

Powering metadata-driven queries

António Dias, Israel Matute, Sara Madeira, Ciro Pappalardo, José Afonso

Introduction

The Atacama Large Millimetre Array, located on the homonymous Chilean desert, is the largest interferometric telescope in the world. Observing radiation in the millimetre and sub-millimetre regions of the electromagnetic spectrum, ALMA offers unprecedented resolution and image solving capabilities that have been explored by a diverse and ever-increasing scientific community studying extragalactic structures, dust cloud emissions, Solar System bodies and stellar formation, among others.

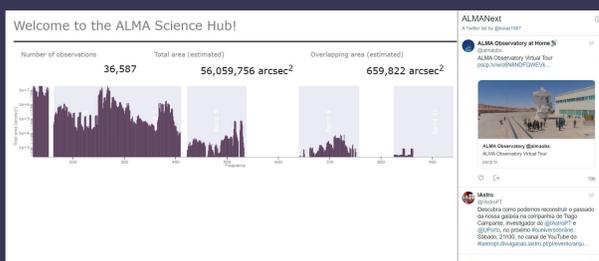


Access to ALMA data is primarily done through the [ALMA Science Archive](#), a web-based tool that allows both amateur and expert users alike obtain information on the observatory's +44000 observation set by filtering through their spatial location, frequency coverage, emission lines, sensitivity and so on. Despite offering a plethora of features and information pertaining to ALMA observations, it is understood that the Archive could be improved with external tools that, apart from increasing public engagement by offering a more robust and appealing visualization component, can also guide the user in new ways to tap into the hidden potential of ALMA data through combining overlapping observations. This poster describes an ongoing project that aims to develop such a tool.

The project

The web service, built on the Django framework, has been designed around two main tools, each facilitating data exploration on a specific level:

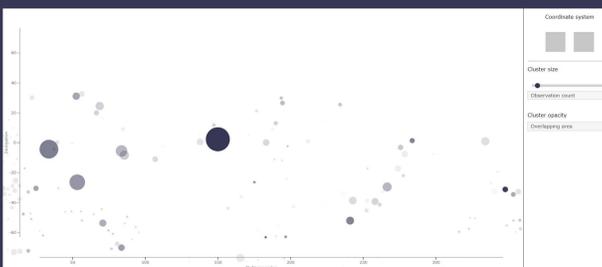
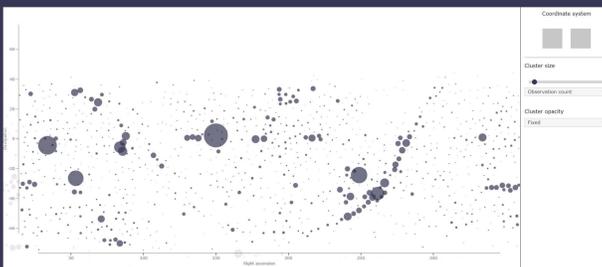
- Full-sky plots that highlight dense observation clusters and other trends
- Regional plots that analyse observations on a per-pixel basis, showing areas where observations overlap and their respective frequency coverage, combined sensitivity and other measurements.



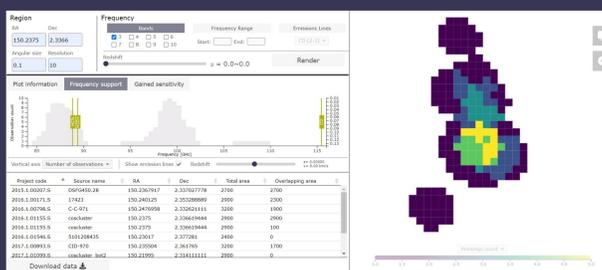
The landing page for the tool currently shows general archive information like observation count, total covered area and total overlapping area measured by the ALMA antennas. The histogram shows area coverage per frequency, which allows the user to quickly understand which spectral bands are the most observed by the facility. The right-hand panel contains a Twitter feed for both ALMA's and IA's accounts, which were added in due to their operational updates and scientific outreach publications.

