CAVITY, Calar Alto Void Integral-field Treasury surveY

July 8, 2024

The Calar Alto Void Integral-field Treasury surveY (CAVITY) is a legacy project aimed at characterizing the population of galaxies inhabiting voids, the most under-dense regions of the Universe. This paper describes the first public data release (DR1) of CAVITY, comprising science-grade optical datacubes for the initial 100 out of a total of ∼300 galaxies in the Local Universe ($0.005 < z < 0.050$)**. These data were acquired using the integral-field spectrograph PMAS/PPak mounted on the 3.5m telescope at the Calar Alto observatory. The DR1 galaxy sample encompasses diverse characteristics in the color–magnitude space, morphological type, stellar mass, and gas ionization conditions, providing a rich resource for addressing key questions in galaxy evolution through spatially resolved spectroscopy. Galaxies were observed with the low-resolution V500 setup spanning the wavelength range 3745–7500 Å with a spectral resolution of 6.0 Å (FWHM). We describe the data reduction and characteristics and data structure of the CAVITY datasets essential for their scientific utilization, highlighting considerations such as vignetting effects, identification of bad pixels, and management of spatially correlated noise. We also provide instructions for accessing the CAVITY datasets and associated ancillary data through the project's dedicated database.**

1 Introduction

The distribution of galaxies in the cosmos is heterogeneous, forming a complex structure known as the Cosmic Web. On megaparsec scales, this web is characterized by elongated filaments, sheet-like walls, and under-dense voids (Bond et al. 1996; van de Weygaert & Bond 2008a,b). Despite occupying approximately 70% of the Universe's volume, voids contain only about 10% of its mass (Cautun et al. 2014; Libeskind et al. 2018). These voids, originating from early density fluctuations, represent regions of weaker gravitational pull, expanding faster than the general Hubble flow and channeling matter into surrounding structures. Understanding void dynamics is pivotal for unraveling the overarching large-scale arrangement of the cosmic web and its impact on galaxy evolution (Courtois et al. 2023).

Galaxies residing within voids provide unique insights into evolutionary processes, distinct from those in denser environments. They undergo different transformation mechanisms, offering clues into galaxy evolution dynamics. Void galaxies exhibit distinct characteristics, such as bluer colors and later morphological types (Rojas et al. 2004, 2005; Park et al. 2007; Constantin et al. 2008), suggesting potential differences in their star formation rates.

With the purpose of investigating systematic differences between void galaxies and those in denser environments, the Void Galaxy Survey (VGS; Kreckel et al. 2011, 2012; Beygu et al. 2016, 2017) has been a multi-wavelength program focusing on 60 void galaxies, primarily utilizing SDSS multiband photometry and HI data. Each galaxy was selected from the innermost regions of identified voids in the SDSS redshift survey using a geometric-topological watershed technique (Platen et al. 2007), without any prior selection based on intrinsic properties of the void galaxies. The project comprehensively studied the gas content, star formation history, stellar content, as well as the kinematics and dynamics of void galaxies and their companions. One of the most intriguing findings of the VGS is the potential evidence for cold gas accretion in several noteworthy objects, including a polar ring galaxy and a filamentary arrangement of void galaxies (Stanonik et al. 2009).

However, the precise impact of void environments on various galaxy properties, including stellar mass assembly and metallicity content, especially on spatially resolved scales, remains contentious and requires further exploration, particularly through methodologies like Integral Field Spectroscopy (IFS).

The Calar Alto Void Integral-field Treasury SurveY (CAVITY) is a comprehensive project aimed at studying galaxies within cosmic voids using IFS data (Pérez et al. 2024). Its primary focus lies in examining the spatially resolved stellar populations, stellar kinematics, and physical properties of ionized gas within these galaxies, alongside discerning their dark matter content, acknowledging the non-linear correlation between baryonic mass and dark halo. CAVITY design ensures thorough investigation of void galaxies, spanning various void sizes and dynamical stages while maintaining adequate spatial coverage across galaxy disks. Moreover, the project seeks to interrogate the predictions of the ΛCDM model regarding galaxy mass assembly across diverse environments, addressing the model limitations in predicting baryonic physics-dominated scales at the galaxy level. Serving as a benchmark for juxtaposing numerical predictions from ΛCDM, CAVITY aims to deepen our understanding of the interplay between large-scale environments and galaxy formation within voids, ultimately elucidating the influence of the cosmic environment on void galaxy mass assembly and evolution.

Several scientific objectives have been explored using the CAVITY parent sample. In one study, Domínguez-Gómez et al. (2023b) utilized Sloan Digital Sky Survey (SDSS) spectra to analyze a well-defined group of galaxies in voids, walls/filaments, and clusters. By applying non-parametric

full spectral fitting techniques, they measured the stellar metallicity in the central regions of these galaxies as covered by the SDSS fiber. They found that void galaxies possess lower stellar metallicities compared to those in filaments and walls, particularly among galaxies with lower stellar mass. Another study by Domínguez-Gómez et al. (2023a), using the same methodology, examined the star formation histories (SFH) of galaxies in various environments. Their findings indicated that, on average, galaxies in voids assemble mass more slowly at a given stellar mass than galaxies in filaments or clusters.

In a more recent investigation utilizing IFS spatially resolved data from the CAVITY survey, Conrado et al. (2024) found that void galaxies generally have a slightly larger half-light radius (HLR), lower stellar mass surface density, and younger ages across all morphological types compared to their counterparts in filaments and walls. Additionally, void galaxies exhibited slightly higher star formation rates (SFR) and specific star formation rates (sSFR), especially in Sa-type galaxies. These differences were most significant in the outer regions of spiral galaxies, where the discs appeared younger and had higher sSFR, suggesting less evolved discs. Early-type spirals in voids also demonstrated a slower transition from star-forming to quiescent states. Overall, the study concluded that void galaxies evolve more gradually, particularly in their outer regions, with environmental effects having a more pronounced impact on low-mass galaxies.

In this article, we introduce the first data release (DR1) of CAVITY, which provides public access to 100 high-quality data cubes observed with the V500 setup. Initially conceived as an IFS survey, CAVITY has undergone expansion to amplify its scientific impact and comprehensively address its proposed objectives. This expansion, named CAVITY+, entails the inclusion of additional observational techniques such as HI, CO, and deep optical imaging of the IFS targets. The current data release focuses solely on IFS data.

This paper is organized as follows. The properties of the galaxies in the DR1 sample are summarized in Section 2. We describe the observing strategy and setup (Sect. 3), processing (Sect. 4), data format (Sect. 5), and data quality (Sect. 6), which comprise essential information for any scientific analysis of the distributed CAVITY data. Access to the CAVITY DR1 data is explained in Sect. 7.

2 The CAVITY DR1 sample

The CAVITY project focuses on identifying and characterizing cosmic voids in the nearby Universe, drawing its sample from the SDSS data catalogued by Pan et al. (2012). Ensuring completeness within the redshift range of 0.005 to 0.050 allows for the inclusion of voids spanning large sizes. From this initial void catalogue (Pan et al. 2012), and taking into consideration the scientific goals of the CAVITY project (Pérez et al. 2024), we established a criterion of at least 20 galaxies per void, culminating in 80 voids containing 19,732 galaxies. Subsequent refinement excluded voids at the edges of the SDSS footprint, leaving 8690 galaxies within 42 voids. Representative voids were then chosen to cover a spectrum of properties, ensuring observability from the Calar Alto Observatory while maintaining the distribution of key galaxy properties. This resulted in a final sample of 4866 galaxies in 15 voids, termed the CAVITY parent sample. Additionally, to guarantee sample purity, galaxies associated with clusters were meticulously removed. A criterion based on the distance to the centre of the voids in units of the void effective radius (within the inner 0.8 × void effective radius) was employed to prioritize galaxies in the innermost regions of voids, covering the full extent of them while avoiding filamentary-like environments. Also, only galaxies with intermediate inclinations (20 to 70 degrees) were selected for observations. This led to a final observable sample of 1115 galaxies after an extra visual refinement process (Pérez et al. 2024). Target galaxies are constantly selected based on observability of galaxies from this parent sample. Since the beginning of the observations, the project has made use of 221 nights distributed between January 2021 and June 2024 and gathered 246 cubes. This initial data release comprises 100 galaxies, chosen based on rigorous quality control measures (Sect. 6).

Figure 1 provides a brief overview of the main characteristics of the three different CAVITY samples (namely, parent sample in light-green, observable sample in violet, and final DR1 sample in red). In order to ensure a 2D analysis of the properties of the galaxies, only those with a d_{25}[1] above ~20" were observed (see top panel), with all galaxies being fully covered by the PMAS/PPak FoV (around 70 arcsecs). Also, in order to allow for a proper study of the stellar and gas kinematics of the galaxies, systems with edge-on and face-on orientations are avoided. Thus, CAVITY DR1 galaxies exhibit only intermediate inclinations, with values ranging from ~ 20 to 70 degrees (middle panel of Fig. 1, by definition), with the bulk of galaxies displaying inclinations between 50 and 60 degrees. Finally, besides observability, galaxies to be observed as part of CAVITY are sorted according to the distance to the centre of the void, prioritising observations of galaxies near the centres first. The number of galaxies within voids greatly decreases as we approach their cores, however, to maximize the scientific impact of the project, we aimed at covering all 'void-centric' distances up to 0.8 void effective radius[2], where the void environment gives way to a more filamentary-like surrounding. The distribution of distances to the void centres is shown in the bottom panel of Fig. 1. For completeness, we also show in that figure the distribution of redshifts as well as the distributions of morphological types and stellar mass (see below). All these quantities are collected and explained in detail in the CAVITY DR1 'master' table (see Table 5, also described in Appendix B).

Figure 2 shows the distribution of galaxies in the color-magnitude diagram (using SDSS magnitudes). The CAVITY DR1 sample mostly follows the overall distribution of the parent and observable samples (see also morphological type panel in Fig. 1), with the exception of the very faint tail (galaxies fainter than absolute magnitude in r-band of -18 hardly enters the DR1 sample). This is mainly due to the

[1] For each galaxy we consider d_{25} to be three times the value of its *petroR90-r* from SDSS, see 'Master' table description in Appendix B.

[2] Computed as the radius of a sphere with the same volume as the actual void.

difficulty of gathering high-quality IFS data for these faint (and generally small) galaxies. Apart from this, we should highlight a subtle excess of galaxies in the bright end of the population (brighter than M_r=-20). We will consider this in the planning of the coming observations, prioritizing the observations of fainter systems. This current lack of faint galaxies and excess of bright galaxies is translated into a slightly unbalanced stellar mass distribution (last panel of Fig. 1).

Figure 3 presents some spectra belonging to two representative CAVITY DR1 galaxies. The left-hand panels correspond to a red, quiescent galaxy, and the right-hand ones to a blue, star-forming galaxy. For each galaxy we show in the top panels its SDSS colour image (left) as well as its integrated light within the covered wavelength range extracted from the CAVITY datacube (right). We represent the spectra of their central spaxel in the middle panels and the integrated spectra up to 2 effective radii (violet ellipse in the images) in the bottom panels. This figure highlights the richness in physical properties covered by the CAVITY sample, as well as exemplifies the quality of the CAVITY DR1 dataset.

3 Observing strategy and setup overview

The observations of the CAVITY survey commenced in January 2021, conducted at the 3.5 m telescope of the Calar Alto observatory using the Potsdam Multi Aperture Spectrograph, PMAS (Roth et al. 2005), in the PPak mode (Verheijen et al. 2004; Kelz et al. 2006). The PPak fiber bundle comprises 382 fibers, each with a diameter of 2.7". Among these fibers, 331 (referred to as the science fibers) are concentrated in a single hexagonal bundle covering a field-of-view (FoV) of 74"× 64", achieving a filling factor of approximately 60%. The sky background is sampled by an additional set of 36 fibers, distributed in 6 bundles of 6 fibers each, positioned along a circle approximately 72" from the center of the instrument's FoV. These sky fibers are interspersed among the science fibers within the pseudo-slit to ensure a comprehensive characterization of the sky, sampled with a similar distortion as the science fibers, while the remaining 15 fibers are utilized for calibration purposes.

