1. Title: The KMOS VVVX-GalCen Spectroscopic Survey

PI: M. Gómez, Instituto de Astrofísica, Universidad Andres Bello, Chile
Co-PI: F. Nogueras-Lara, Instituto de Astrofísica de Andalucía (IAA-CSIC), Granada, Spain

1.1 Background

Spectroscopic surveys of the inner Milky Way, which contains most of the stellar mass of our Galaxy, are challenging. Most objects in this region are either intrinsically red or strongly affected by dust extinction, and their spectral energy distributions peak in the infrared. However, wide-field multiplexed spectroscopic surveys in the infrared remain technologically limited. The KMOS VVVX-GalCen Spectroscopic Survey tackles this difficulty building upon the successful wide-area, multi-epoch VVV/VVVX near-infrared survey (Minniti et al. 2010, New Astronomy, 15, 433; Saito et al. 2024, A&A 689, A148) and the high-resolution GALACTICNUCLEUS imaging of the innermost regions of the Galaxy (Nogueras-Lara et al. 2018, A&A 610, A83; Nogueras-Lara et al. 2019, A&A 631, A20). The proposed diverse multipurpose survey aims to help the broad astronomical community, characterising a range of objects from eruptive young stellar objects, high and low mass stars and brown dwarfs, to various types of complex stellar systems–star clusters, associations and giant star-forming regions.

The wide range of scientific goals requires a flexible strategy to fully utilize the capabilities of KMOS, to cover the necessary spectral regions for each type of object, to carry out data quality control and data processing to fully Phase 3 compliant products. Here we describe our observing strategy and list the requirements on target preparation and calibration. We describe the quality control that we will apply to the raw data, and the data reduction process. We also present the team's organizational structure, defined to meet the obligation of carrying out this project.

2. Survey Observing Strategy

The KMOS VVVX-GalCen Spectroscopic Survey both requires and facilitates a flexible observing strategy because of the diverse nature of our science goals and targets. Therefore, we can easily comply with the P116 visitor mode (VM) allocation. The VM runs during the first year are very important, not only because they set the stage for the whole survey, but also because they help us to fine-tune the strategies making future runs more efficient. For these VM runs we will prepare an observing program with two goals in mind.

First, we would like to collect some observations for each of our science cases. This is important for refining some important points of the data reduction, such as sky subtraction and telluric correction. For example, we will investigate if we can achieve satisfactory sky subtraction across all three KMOS sub-spectrographs, or we have to keep the sky subtraction within each of them or even within each individual IFU. Achieving this goal will stimulate our sub-teams to brace for early data reduction and analysis and to address many unforeseen issues in both the data reduction and in the required human resources – both in FTEs and in expertise. Last but not least, these data may serve as a basis for early publications.

The second goal of the VM run preparation is to have ready backup programs that take into account a variety of observing scenarios, including different kinds of poor observing conditions and for partial instrument failures. For the former issue we will have backup OBs for brighter and isolated targets, that can be executed in cases of poor transparency and for poor seeing, respectively. These targets will come predominantly from the list of YSOs and other variables, because they are typically brighter and not as crowded, as the cluster targets, for example.

In order to successfully carry out this strategy (and the timely at-the-telescope quality control that is discussed in Section 6) we will need two observers during the VM runs, at least during the P116 period, and maybe in the future. We will start with more experienced observers, but we will also take advantage of these 2-member teams to train observers for the future VM runs. We can easily organize these teams because many members of the survey team are based in Chile.

2.1. Scheduling Requirements

We expect a balanced pace in data acquisition according to the total estimated time requested for each of the areas (see below). Optimal target assignment is often a major difficulty for KMOS due to the instrument features. Fortunately, one of the advantages of our proposal is that we have so many targets that this would not be an important issue. In all cases, we will have enough fillers (field M stars, field giants, background galaxies, variable stars, etc.) that would endeavor to leave no single arm empty. Below we outline as examples the observing strategies for some representative areas, namely:

(1) Mass Accretion on Young Stellar Objects: We will prioritise observations of high-variability (ΔKs > 1.5 mag) YSO candidates identified by the VVV/VVVX time series (~2500 sources). We will observe Class I YSOs with JWST observations (~100 sources) to draw a full picture of mass accretion/disc property. We will also include other Class I YSOs identified by cross-matching SPICY and VVV catalogues (~4,000 sources). Additionally, we include members from known young open clusters with existing optical and J/H-band spectra (~600 sources). In total, our YSO sample has 7,200 sources, equivalent to approximately 400 KMOS fields.

(2) The Galactic centre: This survey area will target approximately 40,000 stars in the Nuclear Stellar Disc (NSD), along with two control fields located in the inner Galactic bar. Observations will enable simultaneous spectroscopy of up to 20 stars per pointing. Covering the full sample will require approximately 2,000 pointings. We have defined 18 target fields that are evenly distributed across the survey area, which will serve as reference regions for our observations (see Fig. 1).

To correct for the variable sky background in the K band, we will obtain one sky observation per hour. Because of the extreme crowding in the region, sky frames cannot be taken simultaneously with the science data (except for a few rare cases when a dark cloud falls within the field for view) and will instead be acquired from a nearby dark cloud free of detectable sources (Fig. 2). To use sky observations obtained at different times, we will apply the *sky-tweaking* algorithm described in Davies et al. (2007) and implemented in the KMOS reduction pipeline. This algorithm rescales the intensities of the sky emission lines in the sky cube to match those in the corresponding science cube. We successfully tested this approach in a pilot study carried out for this science case (113.26DP.001 and 115.281U.001, PI: F. Nogueras-Lara). Figure 2 shows an example of a target reduced using two sky frames—one obtained immediately after and another two hours after the science observations. The resulting spectra were fitted with a new tool (*Schultheis et al., in press*), specifically designed for the analysis of Galactic centre spectra and optimised for the cool M giants that dominate this region (e.g. Nogueras-Lara et al. 2020, 2021). The derived line-of-sight velocities, effective temperatures, surface gravities, and α-element enhancements agree within the uncertainties for both reductions, demonstrating that sky frames taken at different times yield consistent results.

We estimate an observing rate of four science pointings plus one sky frame per hour, grouped within a concatenation container. Based on this strategy, the full target sample can be completed in approximately 500 hours. This approach has already been successfully tested in our pilot study during Periods 113 and 115.

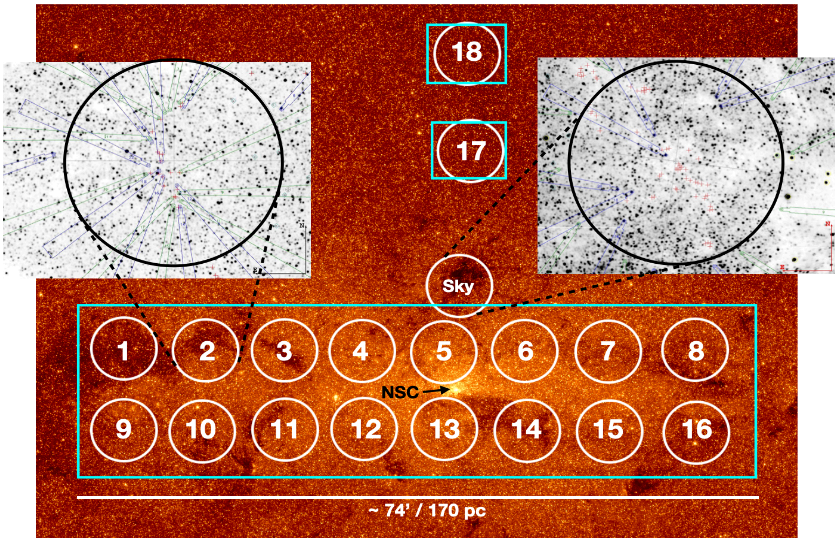


Fig. 1. Target regions (in cyan) and reference KMOS pointings (numbered 1 to 18) overlaid on a Spitzer 3.6 μm image. The sky region is on a dark cloud near the science targets. The zoom-in panels show the IFU configuration obtained using KARMA for a science and a sky pointings. The position of the nuclear star cluster (NSC) is indicated in the image.

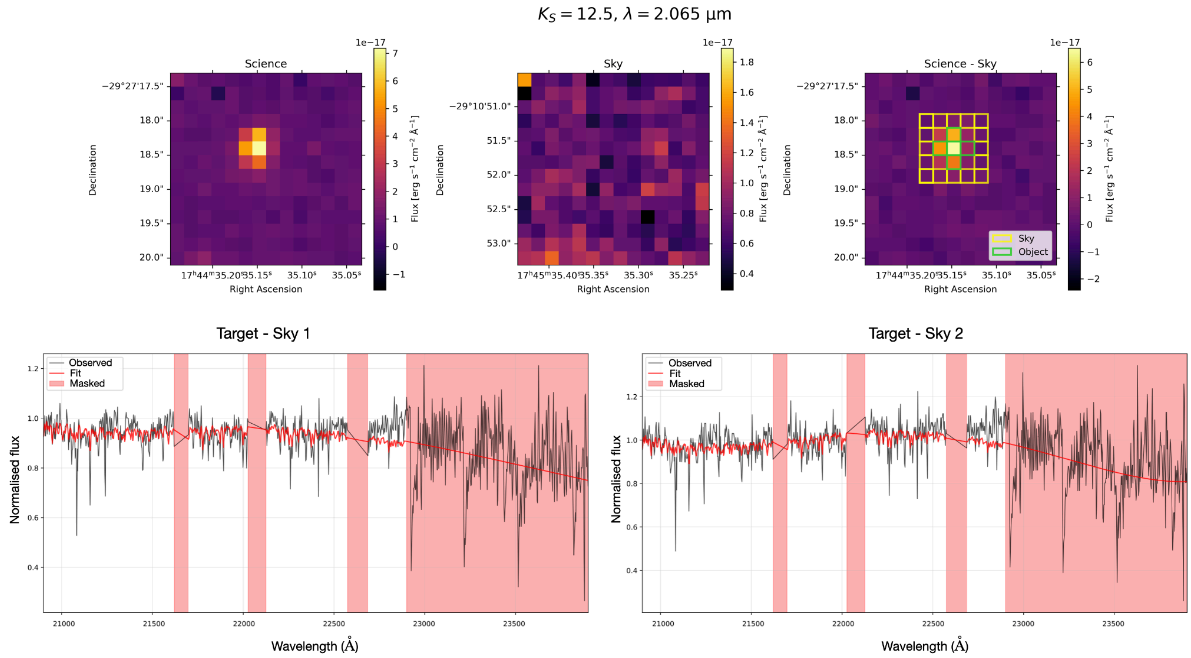


Fig. 2. Preliminary analysis of one of the stars in our pilot study (P113; Nogueras-Lara et al.). Top: target with *Ks*=12.5 mag (left), example of a sky frame observed right after the science exposure (middle), and science frame after sky subtraction, with the regions used to extract the spectrum and background indicated (right). Bottom: extracted spectra using the sky frame shown above (left) and another obtained two hours after the science observations (right), together with their best-fitting models.

Our pilot study for this science case also allowed us to refine the observing strategy. Given the physical limitations in assigning all available KMOS arms to science targets—especially after multiple pointings on the same field—we plan to re-observe a subset of stars to ensure that no arms remain unused and to fully exploit the instrument's capabilities, introducing a time-domain component to the survey. To implement this strategy, we will carry out a preliminary analysis after each observing run to identify the most suitable targets for follow-up and to build a sample for multi-epoch spectroscopy. Visitor mode observations will provide the flexibility required to optimize repeat observations, while service mode runs will follow a system of priorities to guarantee homogeneous coverage of the survey area and to maximise the use of KMOS arms in each period. After every service run, a preliminary analysis will be performed to select the best candidates for multi-epoch follow-up, which will then be incorporated into subsequent field pointings, consistent with the proposed cadence for OB preparation and submission.

