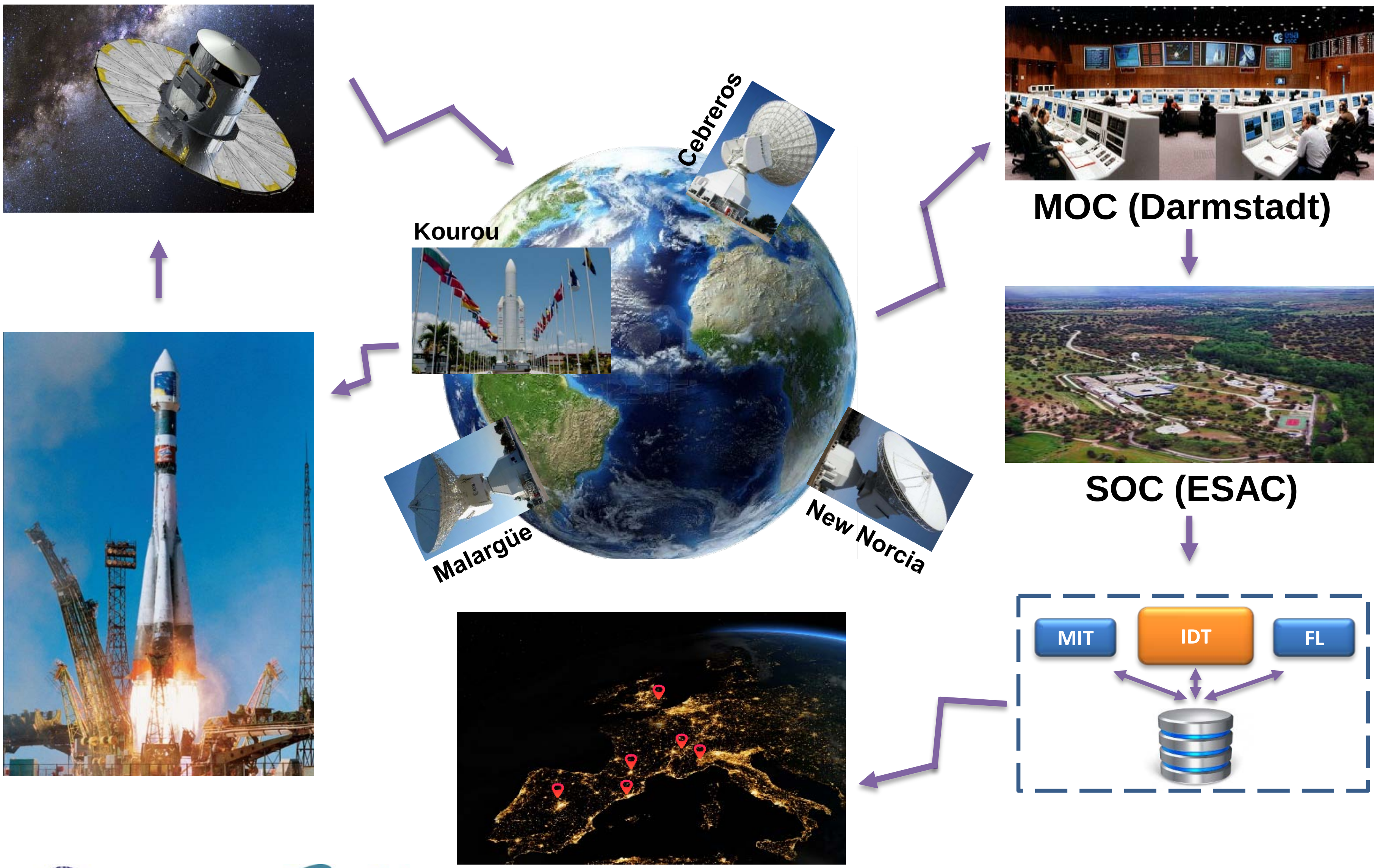
Gaia Cross-Match

El continuo reto del *Cross-Match* (XM) en los datos de Gaia

Juan José González-Vidal, Ph.D.

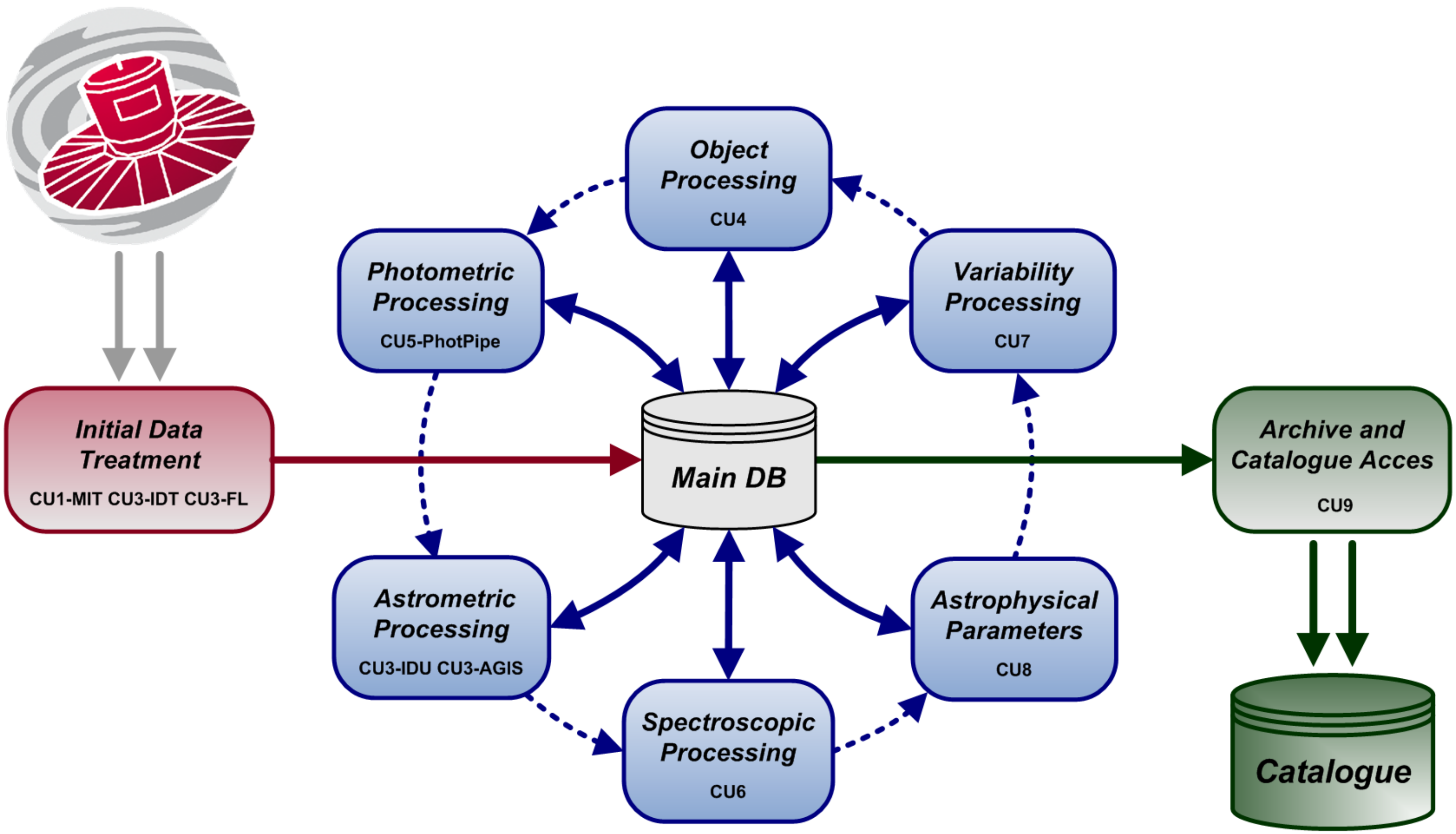
Gaia DPAC IDT Manager and IDU XM Advisor