All DR1 CAVITY galaxies are observed using the V500 grating. It offers a nominal resolution of $\lambda/\Delta\lambda \sim 850$ at $\sim$ 5000 Å with a full width at half maximum (FWHM) of about 6 Å. Paired with the PMAS CCD E2V CCD231 4k×4k, installed in 2009 (Roth et al. 2010), it spans nominally from [o ii] λ 3727 to [s ii] λ 6731 in the rest frame for all objects in the survey. This wavelength coverage facilitates the investigation of the primary scientific objectives of the project, through the examination of stellar populations and the characteristics of the ionized gas, thanks to the diverse range of spectral features and emission line species available in the spectra.

The optical system exhibits some vignetting at the four corners of the CCD. This vignetting effect is illustrated in fig. 4 of Sánchez et al. (2012). It is characterized by a decrease in transmission at the four edges of the map. Approximately 30% of the spectra near the edge of the detector experience some efficiency loss, which gradually worsens towards the corners of the CCD. Vignetting impacts no more than a quarter of the full spectral range, severely affecting less than 15% of the fibers and less than 25% of their wavelength range. The useful wavelength range varies in those fibers affected by vignetting, with the worst cases being reduced to 4240 to 7140 Å. These effects manifest in the FoV as an annular ring at approximately ∼15" from the center of the FoV (see fig. 4, right panel of Sánchez et al. 2012 and fig. 11 of Husemann et al. 2013).

To ensure complete coverage of the central bundle's FoV and enhance spatial resolution, a dithering scheme employing three pointings has been implemented. This approach, akin to the one employed in the CALIFA survey (Sánchez et al. 2016), conducted with the same instrumental setup, aims to mitigate the vignetting effect. The scheme utilizes wider offsets in RA and Dec: 0.00", 0.00"; -5.22", -4.84"; -5.22", +4.84". By adopting this strategy, any spectrum affected by vignetting in the final dithered and rebinned dataset benefits from compensation. This is achieved by incorporating spectra from at least two adjacent, unvignetted fibers positioned closer than 2".

The exposure times were determined based on our previous pilot project of void galaxies using the instrument. These times remain fixed for all targeted objects, which have been categorized into two groups: bright and faint galaxies. To categorize galaxies as faint or bright systems, we study elliptical isophotal profiles from *r*-band, DR12 SDSS photometric data. Galaxies with a surface brightness fainter (brighter) than 22.1 mag/arcsec^2 at `petroR50_r` are considered faint (bright) galaxies. Exposure times for the V500 range between 1.5 and 3.0 hours of total integration, depending on the galaxy brightness. Each dithering position involves two exposures of 900 seconds for the brightest targets, while four 900-second exposures are taken for the faintest galaxies. For the 1.5-hour observations, one set of calibration files is obtained per galaxy. For longer total integration times, a minimum of two sets of calibration exposures are acquired. All observations are conducted at airmasses below 1.4.

Given the simultaneous operation of multiple projects at the observatory on the same telescope, the allotted time for each project within a single night may be fractional. Consequently, due to the relatively lengthy exposure times per galaxy, certain galaxies may be observed across multiple nights. To minimize uncertainties related to positioning and transparency changes between nights, we enforce the completion of a single dithering position within the same night. For faint galaxies, this entails gathering at least the four 900-second exposures of one pointing and their corresponding calibration files on the same night. As a result, a galaxy may be observed within a single night, two nights, or three nights. Approximately 51% of the galaxies are observed entirely within one night, 42% within two nights, and the remainder over three nights, with one pointing conducted each night. Pointings that do not meet minimum quality control standards (see Section 6) are rescheduled and re-observed to ensure complete observation. These pointings are re-observed entirely, with either two or four single exposures. In the following section, we detail the primary stages of the CAVITY data reduction pipeline utilized in generat-

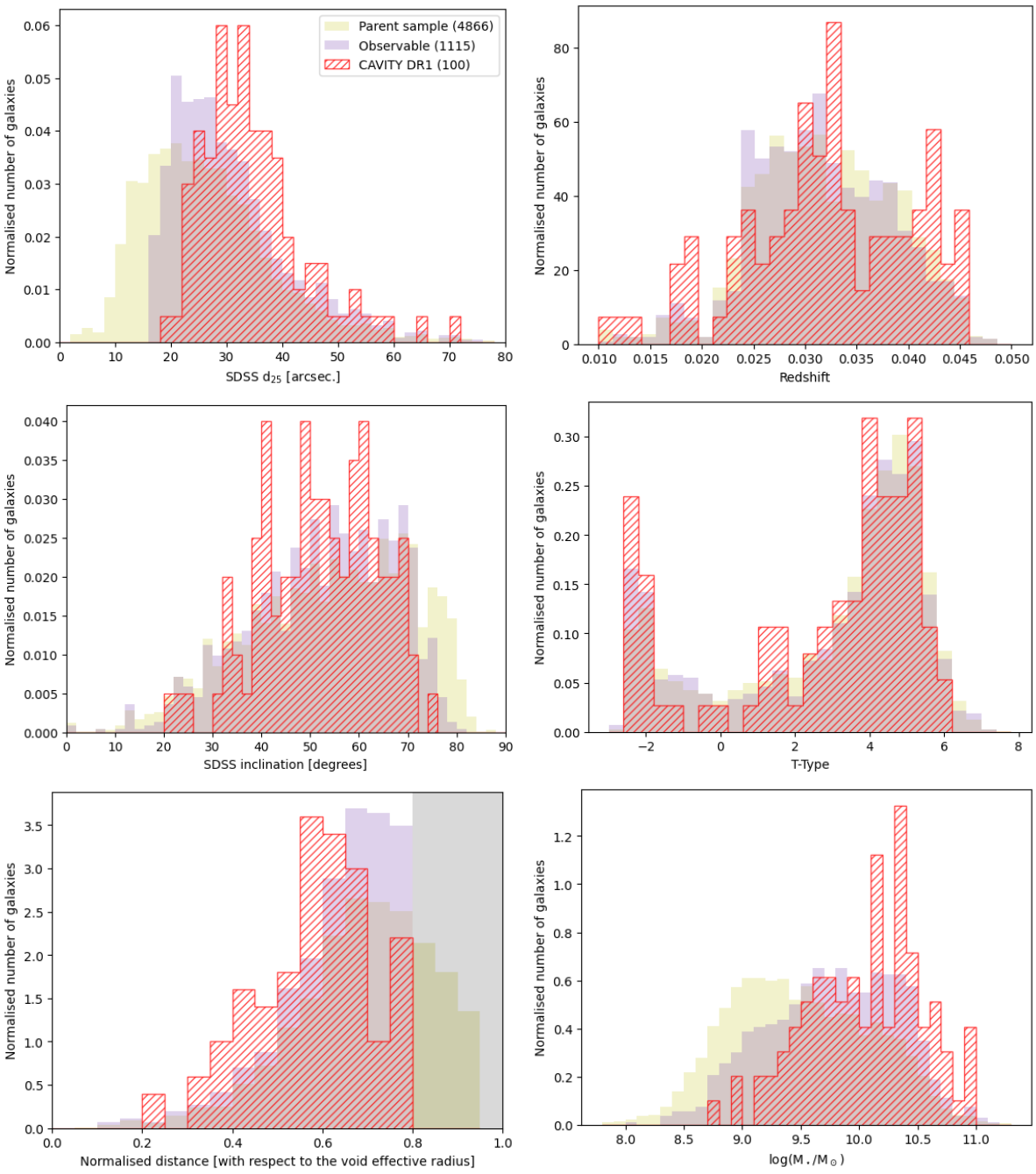

Figure 1: Characterisation of the DR1 CAVITY sample in comparison with the CAVITY parent sample and observable sample from which observed galaxies are drawn (see text for details). From top to bottom and left to right, we show the distribution of d_{25}, the stellar mass, the inclination from SDSS, the morphological T-Type, the normalised distance to the center of the voids and the redshift. In light-green are the histograms corresponding to the CAVITY parent sample, in violet the observable sample, and in red the final CAVITY DR1 sample. The shaded area in the normalised distance to the centre of the void panel histogram (bottom-left) represents the most external parts of the void (area between 0.8 and 1.0 times the void effective radius. This region is avoided as it can be considered as a transition zone from the void to the filamentary-like environment.

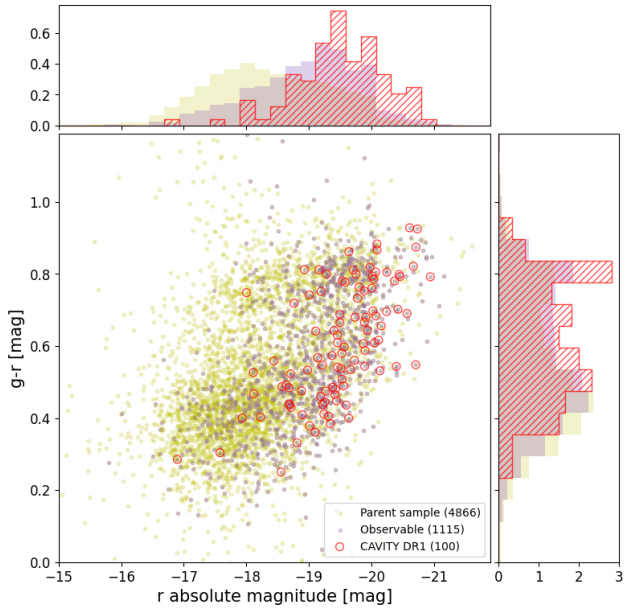

Figure 2: Distribution of CAVITY galaxies in the SDSS g−r vs. M_r color-magnitude diagram. Yellow dots (histograms) denote galaxies in the CAVITY parent sample (4886 galaxies), while violet and red are used for the observable subsample (1115) and the final CAVITY DR1 sample (100). Colour and absolute magnitude distributions are displayed in the top and side panels.

ing the DR1 data.

4 Data processing

The CAVITY data reduction process relies on an automated pipeline that operates independently, requiring minimal human intervention. It produces scientifically significant frames along with quality control metrics to assess the precision of the processed data. The CAVITY pipeline (V1.2) adheres to the fundamental reduction procedures and protocols outlined in the CALIFA survey (Sánchez et al. 2012; Husemann et al. 2013). Tailored to accommodate the specific requirements of the CAVITY survey, it incorporates novel architecture and efficiency enhancements. Implemented in Python 3, the pipeline leverages Cython code (Behnel et al. 2011) to address select computationally intensive tasks. The following outlines the current implementation of the data reduction process.

4.1 CCD readout procedure and frame combination

The reduction workflow encompasses the propagation of Poisson and read-out noise, cosmic rays, defective CCD columns, and vignetting effects. As noted, the CCD employed since October 2009 comprises a 4k×4k E2V detector. The read-out software for the CCD stores each frame in four distinct FITS files, corresponding to the four amplifiers of the detector, each characterized by varying bias and gain levels. The standard procedure since then has been the 4-amplifier read mode, generating four individual files, each with its own bias level and gain. However, in December 2021, CCD quadrant 'a' exhibited abnormally higher readout values, with approximately 2.5 times more noise and a reduction of 2.5 times in counts in well-illuminated fibers. Consequently, the CCD dewar had to be unmounted and shipped to Leibniz-Institut für Astrophysik Potsdam (AIP) for intervention. Suboptimal cable routing, leading to stress, was rectified, and loose connections were addressed during the intervention. Additionally, the read-out procedure was modified to the 2-amplifiers 'bc' mode, which demonstrated better stability and performance post-fixing. The CCD dewar was subsequently returned and re-installed in the instrument by the end of June 2022. During the commissioning and testing phase, various observations were conducted to ensure optimal performance. However, these observations conducted during commissioning did not meet the quality standards required for the survey. Consequently, the official resumption of CAVITY observations occurred in November 2022, following confirmation of system stability and optimal conditions. Thus, observations conducted before December 2021, prior to the CCD failure, were processed using the "classic" 4-amp read mode, which had been in use for over a decade. Meanwhile, observations were carried out after November 2022 were carried out using the 2-amp 'bc' mode, which has since become the standard mode for CAVITY.

Initially, the four (prior to 2022) or two (after 2022) distinct FITS files from the 4k×4k E2V detector are merged into individual frames for each exposure. Subsequently, each frame undergoes pixel cleansing, which includes cosmic ray removal facilitated by PyCosmic (Husemann et al. 2012).