Ideally, we aim to obtain multi-epoch spectra of young massive stars (a few Myr old) to investigate their binary fraction via potential radial velocity variations. Additionally, combining all available spectra for each source will improve the S/N, enhancing our ability to detect and characterise their typically weak spectral lines.

(3) Young Embedded and Massive Star Clusters in our Galaxy: One of our primary goals is clarifying the nature of young embedded (possibly massive) cluster candidates from the lists of Borissova et al. (2018, 2020) and Bica et al. (2019) to trace the recent star formation in our Galaxy. Thus, we selected 52 star cluster candidates on the basis of their J-K_S vs K_S color magnitude diagrams from VVVX photometry and Gaia proper motion (when available). The projected angular diameter of VVV/VVVX and Bica et al. (2019) clusters is between 0.3 and 4.3 arcmin, but a significant fraction of them may have formed in the same parent clouds or at least are projected on the same place in the galactic plane. Thus, we can supply with science targets most of the 20 configurable arms within the 7.2 arcmin KMOS field; any remaining arms will be filled with suitable YSOs or variable stars. An example of such a group is shown in Fig. 3 where the left panels show the group and the right panel the objects targeted for observations.

We are planning the following strategy: In P116 and P117 we will observe in K band most of the selected star cluster candidates with S/N at least 30, necessary to spectroscopically classify the stars, measure the radial velocities and fit the profiles of some spectral lines. In P118 we will finish the observations of the remaining ones. After reducing and analyzing the data in P118 and P119 we will re-observe in K and possibly in H bands some of the most massive cluster candidates to include more cluster members and obtain the fundamental cluster properties such as cluster mass, age and [Fe/H].

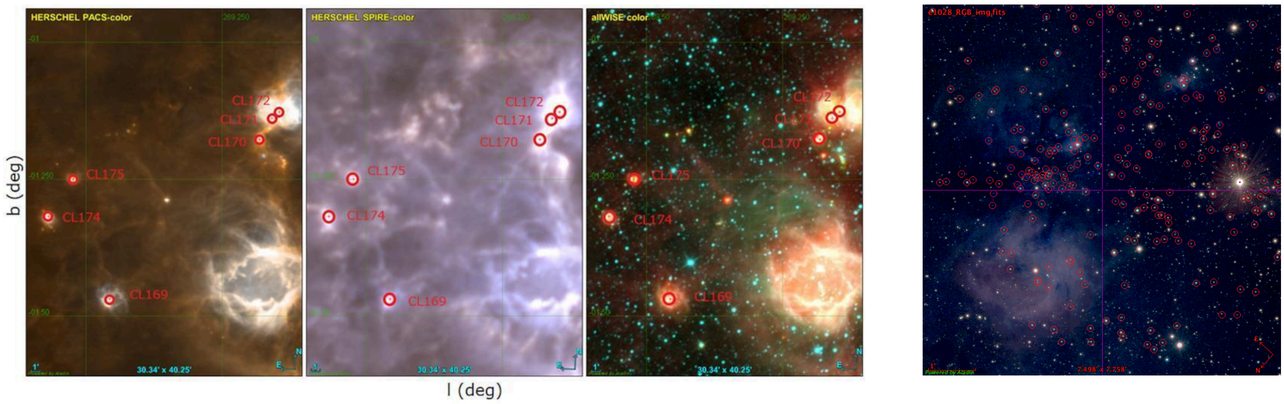


Fig. 3. From left to right: Herschel PACS, Herschel SPIRE, allWISE 3-color composites with overplotted groups of clusters (Borissova et al. 2020), and the objects targeted for observations. North is up, East is to the left.

The DIT and NDIT will be adjusted for every OB, but considering the K band magnitude interval (Table 4 and Fig. 10a), the exposure time calculator gives typical exposure times of 600 sec and thus the estimated execution time is 30 min per field. Some fields will be observed with 2 KARMA configurations to cover enough objects.

Fields for each of the Science areas:

1. Mass accretion on Young Stellar Objects: 400 fields H, K, some IZ, YJ
2. Hidden Galactic Globular Clusters: 100 fields H, K, H+K, some IZ, YJ
3. Free Floating Planets and Brown Dwarfs: 100 fields IZ, YJ, H+K
4. Young Massive Star Clusters: 100 fields H, K, H+K
5. Galaxies and Large-Scale Structure in the Zone of Avoidance: 400 fields H, K, H+K
6. The Galactic Centre: 18 fields K (500h, multiple pointings per field required to cover all targets).

GRAND TOTAL: 140 nights

Cadence for OB submission:

Taking into account the recommendation described in the guidelines about the performance of KMOS, we will submit OBs prior to every Phase 2 deadline and maintain a monthly cadence for SM observations (see Table 3). We aim for steady progress across our science goals, with first-year observations leading to high-impact publications, promoting the survey. For SM observations, we will ensure that a sufficient number of OBs are ready at any time to cover the subsequent 2–3 months, in agreement with the Support Astronomer. OB submissions will be validated using Phase 2 too, for observations that will be carried out in service mode. We will contact the KMOS instrument scientist well in advance to discuss possible improvements and optimisations to the proposed cadence and OBs. Our team includes members with extensive KMOS experience, including former and current KMOS instrument scientists. For VM runs, we will prepare more OBs than strictly needed to allow flexibility in case of unpredictable KMOS pick-off arm performance. For each SM run, we will prepare the corresponding KARMA files containing the information on the guide stars and reference stars. Any necessary additional information will be supplied as requested by the Support Astronomer to better prepare for the SM runs.

Serendipity:
We are all actively working in these areas of research, and are aware that in the immediate future, some unexpected new discoveries would be made - a likely expectation given the multitude of on-going time domain surveys. Therefore, our team plans to incorporate a flexible component, allowing a spectroscopic follow-up of new targets. These would be few, but potentially high-impact observations and it is therefore expected they would have a limited impact on the scheduling/observing plans. We will submit any new targets to ESO for approval, to avoid conflicts with other programmes.

2.2. Observing requirements

Our OBs will have diverse parameters, depending on a combination of target brightness, weather conditions and required set up. For the VM runs we plan to exceed (if possible, taking into account considerations such as airmass limit and change during the execution, necessary observations of telluric standards and sky fields, etc.) the 1-hr OB execution time that is adopted for service mode observations, saving acquisition time and increasing the overall efficiency. This will save us some execution overheads and improve the observing efficiency.

Target distribution and observability:

As it can be seen from the RA distribution for the different Science Areas (Figs. 7, 8 and 9), we will be strongly concentrated on the MW bulge, observable roughly from the first to the third quarter each year. Hence, a paced strategy shall be followed. The allocated observing runs will be populated with targets selected from the corresponding scientific areas, ensuring that each run contributes in proportion to our science goals, so different sub-teams will be supplied with a continuous regular stream of data. Target selection will be guided by scientific priority, feasibility, and visibility during the scheduled observation windows, thereby maximizing the scientific return of each run.

KMOS Arms occupation:

We have extensively used the KMOS Arm Allocation Tool to simulate the number of fields required for each Science Area (e.g., members of the galactic globular cluster VVV-CL160 and galaxies in a background galaxy cluster that we discovered in the zone of avoidance, as shown in Fig. 4, left and right, respectively).

Given the heterogeneous astrophysical nature of all objects in this study, each class of targets will need to be properly sampled from different datasets, spanning the full magnitude range. As an example, for the study devoted to the Galactic centre, the primary targets will be stars within the range $12.5 < K_S < 13$. The survey will also include objects outside this nominal magnitude range that are of particular scientific interest, such as variable and high-velocity stars. These targets may be identified during the course of the public survey and are essential for a comprehensive understanding of the Galactic centre.

Our calculations indicate that around 2,000 KMOS pointings will be needed to observe our ~ 40,000 target stars. Based on the KMOS exposure time calculator, we estimated a total exposure time of 70s to achieve the required signal-to-noise ratio (S/N$\gtrsim$50). Additionally, sky observations will be required once per hour, using the same configuration as the science targets. These will be conducted on a dark cloud located ~ 15′ from the target stars (Fig. 1). Thanks to the brightness of our targets and the exclusion of stars with close neighbours within ~ 0.6″, our observations will not be affected by saturation or persistence effects, even in excellent atmospheric conditions. Using the p2 tool, we estimated the total observing time, including overheads. We determined that four science configurations plus one sky observation, grouped in a concatenation container, can be completed within one hour (DIT=10s, NDIT=7). This setup allows us to observe approximately 80 stars per hour, requiring a total of 500 hours to cover all the targets. As outlined in Sect. 1, if some KMOS arms cannot be assigned to science targets, they will be used to obtain multi-epoch spectra of selected previously observed objects.

For Science Area 5, which focuses on galaxies and large-scale structures in the Zone of Avoidance, a more flexible and adaptive observational strategy is required. Approximately 16% of the target sample is located in fields that can potentially utilize up to 80% of KMOS's field arms. In these high-occupancy fields, integration times suitable for detecting galaxies with *Ks*<15.5-16.5 mag can be employed effectively. In addition, a higher S/N can be reached by observing faint targets continuously throughout the whole nodding cycle.

However, the situation becomes more complex in lower-density regions. In about 80% of the sample, the field occupation drops to around 50%. In such cases, longer integration times pose the risk of saturating bright targets, particularly YSOs (Science Area 1) with an average K_S~14.1 mag,bright stars associated with obscured globular clusters (GCs) and young massive clusters (YMCs), which correspond to Science Areas 2 and 4, respectively. Data integrity will be kept by reducing the detector integration times and increasing NDIT, at the cost of slightly increasing the overheads.

Candidates to free-floating planets and brown dwarfs in our sample will be challenging both in wavelength coverage as well as luminosity. Some of them have K_S~14.0 mag and most are 16<K_S<17 mag. For the latter, the following integration times are proposed:

- Grating HK, INT = 15x120s = 1800s (SNR@2.14um = [41, 43, 50] for each Teff respectively)
- Grating YJ, INT = 3x600s = 1800s (SNR@1.18um = [47, 38, 40] for each Teff respectively)
- Grating IZ, INT = 3x1200s = 3600s (SNR@0.92um = [39, 32, 32] for each Teff respectively)

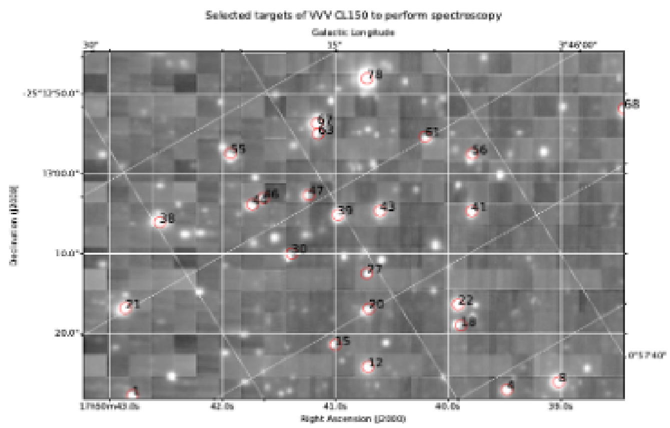
Instead of defining a fixed compromise between integration time and field occupancy across all Science Areas from the outset, we propose a flexible approach. The survey will begin by targeting fields that can be predominantly dedicated to individual Science Areas during the first two observing cycles (P116 and P117) to be carried out mainly in visitor mode. This will allow us to evaluate the scientific return, assess KMOS's actual performance under varying conditions, and identify common observational challenges, especially in terms of S/N under different weather scenarios.

