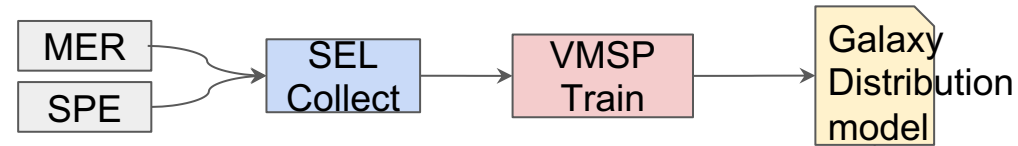
# OU-LE3 Spectroscopic Catalog

- SEL-VMSP (*sample selection & spectroscopic visibility mask*) makes the interface between LE2 and LE3 in the spectroscopic galaxy clustering pipeline.
- The processing functions are distributed between UK, Italy and USA.
  - *Manager*: Sesh Nadathur (UK)
  - *Validation lead*: Lucia F. de la Bella (UK)
  - *VMSP developer*: Ben Granett (INAF)
  - *SEL developer*: Lado Samushia (USA)
  - *Active collaborators*: Coleman Krawczyk (UK), Enzo Branchini, Pierluigi Monaco, Samuele Galeotta (IT)

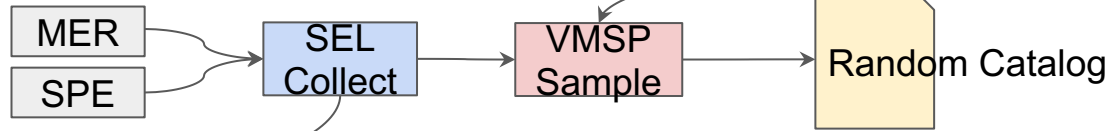
# OU-LE3 Spectroscopic Catalog

- SEL-VMSP is the interface between LE2 and LE3 in the spectroscopic galaxy clustering pipeline.

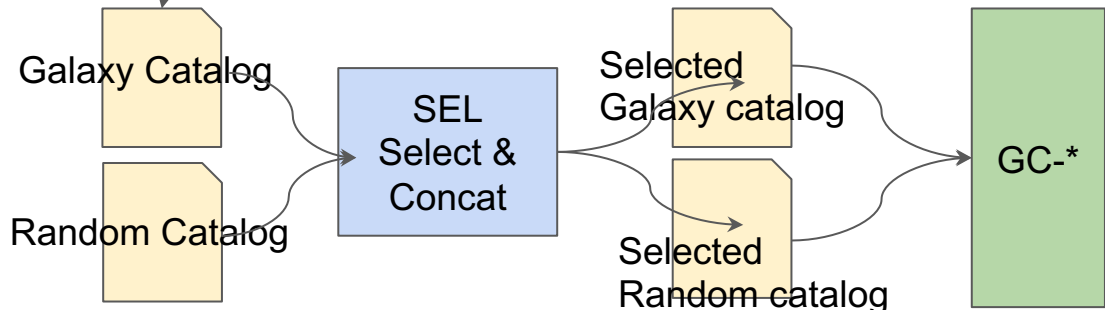
- Deep pipeline



- Wide randoms pipeline



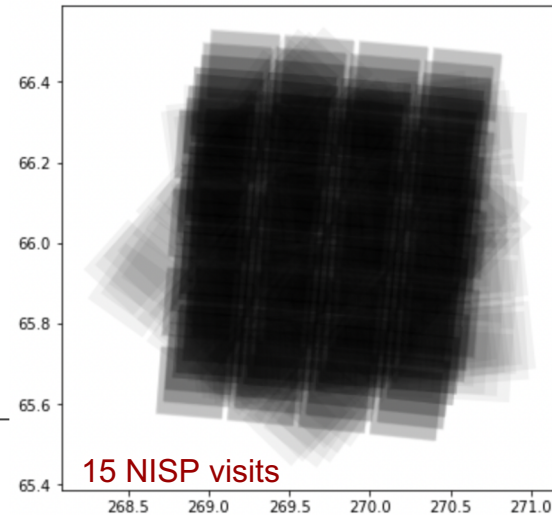
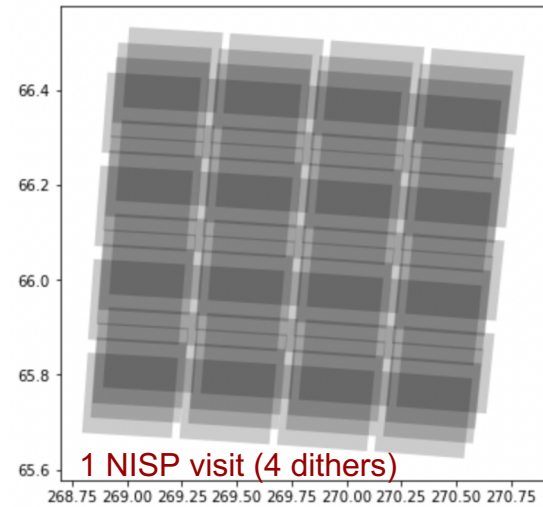
- Wide selection pipeline