4.2 Spectral extraction

The PMAS instrument is susceptible to flexures (Roth et al. 2005), which induce slight shifts of the signal on the CCD along the dispersion and cross-dispersion directions depending on the hour angle of the telescope. To mitigate this effect, calibration frames are acquired within a maximum range of 1.5 hours from the science frames. For longer exposures of low-luminosity galaxies, two calibration files are obtained. These calibration files are paired with science exposures for individual frames based on their temporal proximity. This approach ensures that, in most instances, the flexure pattern of the calibration and science frames is mitigated. Relative flexure offsets are assessed concerning the continuum and arc-lamp calibration frames, and the wavelength solution is adjusted for each individual science frame. The wavelength solution of sky lines is assessed against expected values, with any discrepancies attributed to flexure effects and corrected for individual science frames. These offsets exhibit small values, typically ≤ 0.5 pixels.

The straylight map is reconstructed from fiber trace gaps and subtracted from calibration and science exposures. Fiber profile widths are determined using continuum lamp fiber positions for spectral extraction, with some fibers exhibiting over 15% stray-light contribution depending on throughput and input signal, impacting object SN ratio. Reconstruction involves smoothing data with a 20-pixel

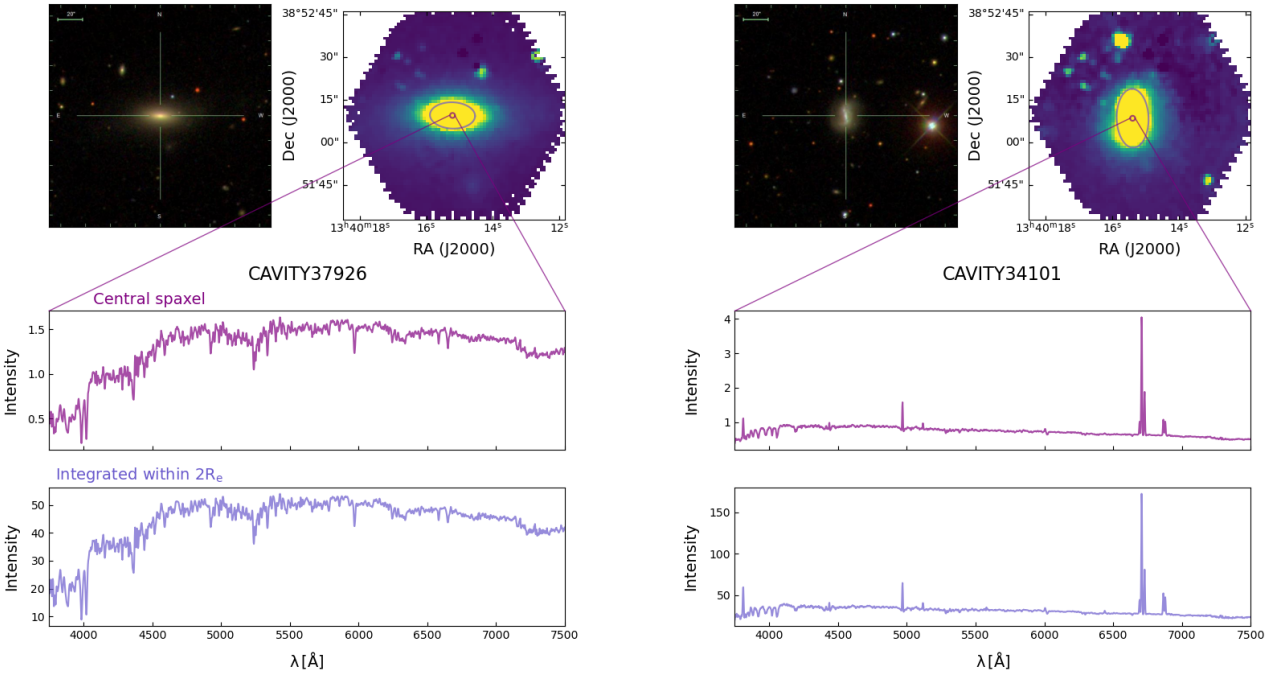

Figure 3: Spectra of two representative CAVITY DR1 galaxies: CAVITY37926 (a red, quiescent galaxy; left column), and CAVITY34101 (a blue, star-forming galaxy; right column). *Top-left panels:* SDSS colour image of the galaxies. *Top-right panels:* Moment 0 image reconstructed from the CAVITY datacube summing all the integrated light within the covered wavelength range (from 4500 to 7000 Å). *Middle* and *bottom panels* display the spectra of their central spaxels and the integrated spectra up to 2 effective radii (violet ellipse in the moment 0 images), respectively. Moment 0 images are in units of 10^{-16}·erg·cm^{-2}·s^{-1}, and spectra in units of 10^{-16}·erg·cm^{-2}·s^{-1}·Å^{-1}.

median filter and fitting sixth-order polynomials to cross-dispersion slices, followed by 2D Gaussian kernel smoothing to suppress high-frequency structure. Straylight subtraction precedes fiber spectrum extraction, ensuring proper inclusion of stray-light effects in the data error budget.

Although adjacent spectra on the CCD can map to different locations in the spatial plane (Kelz et al. 2006), the cross-talk, when employing a modified version of the Gaussian suppression (Sánchez 2006), is minimized to less than 1%. Fiber profile widths are determined by averaging over blocks of 20 fibers, employing χ^2 minimization with integrated fiber fluxes and a common FWHM per block as free parameters, with fixed fiber positions derived from continuum lamp data, including flexure offset adjustments. These measured widths are then interpolated along the entire dispersion axis using a 5th-order polynomial. Spectrum extraction utilizes the optimal extraction algorithm (Horne 1986), with Gaussian fiber profile positions and widths fixed based on previous measurements. Optimal extraction allows for error propagation based on individual pixel errors and exclusion of bad pixels, although unreliable flux extraction occurs if all three central pixels of a fiber are flagged as bad, warranting flagging of the corresponding spectral resolution element to prevent potential data artifacts.

The flux extracted for each pixel along the dispersion direction is stored in a Row-Stacked-Spectrum file (RSS). Wavelength solution and spectral resolution variations are determined for each fiber using HeHgCd calibration lamp exposures within each calibration dataset. This data is employed to resample spectra onto a linear wavelength grid, incorporating flexure offsets in the wavelength solution explicitly to obviate an additional resampling step. Spectral resolution is homogenized using an adaptive Gaussian convolution. Error propagation during Gaussian convolution is performed analytically, while a Monte Carlo method is employed to estimate noise vector post-spline resampling. To address bad pixels, a two-pixel expansion of the bad pixel mask along the dispersion axis is implemented, followed by setting error values of bad pixels to a high value ($\sim 10^{10}$ counts) and replacing data with linear interpolation from the nearest unmasked pixels. Relative fiber-to-fiber transmission differences are addressed by comparing wavelength-calibrated science frames with twilight sky exposures. Additionally, pixels exhibiting a transmission drop below 70%, are masked.

4.3 Flux calibration

Flux calibration was conducted using a master sensitivity curve derived from dozens of observations performed prior to the CAVITY observations, following the one adopted in (García-Benito et al. 2015, Sect. 4.2). For future CAVITY DRs, we plan to compute sensitivity curves corresponding to the years of observations within the project to generate an updated sensitivity function. This approach will enable us to assess any differences compared to the previous computed curve, particularly in light of changes in the readout mode of the CCD as described in Sect. 4.1, and any evolution in fiber transmission.

To facilitate this future analysis, each night, we observe spectrophotometric standard stars listed in the Oke Catalogue[3] (Oke 1990). Whenever feasible, we observe two different calibration stars on the same night. To mitigate potential flux losses stemming from the filling factor of PPak, instead of utilizing a single pointing, we observe stars using the same dithering pattern as for galaxies. This strategy aims to optimize sensitivity across the wavelength range while addressing flux losses inherent to PPak's filling factor when using one single pointing (García-Benito 2009; García-Benito et al. 2010).

The science RSS files for individual frames from each pointing are corrected for atmospheric extinction across the wavelength range. This correction takes into account the airmass and the monitored V band extinction, utilizing the extinction measured by the Calar Alto Visual Extinction monitor (CAVEX) at the moment of the observations and the average extinction curve for the observatory (Sánchez et al. 2007). Subsequently, the sensitivity curve is applied to these corrected frames, converting them from observed counts to intensity, thereby ensuring flux calibration (see Sect. 6.2.3 for details on the accuracy of the spectrophotometric calibration).

4.4 Sky subtraction

Following flux calibration, frames undergo correction for telluric lines and sky subtraction. PPak is equipped with 36 fibers for sampling the sky, arranged around the science fiber-bundle in six small bundles of six fibers each, positioned approximately 75" from the center of the FoV (Kelz et al. 2006). Objects selected for the CAVITY sample occupy only a portion of the FoV of the central PPak bundle, ensuring that all sky fibers remain unaffected by emission from the corresponding target. A sky spectrum is generated by combining the 36 sky fibers of PPak. To mitigate issues arising when a bright field star or neighboring galaxy fills an entire sky fiber bundle, we calculate the median of the 30 faintest sky fibers. Each derived night-sky spectrum is then used to determine the night-sky brightness by convolving the spectrum with the transmission curves of the Johnson V-band filter. This value serves to monitor the actual conditions during data acquisition.

4.5 Cube reconstruction

The individual frames for each pointing are consolidated into a single pointing RSS file. To address potential variations in atmospheric transmission during the science exposures, these three pointings are standardized to a common intensity and response function. This standardization is achieved by comparing the integrated spectra of the fibers within a 30" diameter aperture. Subsequently, the science spectra corresponding to the three dithered exposures are merged into a single frame comprising 993 spectra. The individual position tables corresponding to the PPak central bundle are combined using the offsets provided for the dither. Finally, the spectra undergo correction for Galactic extinction (Schlegel et al. 1998; Cardelli et al. 1989).

[3] BD+25d4655, BD+28d4211, BD+33d2642, Feige 34, and HD 93521, are among the most commonly observed ones. Their corresponding flux-calibrated spectra are publicly accessible on the CAHA's webpage: http://www.caha.es/pedraz/SSS/sss.html.

In this context, in the current implementation of the CAVITY pipeline for the DR1, we depart from other registration methods such as the one detailed in García-Benito et al. (2015). In that registration approach, the flux for the 331 fiber apertures in each pointing is derived from sky-subtracted SDSS broad-band images covering the wavelengths of the observation. These apertures undergo shifts in both right ascension and declination within a search box centered on the nominal pointing coordinates. The optimal registration is determined by comparing χ^2 values between the broad-band aperture-matched fluxes and those obtained from the RSS spectra. The photometric scale factor identified at the best-matching position is then utilized to recalibrate the absolute photometry of each individual RSS pointing.

This approach entails that all IFS observations are accompanied by equivalent broad-band counterpart images for all galaxies. This is not problematic for CAVITY, as its sample is drawn from SDSS, where all galaxies have fully calibrated optical images in DR16 (Ahumada et al. 2020). However, the methodology outlined in García-Benito et al. (2015), was designed for CALIFA observations, where galaxies mostly fill the entire PPak FoV. This facilitated a suitable match, enabling a relatively high number of fibers to capture galaxy light for aperture flux comparison. However, in the case of small and relatively faint galaxies, where only a few fibers are available for comparison, this methodology does not perform effectively. Indeed, for a few galaxies in the CALIFA survey with these characteristics, this registration approach did not work, as noted at the end of Section 4.2 in García-Benito et al. (2015). This scenario mirrors what occurs with most of CAVITY galaxies, which occupy only a fraction of the FoV. In future iterations of the pipeline, we will explore adaptations of this methodology to tailor it to the specific characteristics of the CAVITY data and investigate the feasibility of employing a variant methodology utilizing broad-band imaging registration in individual pointing frames.

After applying the Galactic extinction correction, the RSS is prepared for fiber spatial rearrangement and resampled spatially to form a datacube with a regular grid. We employed a flux-conserving inverse-distance weighting scheme (Sánchez et al. 2012), which is a variation of Shepard's interpolation method (Shepard 1968). The positions of each RSS pointing obtained in the preceding step are utilized for image reconstruction. We set the dispersion of the Gaussian to 0.75″ and constrain the kernel within a radius of 3.50″, resulting in a datacube with a spatial sampling of 1.00″, following García-Benito et al. (2015) and Sánchez et al. (2016).

The effect of differential atmospheric refraction (DAR) on data can be assessed empirically by tracing the centroid of the galaxy along the wavelength axis of the spatially resampled datacube. Typically, each monochromatic image slice is resampled to align the galaxy centroid to a common reference position. To streamline error propagation, a two-stage iteration is employed. First, the datacube is reconstructed, and the DAR effect is estimated. Next, the datacube is reconstructed again with fiber positions adjusted based on empirically measured DAR offsets. While error vectors are well-defined for each spectrum in reconstructed cubes, the inverse-distance weighting scheme introduces signal correlation in the spatial dimension, discussed in more detail in Sect. 4.6.