Following this initial phase, we will be in a stronger position to refine our strategy. The insights gained will enable us to progressively relax the separation between Science Areas, eventually allowing for optimized overlaps where appropriate. Although we are currently modeling the expected field distribution and refining the field selection accordingly, the actual allocation across Science Areas will only become clear as observations proceed during the first two cycles.

Observing conditions needed (depending on area):

We prefer reasonably good seeing (<1") in order to better exploit the instrument's capabilities, but can accommodate a few bad seeing nights to observe the brightest targets. Figure 5 illustrates the expected quality of the science observations in the Science Area 6 of this public survey, based on Martínez-Arranz et al. (2023), which were obtained under comparable conditions (DIMM ≈ 0.7) and with a similar strategy to that planned for the Galactic centre fields (sky observed in a dark cloud at a separate pointing, with a slightly different DIT and the use of mosaic mode). Also we are flexible with respect to moonlight, provided that the Moon is located at least 15 degrees from the target fields, a constraint that can occasionally affect bulge observations. This minimum Moon–target separation should be taken into account during scheduling to avoid the allocation of an excessive number of nights, particularly in visitor mode, where the Moon being too close to the Galactic centre would prevent observations of this part of the survey. Thin cirrus is also tolerable, given that we are working in the near-infrared. Only the standard calibrations provided by the observatory are required.

By combining the observing strategies for the six Science Areas, we estimate the following time distribution across target fields. This layout also ensures sufficient spatial coverage to facilitate telescope scheduling based on target coordinates.



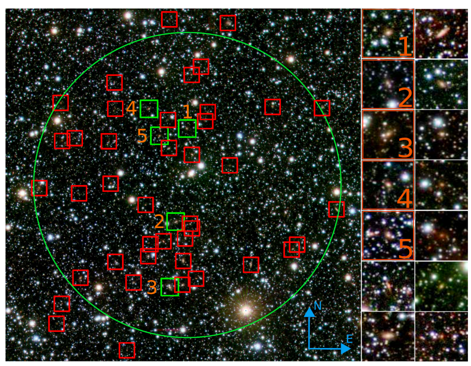


Fig. 4. Left: KMOS targets in the field of VVV-CL160 cluster candidate. Credit: Álvaro Valenzuela. **Right**: False-colour *Z* (blue), *J* (green), and K_S (red) image of a region corresponding to the galaxy group/cluster candidate. The green circle delimits the FOV of the instrument. The green squares show the five galaxies previously observed with Flamingos-2 at Gemini South and the red squares are the interesting targets. In the right side, we zoomed in on some galaxy candidates within the studied area. Each box measures 20” per side.

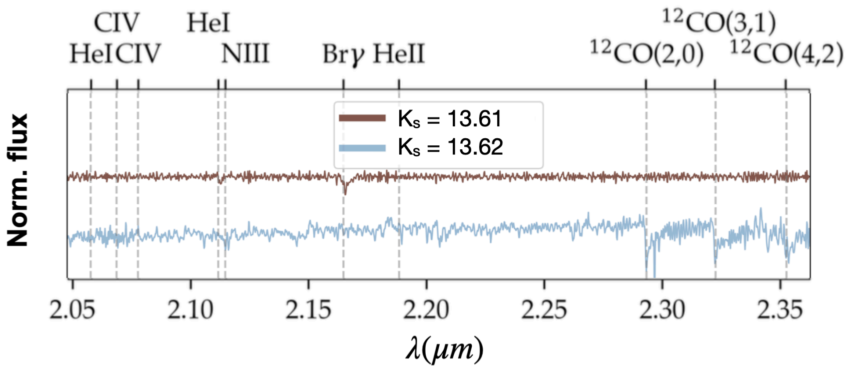


Fig. 5. KMOS spectra of a hot young star (brown) and a cool star showing CO band-head absorption (blue), adapted from Martínez-Arranz et al. (2023).

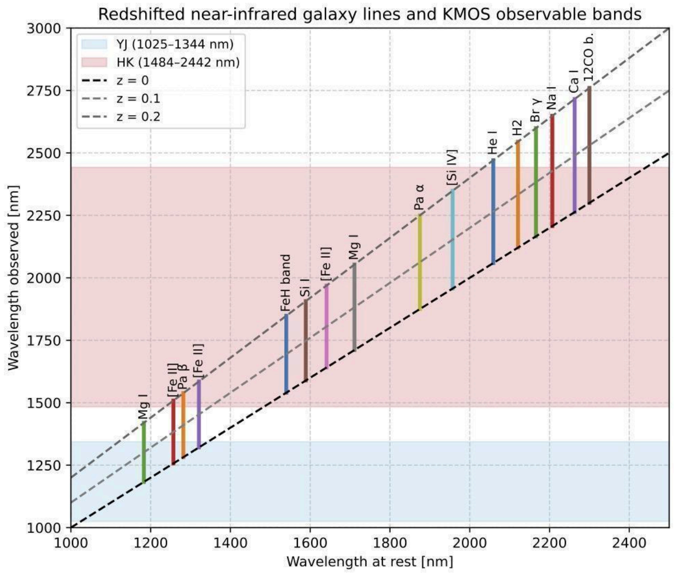


Fig. 6. Most of the interesting spectral features will be covered by the HK band for galaxies in the Science Area 5 within z<0.2.

Table 2: Scheduling requirements for Visitor Mode runs

Period	Start/End date	VM fraction	Total number of nights (VM/SM)	Number of VM runs	Average run length	Requested months
P116	01-10-2025 to 30-04-2026	1.00	10 (10/0)	2	5	mar-apr
P117*	01-05-2026 to 30-04-2027	0.70	45 (32/13)	6	6	may-aug
P118*	01-05-2027 to 30-04-2028	0.50	45 (23/22)	4	6	may-aug
P119*	01-05-2028 to 30-04-2029	0.30	40 (12/28)	2	6	mar-oct

Table 3: Cadence for the submission of the OBs and containers for SM runs

Period	Frequency biweekly/monthly etc.	Number of OBs	Average OBs properties
P116			OB properties will depend on Science Area (please check text)
P117*	monthly	500	
P118*	monthly	900	
P119*	monthly	1200	

Table 4: Observing requirements for Service Mode

Period	Requested time (hrs) Including overheads	Exp. Time (hrs)	Mean RA or RA range	Priority	Moon	Seeing (“)	Spectral bands	Transparency	Containers
P116	0	0	8h	1	Full	0.6 -1.0	H, K, H+K, IZ, YJ	Thin	#concat #group
P117*	105	60	08-10 10-18	3 2	Full	0.6 -1.0	H, K, H+K, IZ, YJ	Thin	#concat #group
P118*	180	100	08-18	3	Full	0.6 -1.0	H, K, H+K, IZ, YJ	Thin	#concat #group
P119*	225	125	08-18	1	Full	0.6 -1.0	H, K, H+K, IZ, YJ	Thin	#concat #group

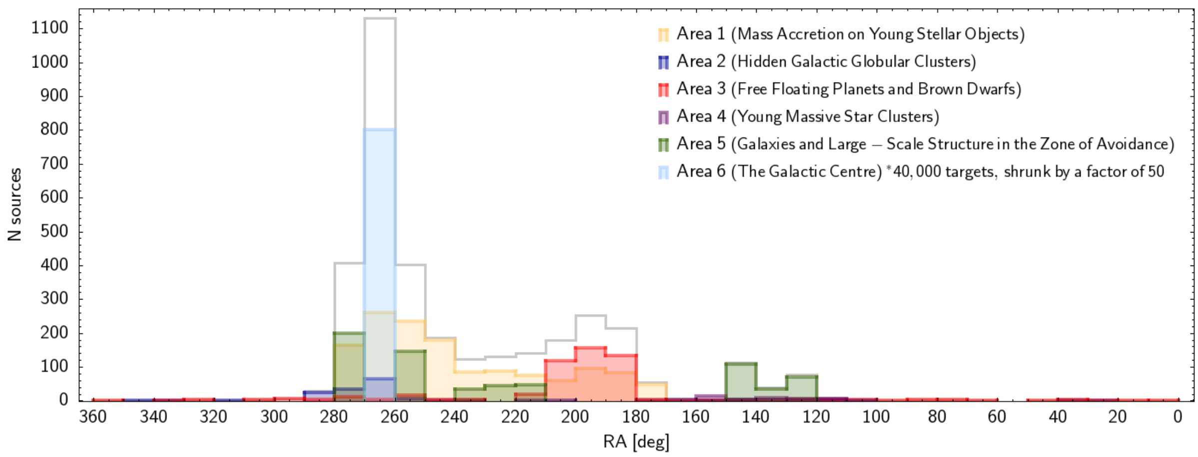


Fig. 7. Targets for high-priority observations from areas 1 to 6, distributed along RA. Area 6 has a total of 40,000 targets. Its histogram has been shrunk by a factor of 50 for visualisation purposes. The need for a higher allocation of nights to accommodate targets close to the MW centre is evident.

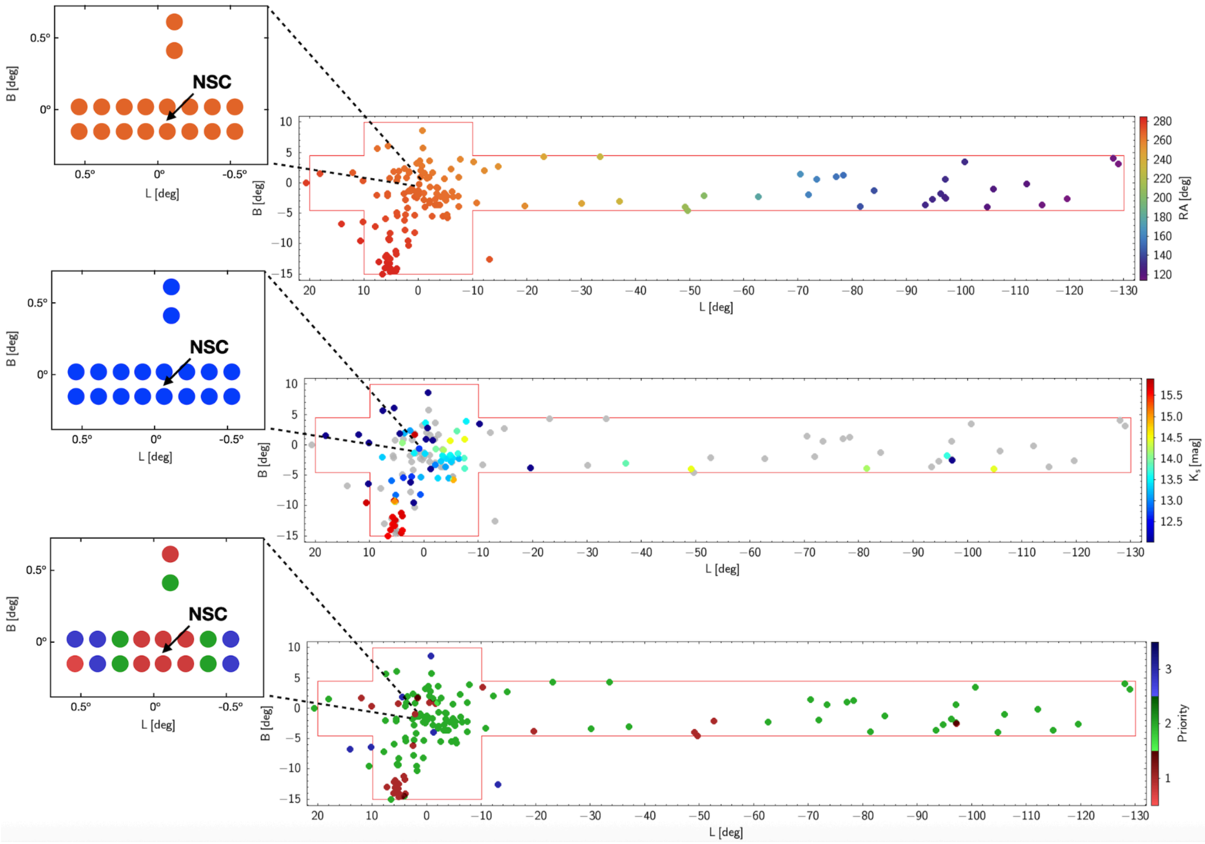


Fig. 8. Distribution of high priority targets for GC candidates (Science Area 2) and the 18 fields for the Galactic centre stars (Science Area 6) within the VVV+VVVX footprint.