Gaia Data Reduction Concept

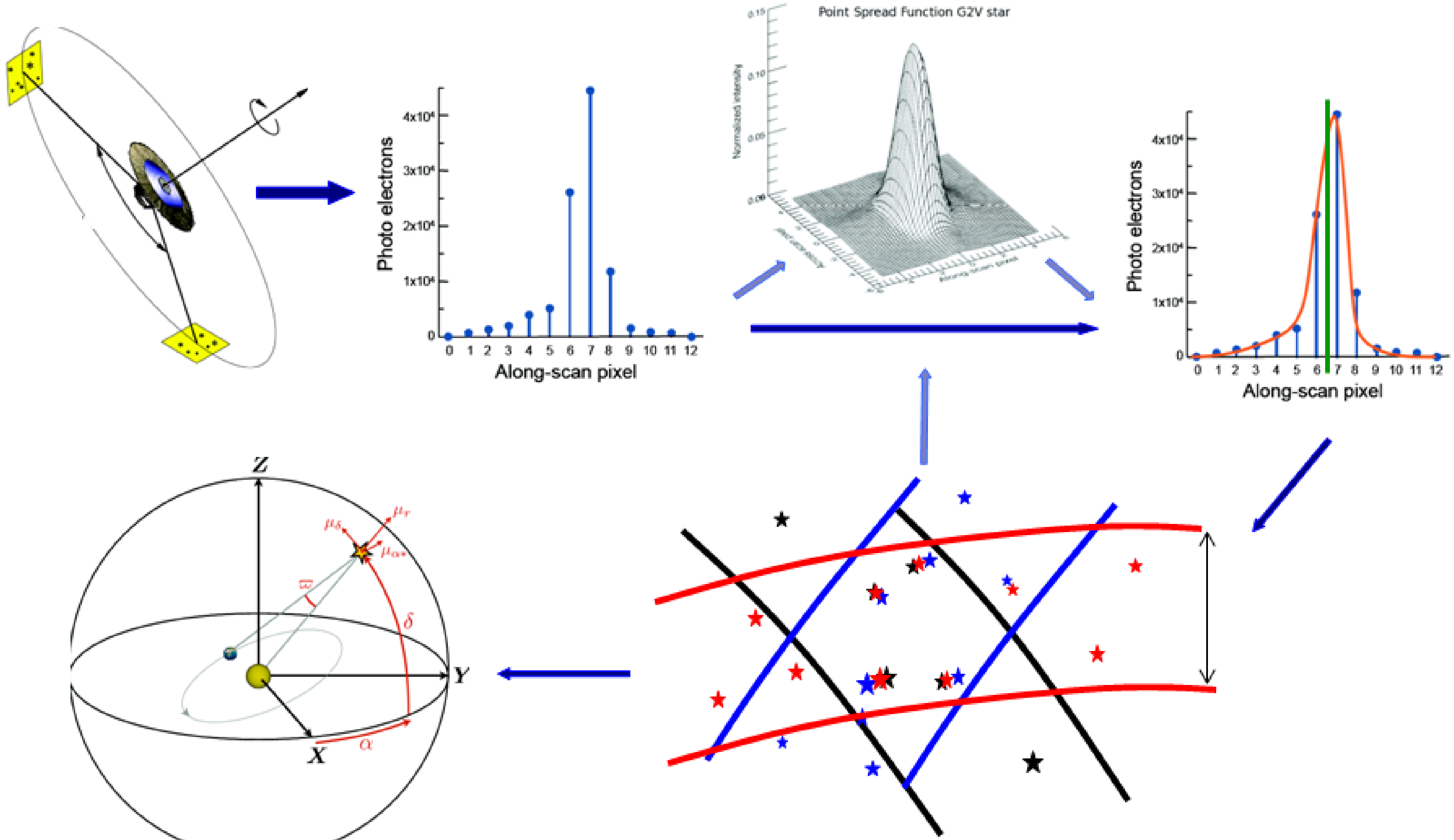


Gaia Data Reduction Concept

Gaia Data Reduction Concept



Gaia Data Reduction Concept



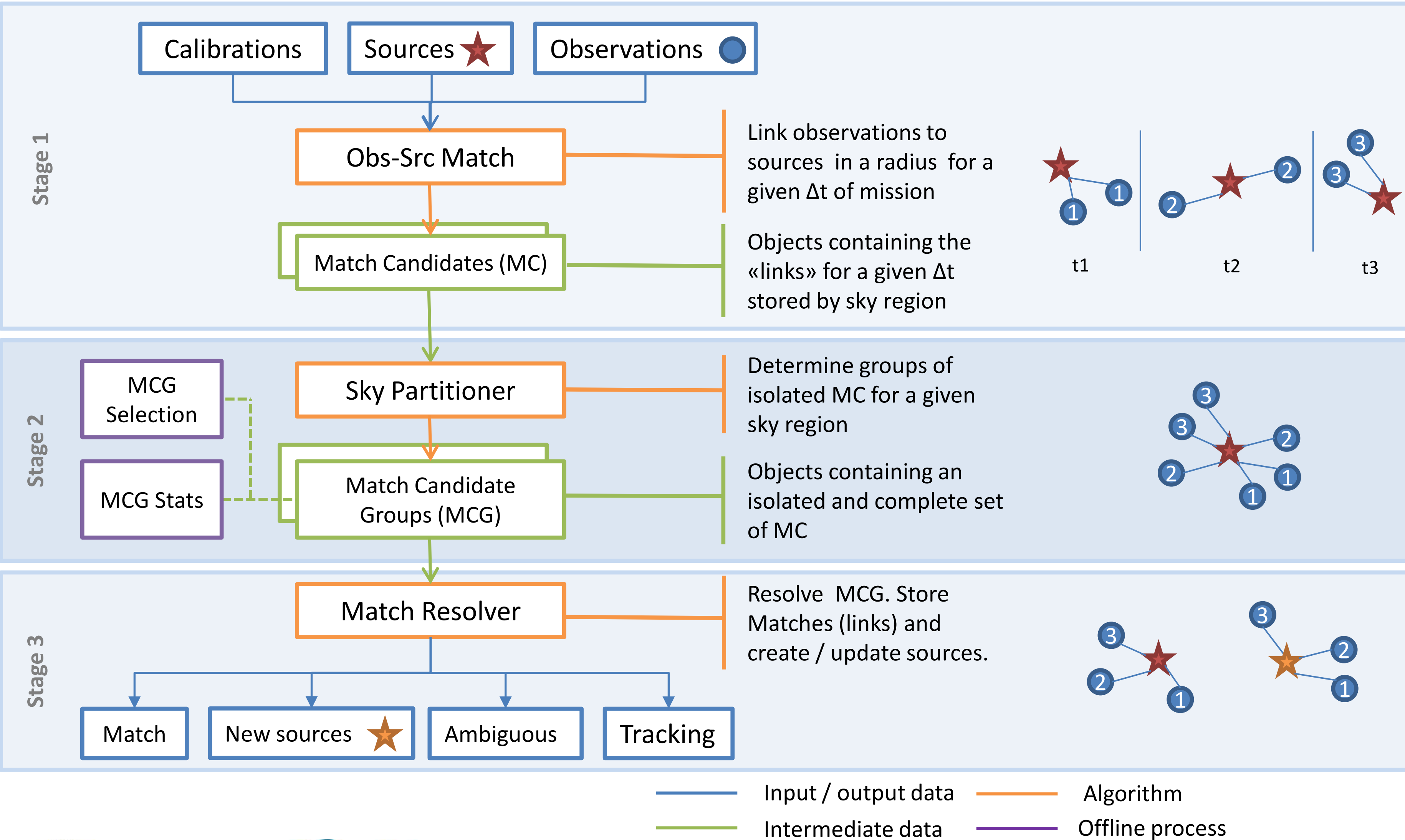
Crèdits: B. Holl, L. Lindegren

Cross Match in Gaia DPAC

- Daily in IDT
  - Limited in terms of available inputs (delayed telemetry, …)
  - Provides the input for the daily systems across DPAC: First Look, Science Alerts, SSO-daily, …
  - Has to be “quick” in terms of computing time requirements

- Cyclic in IDU
  - Is the Cross-Match that ends up in the different data releases
  - All the inputs for a given data segment are available beforehand
  - Uses the latest updates on attitude, calibrations, etc
  - Provides the *links* between observations and sources for all the cyclic processing software
  - Less restrictions on computing requirements, therefore more advanced algorithms can be used

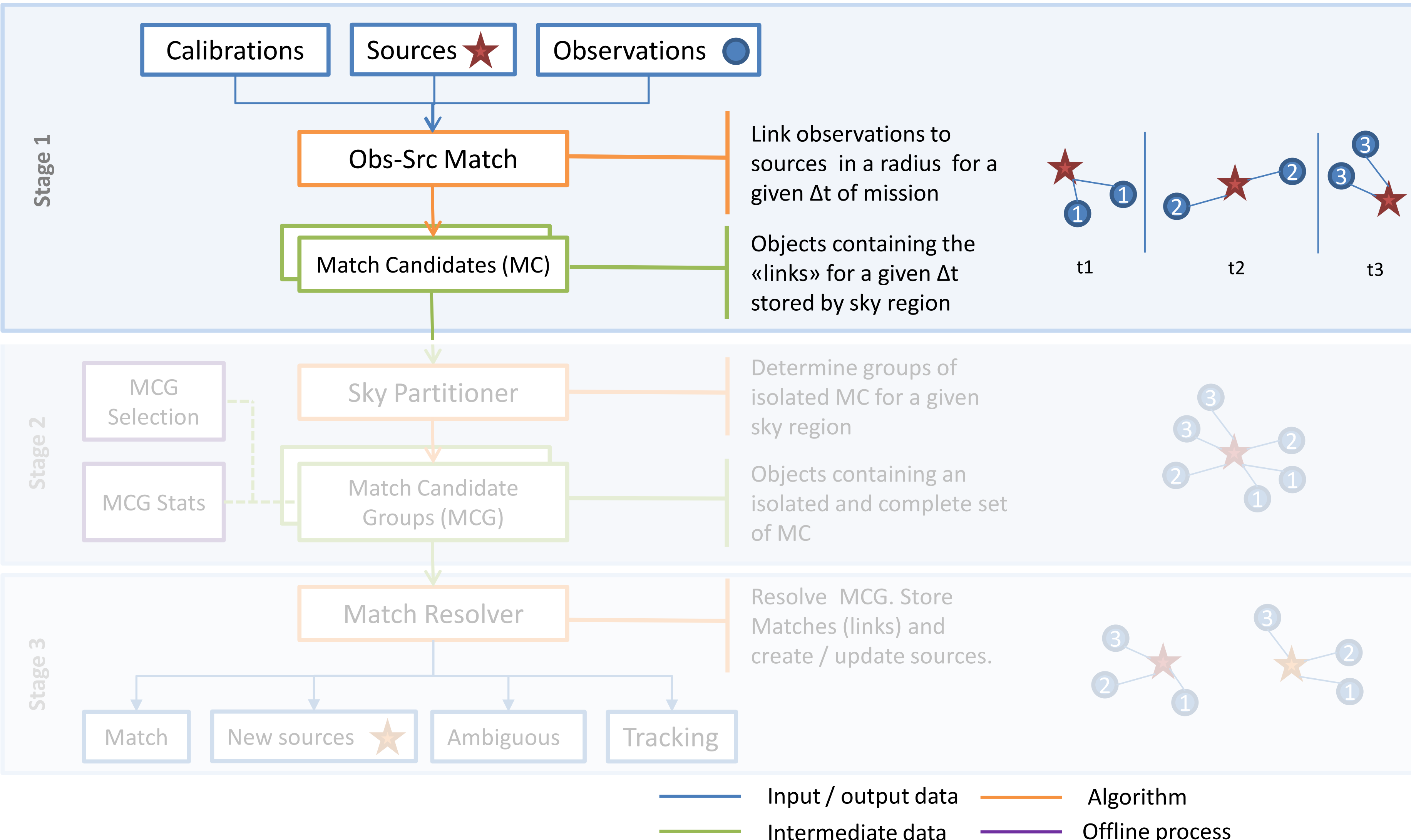
IDU Cross-Match



DAPCOM Data services

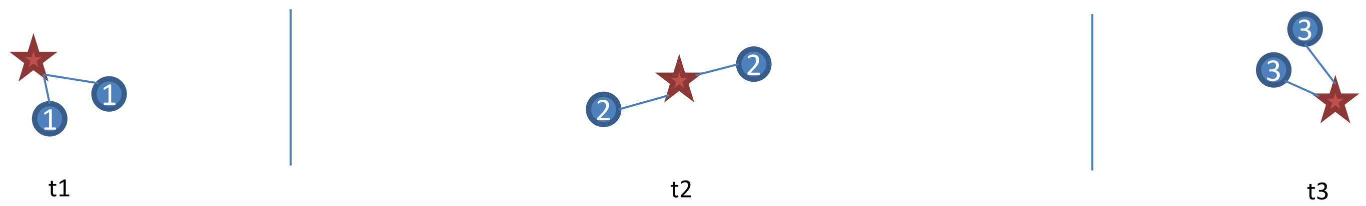
Gaia DPAC Data Processing & Analysis Consortium