Sky map

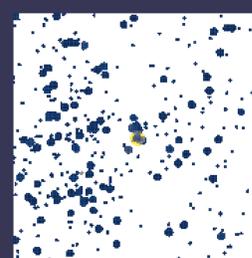
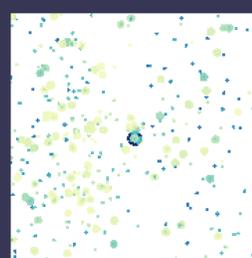
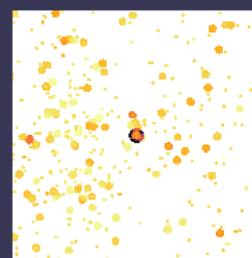
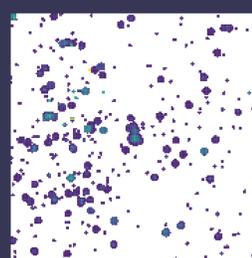
Aiming to support large-scale queries of the ALMA archive, the sky map page shows a scatterplot of all observations through location-based, multi-level data clustering.



Field plot



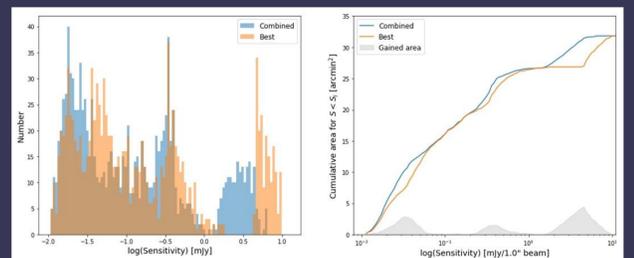
This tool was developed as a means to visualize specific regions or fields in higher detail through pixelization. Apart from spatial location, the user can query observations based on the desired frequency coverage at any redshift factor. The page provides a direct, user-friendly assessment of a field's overlapping observations, but also shows information regarding frequency coverage and estimated combined sensitivity. Multiple rendering modes can also be chosen:



Data combination

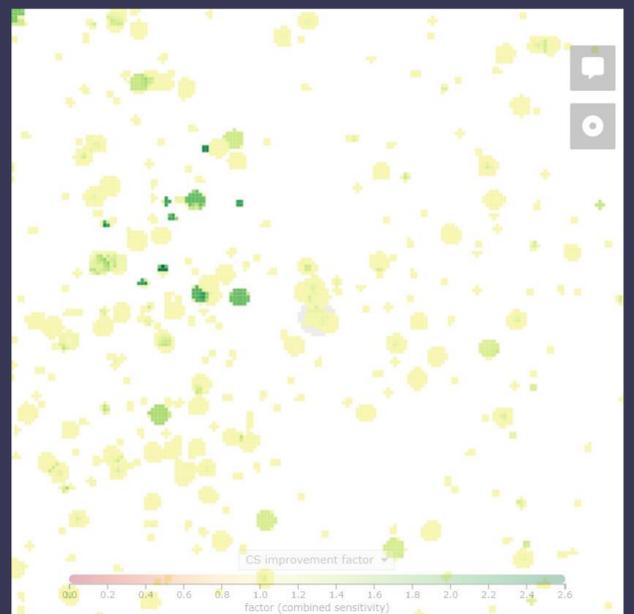
By combining two or more different observations, it is possible to obtain images with better sensitivity that each could provide individually. The line sensitivity of an observation can be approximated by a Gaussian function, the parameters of which depending primarily on the used arrays but also, for instance, the observed frequency.

It is important that the user is able to obtain an estimate of what can be achieved through data cube combination, which can be done just by resorting to the observations' metadata without need for the cubes themselves. The following graphs represent such an estimate for a region within the COSMOS field on band 3. Left side is the total number of pixels, and the right histogram highlights the gained area for different sensitivities:



Sensitivity improvement factors obtained for a pixel size of 10 arcsecs and beam sizes scaled to 1". Mosaic information included. Mean improvement of 1.8 ± 0.3 .

Both the sky map and the field plot tool were design to facilitate the data combination process by highlighting regions where overlapping observations can be found. The latter also estimates the sensitivity improvement factor on each pixel of the rendered region, in this case a 1 square degree section of the COSMOS field:



Regions highlighted in green offer the most potential for data combination, whereas yellow represents little-to-no gain in sensitivity - either because of fewer overlapping observations on those pixels or no overlap whatsoever. Also shown is part of the plot's own GUI.

Acknowledgements

IM and JA gratefully acknowledge support from the Science and Technology Foundation (FCT, Portugal) through the research grants PTDC/FISAST/29245/2017, UIDB/04434/2020 and UIDP/04434/2020