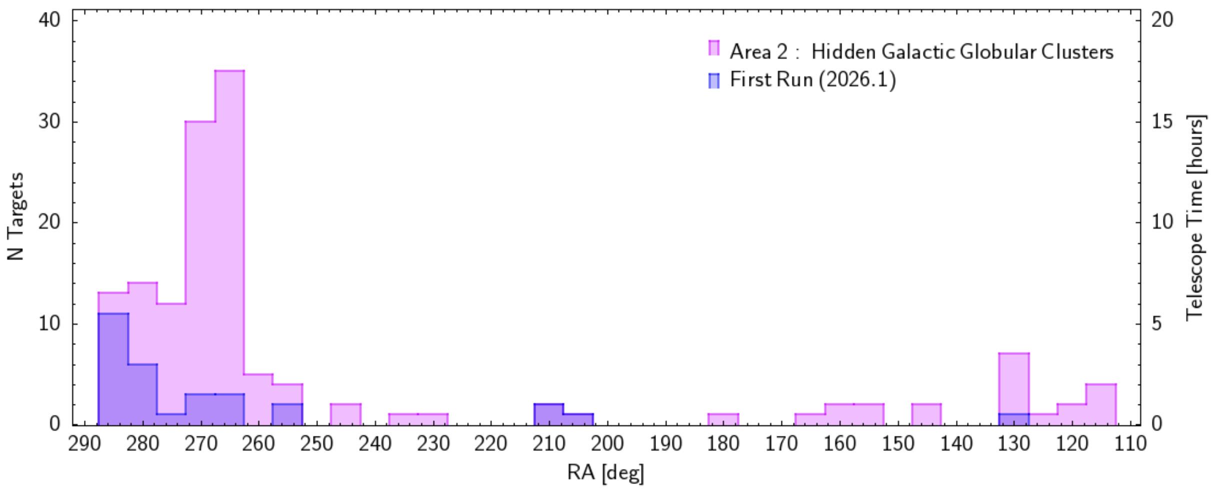


Fig. 9. The magenta histogram shows the distribution of targets (left) and hours (right) for the Science Area 2. In purple, those targets tentatively selected for P116 observations.

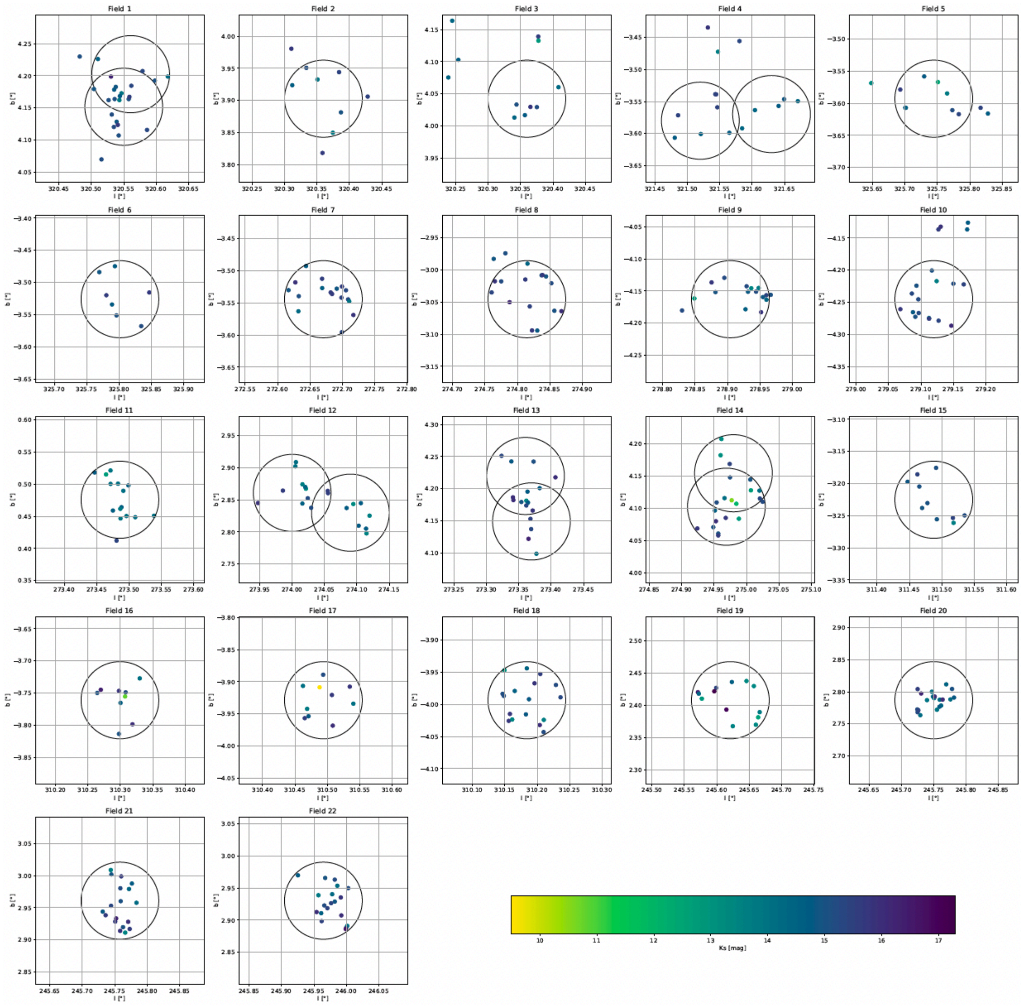


Fig. 10. Distribution of the KMOS fields for the Science Area 5, which will be tentatively covered during our first two P116 VM runs.

Table 5: Some selected fields for observations during our first VM runs from the Science Area 4. The list is not complete but illustrates that the KMOS arms occupation will be very high.

Name	RAJ2000	DEJ2000	Priority	Magnitude Range	No. Objects	Band	Exposure time (s)	Execution time (min)
Bica_376_377_378_BDSB9	112,9591	-19,3313	1	10.0-15.0	44	K	1200	60
Bica_413_414_415_416	155,0625	-58,0561	1	10.0-15.0	22	K	600	30
Cl170_171_172	135,8583	-48,4861	1	10.0-15.0	44	K	1200	60
Cl174_175	136,1042	-48,8692	2	10.0-15.3	44	K	1200	60
CL176	137,7833	-48,2656	2	10.0-15.0	22	K	600	30
CL204	128,4525	-44,4484	1	10.0-15.1	44	K	1200	60
CL206	137,4686	-48,8665	1	10.0-15.3	44	K	1200	60
CL207	145,6299	-52,4305	1	10.0-15.2	44	K	1200	60
CL208	148,5833	-56,4236	1	10.0-15.4	44	K	1200	60
DBS38	141,1375	-52,0122	2	10.0-15.0	44	K	1200	60
FSR2007_1461	135,5583	-48,6953	2	10.0-15.4	22	K	600	30

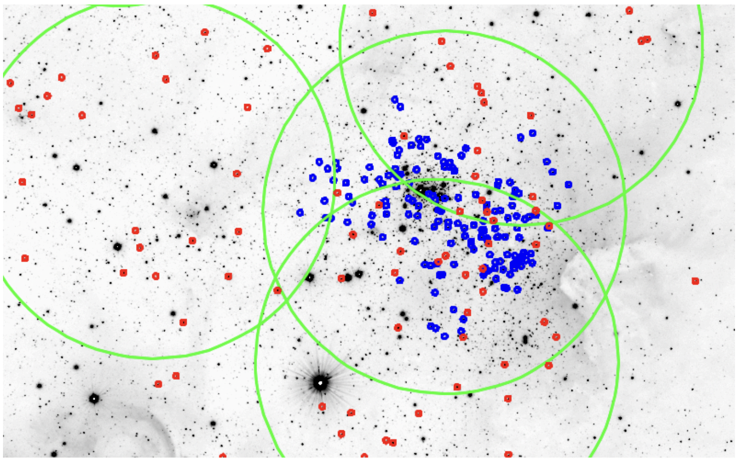


Fig. 11. Distribution of 4 KMOS fields (7.2 arcmin, in green) for the Science Area 1, which will be tentatively covered during our first two P116 VM runs. There are ~700 high-priority targets located in the dense regions of Carina, with some of the fields containing over 100 targets. North is up, East right.

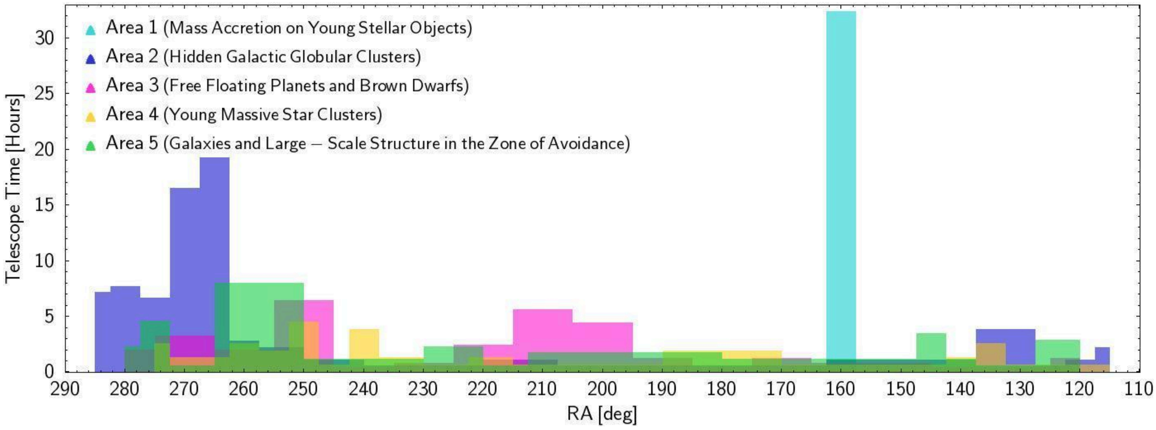


Fig. 12. Complementary to Figs. 7 and 9, the histogram shows the distribution of hours as function of RA for high-priority targets from Science Areas 1-5 that could be scheduled within the upcoming VM runs during P116.