IDU Cross-Match: Obs-Src Match



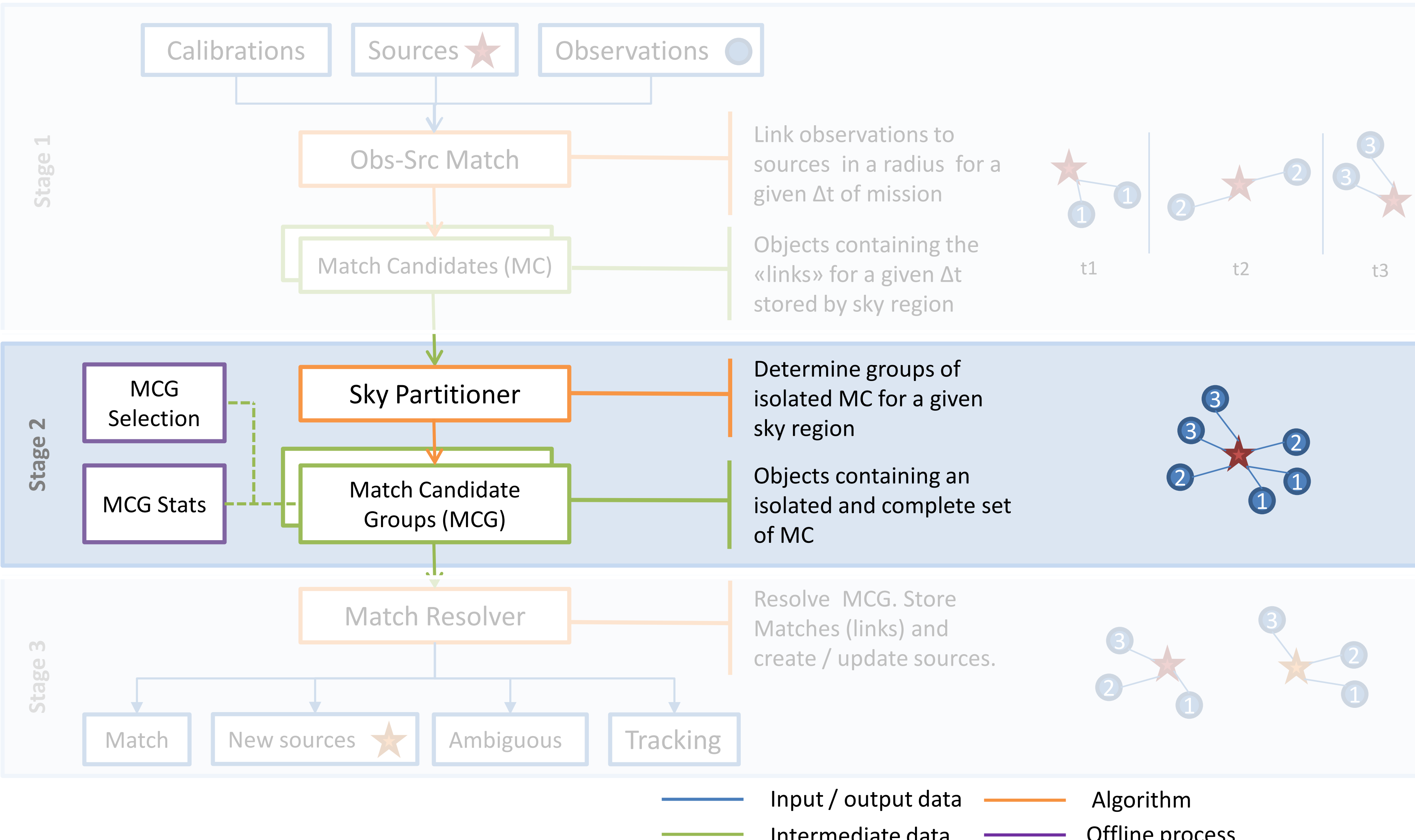
IDU Cross-Match: Obs-Src Match

- In this first step, we process the input observations in time order to compute the detection sky coordinates and obtain the preliminary source candidates for each individual detections.



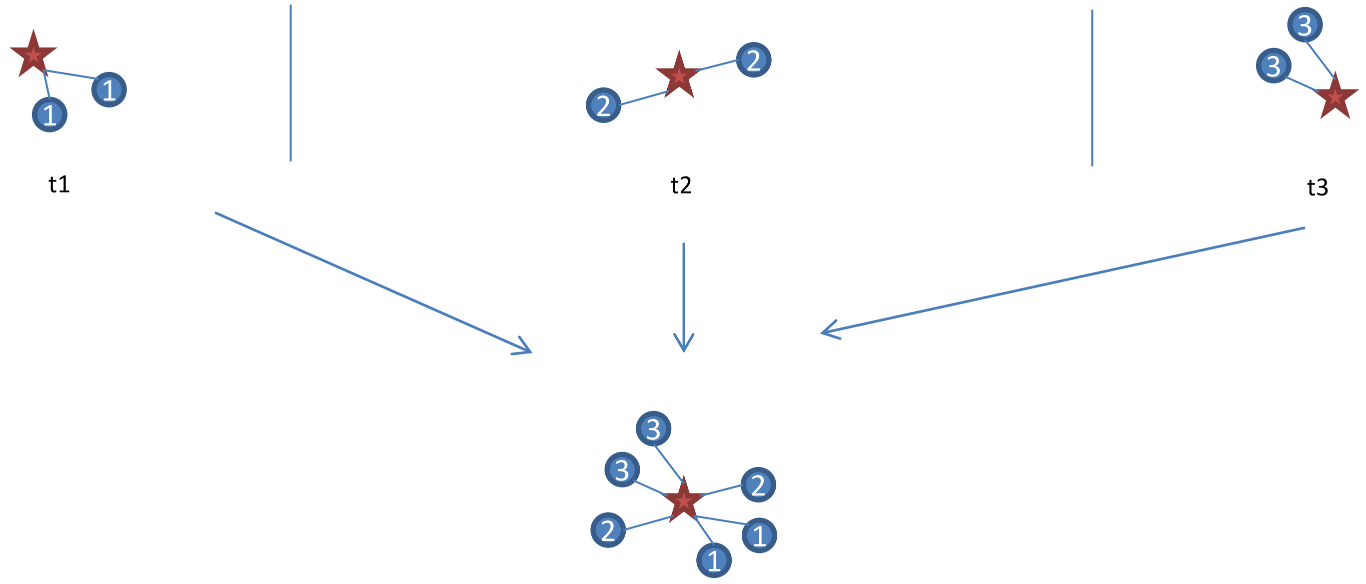
- Process:
  1. Sources are propagated to each relevant epoch (for each task)
  2. Each observation is compared to the nearby sources (distance criterion)
  3. All sources that fall within the given preconfigured radius are stored as *candidates*