For optimal absolute flux calibration, considering the impracticality of individually registering pointing frames on a fiber basis, we conducted recalibration using large aperture SDSS photometry, as detailed in the methodologies of Sánchez et al. (2012) and Husemann et al. (2012). The broad-band SDSS photometry is available for all targets as per the CAVITY sample design. Notably, the V500 grating data comprehensively covers the g (λ_{eff} = 4728 Å) and r (λ_{eff} = 6200 Å) filters (Doi et al. 2010) among the five SDSS filters (*ugriz*). To carry out the primary recalibration, we first measured the counts for each galaxy within a 30" diameter aperture in the SDSS images. These counts were then converted to flux density using the counts-to-magnitude formula provided in the SDSS documentation[4]. We then extracted spectrophotometric data from the reduced V500 datacubes. This involved summing the flux from individual spectra within a 30" diameter aperture and convolving the resulting spectrum with the SDSS g and r filter passbands. We proceeded to derive a scaling solution by averaging the flux ratios in both bands and adjusted the datacube flux and error accordingly.

Similar to the registration process at the individual pointing level, a second-order correction can be applied at the cube level to align the cube spectrophotometry with the corresponding broad-band images. The ratio between broad-band and reconstructed cube images can be computed to adjust fluxes across the datacube without altering spectral shape. The FLAT correction, first introduced in CALIFA DR3 (Sánchez et al. 2016) and subsequently applied in the eCALIFA remastered sample (Sánchez et al. 2023), was incorporated into the CALIFA datacubes as an additional extension. However, this method is effective only if the galaxy covers a substantial portion of the cube and there are sufficient S/N spaxels for the matching procedure. In the case of the CAVITY sample, similar to individual pointing registration, these conditions are not met for a significant portion of the targets, rendering the method inapplicable. In future pipeline versions, we will explore alternative methodologies tailored to the characteristics of the CAVITY dataset.

4.6 Noise covariance

The interpolation procedure used to achieve a regular grid causes the output pixels in the final datacube to be interdependent. The Gaussian interpolation method distributes flux from a given fiber across multiple pixels, creating correlated noise between adjacent pixels. This correlation leads to an underestimation of noise in a stacked spectrum of N pixels, affecting statistical error estimates when coadding spectra in datacubes. Therefore, it is crucial to account for this noise covariance to ensure correct error propagation when combining or spatially binning spectra to enhance the SN ratio, which is often necessary in analyzing IFS data.

One approach to addressing noise covariance is to compute the covariance matrix (Sharp et al. 2015). In contrast,

[4] https://live-sdss4org-dr16.pantheonsite.io/algorithms/fluxcal

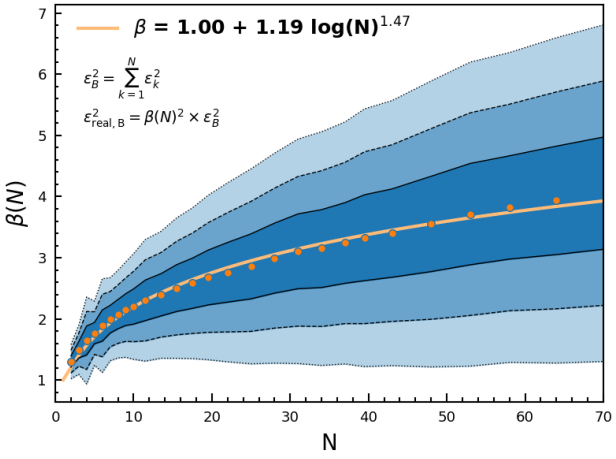

Figure 4: Correction factor to the noise introduced by the spatial covariance, β(N) (ratio of the real estimated error to the analytically propagated error), as a function of the number of co-added adjacent spaxels for all the DR1 V500 data. Shaded regions indicate the 1σ, 2σ, and 3σ width areas around the average value, represented by solid circles. The solid line shows the best-fitted model, following logarithmic parameterization, as shown in the legend.

following the methodology of García-Benito et al. (2015), we adopted an empirical approach to account for noise correlation by introducing the noise correction factor β(N). This factor is the ratio between the error estimated on the combined spectrum and the error derived by adding the individual errors of each spectrum quadratically.

To estimate this parameter, we randomly generated a set of circular apertures of various sizes (N spaxels) on each datacube, ensuring at least 20 apertures of the same size at different center positions on spaxels with a S/N > 3. For each aperture, we empirically estimated the error of the coadded spectrum as the standard deviation from the detrended spectrum within the 5590–5710 Å range in the rest frame. We also calculated the expected error, assuming no covariance, by propagating the individual empirical errors (measured in the same window and in the same way) within the aperture. The ratio between these two errors provides the β value for the corresponding spectrum.

Figure 4 illustrates the distribution resulting from over 100k estimations of β(N) on the CAVITY DR1 data as a function of the number of co-added spaxels/spectra (N). Following García-Benito et al. (2015), we parameterised this distribution using a logarithmic function. However, unlike previous estimations, we included an exponent term to better fit the shape of the median distribution:

$$\beta(N) = 1 + 1.19 \log(N)^{1.47} \quad (1)$$

This function can then be applied to the noise spectrum from the data cube by multiplying the β(N) parameter by the error of ϵ_B, the quadratic sum of the vector noise of each spectrum in the aperture.

5 CAVITY data format

The final reduced data for each observed object are consolidated into a single file, following the standard binary FITS format, and organized as datacubes, each comprising four FITS header/data units (HDU). The format mirrors that of CALIFA (Husemann et al. 2013) and other released IFS galaxy surveys. This multi-extension FITS file comprises various outcomes of the reduction process, detailed in Table 1. These datacubes encompass (1) measured flux densities, corrected for Galactic extinction, measured in units of 10^{-16} erg s^{-1} cm^{-2} Å^{-1} (primary datacube), (2) associated errors, (3) error weighting factors, (4) flags denoting bad pixels, and (5) fiber coverage. These data structures enable users to appropriately consider dataset characteristics such as bad pixels and noise for their specific analyses.

The bad pixel FITS extension identifies pixels interpolated due to insufficient data from cosmic rays, defective CCD columns, or vignetting, and assigns them a large error value (~10^{10}) in the error vector stored in the error HDU extension. These bad pixels are unevenly distributed within the datacube, with four bad CCD columns reducing the number of valid pixels in specific wavelength regions. The hexagonal PPak FoV is resampled to a rectangular grid, where uncovered corners are flagged as bad pixels, while residuals from bright night-sky emission lines remain unflagged.

The spatial dimensions of the datacubes are delineated along right ascension and declination in the first two axes, each sampled at 1"× 1". The third axis represents the wavelength, sampled on a linear grid with 2 Å steps. Table 2 provides a summary of the dimensions of each datacube (Nα, Nδ, and Nλ), as well as the spectral sampling (dλ) and resolution (δ_λ) across the entire wavelength range.

These files are named after the official object names from the CAVITY mother sample, adopting the format CUBENAME.V500.rscube.fits, with CUBENAME corresponding to the OBJECT header.

The FITS header encompasses the standard keywords that map the spatial axes onto the standard World Coordinate System (Greisen & Calabretta 2002), and the wavelength onto the spectral axis within a linear grid. Within each CAVITY datacube resides the complete FITS header information of all constituent raw frames, wherein each header entry is augmented with a distinct prefix corresponding to a specific pointing/frame. This prefix comprises the grating utilized (`GRAT`; V500 in this DR), followed by "PPAK" and the identification of the pointing frame $\mathtt{P_iF_j}$, where P_i denotes the pointing (ranging from 1 to 3), and F_j represents the frame within a particular P_i pointing (with j varying from 1 to 4 based on the luminosity of the galaxy; 2 frames for bright galaxies and 4 frames for low-luminosity galaxies). Information pertaining to the combined pointing, where all individual frames are amalgamated into a single pointing file, follows the structure `GRAT` PPAK P_i. Moreover, the reduction pipeline aggregates supplementary details such as sky brightness, Galactic extinction, etc., and integrates them into the FITS header. Header keywords that hold potential significance for data analysis and/or assessment are succinctly outlined in Table 3 for ease of reference.

Table 1: CAVITY FITS file structure.

HDU	Extension name	Format	Content
0	Primary	32-bit float	Flux density in units of 10^{-16} erg s^{-1} cm^{-2} Å^{-1}
1	ERROR	32-bit float	1σ error on the flux density
2	ERRWEIGHT	32-bit float	Error weighting factor
3	BADPIX	8-bit integer	Bad pixel flags (1=bad, 0=good)
4	FIBCOVER	8-bit integer	Number of fibers used to fill each spaxel

Table 2: Dimension and sampling of CAVITY datacubes.

Setup	N_αa	N_δa	N_λa	$\lambda_{\rm start}$b	$\lambda_{\rm end}$c	d_λd	δ_λe
V500	78	73	1877	3749Å	7501Å	2.0Å	6.0Å

aNumber of pixels in each dimension. bWavelength of the first pixel on the wavelength direction. cWavelength of the last pixel on the wavelength direction. dWavelength sampling per pixel. eHomogenized spectral resolution (FWHM) over the entire wavelength range.

6 Data quality

In order to select the 100 galaxies to be released in the CAVITY DR1, we have run a careful quality control (QC) on the resulting datacubes to ensure that the released data comply with the standard requirements for their use by the scientific community. In this section we detail the QC checks that have been performed for such purpose. These checks are first based on a visual inspection of the datacubes by the QC team (Sect. 6.1), which is followed by a final quality assessment based on a set of measured parameters (Sect. 6.2).

6.1 Visual quality checks

A visual inspection of each final reduced datacube is performed to estimate the quality of the data and identify any problem with the observations or the reduction that highly affects the cube reconstruction and/or the individual spectra. This QC is carried out based on a series of graphics that are automatically generated for each datacube after one round of observations is reduced. To quickly assess the quality of the data and their suitability for scientific use, four flags are defined:

1. CUBE. This flag allows to identify problems related to the cube reconstruction, such as the presence of a grid associated with the dithering scheme (which might uncover issues with a particular pointing) or a target with very low surface brightness unsuitable for our observing strategy. To flag these problems, a white image of the galaxy (from 4500 to 7000 Å) is provided together with a map of its negative values (low-transmission fibers clearly stand out in this plot, see Sect. 6.3.1).

2. SPECTRA. We indicate here if the spectra present any distortion or odd features, or an atypical shape in the continuum emission. For that, the reviewers are given for each galaxy the central spectrum, together with the integrated spectrum within 0.5, 1, and 2 effective radii.

3. VIGNETTING. Overall, CAVITY galaxies tend to be small ($d_{25} < 40$ arcsec), so the vignetting effect (concentrated in an annular ring at approximately ~15" from the center of the FoV, see Sect. 3 for more details) only affects the observations in the very outer parts of the galaxies. However, for larger galaxies, or if there is an offset of its center with respect to the center of the FoV, this effect can impact the spectra of the inner parts and hamper for instance the analysis of the stellar populations. This flag allows us to identify such cases, by means of observing a clear jump at the ends of the integrated spectra within 0.5 or 1 effective radii or in the white image if the galaxy extends across the vignetting area marked with a ring of semi-transparent dots (a contour indicating the 23 mag/arcsec^2 level of the galaxy surface brightness allows to delimit the galaxies).

4. JUMPCCD. The reported issue with the CCD quadrant 'a' (see Sect. 4.1 for details) produced for some observations the presence of a jump in the spectra at around 5800 Å and a sharp step in the sky level between the top and bottom halves of the instrument FoV (see also Sect. 6.3.2). All observations affected by this issue were discarded and the galaxies were re-observed once the problem was fixed. This flag helps to identify these cases after inspecting (i) a zoom-in image of the integrated spectra within 0.5, 1, and 2 effective radii in the spectral range covering from 5500 to 6100 Å (where the spectral jump stands out), and (ii) an image displaying the average flux in each row of pixels when scanning the FoV from north (top) to south (bottom) after removing the central region where the galaxy is located (where the spatial jump is clearly observed).

Besides, any other spotted issue not covered by these flags is reported as additional comments.

The visual assessment of each datacube is done by two different members of the team. When any of the described flags are marked as BAD by the two members, we consider the datacube unsuitable, and the galaxy is discarded from the DR1 candidate sample. When the two reviewers disagree in any of the flags, the examination of a third member of the QC team breaks the discrepancy and decides whether the galaxy passes the QC visual inspection and enters the set of candidates from which the DR1 sample will be drawn or is discarded from further QC checks.