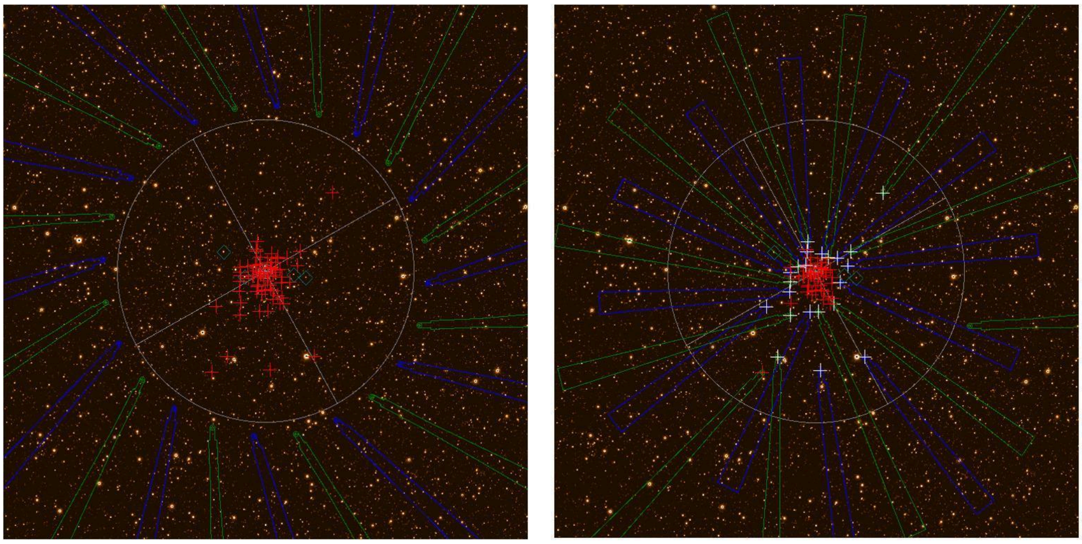


Fig. 13. An example of KARMA usage for Science Area 2 observations. The background image is from the VVV survey in the Ks band and shows the recently discovered open cluster VVV-CL-160 (RA/Dec 18:06:56.88, -20:00:39.6, L/B=10.151,0.302), with selected targets for observation, all stars with 12<*Ks*<15. The cluster is on the observation list for the first run and will be used as a proxy for future observations.

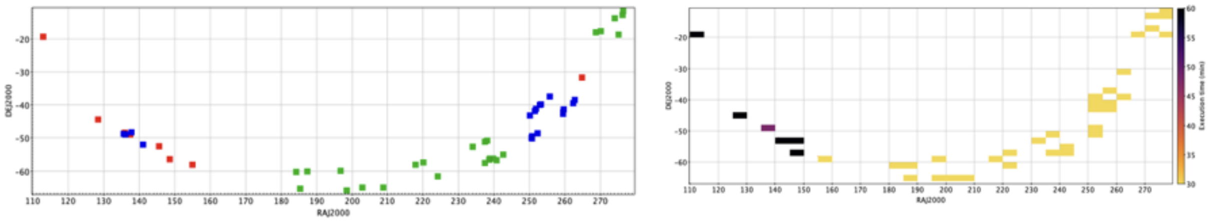


Fig. 14. Distribution of priority targets and calculated exposure times for Science Area 4. Red, blue and green colors stand for Priority 1, 2 and 3.

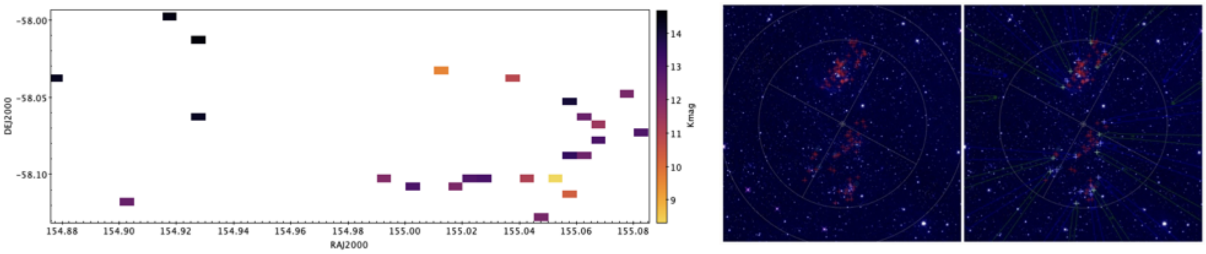


Fig. 15. First run (P116) KMOS targets in the field of Bica376_377_378_BDSB9 cluster candidates. The left panel shows the K mag distribution, and the right one is a KARMA configuration of the targets.

3. Survey data calibration needs

The standard calibration provided by ESO is sufficient for our needs, as confirmed by previous studies (e.g., Feldmeier-Krause et al. 2017; Martínez-Arranz et al. 2023). In particular, Martínez-Arranz et al. (2023) showed that the telluric absorption correction—using the calibration star provided by the observatory together with molecfit—meets our scientific requirements. Furthermore, the same study demonstrated that our proposed strategy for the Galactic centre region, namely observing the sky in a separate OB on a nearby dark cloud, can be successfully implemented.

For the Galactic centre section of our survey, we plan to validate the derived stellar parameters, including metallicities, surface gravities, and effective temperatures, by observing a set of Galactic bulge stars with existing high-resolution spectra from APOGEE (e.g., Majewski et al. 2017). These benchmark stars will allow us to directly compare the parameters obtained from our KMOS spectra with well-calibrated measurements obtained with high-resolution spectroscopy, ensuring the reliability and accuracy of our analysis. We will also observe stars in the NSD that have been previously studied with high-resolution spectroscopy (Ryde et al. 2025) to further assess our results.

During our Galactic centre pilot study (PI Nogueras-Lara, 113.26DP.001, 115.281U.001), we found that 10–20% of the science targets were not properly centred and, in some cases, were entirely outside the IFU fields of view. Following an investigation, observatory staff concluded that the problem was related to the calibration of the adapter rotator. It is therefore essential to monitor pointing accuracy to ensure valid observations and to request recalibration if necessary. We also emphasise that, to guarantee proper performance, the calibration of the adapter rotator should be carried out by ESO at least once a year, and ideally every six months.

4. Data reduction process

Given the broad scientific scope of our survey, the observations for each science area will be tailored to its specific requirements. While the standard ESO pipelines will be employed for most of the data reduction (see below), certain datasets will require additional processing steps or modifications to the pipeline. In particular, Galactic centre observations cannot always include simultaneous clear sky measurements with the science targets, necessitating separate sky observations on a nearby dark cloud located several tens of arcminutes away. Consequently, the standard reduction pipeline will be adapted to incorporate these dedicated sky frames. The team has the necessary skills and experience to do this.

The team will use the ESO KMOS pipeline (currently version 4.5.3), because it has been tested and debugged for many years by a multitude of users, including some team members. The pipeline is a collection of recipes that can be executed in various environments, including Esorex, EsoReflex and as of recently, EDPS. The latter is the most advanced but it is still in the early stages of development of its graphical user interface, so at the beginning of the survey we will adopt EsoReflex, because we plan to optimize and refine the data reduction process and interactivity will help during this initial step. EsoReflex is preferred over Esorex, because it is optimized to repeat after each parameter change only the recipes downstream of the change, minimizing the reduction time. We expect to revert to EDPS as soon as a stable interactivity version and proper documentation become available and we will compare the output from the reduction in these two environments to ensure consistent processing of the all survey data.

The data reduction steps can be grouped into the following major stages:

(1) Removal of instrument signatures (e.g., dark, flat fielding corrections) with the ESO KMOS pipeline; the final products of this step are astrometrically and wavelength calibrated data cubes for each of the 24 IFUs. They are Phase 3 compliant and will be submitted to the archive as part of the delivery package.

(2) Extraction of 1-dimensional spectra from the 3-dimensional data cubes with the Three-Dimensional Optimal Spectral Extraction (TDOSE) from integral field spectroscopy by Schmidt et al. (2019, A&A, 628, 22). Note that TDOSE was originally created for processing MUSE observations but it can handle any 3-dimensional spectroscopic data and it is publicly available (https://github.com/kasperschmidt/TDOSE). For demonstrating the feasibility a 1-D spectrum extracted from KMOS data cube is shown in Fig. 12. We are aware that the KMOS pipeline can extract spectra but TDOSE offers a critical possibility to carry out optimal extraction following Horne (1986, PASP, 98, 609) that will significantly improve the scientific output of the survey. The “optimal” method option of the recipe *kmos_extract_spec* can only be applied if the KMOS IFU contains a single object; the KMOS pipeline can not combine the optimal extraction with masking multiple objects, the latter is only possible with aperture extraction. This lack of feasibility forces us to break from the KMOS pipeline routine and to resort to an external tool. The use of an external tool will require special measures ensuring that the TDOSE 1-D spectra comply with the ESO Science Data Products Standard (https://www.eso.org/sci/observing/phase3/p3sdpstd.pdf). This includes file format (binary FITS table) and header keywords (as described in Sec. 8.1 of the requirements). We will apply as a matter of course also the regular aperture extraction, as a consistency check. The latter 1-dimensional spectra are Phase 3 compliant, so our strategy guarantees that we will be able to populate the archive with products.

(3) Removal of telluric absorption with Molecfit (Smette et al. 2015, A&A, 576, 77; Kausch et al. 2015, A&A, 576, 78) or with telluric standards observed (nearly) simultaneously with the targets. The KMOS pipeline does include Molecfit recipes (*kmos_molecfit_model*, *kmos_molecfit_calctrans* and *kmos_molecfit_correct*) and the pipeline native correction will be carried out on the pipeline aperture extracted spectra. The TDOSE optimal extraction spectra will be corrected externally: either with the pipeline derived telluric correction, or if the optimal extraction significantly improves the data quality, we will run on them the stand-alone Molecfit workflow.

(4) Combination of multi-epoch observations, if necessary to boost the signal-to-noise for faint objects. The KMOS recipe *kmos_combine* can combine reconstructed cubes and there is the generic

spec_combination recipe for adaptive resampling and combination of spectra, but most likely we will use a custom built python script to combine the extracted and telluric corrected 1-D spectra.

(5) Conversion to IDP compliant products. Normally, the KMOS pipeline delivers Phase 3 compliant data and we will take advantage of this for the data cubes and for aperture extracted 1-dimensional spectra. However, the introduction of a step performed with an external tool (TDOSE) will affect this. Therefore, we expect to require a custom built python script to update the headers, whenever necessary.

Concerning the errors: TDOSE calculates the errors according to the Horne et al. (1986, PASP, 98, 609), based on the actual variance that is calculated by the pipeline for every pixel in the 2-D spectrum. Furthermore, it has been widely used and tested by the community. Nevertheless, we will verify the errors on our own and compare them with the regular aperture extraction errors.

Once we revert to EDPS, we will investigate if we can incorporate the TDOSE optimal extraction (TDOSE is python-based) in the workflow. This would allow us to streamline the processing, combining the first three stages into one. Should we find ways to significantly improve the data products in the course of our survey, we will consider reprocessing the previously reduced data, depending on the schedule of the data releases.