IDU Cross-Match: Sky Partitioner

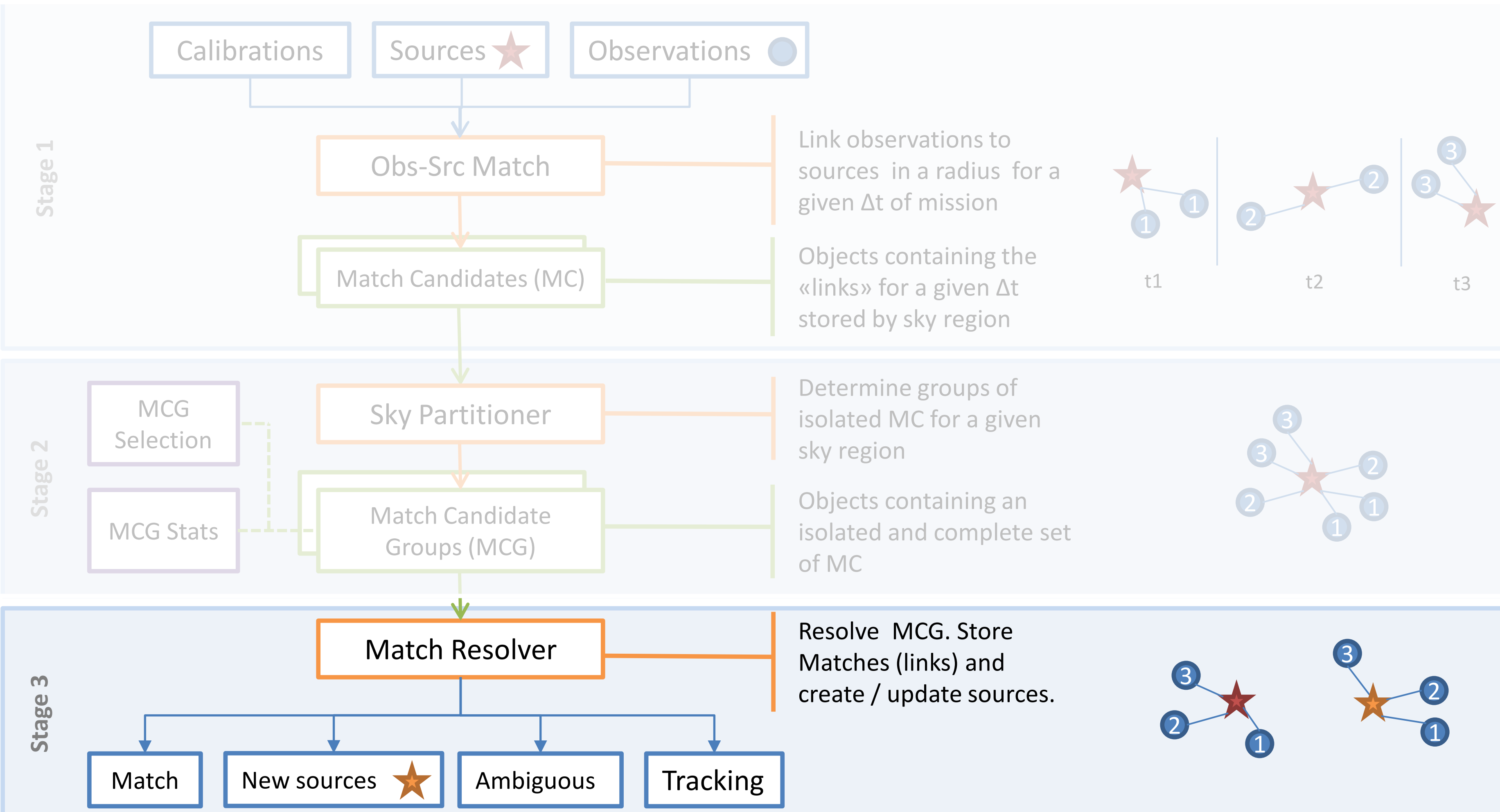


IDU Cross-Match: Sky Partitioner

- This second step groups the results according to the source candidates of each individual detection
- The objective is to determine isolated groups of detections, avoiding spatial boundary issues (bridge between the time-based and the spatial-based processing)



IDU Cross-Match: Match Resolver



Input / output data
Intermediate data
Algorithm
Offline process

DAPCOM Data services

Gaia DPAC Data Processing & Analysis Consortium

IDU Cross-Match: Match Resolver

- The final step is a spatial-based processing where all detections from a given isolated sky region are resolved together, thus taking into account all observations and sources of that region

- This step provides a *match* for each detection: a *link* between the observation and a source

- Resolve conflictive scenarios, for example:
  - Two simultaneous observations but just one source in the current catalogue. New source has to be created
  - There are observations but no sources in the catalogue. New source has to be created
  - Other conflicts, such as source merging, splitting, …

IDU CROSS-MATCH IN DR1

IDU Cross-Match in DR1

- Simplistic approach based on IDT Cross-Match
  - Local (two-by-two) conflicts resolution
  - Iterative two step process:
  1. Clean conflicts: duplicate matches
  2. Create new sources for unmatched observations

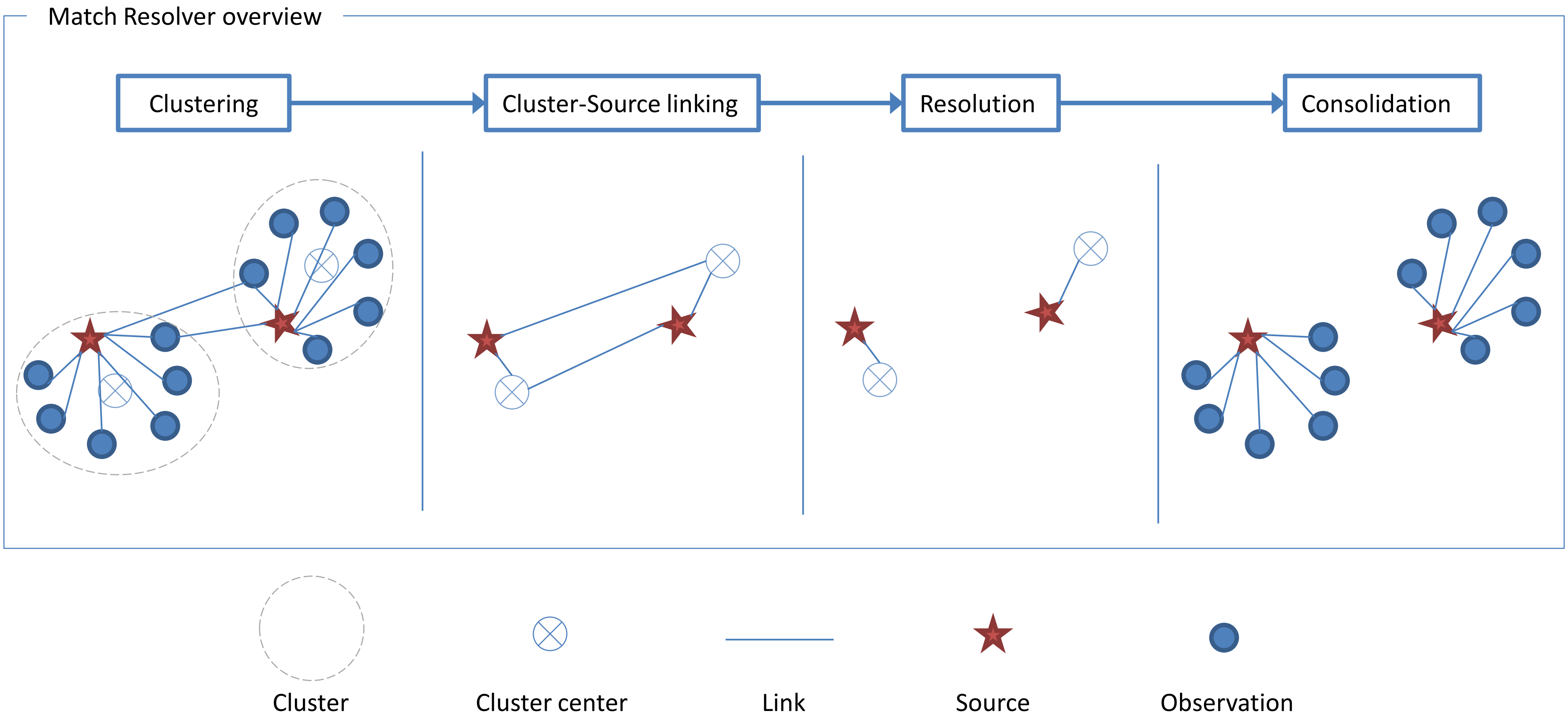
- Advantages:
  - Very fast as no input catalog is required

- Limitations:
  - No source merging or source split
  - No new source minimization
  - Observations scattered between sources
  - New sources created from a single unmatched observation

IDU CROSS-MATCH IN DR2

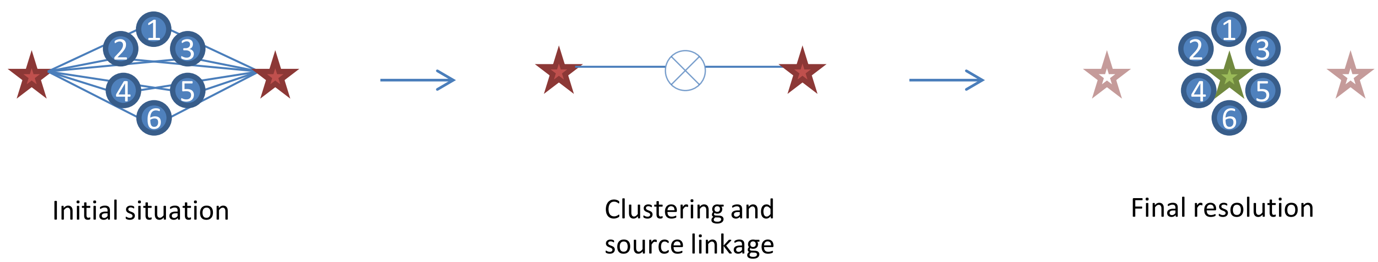
IDU Cross-Match in DR2

- Algorithm split into 3 different stages: Clustering, Cluster-Source linking and Resolution, plus a final consolidation process.

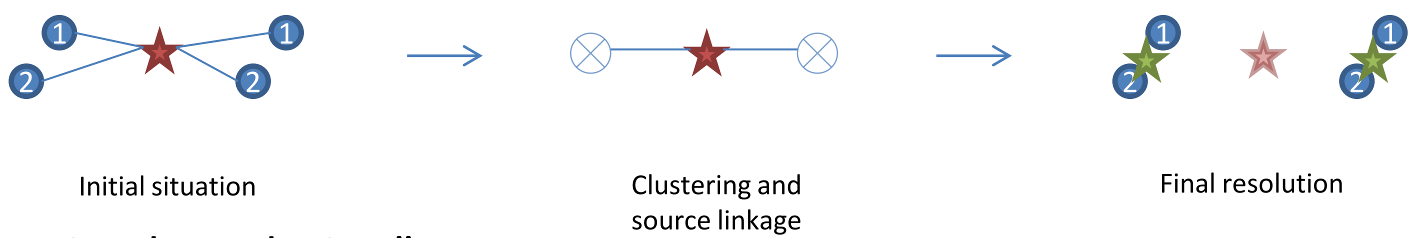


IDU Cross-Match in DR2

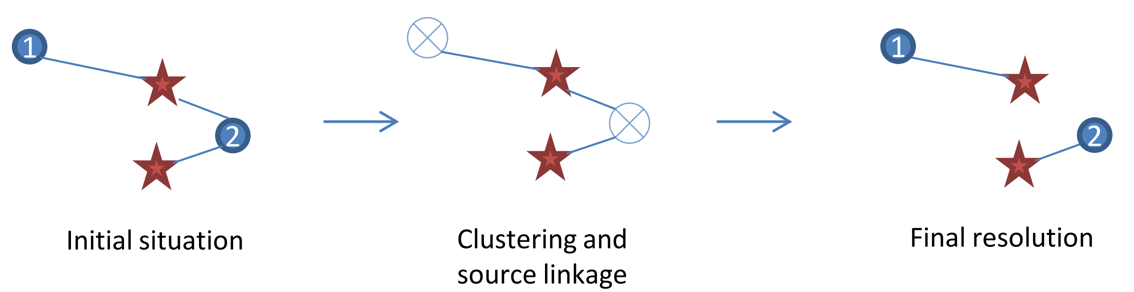
- Source merging:



- Source splitting:

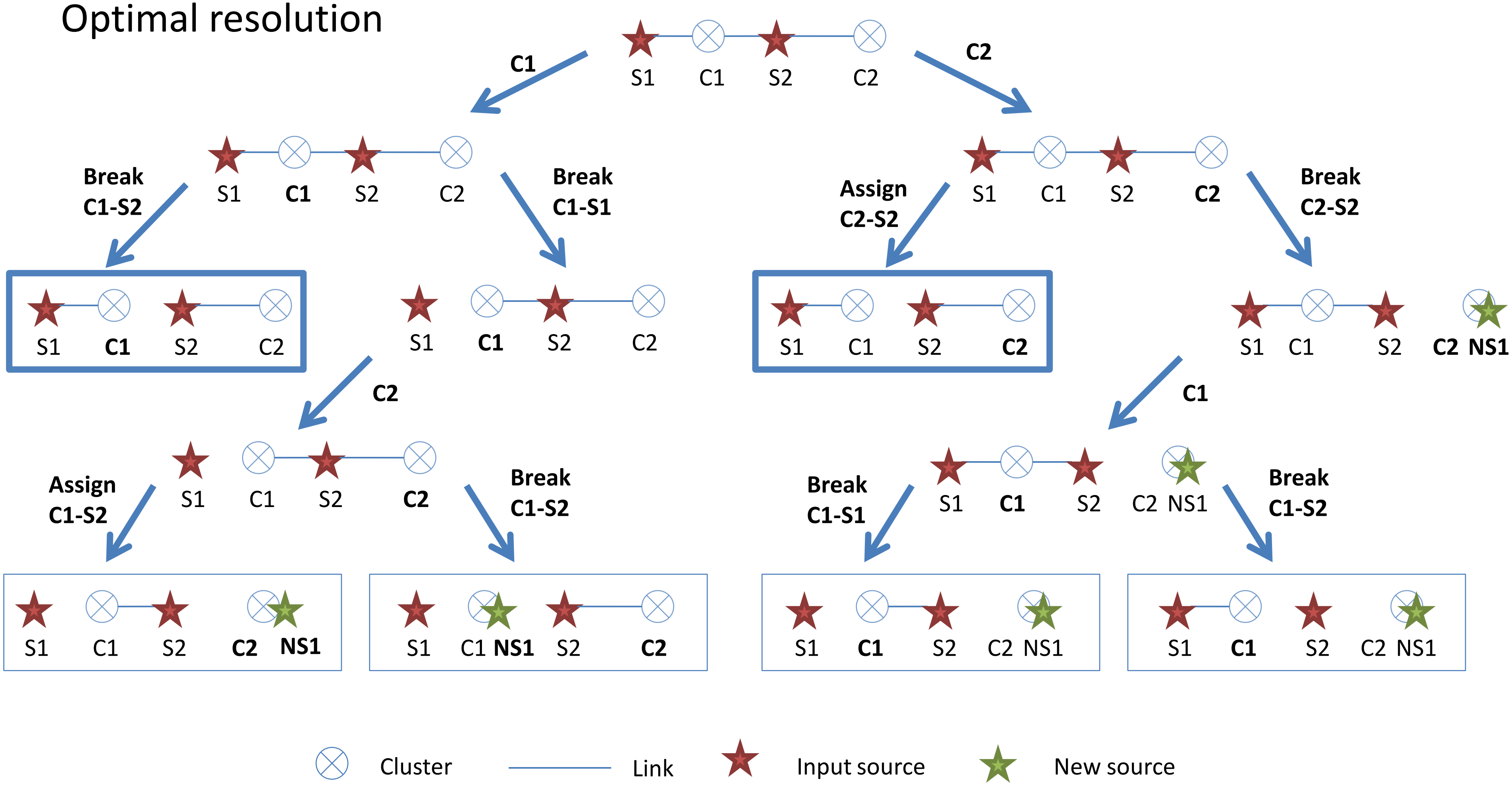


- “Optimal resolution”