All the galaxies with GOOD values of the four described flags pass to the second stage of the QC assessment, detailed in the next subsection.

Table 3: Main FITS header keywords

Keyword	data type	Meaning
OBJECT	string	Name of the target galaxy (CAVITY ID)
NAXIS1	integer	Number of pixels along right ascension axis (N_α)
NAXIS2	integer	Number of pixels along declination axis (N_δ)
NAXIS3	integer	Number of pixels along wavelength axis (N_λ)
CRPIX1	float	Reference pixel of the galaxy center in right ascension
CRVAL1	float	Right ascension α (J2000) of the galaxy center in degrees
CDELT1	float	Sampling along right ascension axis in arcsec
CRPIX2	float	Reference pixel of the galaxy center in declination
CRVAL2	float	Declination δ (J2000) of the galaxy center in degrees
CDELT2	float	Sampling along declination axis in arcsec
CRVAL3	float	Wavelength of the first pixel along the wavelength axis in Å
CDELT3	float	Sampling of the wavelength axis in Å
hierarch CAVITYVER	string	Version of the CAVITY reduction pipeline
hierarch PIPE UNITS	string	Units of the flux density
hierarch PIPE GALEXT COR	integer	Applied Galactic Extinction (0 No, 1 Yes)
hierarch PIPE GALEXT AV	float	Galactic V band extinction along line-of-sight
hierarch PIPE REDUDATE	string	Date/time the data were reduced
hierarch GRAT PPAK P_iF_j DATE-OBS	string	Date/time of the observation
hierarch GRAT PPAK P_iF_j UT_START	float	Start of the observation in seconds UT time
hierarch GRAT PPAK P_iF_j UT_END	float	End of the observation in seconds UT time
hierarch GRAT PPAK P_iF_j AIRMASS	float	Airmass during the observation
hierarch GRAT PPAK P_iF_j EXPTIME	float	Exposure time of the observation
hierarch GRAT PPAK P_iF_j HVCOR	float	Line-of-sight difference to heliocentric velocity
hierarch GRAT PPAK P_iF_j EXT_V	float	V band atmospheric extinction in magnitudes
hierarch GRAT PPAK P_iF_j PIPE SPEC RES	float	Homogenized spectral resolution (FWHM) in Å
hierarch GRAT PPAK P_iF_j PIPE NSKY FIB	integer	Number of averaged sky fibers
hierarch GRAT PPAK P_iF_j PIPE SKY MEAN	float	V Johnson mean surface brightness ($\mu_{\rm sky}$) during observation
hierarch GRAT PIPE P_i RELFLUX RATIO MEDIAN	float	Absolute flux scaling of the combined pointing

6.2 Automatic quality checks

This second phase of the QC assessment focuses on inspecting a set of measured parameters, extracted by the pipeline at different stages of the reduction process, that will help to characterise different aspects in the quality of the released datacubes. These parameters are categorised in three groups: those related to the observing conditions (denoted by the OBS prefix), those associated with the instrumental configuration and the performance of the data reduction pipeline (RED), and those evaluating the accuracy and quality of the final calibrated data products (CAL). For the assessment, different thresholds are defined for each parameter mainly following the procedure adopted for the CALIFA DR3. This decision is based on the use of the same instrument by both CALIFA and CAVITY projects, and similar observing strategies and reduction pipelines. Below we describe the explored parameters and the corresponding flags, but we encourage the reader to consult Sánchez et al. (2016) for more detailed information. Since the parameters are measured normally on the individual pointings and/or fibers (depending on the parameter), the thresholds are applied to different measurements obtained when combining the individual estimations. The defined flags will involve in general the mean value (mean), the maximum value (max), and the standard deviation σ (std) of the different estimations of each quantity. For convenience, the used flags and their corresponding thresholds are summarised in Table 4. The values of these flags for the entire DR1 sample are provided in Appendix A (Tables 6-8). This sample comprises 100 galaxies randomly selected from all galaxies with no BAD flags in any of the analysed parameters, ensuring a proper coverage of the color-magnitude diagram (see Fig. 2 and related information provided in Sect. 2).

6.2.1 Quality of the observing conditions

To evaluate the quality of the observing conditions we focus on three main quantities: the airmass, the sky brightness, and the atmospheric extinction. As explained in CALIFA DR3, observing conditions alone do not qualify to consider a datacube unsuitable for science, but they raise a warning when the defined thresholds are exceeded, which indicates the presence of minor issues slightly affecting the quality of the data.

For the airmass we consider three different flags based on the average (OBS_AIRMASS_MEAN), the maximum (OBS_AIRMASS_MAX), and the σ (OBS_AIRMASS_STD) values over all contributing pointings. When planning the observations, we request a minimum elevation of $\sim$ 45° for the targets, which corresponds to an airmass of 1.4 (well below the set warning threshold). This implies that only the STD flag can raise a warning when its threshold is exceeded, which happens in only 3% of the cases.

The sky brightness is measured in the V-band in each pointing using the 30 faintest sky fibers (see Sect. 3). In this case, the mean (OBS_SKYMAG_MEAN) and the σ (OBS_SKYMAG_STD) values over all pointings are considered to define the corresponding flags. The threshold of 19.5 mag arcsec^{-2} for the average is never passed; however, the limit of 0.1 mag arcsec^{-2} for σ is exceeded in 36 galaxies. This is presumably related to the sky gradient that is observed in several galaxies (more details can be found in Sect. 6.3.2), as the average value of the sky fibers is used to define and subtract the sky level (see Sect. 4.4).

Finally, the V-band extinction at the time of each observed pointing is monitored by the observatory and allows us to characterise the transparency of the sky. We provide thresholds for the mean (OBS_EXT_MEAN), maximum (OBS_EXT_MAX) and σ (OBS_EXT_STD) values across the pointings (reported in Table 4). A warning is raised in 13% of the cases, indicating inhomogeneous observing conditions.

6.2.2 Quality of the instrumental/data reduction performance

The performance of the instrument and the data reduction is assessed based on four properties: straylight, spectral dispersion, cross dispersion, and residuals from the subtraction of bright skylines. These quantities are measured on the reduced data (individual fiber spectra), before combining them to create the final datacubes.

The straylight (see Sect. 4.2) is subtracted from the spectra by the reduction pipeline, but high residuals can affect the final quality of the data. To ensure that this is not the case for the CAVITY DR1 galaxies, we set some thresholds on the mean (RED_MEANSTRAYLIGHT_MAX), maximum (RED_MAXSTRAYLIGHT_MAX), and σ (RED_STDSTRAYLIGHT_MAX) values for all the contributing frames (and thus the _MAX suffix). No warnings in the DR1 sample have been raised due to low levels of these quantities.

The spectral dispersion (DISP) and cross dispersion (CDISP) are measured on the arc-lamp calibration frames and the continuum lamp data as the FWHM of the line and the FWHM of the spectral trace, respectively (Sect. 4.2). Thresholds are set on the mean (RED_DISP_MEAN, RED_CDISP_MEAN), maximum (RED_DISP_MAX, RED_CDISP_MAX), and σ (RED_DISP_STD, RED_CDISP_STD) values to avoid significant departures of both parameters from the nominal target specifications. 12 and 7 galaxies of the sample raise a warning because of the requirements on the spectral dispersion and cross dispersion values, respectively.

The accuracy of the sky subtraction is quantified by measuring the flux residuals of the 5577 Å [o i] skyline. We define the flags based on the minimum (RED_RES5577_MIN) and the maximum (RED_RES5577_MAX) values over all pointings of the fiber-average residuals, and the maximum (over all pointings) of the σ of the residuals (RED_STDRES5577_MAX). As explained in Sánchez et al. (2016), average residuals that are too negative or too positive are indicative of systematic bias in the sky subtraction, whereas large σ can reflect localised failures or noisy data. Large enough residuals and/or σ appear in 24% of the cases as to raise a warning, but in any case they are below $\sim$ 0.5 and $\sim$ 1.5 counts, respectively.

6.2.3 Quality of the calibrated data products

The quality of the calibrated data products is determined by checking the global spectrophotometry and the stability of the wavelength calibration across the spectral range.

Table 4: CAVITY DR1 defined quality control flags and thresholds for the automatic checks described in Sect. 6.2.

Parameter involved	QC flags	WARNING condition	BAD condition
Airmass	OBS_AIRMASS_MEAN	> 2	–
	OBS_AIRMASS_MAX	> 2.5	–
	OBS_AIRMASS_STD	> 0.15	–
Sky brightness	OBS_SKYMAG_MEAN	< 19.5 mag arcsec^{-2}	–
	OBS_SKYMAG_STD	> 0.1	–
Atmospheric extinction	OBS_EXT_MEAN	> 0.3 mag	–
	OBS_EXT_MAX	> 0.35	–
	OBS_EXT_STD	> 0.1	–
Straylight	RED_MEANSTRAYLIGHT_MAX	> 50 counts	> 100 counts
	RED_MAXSTRAYLIGHT_MAX	> 75	> 150
	RED_STDSTRAYLIGHT_MAX	> 15	> 30
Spectral dispersion	RED_DISP_MEAN	> 5.5 Å (FWHM)	> 10
	RED_DISP_MAX	> 10	–
	RED_DISP_STD	> 0.5	–
Cross dispersion	RED_CDISP_MEAN	> 3 pixels (FWHM)	–
	RED_CDISP_MAX	> 4	–
	RED_CDISP_STD	> 0.25	–
Sky subtraction	RED_RES5577_MIN	< −0.1 counts	–
	RED_RES5577_MAX	> 0.1	–
	RED_STDRES5577_MAX	> 1.0	–
Spectrophotometric calibration	CAL_SPECPHOT_FLUXRATIO	–	< 0.7
			> 1.6

As described in Section 4.5, each V500 datacube in the CAVITY pipeline was rescaled to match the absolute flux level of the SDSS DR16 broad-band photometry within a 30" diameter aperture. This flux rescaling process for absolute re-calibration is not perfect, as the V500 is slightly affected by vignetting in some fibers (see Section 3). Additionally, artifacts and minor contamination from field stars, which vary due to seeing conditions and PSF differences between the DR16 images and the datacube, also may affect slightly the flux within the aperture in different ways. We constrained the median scale factors (CAL_SPECPHOT_FLUXRATIO) to be between 0.7 and 1.6, removing objects with values outside these limits. Such outliers were primarily due to the partial coverage of the circular aperture in the DR16 images.

The distribution of photometric scale factors for the g and r bands is shown in Fig. 5 (left panel). The median SDSS/-CAVITY g and r band ratios after recalibration are 0.99 ± 0.02 and 1.01 ± 0.02, respectively, indicating that the absolute photometric calibration of the CAVITY data is accurate to within 5%. The right panel of Fig. 5 shows the distribution of the $\Delta(g - r)$ color difference between the SDSS and CAVITY data. We observe a slight systematic offset of $\Delta(g - r) = 0.03$ mag (median) with a scatter of only 0.04 mag. Thus, the spectrophotometric accuracy across most of the covered wavelength range is ∼ 5% for the CAVITY data.

Finally, in order to evaluate the stability of the wavelength calibration over the full spectral range, we look at the position of several sky lines, comparing the nominal wavelength with the measured value in the observed spectra. In particular, we check the median, mean and σ values across the FoV. In all cases, we find that the centroids are fully consistent with zero offsets from the nominal wavelength, and the σ is consistent with pure measurement errors and the absence of any detectable systematic spatial variation.

6.3 Datacubes cosmetics

Some of the CAVITY datacubes present aesthetic features that we describe and characterise in detail in this section. These include i) the presence of some fibers with variable transmission that are reflected in the reconstructed datacubes as bright/faint patches; ii) a slight gradient visible in the lowest surface brightness regions (normally oriented north to south); and iii) a hexagonal pattern in the central parts of the cubes of the smallest galaxies. Figure 6 gives some examples of all cosmetic issues detected in CAVITY DR1 datacubes.