# Open issues - purity and completeness

- SEL will receive a full depth CPC catalog (10 visits).
- And one or more wide-like catalogs derived from single-pass reductions of the CPC field.
- SEL will compute purity and completeness of the wide-like catalog with respect to the full depth catalog.
- It is only possible to estimate **purity** where we have a deep catalog for reference (in the CPC and Deep fields).
- Completeness can be estimated everywhere in the Wide survey using the visibility mask random catalog.
- There is no plan in LE3 to estimate purity and completeness of the CPC or Deep fields.
- The data products needed to estimate purity and completeness were not produced in SC8 and the pipeline has not been run.



# Open issues - photometric mask

- The spectroscopic selection uses  $VIS, Y, J, H$  photometry (eg  $H < 24$ , star-galaxy separation, etc) and may also adopt a more complex color pre-selection.
- The photometric selection function must be tracked with mask data products from VMPZ.
- Required inputs:
  - Depth mask on healpix grid in  $VIS, Y, J, H$ .
  - Cutout mask around bright stars.
  - Cutout mask around large galaxies.
- The interface between photo and spectro visibility masks was not tested in SC8.

# Open issues - redshift PDF

- SPE will produce a redshift PDF for each source based on template fitting.
- The z-PDF may be very useful for galaxy clustering:
  - To build statistical weights based on the probability of redshift solutions.
  - To apply a-posterior constraints to choose a redshift solution given information not used by SPE (eg colors).
  - To discriminate spurious detections based on the PDF.
- The format of the PDF has not been decided. Compression is needed since it is a large amount of data.
- No plan currently to incorporate the PDF into SEL processing.
- From the SWG side there is a request to SPE to extract the PDF for SC8 Deep sources so that tests can be done.

# Update on CL PFs with Italian (co)-implementation

## **DET-CL AMICO: P2 - IT+FR**

See table in slide 2. In a few weeks the projects of the 2 DET algorithms will be merged into a single project (2 modules, same repository and same documents) that will be submitted to MA aiming at 3A

## **SEL-CL: P3 - IT+FR**

Sinfonia, the part under italian integration responsibility, has been integrated in EDEN, although the DM has not been implemented yet. Further development will be made as soon as the new tiling will be defined

## **Z-CL: P3 - IT**

ML0 close to ML1, although it has already been integrated in the Cluster Catalog Pipeline and run in SC8. It is integrated in EDEN and the DM is defined and implemented

## **SIGV-CL: P3 - IT**

Undergoing studies to define the algorithm

## **CCP and tiling**

See next slide



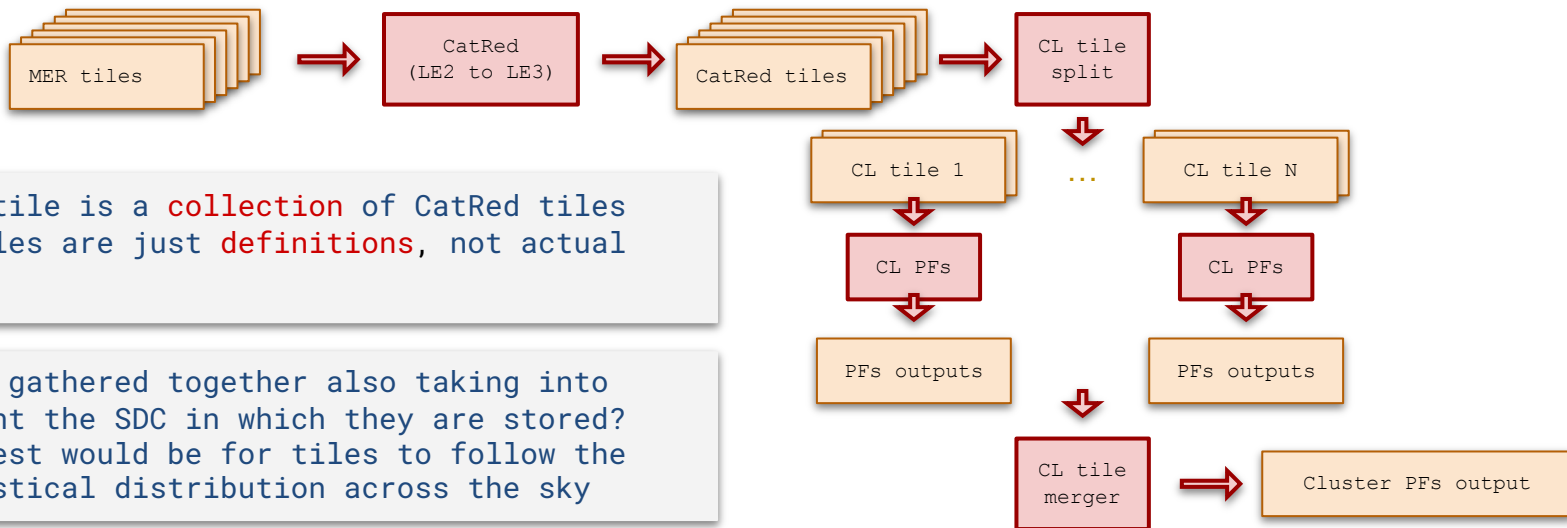
# New OU-LE3 CL data flow?