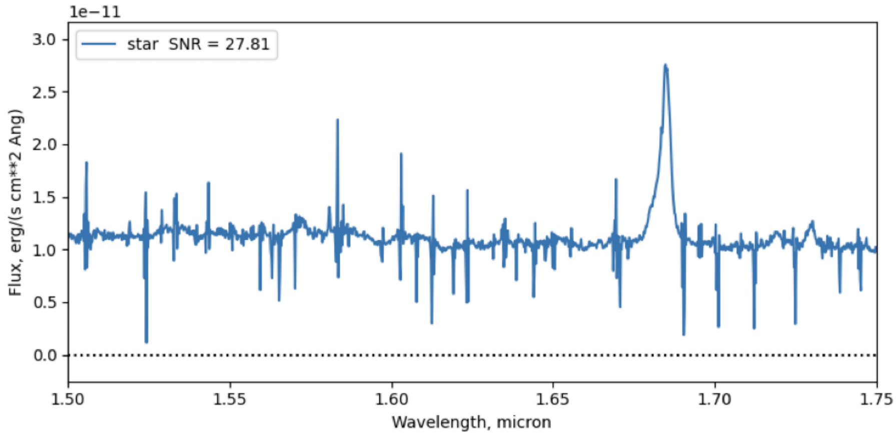


Fig. 12. An example of TDOSE optimally extracted 1-D spectrum of an object from the KMOS.2015-05-18T05:38:12.110 H band data set.

5. FTEs and hardware capabilities devoted to data reduction and quality assessment

Name	Function	Affiliation	Country	FTE allocated to project
M. Gomez	PI	UNAB	Chile	0.35
F. Nogueras-Lara	Co-PI and Leader for Science Area #6	IAA-CSIC	Spain	0.35
V. Ivanov	Quality Control	ESO	Germany	0.15
Z. Guo	Responsible for Science Area #1	Univ. of Valparaiso	Chile	0.15
P. Lucas	Responsible for Science Area #1,3	U. Hertfordshire	UK	0.15

G. Navarro	Responsible for Science Area #1	INAF - Rome Astronomical Observatory	Italy	0,15
J. Alonso-García	Responsible for Science Area #2	U. Antofagasta	Chile	0.15
J. G. Fernandez-Trincado	Responsible for Science Area #2	Univ. Católica del Norte	Chile	0.15
E. R. Garro	Responsible for Science Area #2	ESO	Chile	0.15
R. K. Saito	Responsible for Science Area #2	UFSC	Brasil	0.15
C. Cáceres	Responsible for Science Area #3	UNAB	Chile	0.15
R. Kurtev	Responsible for Science Area #3,4	Univ. of Valparaiso	Chile	0.15
D. Minniti	Responsible for Science Area #3	UNAB	Chile	0.15
J. Borissova	Responsible for Science Area #4	Univ. of Valparaiso	Chile	0.15
A. N. Chene	Responsible for Science Area #4	Gemini Observatory	USA	0.15
L. D. Baravalle	Responsible for Science Area #5	IATE–CONICET OAC-UNC	Argentina	0.15
D. Galdeano	Responsible for Science Area #5	FCEFN-CONICET-UNSJ	Argentina	0.15
J. Corral-Santana	KMOS Specialist	ESO	Chile	0.10
A. Luna	KMOS Specialist	ESO	Chile	0.10
J. M. Gavilán Sánchez	PhD student	IAA-CSIC	Spain	0.50
M. Álvarez-Santiago	PhD student	IAA-CSIC	Spain	0.40
M. Schultheis	Spectroscopy specialist	OCA	France	0.10
A. Feldmeier-Krause	KMOS specialist	U. Vienna	Austria	0.10
~90 Team members				nominally 0.10 but up to 0.50 for postdoctoral fellows

Now that this programme is formally approved, we are applying for funding for 5 postdoctoral positions, dedicated to this survey (FTE 50%). These applications have been submitted to ANID, ESO-Chile Joint Committee, and participating Universities, where we expect a good reception given the breadth of our survey (even though we already have the necessary manpower, we are hopeful that these hires would be made within the first year of the survey, allowing us to make faster progress).

Hardware: Reduced datacubes will be archived in the ESO Science Archive Facility for long-term storage and community access. For data that require non-standard processing, we have privileged access to HPC resources at UNAB, HPC resources at Univ. Católica del Norte, IAA-CSIC and collaborating institutions. This infrastructure will enable us to perform the more demanding reduction and analysis tasks efficiently and distributed among the different Science Areas.

At UNAB our HPC cluster is composed of 6 nodes with 136 CPU cores, 1.38 TB RAM and approximately 100 TB of storage capacity. This cluster is currently being enhanced, but even as is, it can provide a robust and reliable laboratory for data processing, testing and analysis. A dedicated server at the IAA-CSIC, with 40 TB of storage, 24 processing cores, and 64 GB of RAM, will be used for reducing and analysing the Galactic centre survey data.

In addition to local computational resources, we benefit from high-speed secure internet connection through REUNA (https://www.eso.org/public/chile/announcements/annlocal24003-es-cl/) which ensures efficient data transfer between ESO and UNAB. This connectivity is critical for timely access to raw and processed data, and collaborative efforts between our team and ESO.

The Institute of Theoretical and Experimental Astronomy (IATE, CONICET) has a computing cluster consisting of a head node with 6 cores and 32 GB of RAM, and six additional nodes, each with 28 CPU cores, 56 threads, and 128 GB of RAM each. Four additional independent computing nodes are also available: two with 16-core processors, 32 virtual SMT cores, and 64 GB of RAM each; and two with 64 CPU cores, 128 threads and 256 GB of RAM each. All the equipment is connected via Infiniband and two storage nodes provide 252 TB of shared storage. Furthermore, the institute is a member of the High Performance Computing Centre at the National University of Córdoba.

6. Data quality assessment process

The quality control (QC) process of the survey is a balance between the need for rapid detection of problems and thorough analysis of the new data. To address these (sometimes conflicting) requirements we adopt a combined multistage QC process, spanning different time scales: fast, at the telescope, aiming to allow for nearly real time response, but at the most basic level; on a longer term we will carry QC during the data reduction and analysis, building a data base - not unlike the ESO QC database - containing more sophisticated QC parameters that are possible to calculate only in the course of the real data processing. We will work towards a scheme that considers QC descriptors within the headers of our products. Final details will be part of the data releases.

QC at the telescope:

(1) Pipeline based QC: it will take advantage of the QC parameters routinely calculated by the pipeline. The first few runs will be carried out in visitor mode, therefore, we will be able to use the output from the pipeline that is run in real time on the data on Paranal. The processing by this pipeline is template triggered, so it begins only after a template is completed, introducing some delay, e.g. as long as tens of minutes. The QC parameters include positions and FWHM of certain arc lines, spectral resolution (on the arc frames), number of saturated pixels, throughput and spectral curvature solution check (on the standard star observations). The information from these parameters is useful although not complete, especially when it comes to evaluating the quality of the science frames.

(2) Purpose-built python scripts based QC: this is necessary to fill in the gaps from the pipeline QC and to measure QC parameters on an exposure-by-exposure base in nearly real time. Our script will calculate for each IFU: sky level at least the locations across the spectral region (near the blue and red edges and in the center), number of saturated pixels, positions of some prominent sky lines and image quality on images, reconstructed from a limited number of wavelength slices in the regions of the best sky transparency for the given spectral setting. The sky levels, scaled with the detector integration time (DIT) will be compared with the expected sky contribution and the other parameters will be compared with predefined reference values. Here we mention the bare minimum of QC parameters, and they were selected as a compromise between the requirements for speed and thoroughness.

Long-term QC:

Scientific analysis is the ultimate QC and, for the purpose of evaluating data quality and tracking changes in the instrument status, we will keep track of further parameters that are possible to calculate only when the entire data reduction is carried out fully.

The purpose of this step will go beyond the estimation of the data quality; having at our disposal a large data set may allow us to calibrate any systematic effects and to attempt correcting them, thus, improving the quality of the data beyond what is possible with a single or a small number of observations.

These will include (in random order, that is expected to be expanded as we gain experience and insight):

- Stability of the instrument response with time, from comparing response functions derived from standard star observations.
- Stability of the wavelength calibration with both time and spatially across the IFUs and within each IFU from following the various wavelength solutions.
- Stability of the flat fielding, including spatial from a comparison with referencemaster flats.
- Nightly, seasonal, airmass, etc. variations of the telluric correction from a comparison of the various telluric transmission curves.
- Quality of the astrometric solution from comparing the KMOS derived object coordinates and Gaia, 2MASS, VVV/VVVX coordinates.
- Spatial and temporal variations of the sky subtraction from measurements of the residual sky on the processed spectra.

This is an incomplete list, and we expect to expand it in the course of the survey.

The delivery of the data products to the ESO archive will be the responsibility of the PI, who certifies the scientific quality and accuracy of the data products. We will have a QC team that will include one representative from each Science Area to spread the work. To ensure uniformity of the QC process, we will have a common procedure for the QC that will be followed by all team members. We acknowledge that the QC and the Phase 3 delivery are linked and therefore, the QC team will ensure the Phase 3 compliance. The team will have a coordinator – Valentin D. Ivanov (ESO) – who will be responsible for the preparation, implementation and – if necessary – refinement of the QC procedure for the Phase 3 delivery.

We will produce the initial catalogue of selected known sources to be observed for this proposal, and for 3 years ESO will collect about 5.6 million KMOS spectra. ESO also serves as the main source where the public data will be archived.

Our team will be responsible for handling the data quality control and producing the Level 1-2 data, and delivering those to the ESO archive, as described below. The rigorous data quality control is essential to guarantee properly processed and calibrated data, and to discard obvious **problems** like missing targets, saturation, poor spectra, as well as to flag more subtle effects like cosmic rays, spikes, ghosts, persistence, halos, non-linearity, edge effects, etc. We will endeavour to provide a fast look up for all the spectra to fix any problems that may arise.

The **Level 1 data** would be the spectra that pass our quality control (QC). At this stage the processing would be divided into the 6 Science Area**s** as specified in Section 2. To maximise efficiency those data can be grouped in suitable catalogues according to the target characteristics for the different scientific areas (for example, a galaxy catalogue would be different from a brown dwarf's catalogue, or processing of compact vs binary vs diffuse source extraction). We expect to deliver these Level 1 data into the ESO archive within a year of acquisition. Additionally, at this stage we envision publishing the first main paper describing the Survey with a few real data examples.

The **Level 2 data** would add products like measurements of physical parameters of interest (radial velocities, spectral indices, fluxes, etc.), as well as matches with some relevant catalogues. These Level 2 data are suitable for follow-up projects. We expect to feed these Level 2 data into the ESO Archive (Phase 3) within 2 years of acquisition. Intermediate data releases will be submitted to Phase 3 on a periodic basis until the complete catalogue is available, rather than waiting for its full completion. These calibrated publication quality data would allow immediate use by the astronomical community.

The **Level 3 data** would approach to our final products, where we release what we call our public analysis tools, along with the advanced data:

- Mass accretion on Young Stellar Objects: classifications, fluxes, physical parameters, etc.
- Hidden Galactic globular clusters: confirmation, orbits, metallicities, sizes, masses, IMF parameters, etc.
- Free-floating planets and brown dwarfs: temperatures, binarity, etc.
- Young embedded and massive star clusters in our Galaxy: distances, masses, etc.
- Galaxies and large-scale structure in the zone of avoidance: redshifts, classification, etc.
- Galactic centre stars: derived stellar properties such as metallicities, temperatures, and log g.

The catalogs from each area will be somewhat different, but many of them will have similar structure – e.g. stellar (or brown dwarf of free floating planets) parameter catalogs that contain stellar physical parameter (T_{eff}, log g, [Fe/H], V_{rad}, etc – may not have all of those but for compatibility they can all contain the same columns). Some catalogs will be time series of measurements, some will contain galaxy parameters.