Before proceeding with the description and characterisation of all these features we must assure the reader about the effect of these features on the science to be carried out. During the CAVITY survey, larger galaxies similar to those targeted by CALIFA were also observed by other PMAS programs involving collaboration members and reduced using the CAVITY pipeline described in this work. The reconstructed cubes from these larger galaxies do not exhibit the cosmetic effects described in this section. This indicates that these effects are only noticeable at very low surface brightness in small galaxies, where a significant portion of the FoV is dominated by sky emission. Thus, al-

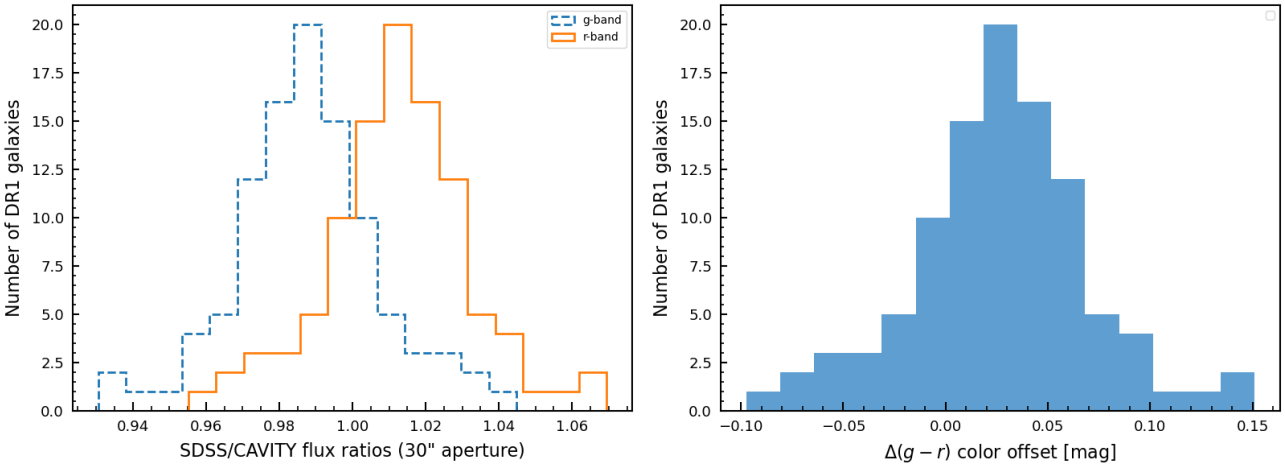

Figure 5: *Left panel:* Distribution of the scale factor for 30" aperture photometry between the SDSS DR16 images and the recalibrated CAVITY data. *Right panel:* Distribution of the color offset between the SDSS DR16 images and the synthetic broadband images derived from cubes in CAVITY.

though striking to the eye, these datacube cosmetics are mainly low-order features that do not affect the main science to be carried out with the datacubes, with the exception of variable transmission fibers that should be masked out (see next).

6.3.1 Variable transmission fibers

As we can observe from Fig. 6, CAVITY DR1 galaxies present some regions with very low or high transmission that are reflected in the reconstructed datacubes as faint or bright patches. Each fiber affected with poor transmission (high and low in the case of CAVITY 16769 or low for the rest), is repeated three times following the dithering scheme used for the observations (see Sect. 3). These three examples highlight the variability of the phenomenon (not every galaxy display the same distribution of ill-behaved fibers).

Figure 7 displays the ubiquity of this issue as the fraction of galaxies for which a given pixel is affected by a fiber with variable transmission. For this, first we inspect, in a galaxy-by-galaxy basis, the presence and location of such fibers. Then we evaluate, for each pixel in the datacube, the number of galaxies out of the 100 CAVITY DR1 galaxies for which such pixel is affected by a transmission variance. Given the variable nature of the problem, this figure should not be interpreted as an effective mask to deal with this phenomenon, but rather as a statistical description of where, and how often, ill-behaved fibers are found in CAVITY datacubes. Plans exist for future reduction pipeline versions and future data releases to minimise this feature via the creation of masterflat fields tailored to specific observing epochs.

6.3.2 Overall gradient

The relatively small sizes of the CAVITY galaxies imply that a good percentage of the PMAS/PPak area is not filled with the galaxy. For a small number of systems, these areas, mainly dominated by sky emission, display a low-amplitude background gradient. This gradient is usually oriented from north to south (top to bottom), although some exceptions are found.

It displays a north-to-south median flux profile of the datacube (vertical scan of all pixels after the galaxy is removed, see Sect. 6.1, point *4*). In general, the low amplitude of these gradients have a negligible effect in the regions where we detect light from the observed galaxies (a factor of $\gtrsim$100 brighter), however, it is of the order of the flux in the sky-dominated areas after sky subtraction. Thus, although visible in a few galaxies, it does not affect the regions of scientific interest.

To fully assess the stability of this feature, we analyse its evolution over time in Fig. 8. For doing that, we measure the median flux in a northern (yellow) and southern (green) background area (avoiding light from the galaxy) from datacubes observed during the different observing campaigns (x-axis). We highlight the stability of the PMAS observations from 2019 to December 2021 (low difference in flux between north and south). It is then when a clear difference between the northern and southern areas starts to appear (signs of a high amplitude gradient), coinciding with the CCD failure reported before. From November 2022 onwards (official resumption of the CAVITY observations) this difference was hugely minimised, with some epochs of slightly larger gradients coinciding with winter observations.

From this analysis we can conclude that most of the gradients (and all of the steep gradients) detected during the assessment of the quality of the CAVITY data are associated with the CCD failure reported before. These datacubes will not be part of any CAVITY data release as they do not fulfill the minimum quality requirements. Also, we would like to highlight that galaxies observed during this period have been re-observed over the past year to ensure they were observed under optimal observing conditions. As a consequence, only a handful of CAVITY DR1 galaxies present low amplitude gradients.

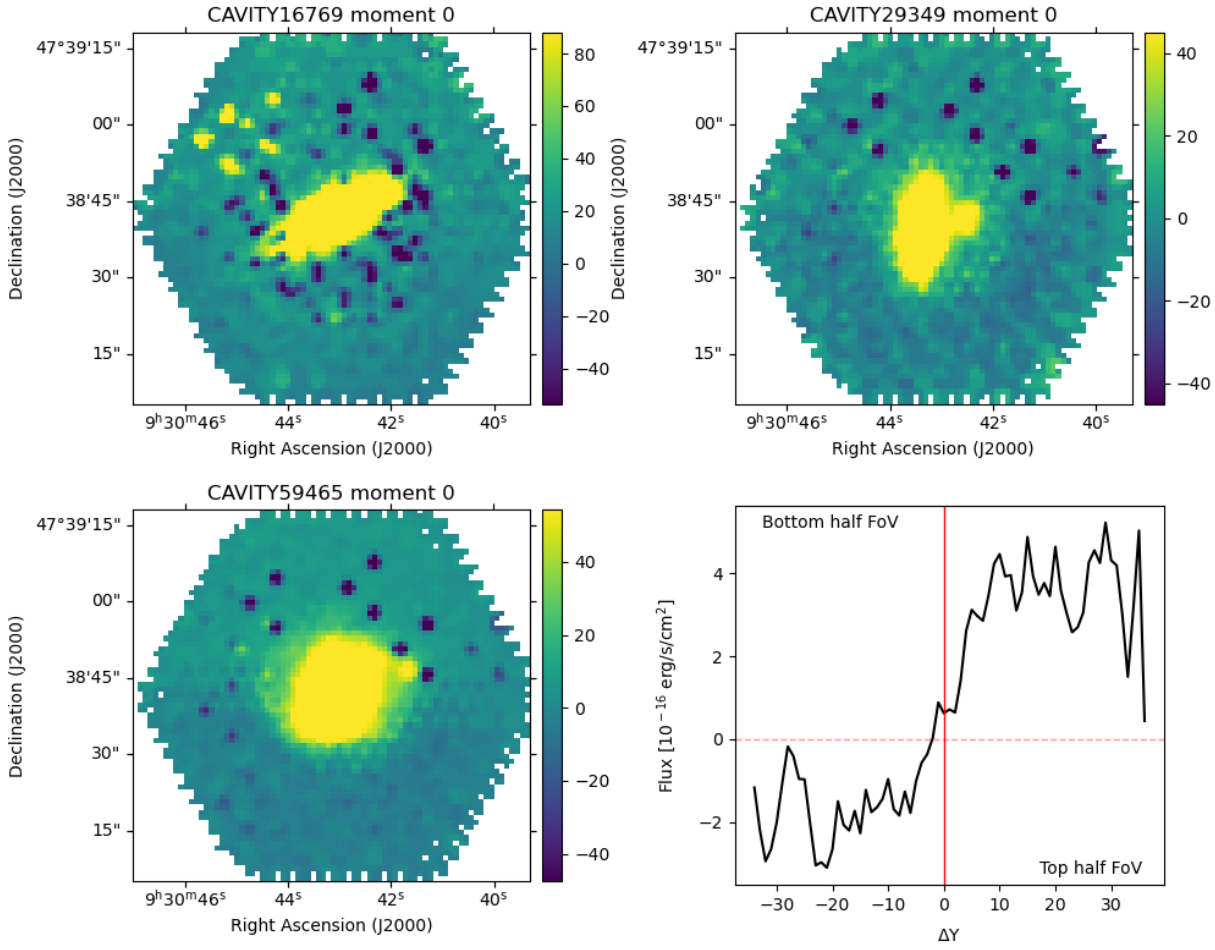

Figure 6: Examples of CAVITY DR1 datacubes affected by cosmetic issues. *Top-left panel:* CAVITY 16769 displays clear examples of bright and faint patches due to variable fiber transmission. *Top-right panel:* CAVITY 29349 shows a patchy, hexagonal pattern in the central regions consequence of the dithering scheme. *Bottom panels* exemplify the gradient in the faintest regions for CAVITY 59465. In the left we show the moment 0 image while the right-hand panel displays bottom-to-top profile of the background flux. All moment 0 images (from 4500 to 7000 Å) are in units of 10^{-16}·erg·cm^{-2}·s^{-1}.

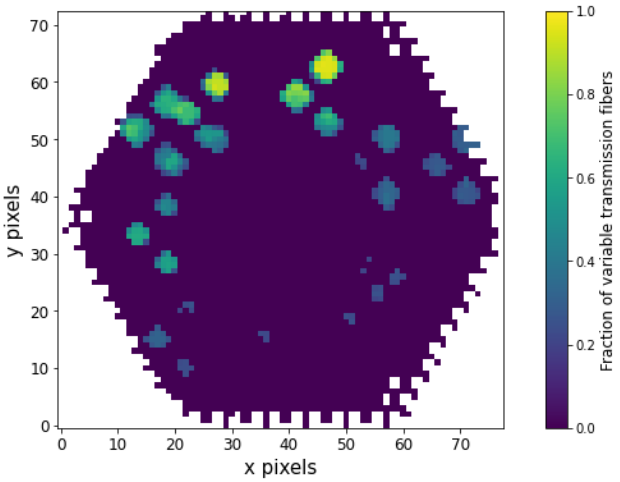

Figure 7: Fraction of datacubes in which a given pixel presents variable transmission along the FoV of the reconstructed datacubes. Only those present in more than 20% of galaxies are shown.

6.3.3 Central low-surface brightness pattern

Another low-order cosmetic artifact observed in a minority of DR1 CAVITY datacubes is a hexagonal-like patchy pattern, exemplified by CAVITY 29349 (see Fig. 6). This pattern results from the dithering scheme used during observations and the slight flux variations between pointings. Although all pointings are flux-matched as described in section 4.5, these subtle differences become more apparent when the surface brightness is low. This artifact is seen only in the central regions of the lowest surface brightness, underscoring its low amplitude. Consequently, it is mainly observed in the smallest CAVITY galaxies.

7 Access to the CAVITY DR1 data

CAVITY DR1 comprises cubes and catalogues containing essential variables and derived physical properties for 100 CAVITY galaxies within 15 voids. All the cubes and catalogues of this public data release are available on the survey's dedicated web page, under the URL: `https://cavity.caha.es`. This web page serves as the primary hub for the survey, where all updates, news, and future data releases related to both CAVITY and CAVITY+ will be provided (for more information on CAVITY+ see Pérez et al. 2024). Constructed using the Open Source Daiquiri framework (Galkin et al. 2020)[5], the data release web page facilitates SQL query storage, blog notifications, and commenting functionality, with optional login features. For convenience, all CAVITY catalogues are provided in FITS and CSV formats. The tables are also compatible with the Virtual Observatory (VO) and are accessible using Table Access Protocol (TAP; Dowler et al. 2011) at the following URL: `https://cavity.caha.es/tap`. In Fig. 9, a screenshot captured in July 2024 showcases the project home page. A brief description of each section of the webpage is provided. In what follows, we explain the properties and information contained in each part of the web page.