IDU Cross-Match in DR2

Optimal resolution

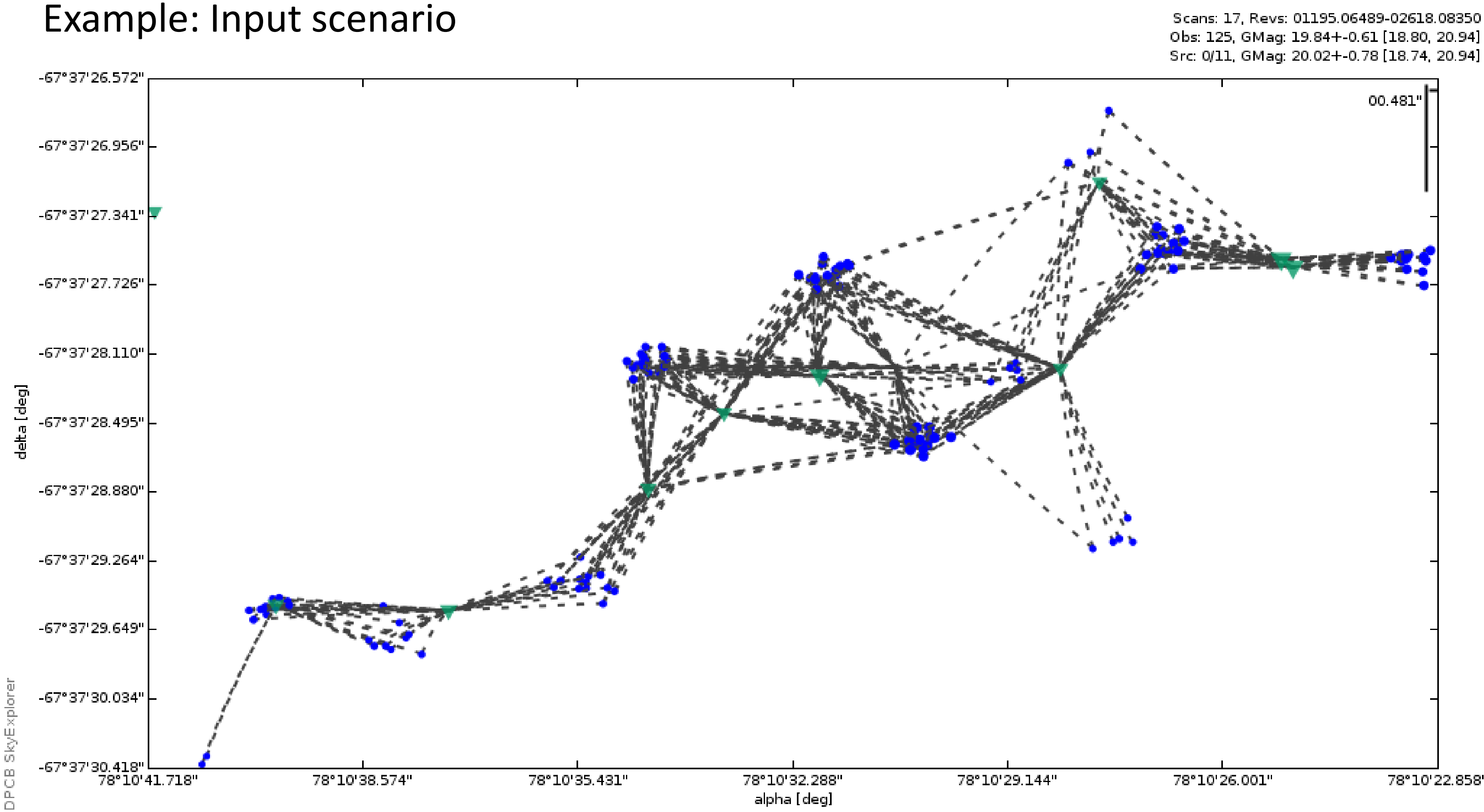


DAPCOM Data services

Gaia DPAC Data Processing & Analysis Consortium

IDU Cross-Match in DR2

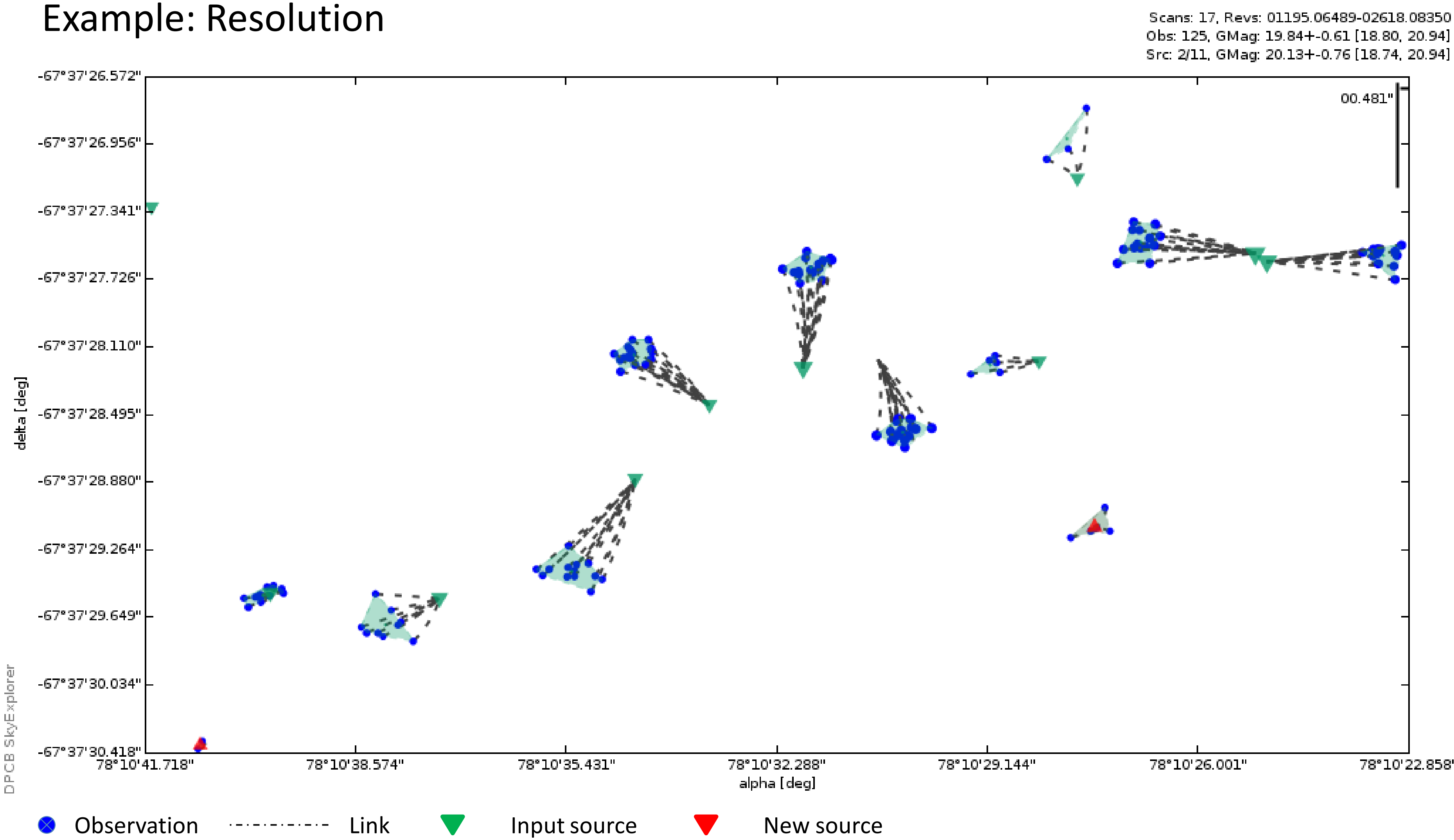
Example: Input scenario



Observation Link Input source New source

IDU Cross-Match in DR2

Example: Resolution



IDU CROSS-MATCH IN DR3

IDU Cross-Match in DR3

- The clustering algorithm model has been modified to take into account the proper motion

$$u(t) = u_0 + u_1 t$$

where u_0 is the mean position and u_1 is the proper motion.

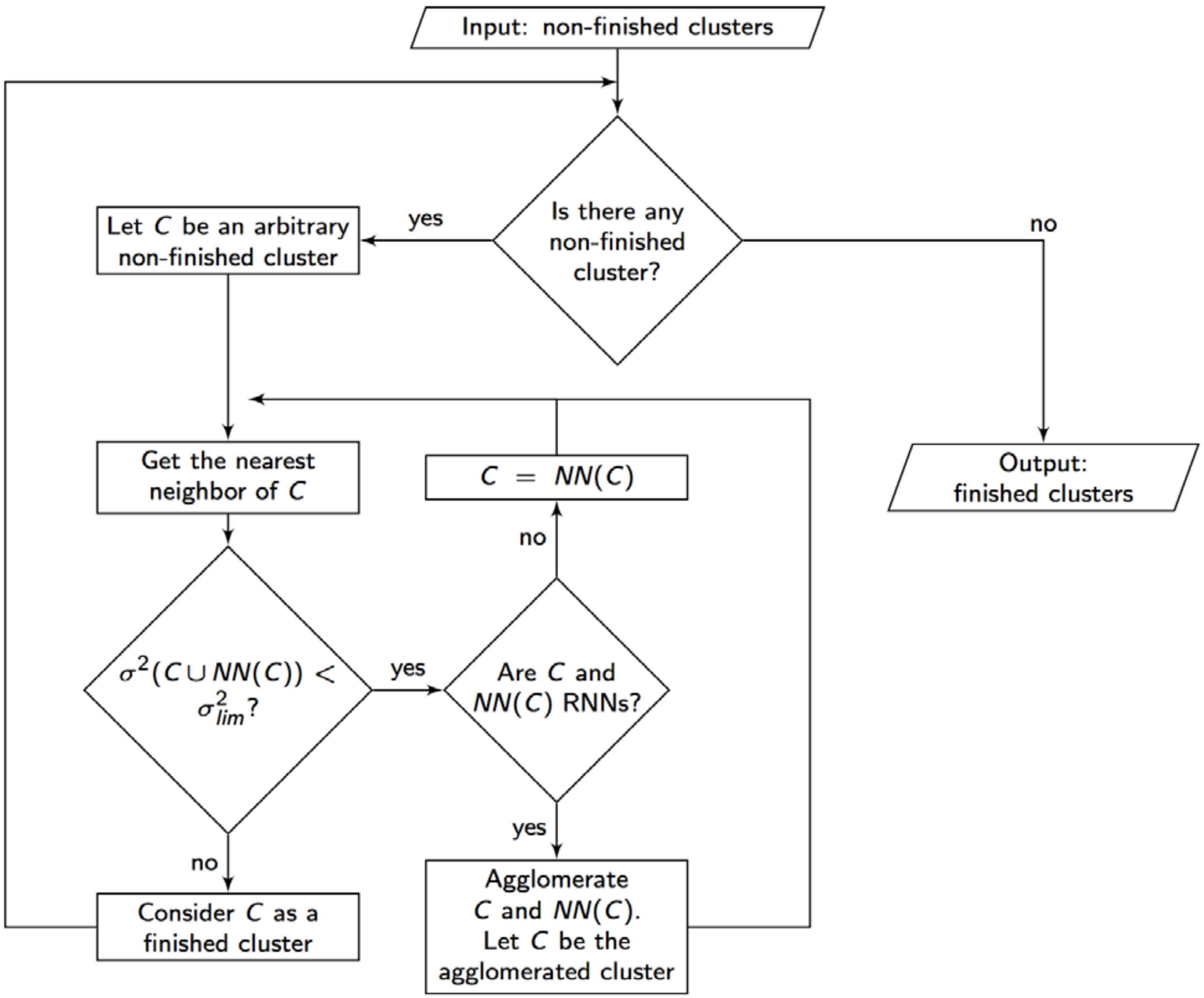
- Dissimilarity between clusters C_i and C_j will be defined as:

$$\Delta(C_i, C_j) = \begin{cases} \frac{n_i n_j}{n_i + n_j} \left\| \boldsymbol{x}(C_i) - \boldsymbol{x}(C_j) \right\|^2, & n_i + n_j \leq 3, \\ \Delta_\alpha(C_i, C_j) + \Delta_\delta(C_i, C_j) + w_m \frac{n_i n_j}{n_i + n_j} \left(m(C_i) - m(C_j) \right)^2, & n_i + n_j > 3, \end{cases}$$

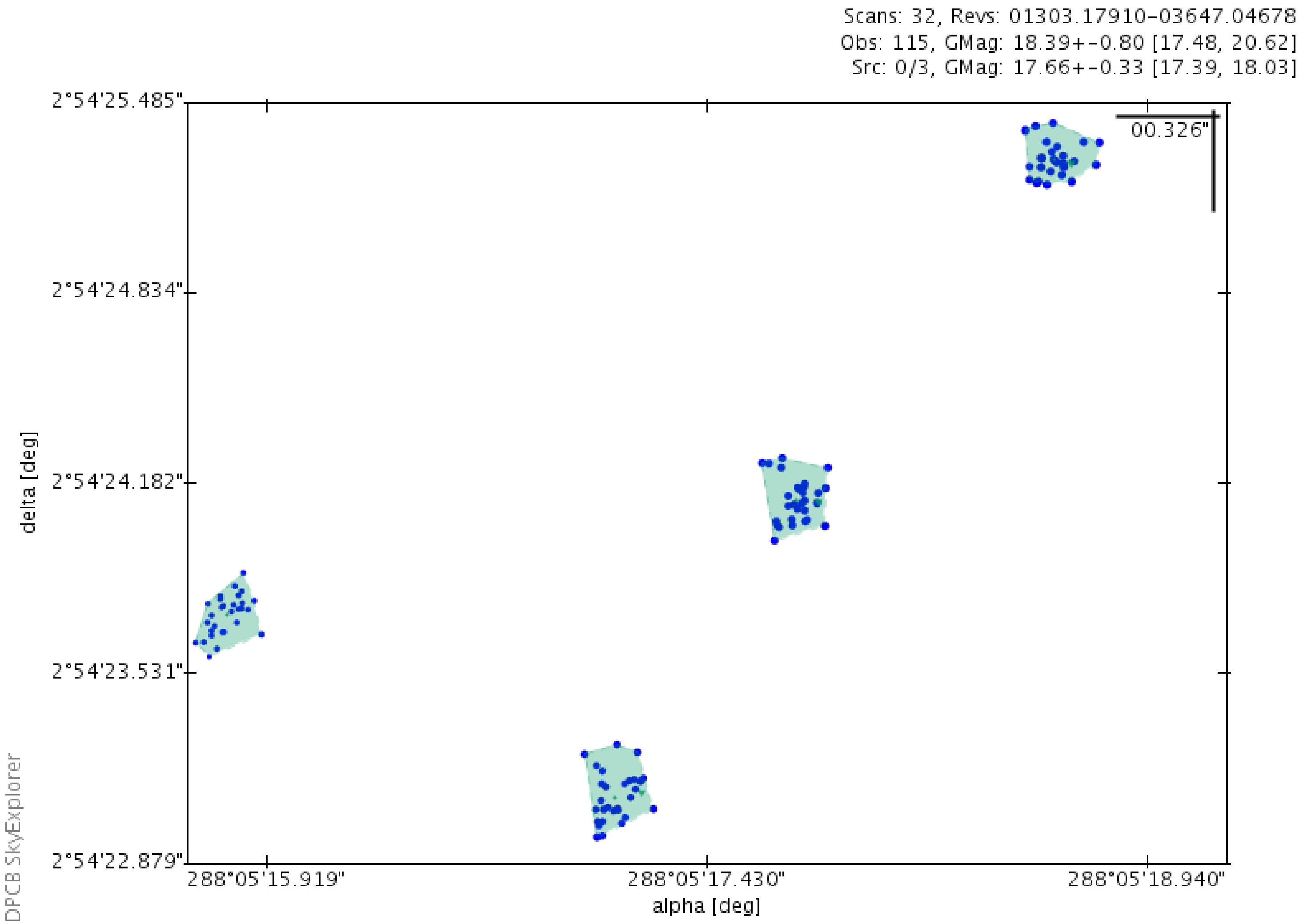
where:

- The proper motion model is only applied when enough samples are available
- $m(C)$ is the average magnitude of the cluster
- ω_m is a weight factor for the magnitude term
- Dissimilarity limit is determined for each cluster

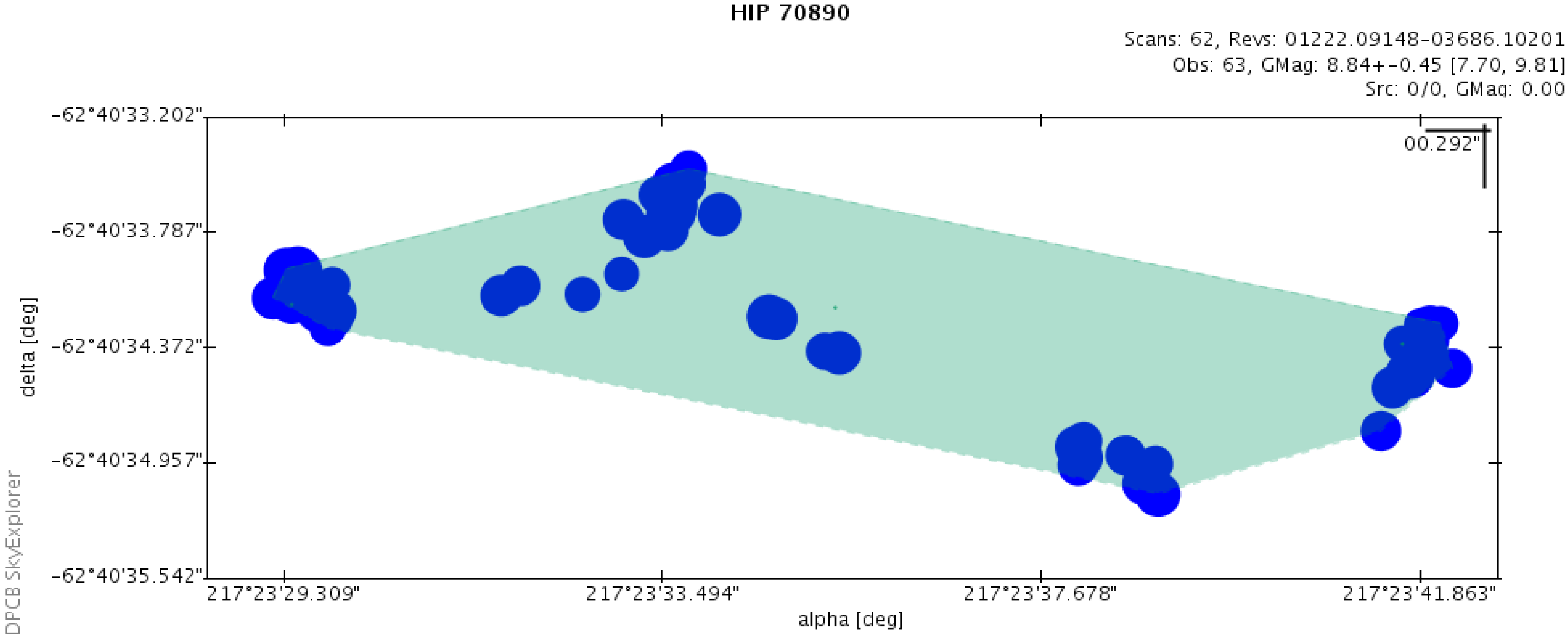
IDU Cross-Match in DR3



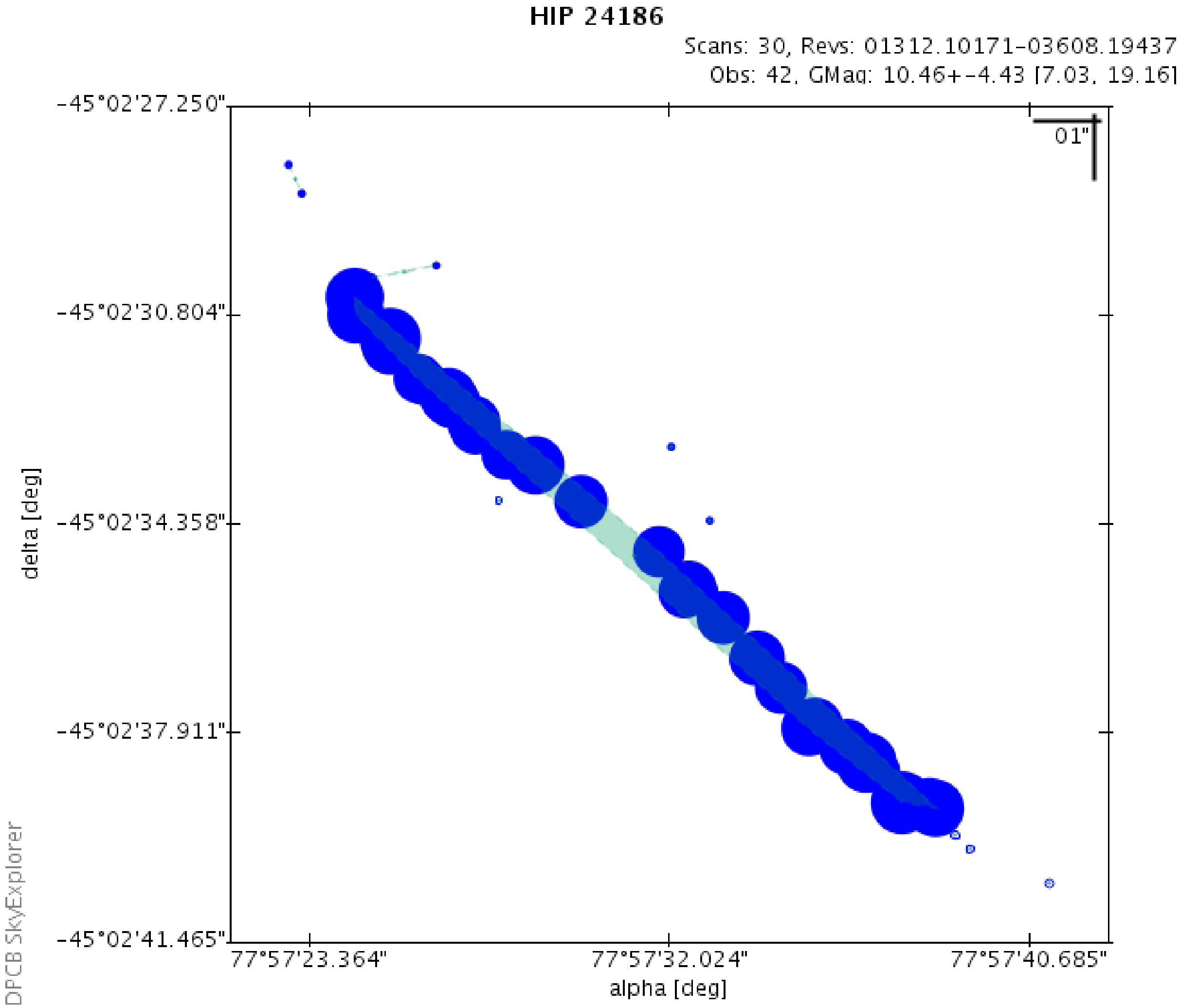
IDU Cross-Match in DR3



IDU Cross-Match in DR3



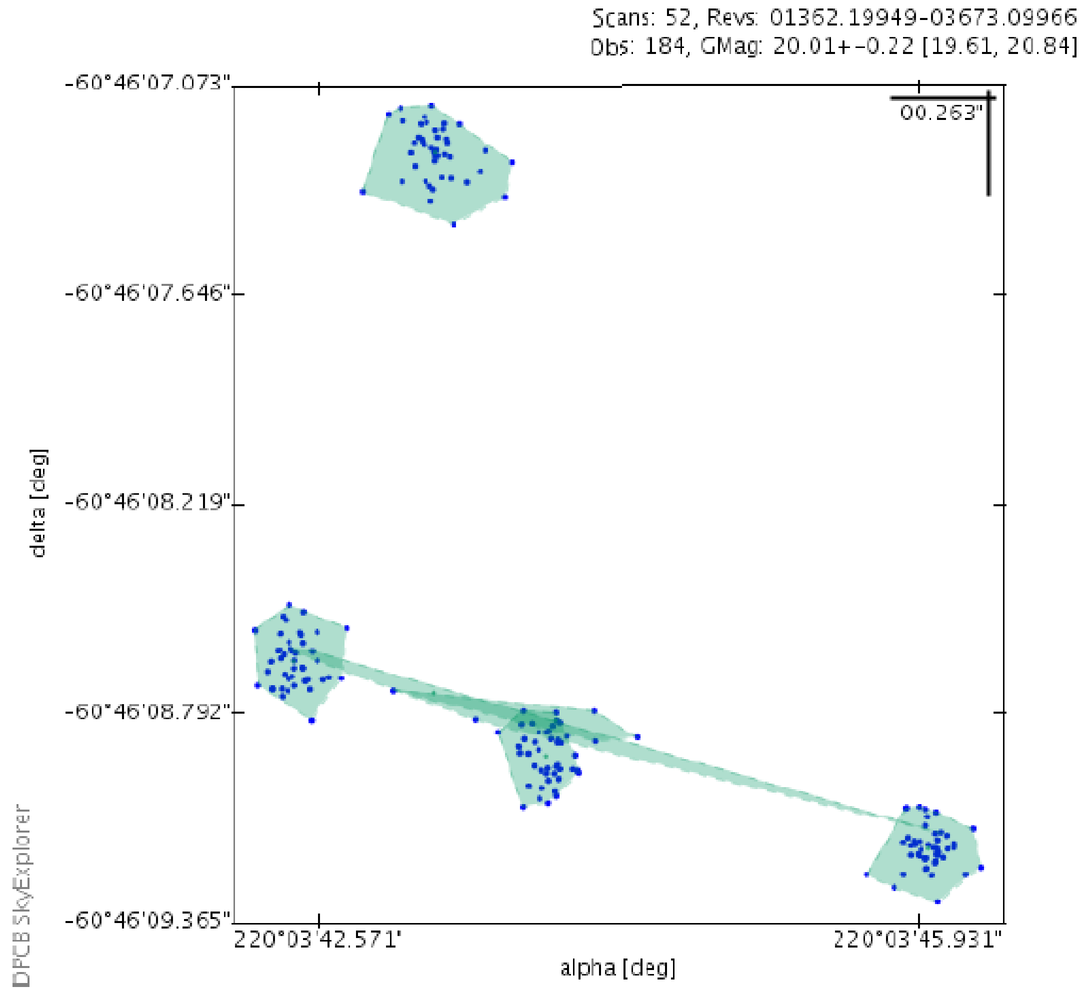
IDU Cross-Match in DR3



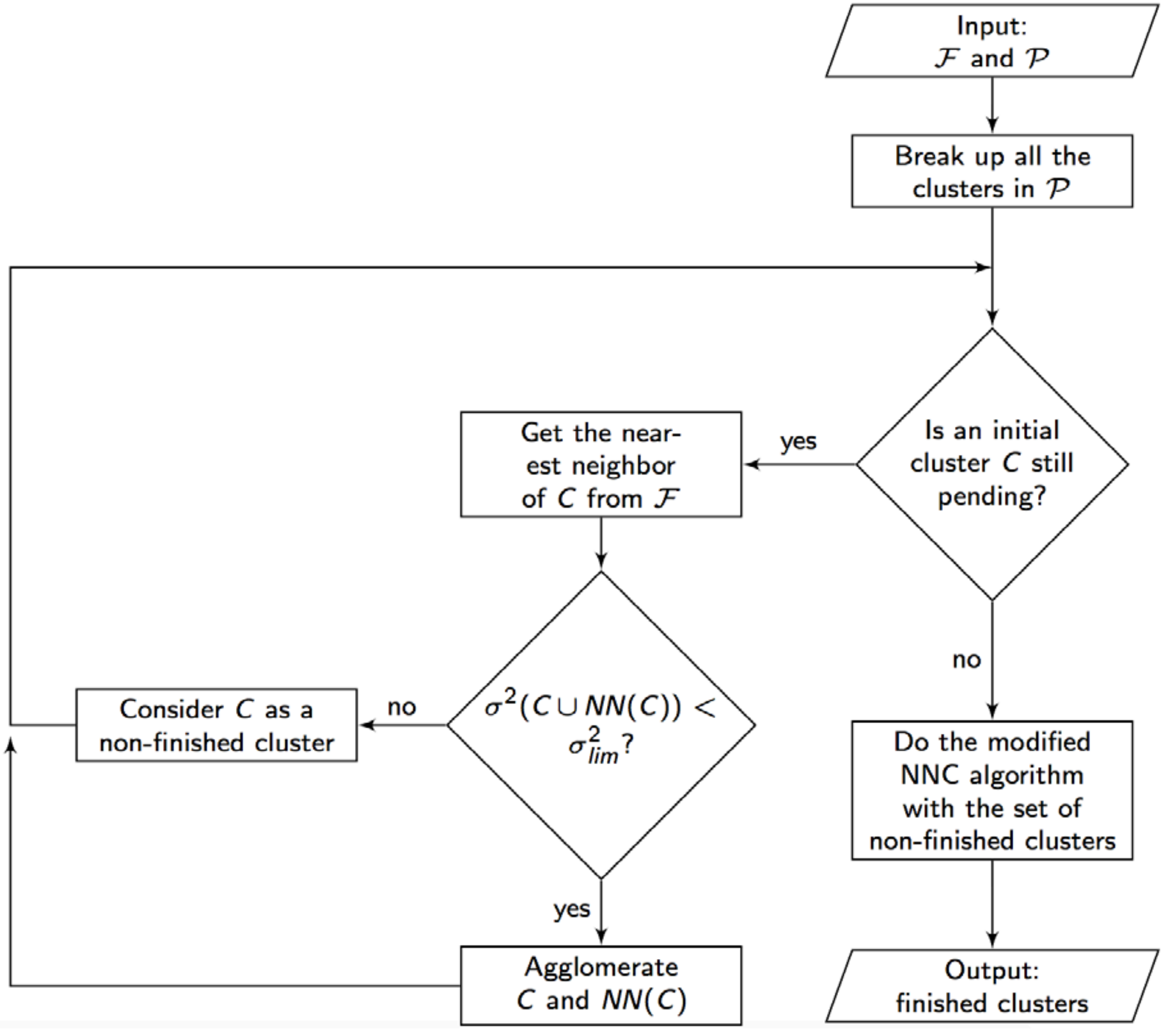
IDU Cross-Match in DR3