## A draft under discussion

OLD



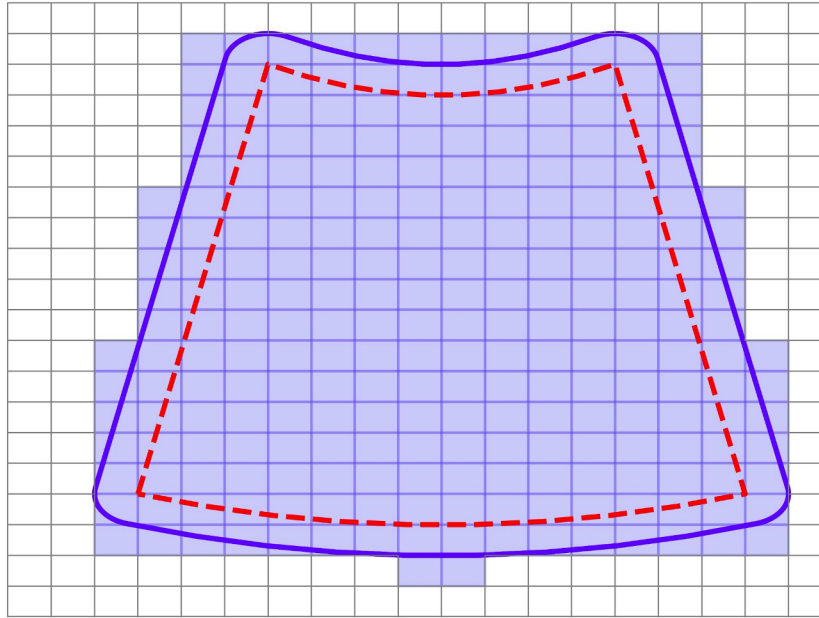
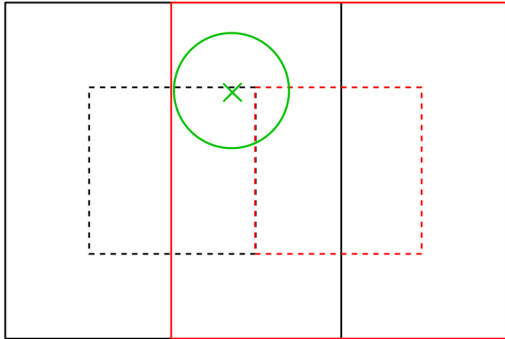
**Draft** after discussion with Thomas Vassallo, Erik Romelli, Samuele Galeotta and after a meeting with CL PFs developers. In ~1 month the structure will be finalized and the implementation will begin



- A CL tile is a **collection** of CatRed tiles
- CL tiles are just **definitions**, not actual files

- Tiles gathered together also taking into account the SDC in which they are stored?
- The best would be for tiles to follow the statistical distribution across the sky

# CL tiling



A CL tile could be  $\sim 50 \text{ deg}^2$ , while a CatRed tile (Wide Field) is  $\sim 0.25 \text{ deg}^2$  ( $30' \times 30'$ )  
The border must be large enough to contain the largest expected cluster:  
 $\sim 0.4 \text{ deg} = 5 \text{ Mpc}$  @  $z=0.2$  which corresponds to  $\sim 1$  CatRed tile

Each MER tile has a polygon delimiting it. Easy to compute intersections of polygons (e.g. Python library Shapely)

A big thank to all people, including  
SDC and SWG members,  
who have contributed to OU-LE3-IT  
activities.