7.1 Project & Data tabs

In the "Project" tab, users can find general information about the project and CAVITY team members. The "Data" tab provides news and details about CAVITY data releases. As outlined in Pérez et al. (2024), the CAVITY project extends to different wavelengths by taking advantage of available data from public surveys and conducting dedicated observing campaigns. Value-added catalogues, products of various projects performed by the team members over the CAVITY and CAVITY+ datasets, are published and accessible through this tab.

7.2 Database Tables

The metadata description for DR1 can be found under the "Database Tables" drop-down menu on the "CAVITY DR1" page. This page includes a list of table icons that provide various information and parameters for CAVITY galaxies. The "PPAK FITS files" table contains data cubes of the released CAVITY galaxies. Users can access the footprint and radar imagery of the CAVITY galaxies in the "PPAK-images" table. These images show the spatial coverage of the PPAK hexagonal footprint relative to the galaxy's effective radius. The other tables on this page include additional related parameters for CAVITY galaxies, gathered through cross-correlation with other large multi-wavelength surveys. Detailed descriptions of all tables on the "CAVITY DR1" page are presented in Appendix B.

In each sub-page of tables, users can find comprehensive information about the columns of tables and references for the data sources. Each page starts with a short introduction to the table and its references, followed by a table listing all the physical parameters available, their units, UCDs[6] (Unified Content Descriptors, from the International Virtual Observatory Alliance; Derriere et al. 2004), and brief descriptions. These catalogues can be accessed via SQL query in both CSV and FITS formats (see Section 7.3 for an example).

The "CAVITY Master catalogue" serves as a comprehensive summary of the main physical properties of the galaxies, extracted from all the catalogues accessible on the "CAVITY DR1" page. As such, it stands as an excellent initial resource for users aiming to get familiar with the sample and pursue additional scientific inquiries. Specifics regarding the master table are outlined in Table 5.

7.3 SQL Query page

Users can access all the DR1 catalogues and CAVITY PPak datacubes published through the "Query" page (see Fig. 9) by performing SQL or ADQL (Astronomical Data Query Language, IVOA standard language) queries. Narrowing down the sample selection by filtering or cross matching with multiple catalogues on the database or uploaded data is possible. Cross correlations with Simbad and VizieR services are also the available options on this page. It should

[5] `https://github.com/aipescience/django-daiquiri`

[6] `https://ivoa.net/documents/UCD1+/`

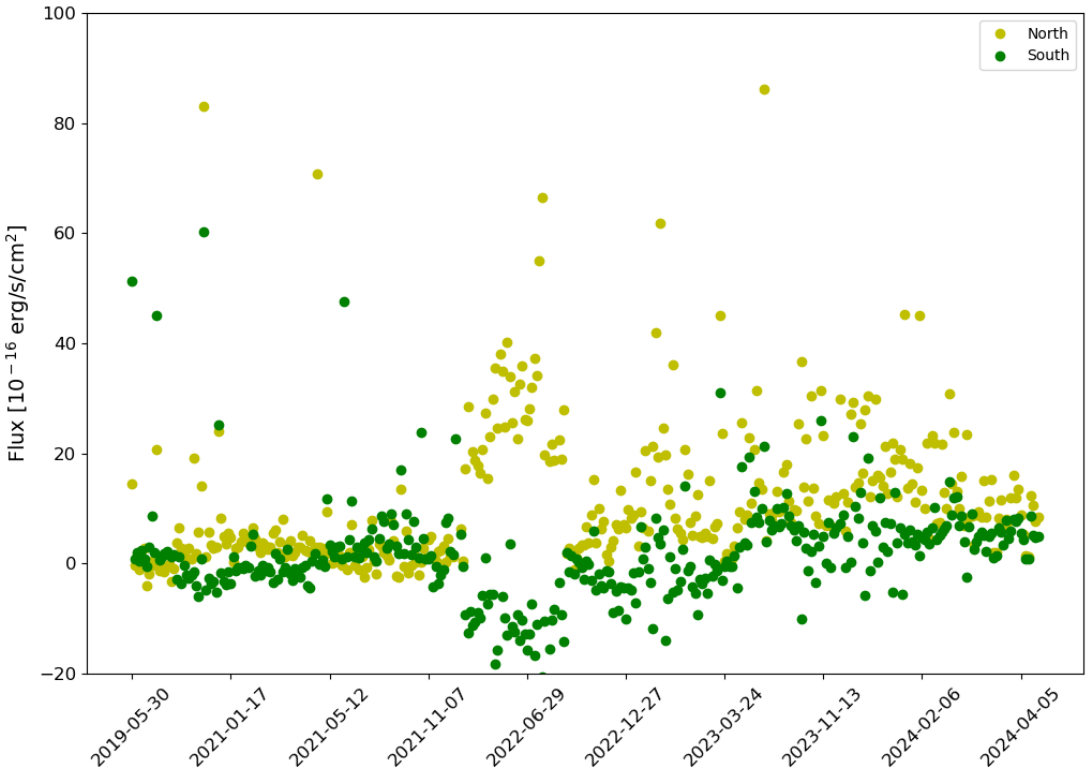

Figure 8: Time evolution of the median flux in background regions located towards the north (yellow) and south (green) of the CAVITY DR1 datacubes (avoiding spaxels with contribution from the observed galaxy).

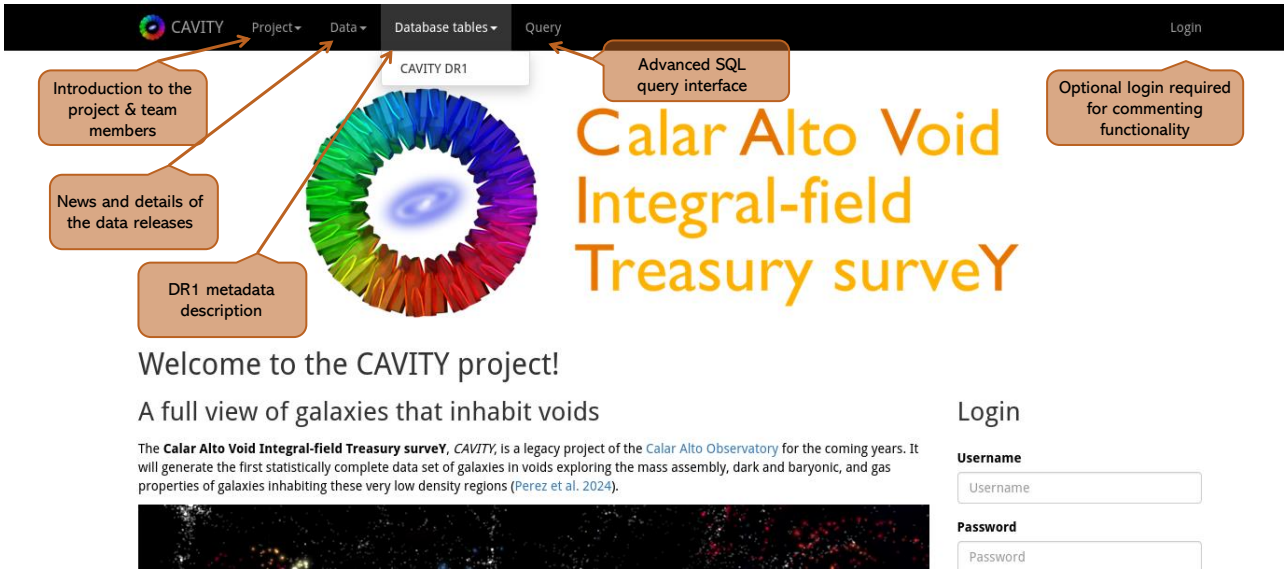

Figure 9: Screenshot of the landing page of CAVITY website captured in July 2024. CAVITY data and catalogues are accessible through various tabs and an advanced SQL query interface.

Table 5: The CAVITY Master Table

Physical parameter	data type	Description
galaxy	integer	Name of the target galaxy (CAVITY ID)[a]
void	integer	The hosting void number[a]
ra	float	Right ascension of the galaxy [degree][b]
dec	float	Declination of the galaxy [degree][b]
redshift	float	Redshift[b]
effr-frac	float	Galaxy's distance to the center of its host void, normalised to the effective radius of the void[a]
morphology	float	Morphology type[c]
d25	float	Major axis [arcsec][a]
incl	float	Inclination [degree][a]
pa	float	Position angle [degree][d]
rabsmag	float	Absolute r-band magnitude [mag][b]
sb	float	Surface brightness [mag.arcsec^{-1}][a]
g-r	float	Color g-r [mag][b]
Av	float	Galactic r band extinction along line-of-sight [mag][a]
M-star	float	Log of stellar mass in solar masses [e]
fluxHI	float	HI line flux density [Jy.km.s^{-1}][f]
vmaxg	float	Rotational velocity [km.s^{-1}][d]
log(SFR)	float	Log of star formation rate [log(M$_\odot$.yr^{-1})][e]
SDSS-ID	float	The SDSS ID[g]

a: Pérez et al. (2024), *b*: Pan et al. (2012), *c*: Domínguez Sánchez et al. (2018), *d*: Makarov et al. (2014)
e: The MPA-JHU survey, *f*: Durbala et al. (2020), *g*: Public Sloan Digital Sky Survey Catalogue-DR8

be noted that keywords and unquoted identifiers are case insensitive in the SQL. Hence, in general, lower-case identifiers are used throughout the database, and 'galaxy' is the main identifier of target galaxies in this survey. An example of SQL request on how to select all the late type galaxies (morphological T-type greater than 0) at a redshift lower than 0.030 and belonging to voids 482 or 487 is outlined below:

```
SELECT galaxy, file FROM dr1.PPAK_FITS
WHERE galaxy IN (
    SELECT galaxy FROM voids_15.Master
    WHERE morphology > 0
    AND redshift <= 0.030
    AND void IN (482, 487)
)
```

8 Summary

In this article, we detail the data reduction, primary characteristics, and data access of the first public data release (DR1) from the Calar Alto Void Integral-field Treasury SurveY (CAVITY). DR1 includes 100 datacubes obtained with the V500 setup of the PMAS/PPak spectrograph, covering a wavelength range of 3745–7500 Å and offering a spectral resolution of 6.0 Å (FWHM). This initial release features 100 void galaxies, which represent one-third of the total anticipated sample. These galaxies span a range of masses, morphological types, and colors similar to those of the full sample. The randomly selected subset forms a statistically representative sample of void galaxies, enabling the constraint of the evolution of galaxies residing in cosmic voids across the color–magnitude diagram. CAVITY DR1 delivers science-grade, quality-checked integral-field spectroscopic data to the research community[7].

[7] cavity.caha.es

We have outlined the main quality parameters analyzed during the validation process, providing users with complete tables to aid in selecting objects for their science cases. The data were reduced using the first official version of the CAVITY pipeline (V1.2), which operates automatically. We plan to enhance this pipeline and consistently offer re-reduced complete datasets to the community in the next data release.

The quality of each datacube underwent cross-checking using a set of defined figures of merit, enabling us to select and distribute data of the highest quality. Parameters such as wavelength calibration accuracy, spectral resolution, sky subtraction, and the stability of observations and reduction processes were among those cross-checked.

The data distributed in DR1 already includes more than 100k independent spectra. While this constitutes a sizable sample, it represents only one-third of the total number of spectra anticipated upon completion of the CAVITY survey. Currently, data has been acquired for over 200 galaxies, for which we will continue to conduct the quality control tests implemented for this data release. It is anticipated that the second and final CAVITY data release will occur after 2026, by which time we aim to have collected high-quality science-grade data for more than 300 galaxies in voids.