We emphasize that IFS cubes will be generated by the KMOS pipeline and archival users may be able to use or process the data in ways that the survey team had not intended, including forced extraction on the position of extremely faint or emission line-only objects, potentially increasing the scientific impact of the survey.

Observing report:

As established in the guidelines, we will submit to OPC a progress report at the end of the observing periods (including a list of the OBs that are deemed to be category "D"). For this purpose, along with our dedicated website we will keep a log of the observed OBs and a QC flag.

Software tools:

This Public Survey will primarily rely on the standard ESO pipeline and data reduction tools, which are well-established and supported by a comprehensive body of documentation. These resources will serve as our main reference for the initial stages of data processing and quality control.
However, we anticipate that a subset of the data will not conform to the standard pipeline parameters and will therefore require tailored strategies. For these cases, local storage combined with carefully optimized data reduction procedures are expected to yield significantly improved results compared to a fully generic approach. The precise fraction of the dataset that will require such custom treatment is not yet known and will depend on the specific characteristics of the incoming observations. Nevertheless, based on our prior experience with high-quality KMOS datasets (e.g., Feldmeier-Krause et al. 2017, 20; Martínez-Arranz et al. 2024), we foresee that:

- Visual inspection of the stacked IFU images for each KMOS pointing will be carried out to verify both the pointing accuracy and the correctness of the target coordinates. This step is particularly important given the previously mentioned pointing issues detected thanks to our Galactic centre pilot study (see Sect. 3).
- Visual inspection of the extracted spectra that will be essential in ensuring the quality of the data. This can be quickly done through the preliminary reduced files provided by ESO.
- Multiple passes through EsoRex/EsoReflex/EDPS or equivalent tools may be required, particularly for the more challenging or non-standard observations.
- We have developed a new tool (Schultheis et al., *in press*) specifically adapted for the analysis of Galactic centre spectra, optimised for the cool M giants that dominate this region (e.g., Nogueras-Lara et al. 2020, 2021). Building on the Bayesian full spectral fitting code Starkit (Kerzendorf & Do 2015), we employ a newly constructed synthetic grid tailored to M giants, improving over the Phoenix models used previously (e.g., Feldmeier-Krause et al. 2017, 2020). The grid incorporates updated atomic and molecular line lists, as well as non-local thermodynamic equilibrium (non-LTE) effects for key elements, which account for the influence of the radiation field on spectral line formation and enhance the accuracy of derived stellar parameters and abundances. For the first time, this method will allow us to derive robust α-element abundances from KMOS data for the Galactic centre in addition to standard stellar parameters. We have validated the approach by degrading high-resolution

spectra to the KMOS resolution, confirming that it can reliably recover stellar parameters and chemical abundances at the survey's spectral quality (Schultheis et al., *in press*). Figure 2 (lower panels) shows an example of a fit obtained using this approach for one of the target stars of our pilot study (P115).

Several members of our Survey have substantial expertise in developing and adapting custom software for specialized reduction steps and rapid data visualization (see examples above). We are committed to making any such custom codes publicly available, accompanied by thorough documentation, to ensure transparency and reproducibility within the community.

7. External data products and Phase 3 compliance

Our team understands the legacy value of the survey products, distributed through the ESO science archive. We will make every effort, in collaboration with the ESO Science Archive Group, to facilitate community access to both Level 0 and Level 1 data products. We will deliver cubes and 1-D spectra with a delay of 1.5 years as we will be refining the processing and analysis. There will be 6 catalogs, related for each Science Area, but for Level 1 products these will be only source catalogs with positions and positional uncertainties, (preliminary) source classification and cross-identifications with other catalogs. High level products like comprehensive catalogues of stellar properties, including fundamental parameters, kinematics and variability will require at least 50% of the necessary data and therefore will be released no earlier than two years after the first observations.

The standard ESO KMOS pipeline will be used for the basic data reduction (instrument signature removal and atmospheric telluric correction). It produces Phase 3 compliant data cubes and 1-dimensional aperture extracted spectra.

External tools (most likely, TDOSE, see Sec. 4) will be used for optimal extraction of 1-dimensional spectra from the data cubes. By default the products of these tools are not Phase 3 compliant and will require custom work to ensure this compatibility. Our team is in possession of the necessary expertise to do this and will collaborate with the ESO Science Archive Group to this end.

A special case is the multi-epoch observations. For individual epochs the products described above will be available by default. However, in addition to them we will use the KMOS pipeline recipe *kmos_combine* to obtain stacked cubes. These will also be Phase 3 compliant and will be submitted to the archive.

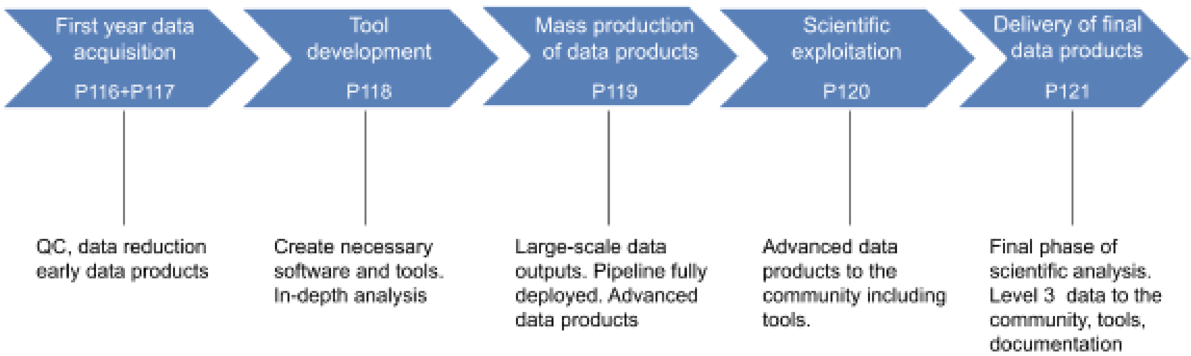
		2026				2027				2028		2029		2030	
Area	Step	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	H1	H2	H1	H2	H1	H2
Data acquisition & analysis	Data acquisition	■	■	■	■	■	■	■	■	■	■	■	■		
	Initial QC	■	■	■	■	■	■								
	Processing of Level 1 and Level 2 data for the H, K, and H+K bands.	■	■	■	■	■	■								
	Development of custom data processing tools	■	■	■	■	■	■	■	■	■	■				
	Acquisition and processing of Level 1 and Level 2 data for the YZ and YJ.					■	■	■	■	■	■				
	Initial QC and data reduction for YZ and YJ			■	■	■	■	■	■	■	■				
	Refinement of analysis and tools							■	■						
	Delivery of P116 data cubes and 1-D spectra							■	■						
	Delivery of P117 data cubes and 1-D spectra									■					
	Delivery of P118 data cubes and 1-D spectra										■				
	Delivery of P119 data cubes and 1-D spectra											■			
	Start of production of Level 2 and 3 data for the H, K, and H+K bands.								■	■	■	■	■	■	■
	In-depth scientific analysis of the processed datasets from both observing cycles.								■	■					
	Processing pipeline including QC fully developed									■	■				
	Start of production of Level 2 and 3 data for the YZ and YJ bands									■					
	Advanced data products, including advanced processing and calibration								■		■				
	Continued mass production and analysis of Level 2 and Level 3 data for all bands										■	■	■	■	■
	Scientific exploitation, including publication of key results									■	■	■	■	■	■
	Delivery of advanced products to the community, with documentation and tools												■	■	■
	Final mass production and validation of Level 3 data for all bands												■	■	■
management, outreach & collaboration	Survey Workshop					■					■		■		■
	Hands-on tutorials	■	■			■	■			■					
	Application to external funding opportunities	■	■	■	■	■			■	■					
	Press releases	■			■	■			■		■		■		■

Level 2 products will be generated with custom analysis tools for measuring the relevant parameters for the specific Science Area: redshifts, line-of-sight velocities, extinction, spectral classification, etc. For most of these products the final format will be fits tables, but for some -metallicity and radial velocity maps, for example- can be fits images. The ESO Archive already contains data of these types and we will follow the Phase 3 standards to ensure our products are compatible. If appropriate, we will combine our KMOS catalogs with existing external photometric, proper motion, extinction, and variability from other sources to increase the legacy value of our products. For example, for the Galactic centre, this will yield the most comprehensive catalogue of stellar properties, including fundamental parameters, kinematics, and variability, enabling detailed studies of the region's structure, dynamics, and stellar populations. Similarly, our catalogs of variable stars will be complemented with variability information from other surveys, with any available optical and mid-infrared colors, etc.

The primary legacy of our Survey will be a series of refereed publications. In addition, the KMOS data will provide a valuable foundation for student theses, contributing to the training of the next generation of researchers.

Other remarks:
We would like to stress that most importantly, we will produce numerous refereed papers, impacting the six diverse areas of research covered by our science team. But not only that: building upon the successful experience with our previous surveys, we expect to establish a solid international scientific network, to train students at all levels (PhD, Master and undergraduate), to organise annual workshops, present results at conferences, and to produce international press releases, short movies and popular articles highlighting our results.

8. Timeline delivery of data products to the ESO archive



To ensure efficient coordination and rigorous data handling, each scientific area of the public survey will be organised into dedicated work packages. This structure allows us to distribute the workload effectively, assign clear responsibilities, and guarantee careful management of the data throughout the project. Whenever possible, the work packages will be executed in parallel, sharing experience, thereby accelerating progress and maximising efficiency across the different scientific areas.

We expect that each single area would have produced at least a couple of scientific publications (and follow-up proposals) based on our survey data after the first year of observations. This is more difficult to predict, but we boldly expect to feed these final Level 3 data ingested into the ESO Archive after survey completion, the timing estimated to be about 1 year after the end of observations. Adding to the legacy, this release would be followed by a comprehensive end-of-survey publication.

The team will keep an updated webpage, with mirrors in Europe and Chile. The relevant software and analysis tools for search, visualisations and measurements (mosaics, IFS cubes, variability, simulations, complex background source subtraction, etc.) developed throughout the survey would be made available as part of the releases.

Work Plan:

Year 1 (P116 - P117)

- Acquisition and processing of Level 1 and Level 2 data for the H, K, and H+K bands.
- Initial QC and data reduction for these bands.
- Development of custom data processing tools.
- Preliminary scientific analysis of early data products to validate pipeline performance and guide future observations.

Year 2 (P118)

- Acquisition and processing of Level 1 and Level 2 data for the YZ and YJ bands, including QC and data reduction.
- Continued tool development and refinement based on Year 1 experience.
- Start of mass production of Level 2 and Level 3 data for the H, K, and H+K bands.
- In-depth scientific analysis of the processed datasets from both observing cycles.

Year 3 (P119)

- Full processing pipeline in place for Level 1, 2, and 3 data products for the H, K, and H+K bands, including comprehensive QC, reduction, and analysis.
- Begin mass production and scientific exploitation of Level 2 and Level 3 data for the YZ and YJ bands.
- Advanced data products, including advanced processing and calibration.

Year 4 (P120)

- Continued mass production and analysis of Level 2 and Level 3 data for the H+K, YZ, and YJ bands.
- Ongoing scientific exploitation, including publication of key results.
- Delivery of advanced science-ready products to the community, along with relevant documentation and tools.

Year 5 (P121)

- Final mass production and validation of Level 3 data for both H+K and YZ/YJ bands.
- Final phase of scientific analysis.
- Delivery of the final data products, including all Level 3 products, associated tools, and scientific documentation.