- The source model might fit inconsistent clusters:
  - Depending on the initial conditions
  - Unavoidable as the initial situation is always unknown
  - These clusters will fit very few observations
- We post process the cluster with few observations:

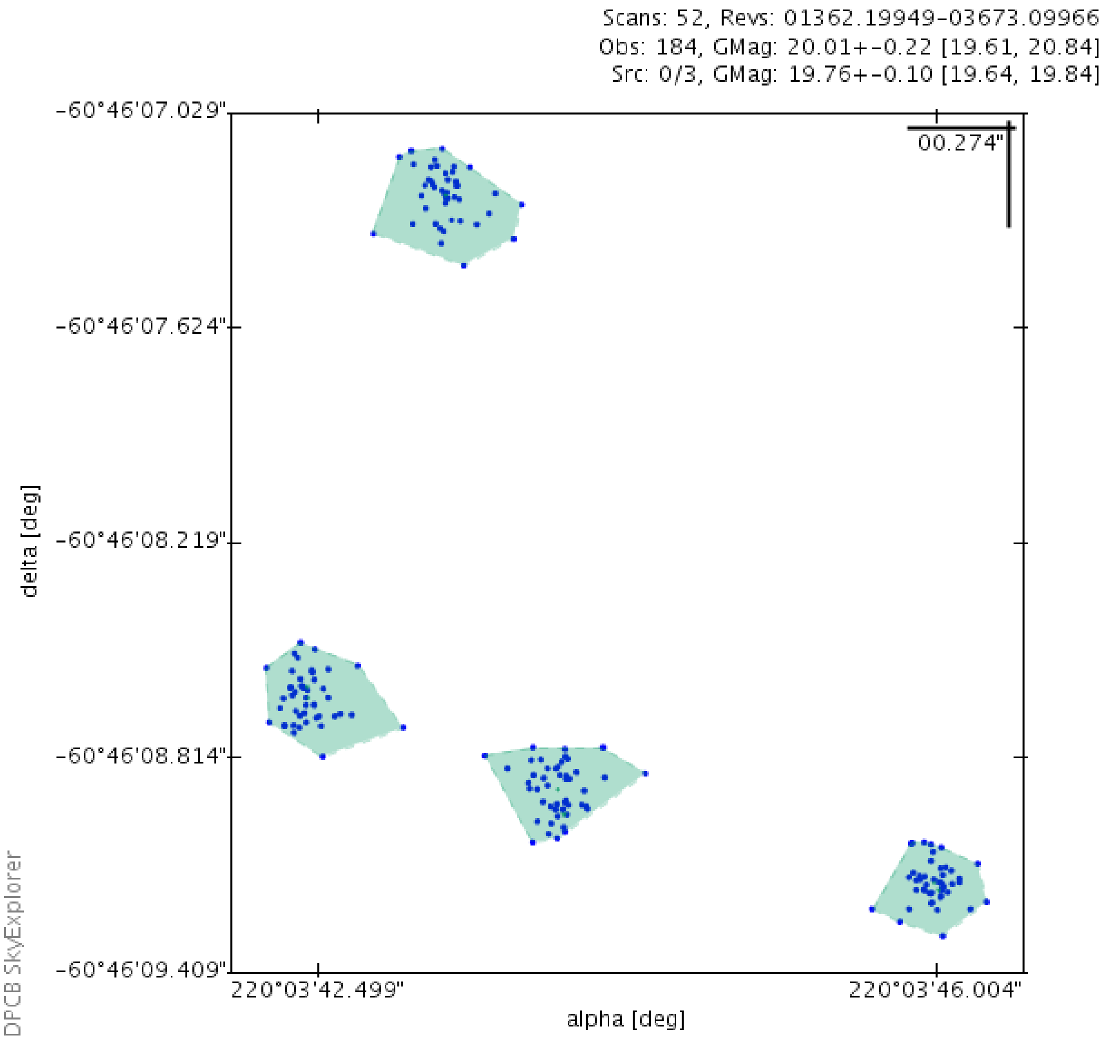
$$\mathcal{P} = \left\{ C \in \mathcal{C} \mid n \leq \frac{n(scans)}{2} \right\}$$



IDU Cross-Match in DR3



IDU Cross-Match in DR3



Thank you

Juan José González-Vidal, Ph.D.
Gaia DPAC IDT Manager and IDU XM Advisor