References:

1. Minniti et al. 2010, New A, 15, 433 — 2. Saito et al. 2024, A&A 689, A148 — 3. Nogueras-Lara et al. 2018, A&A 610, A83 — 4. Nogueras-Lara et al. 2019, A&A 631, A20 — 5. Fischer et al., 2023, PPVII — 6. Contreras Peña et al., 2017, MNRAS, 465, 3011 — 7. Lucas et al., 2024, MNRAS, 528, 1789 — 8. Contreras Peña 2024, MNRAS, 528, 1823 — 9. Kuhn et al., 2021, ApJS, 254, 33 — 10. Kospal et al., 2023, ApJ, 945, 7 — 11. Manara et al., 2023, PPVII — 12. Minniti et al., 2021, A&A 647, L4 — 13. Palma et al., 2019, MNRAS 487, 3140 — 14. Garro et al. 2024, A&A, 687, A214 — 15. Massari, 2019, A&A 630, L4 — 16. Vasiliev Galaxies 11, 59 —17. Belokurov et al. 2024, MNRAS, 528, 3198 — 18. Fernández-Trincado et al. 2022, A&A, 658, A116 —19. Minniti et al., 2024, A&A 683, A150 — 20. Mejías et al. 2022, A&A, 660, A131 — 21. Sarro et al 2023, A&A 669, 139 — 22. Caceres et al. 2015, ApJ, 806, L22 — 23. Mejías et al. 2022, 44th COSPAR Scientific Assembly — 24. Marocco et al. 2015, MNRAS, 449, 3651 — 25. Ascenso 2018, Astrophysics and Space Science Library, 424, 1 — 26. Borissova et al. 2011, A&A, 532, A131 — 27. Borissova et al. 2014, A&A, 569, A24 — 28. Borissova et al. 2018, MNRAS, 481, 3902 — 29. Borissova et al. 2020, MNRAS 499, 3522–3533 — 30. Solin et al. 2014, A&A, 562, A115 — 31. Barbá et al. 2015, A&A, 581,

A120 — 32. Ivanov et al. 2017, A&A, 600, A112 — 33. Froebrich 2017, MNRAS, 469, 1545F — 34. Bica et al. 2019, AJ, 157, 12 — 35. Martins et al. 2019, A&A, 627A,170M — 36. Ramirez Alegria et al. 2016, A&A, 588A, 40R — 37. Chene et al. 2012, ApJ, 911, 91C — 38. Chene et al. 2013, A&A, 549, A98 — 39. Chene, A. N. et al. 2015, A&A, 584A,31C — 40. Baravalle et al. 2021, MNRAS 502, 601 — 41. Daza-Perilla et al. 2023, MNRAS 524, 678 — 42. Duplancic et al. 2024, A&A 682, 153 — 43. Alonso et al. 2025, submitted — 44. Galdeano et al. 2021, A&A 646, 146 — 45. Mason et al., 2015, ApJS, 217, 13 — 46. Kruijssen et al. 2013, MNRAS, 435, 2598 — 47. Launhardt et al., A&A, 384, 112 — 48. Sormani et al. 2022, MNRAS, 512, 2 — 49. Gallego-Cano et al. 2020, A&A, 634, A71 — 50. Nogueras-Lara et al. 2020, Nat. Astron., 4, 377 — 51. Mezger et al. 1996, A&A Rev., 7, 289-388 — 52. Matsunaga et al. 2011, Nature, 477, 188-190 — 53. Crocker et al. 2011, MNRAS, 431, 763-788 — 54. Mauerhan et al. 2010, ApJ, 710, 706 — 55. Morris et al. 1996, A&A Rev., 34, 645 — 56. Dahmen et al. 1998, A&A, 331, 959 — 57. Pierce-Price et al. 2000, ApJ, 545, L121 — 58. Yamauchi et al. 1990, ApJ, 365, 532 — 59. Spergel et al. 1992, Nat, 357, 6380 — 60. Muno et al. 2004, ApJ, 613, 1179 — 61. Pillai et al. 2015, ApJ, 799, 1 — 62. Schoedel et al. 2002, Nature, 419, 694 — 63. Feldmeier-Krause et al. 2017, MNRAS 464,194 — 64. Lohr et al. 2018, A&A 617, A66 — 65. Clark et al. 2018, A&A 618, A2 — 66. Bressan et al. 2012, MNRAS, 427, 127 — 67. Smith et al. 2018, MNRAS, 474, 1826 — 68. Ginsburg et al. 2024, APJL, 968, L11 — 69. Nogueras-Lara et al. 2023, A&A 671, L10 — 70. Sanders et al. 2024, MNRAS 530, 2972 —71. Thorsbro et al. 2023, ApJL, 958, L18 — 72. Schultheis et al. 2021, A&A 650, A191 — 73. van-Donkelaar et al. 2023 — 74. Bittner et al. 2020, A&A 643, A65 — 75. Gadotti et al. 2020, A&A, 643, A14 — 76. Fritz et al. 2021, A&A, 649, A83 — 77. Libralato et al. 2021, MNRAS, 500, 3 — 78. Shahzamanian et al. 2022, A&A 662, A11 — 79. Nieuwmunster et al. 2024, A&A 685, A93 — 80. Nogueras-Lara et al. 2023, A&A 680, A75 — 81. Guo et al. 2024, MNRAS, 528, 1769 — 82. Tobin et al. 2024, ARA&A, 62, 203.

APPENDIX 1: Content of the fits header for example TDOSE product. Extension 1 contains the optimally extracted 1D spectrum and Extension 2 contains the input data cube with its full header.

```
====> file tdose_spectrum_mxdf_gauss_0000000001.fits (main) <====
SIMPLE  =                    T / conforms to FITS standard
BITPIX  =                    8 / array data type
NAXIS   =                    0 / number of array dimensions
EXTEND  =                    T
END
===> xtension 1
XTENSION= 'BINTABLE'           / binary table extension
BITPIX  =                    8 / array data type
NAXIS   =                    2 / number of array dimensions
NAXIS1  =                   32 / length of dimension 1
NAXIS2  =                 2048 / length of dimension 2
PCOUNT  =                    0 / number of group parameters
GCOUNT  =                    1 / number of groups
TFIELDS =                    4 / number of table fields
EXTNAME = 'SPEC1D  '           / cube containing source
TTYPE1  = 'wave    '
TFORM1  = 'D       '
TUNIT1  = 'ANGSTROMS'
TTYPE2  = 'flux    '
TFORM2  = 'D       '
TTYPE3  = 'fluxerror'
TFORM3  = 'D       '
TTYPE4  = 's2n     '
TFORM4  = 'D       '
END
===> xtension 2
XTENSION= 'IMAGE   '           / Image extension
BITPIX  =                  -64 / array data type
NAXIS   =                    3 / number of array dimensions
NAXIS1  =                   10
NAXIS2  =                    9
NAXIS3  =                 2048
PCOUNT  =                    0 / number of parameters
GCOUNT  =                    1 / number of groups
EXPTIME =                380.0
CTYPE1  = 'RA---TAN'           / RA---TAN
CTYPE2  = 'DEC--TAN'           / DEC--TAN
CRPIX1  =                 -2.5
CRPIX2  =                 -1.5
CRVAL1  =      253.24569166667
CRVAL2  =      2.4005105555556
CD1_1   = -5.5080270076322E-05
CD2_1   =  7.2514551233362E-06
CD1_2   =  7.2514551233362E-06
CD2_2   = 5.50802700763228E-05
CRVAL3  =     1.42499995231628
CRPIX3  =                  1.0
CDELT1  =                  1.0
CDELT2  =                  1.0
CDELT3  = 0.000215820327866822
CTYPE3  = 'WAVE    '           / WAVE
CUNIT1  = 'deg     '           / deg
CUNIT2  = 'deg     '           / deg
CUNIT3  = 'um      '           / um
CD1_3   =                  0.0
CD2_3   =                  0.0
CD3_1   =                  0.0
CD3_2   =                  0.0
CD3_3   = 0.000215820327866822
INHERIT =                    T
CHECKSUM= '46A6657445A44554'   / 46A6657445A44554
DATASUM = '4284350671'         / 4284350671
HIERARCH ESO INS FILT2 ID = 'H       ' / H
HIERARCH ESO INS GRAT2 ID = 'H       ' / H
HIERARCH ESO DET CHIP DATE = '2006-07-20' / 2006-07-20
HIERARCH ESO DET CHIP GAIN = 2.1
```

```
HIERARCH ESO DET CHIP ID = 'KMOS-H2RG-212' / KMOS-H2RG-212
HIERARCH ESO DET CHIP INDEX = 2
HIERARCH ESO DET CHIP LIVE = T
HIERARCH ESO DET CHIP NAME = 'Hawaii2RG' / Hawaii2RG
HIERARCH ESO DET CHIP NX = 2048
HIERARCH ESO DET CHIP NY = 2048
HIERARCH ESO DET CHIP PSZX = 18.0
HIERARCH ESO DET CHIP PSZY = 18.0
HIERARCH ESO DET CHIP PXSPACE = 1.8E-05
HIERARCH ESO DET CHIP RGAP = 0.0
HIERARCH ESO DET CHIP RON = 0.0
HIERARCH ESO DET CHIP TYPE = 'IR      ' / IR
HIERARCH ESO DET CHIP X = 2
HIERARCH ESO DET CHIP XGAP = 8.0
HIERARCH ESO DET CHIP Y = 1
HIERARCH ESO DET CHIP YGAP = 0.0
HIERARCH ESO DET EXP UTC = '2015-05-18T05:51:02.2297' / 2015-05-18T05:51:02.2297
HIERARCH ESO DET FRAM INITCNT = 0
HIERARCH ESO DET FRAM NO = 1
HIERARCH ESO DET FRAM STRX = 1
HIERARCH ESO DET FRAM STRY = 1
HIERARCH ESO DET FRAM TYPE = 'INT     ' / INT
HIERARCH ESO DET FRAM UTC = '2015-05-18T05:51:02.8166' / 2015-05-18T05:51:02.816
HIERARCH ESO DET SOFW MODE = 'NORMAL  ' / NORMAL
HIERARCH ESO QC ARC NE POS STDEV = 1.69810831951245
HIERARCH ESO QC ARC AR POS STDEV = 1.66979032894785
HIERARCH ESO QC ARC NE FWHM MEAN = 71.6215668824434
HIERARCH ESO QC ARC AR FWHM MEAN = 75.8840486044585
HIERARCH ESO QC ARC NE VSCALE = 37.6902278713086
HIERARCH ESO QC ARC AR VSCALE = 38.6401028257233
HIERARCH ESO OCS ARM14 NAME = 'ARM14_SCI' / ARM14_SCI
HIERARCH ESO OCS ARM14 ALPHA = 165258.966
HIERARCH ESO OCS ARM14 DELTA = 22401.838
HIERARCH ESO QC CUBE_UNIT = 'erg.s**(-1).cm**(-2).angstrom**(-1)' / erg.s**(-1).
WCSAXES =                    2
PC1_1   = -5.5080270076323E-05
PC1_2   =  7.2514551233362E-06
PC2_1   =  7.2514551233362E-06
PC2_2   =  5.5080270076323E-05
LONPOLE =                180.0
LATPOLE =     2.4005105555556
MJDREF  =                  0.0
RADESYS = 'ICRS    '           / ICRS
OBJID   = '0000000001'         / ID of object
SRCIDS  = '[0]     '           / IDs of sources combined in object
EXTNAMEC= 'IFU.14.DATA'        / EXTNAME of original source cube
EXTNAME = 'SOURCECUBE'         / cube containing source
END
```