Based on observations collected at the Centro Astronómico Hispano en Andalucía (CAHA) at Calar Alto, operated jointly by Junta de Andalucía and Consejo Superior de Investigaciones Científicas (IAA-CSIC), under the CAVITY legacy project. We acknowledge financial support by the research projects AYA2017-84897-P, PID2020-113689GB-I00, and PID2020-114414GB-I00, financed by MCIN/AEI/10.13039/501100011033, the project A-FQM-510-UGR20 financed from FEDER/Junta de Andalucía-Consejería de Transformación Económica, Industria, Conocimiento y Universidades/Proyecto and by

the grants P20-00334 and FQM108, financed by the Junta de Andalucía (Spain). RGB, AC, and RGD acknowledge financial support from the Severo Ochoa grant CEX2021-001131-S funded by MCIN/AEI/ 10.13039/501100011033 and to grant PID2022-141755NB-I00. LSM acknowledges support from Juan de la Cierva fellowship (IJC2019-041527-I). TRL acknowledges support from Juan de la Cierva fellowship (IJC2020-043742-I). SDP acknowledges financial support from Juan de la Cierva Formación fellowship (FJC2021-047523-I) financed by MCIN/AEI/10.13039/501100011033 and by the European Union 'NextGenerationEU'/PRTR, Ministerio de Economía y Competitividad under grants PID2019-107408GB-C44 and PID2022-136598NB-C32, and is grateful to the Natural Sciences and Engineering Research Council of Canada, the Fonds de Recherche du Québec, and the Canada Foundation for Innovation for funding. MAF acknowledges support from the Emergia program (EMERGIA20_38888) from Consejería de Universidad, Investigación e Innovación de la Junta de Andalucía. SBD acknowledges financial support from the grant AST22.4.4, funded by Consejería de Universidad, Investigación e Innovación and Gobierno de España and Unión Europea — NextGenerationEU, also funded by PID2020-113689GB-I00, financed by MCIN/AEI. DE acknowledges support from a Beatriz Galindo senior fellowship (BG20/00224) from the Spanish Ministry of Science and Innovation. GTR acknowledges financial support from the research project PRE2021-098736, funded by MCIN/AEI/10.13039/501100011033 and FSE+. MIR acknowledges financial support by the research projects PID2020-113689GB-I00, and PID2020-114414GB-I00, financed by MCIN/AEI/10.13039/501100011033, the project A-FQM-510-UGR20 financed from FEDER/Junta de Andalucía-Consejería de Transformación Económica, Industria, Conocimiento y Universidades/Proyecto and by the grants P20-00334 and FQM108, financed by the Junta de Andalucía (Spain) and Grant AST22.4.4, funded by Consejería de Universidad, Investigación e Innovación and Gobierno de España and Unión Europea – NextGenerationEU. HMC acknowledges support from the Institut Universitaire de France and from Centre National d'Etudes Spatiales (CNES), France. JFB acknowledges support from the PID2022-140869NB-I00 grant from the Spanish Ministry of Science and Innovation. LG acknowledges financial support from AGAUR, CSIC, MCIN and AEI 10.13039/501100011033 under projects PID2020-115253GA-I00, PIE 20215AT016, CEX2020-001058-M, and 2021-SGR-01270. AFM acknowledges support from RYC2021-031099-I and PID2021-123313NA-I00 of MICIN/AEI/10.13039/501100011033/FEDER,UE,NextGenerationEU/PRT. IMC acknowledges support from ANID programme FONDECYT Postdoctorado 3230653 and ANID, BASAL, FB210003. Funding for this work/research was provided by the European Union (MSCA EDUCADO, GA 101119830). PMSA acknowledges grants PID2019-105602GBI00 and PID2022-136505NB-I00. JR acknowledges financial support from the Spanish Ministry of Science and Innovation through the project PID2022-138896NB-C55. PSB acknowledges support from grant PID2022-138855NB-C31 funded by MCIN/AEI/ 10.13039/501100011033. PVG acknowledges that the project that gave rise to these results received the support of a fellowship from "la Caixa" Foundation (ID 100010434). The fellowship code is B005800. This research made use of astropy, a community-developed core python (http://www.python.org, Van Rossum & Drake 2009) package for Astronomy (Astropy Collaboration et al. 2013, 2018, 2022); IPython (Pérez & Granger 2007); Matplotlib (Hunter 2007); NumPy (van der Walt et al. 2011); SciPy (Virtanen et al. 2020; Gommers et al. 2024); Pandas (McKinney 2010) and TOPCAT (Taylor 2005).

References

Ahumada, R., Allende Prieto, C., Almeida, A., et al. 2020, ApJS, 249, 3

Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, ApJS, 219, 12

Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33

Behnel, S., Bradshaw, R., Citro, C., et al. 2011, Computing in Science and Engineering, 13, 31

Beygu, B., Kreckel, K., van der Hulst, J. M., et al. 2016, MNRAS, 458, 394

Beygu, B., Peletier, R. F., van der Hulst, J. M., et al. 2017, MNRAS, 464, 666

Bond, J. R., Kofman, L., & Pogosyan, D. 1996, Nature, 380, 603

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245

Cautun, M., van de Weygaert, R., Jones, B. J. T., & Frenk, C. S. 2014, MNRAS, 441, 2923

Conrado, A. M., González Delgado, R. M., García-Benito, R., et al. 2024, A&A, 687, A98

Constantin, A., Hoyle, F., & Vogeley, M. S. 2008, ApJ, 673, 715

Courtois, H. M., van de Weygaert, R., Aubert, M., et al. 2023, A&A, 673, A38

Derriere, S., Gray, N., McDowell, J. C., et al. 2004, in Astronomical Society of the Pacific Conference Series, Vol. 314, Astronomical Data Analysis Software and Systems (ADASS) XIII, ed. F. Ochsenbein, M. G. Allen, & D. Egret, 315

Doi, M., Tanaka, M., Fukugita, M., et al. 2010, AJ, 139, 1628

Domínguez-Gómez, J., Pérez, I., Ruiz-Lara, T., et al. 2023a, A&A, 680, A111

Domínguez-Gómez, J., Pérez, I., Ruiz-Lara, T., et al. 2023b, Nature, 619, 269

Domínguez Sánchez, H., Huertas-Company, M., Bernardi, M., Tuccillo, D., & Fischer, J. L. 2018, MNRAS, 476, 3661

Dowler, P., Gaudet, S., & Schade, D. 2011, in Astronomical Society of the Pacific Conference Series, Vol. 442, Astronomical Data Analysis Software and Systems XX, ed. I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, 603

Durbala, A., Finn, R. A., Crone Odekon, M., et al. 2020, AJ, 160, 271

Galkin, A., Klar, J., Matijevic, G., Riebe, K., & Enke, H. 2020, in Astronomical Society of the Pacific Conference Series, Vol. 527, Astronomical Data Analysis Software and Systems XXIX, ed. R. Pizzo, E. R. Deul, J. D. Mol, J. de Plaa, & H. Verkouter, 395

García-Benito, R. 2009, PhD thesis, Autonomous University of Madrid, Spain

García-Benito, R., Díaz, A., Hägele, G. F., et al. 2010, MNRAS, 408, 2234

García-Benito, R., Zibetti, S., Sánchez, S. F., et al. 2015, A&A, 576, A135

Gommers, R., Virtanen, P., Haberland, M., et al. 2024, scipy/scipy: SciPy 1.13.1

Greisen, E. W. & Calabretta, M. R. 2002, A&A, 395, 1061

Horne, K. 1986, PASP, 98, 609

Huertas-Company, M., Aguerri, J. A. L., Bernardi, M., Mei, S., & Sánchez Almeida, J. 2011, A&A, 525, A157

Hunter, J. D. 2007, Computing In Science & Engineering, 9, 90

Husemann, B., Jahnke, K., Sánchez, S. F., et al. 2013, A&A, 549, A87

Husemann, B., Kamann, S., Sandin, C., et al. 2012, A&A, 545, A137

Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 33

Kelz, A., Verheijen, M. A. W., Roth, M. M., et al. 2006, PASP, 118, 129

Kreckel, K., Platen, E., Aragón-Calvo, M. A., et al. 2012, AJ, 144, 16

Kreckel, K., Platen, E., Aragón-Calvo, M. A., et al. 2011, AJ, 141, 4

Libeskind, N. I., van de Weygaert, R., Cautun, M., et al. 2018, MNRAS, 473, 1195

Lintott, C. J., Schawinski, K., Slosar, A., et al. 2008, MNRAS, 389, 1179

Makarov, D., Prugniel, P., Terekhova, N., Courtois, H., & Vauglin, I. 2014, A&A, 570, A13

McKinney, W. 2010, in Proceedings of the 9th Python in Science Conference, ed. S. van der Walt & J. Millman, 51 – 56

Oke, J. B. 1990, AJ, 99, 1621

Pan, D. C., Vogeley, M. S., Hoyle, F., Choi, Y.-Y., & Park, C. 2012, MNRAS, 421, 926

Park, C., Choi, Y.-Y., Vogeley, M. S., et al. 2007, ApJ, 658, 898

Pérez, F. & Granger, B. E. 2007, Computing in Science and Engineering, 9, 21

Pérez, I., Verley, S., Sánchez-Menguiano, L., et al. 2024, arXiv e-prints, arXiv:2405.04217

Platen, E., van de Weygaert, R., & Jones, B. J. T. 2007, MNRAS, 380, 551

Rojas, R. R., Vogeley, M. S., Hoyle, F., & Brinkmann, J. 2004, ApJ, 617, 50

Rojas, R. R., Vogeley, M. S., Hoyle, F., & Brinkmann, J. 2005, ApJ, 624, 571

Roth, M. M., Fechner, T., Wolter, D., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7742, High Energy, Optical, and Infrared Detectors for Astronomy IV, ed. A. D. Holland & D. A. Dorn, 774209

Roth, M. M., Kelz, A., Fechner, T., et al. 2005, PASP, 117, 620

Sánchez, S. F. 2006, Astronomische Nachrichten, 327, 850

Sánchez, S. F., Aceituno, J., Thiele, U., Pérez-Ramírez, D., & Alves, J. 2007, PASP, 119, 1186

Sánchez, S. F., Galbany, L., Walcher, C. J., García-Benito, R., & Barrera-Ballesteros, J. K. 2023, MNRAS, 526, 5555

Sánchez, S. F., García-Benito, R., Zibetti, S., et al. 2016, A&A, 594, A36

Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012, A&A, 538, A8

Schlafly, E. F. & Finkbeiner, D. P. 2011, ApJ, 737, 103

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Sharp, R., Allen, J. T., Fogarty, L. M. R., et al. 2015, MNRAS, 446, 1551

Shepard, D. 1968, in Proceedings of the 1968 23rd ACM National Conference, ACM '68 (New York, NY, USA: Association for Computing Machinery), 517–524

Stanonik, K., Platen, E., Aragón-Calvo, M. A., et al. 2009, ApJ, 696, L6

Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29

Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898

van de Weygaert, R. & Bond, J. R. 2008a, in A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure, ed. M. Plionis, O. López-Cruz, & D. Hughes, Vol. 740, 335

van de Weygaert, R. & Bond, J. R. 2008b, in A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure, ed. M. Plionis, O. López-Cruz, & D. Hughes, Vol. 740, 24

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science and Engineering, 13, 22

Van Rossum, G. & Drake, F. L. 2009, Python 3 Reference Manual (Scotts Valley, CA: CreateSpace)

Verheijen, M. A. W., Bershady, M. A., Andersen, D. R., et al. 2004, Astronomische Nachrichten, 325, 151

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Willett, K. W., Lintott, C. J., Bamford, S. P., et al. 2013, MNRAS, 435, 2835

A Automatic QC parameters

Tables 6-8 show the values of all the QC flags described in Sect. 6.2 for the galaxies in the DR1 sample. These numbers are compared with the defined thresholds for each flag (see Table 4 for a summary) in order to decide whether the galaxy is suitable for science and therefore can be included in the sample or should be discarded. In some cases a warning is raised when some minor issues affecting the quality of the data are found. We refer the reader to Sect. 6.2 for more details.

Table 6: Values of the quality control flags on the observing conditions defined for the CAVITY DR1 sample based on the automatic checks described in Sect. 6.2.

Name	OBS_AIRMASS_MEAN	OBS_AIRMASS_MAX	OBS_AIRMASS_STD	OBS_SKYMAG_MEAN [mag arcsec^{-2}]	OBS_SKYMAG_STD [mag arcsec^{-2}]	OBS_EXT_MEAN [mag]	OBS_EXT_MAX [mag]	OBS_EXT_STD [mag]
CAVITY10595	1.14	1.38	0.16	21.11	0.2	0.18	0.32	0.1
CAVITY11248	1.26	1.35	0.06	20.64	0.06	0.3	0.32	0.03
CAVITY12190	1.04	1.12	0.04	21.39	0.18	0.17	0.33	0.08
CAVITY16768	1.03	1.06	0.02	21.08	0.02	0.13	0.14	0.01
CAVITY16769	1.24	1.37	0.09	20.54	0.12	0.28	0.28	0.0
CAVITY16881	1.07	1.13	0.05	21.32	0.02	0.13	0.13	0.0
CAVITY17302	1.13	1.35	0.12	21.03	0.05	0.14	0.15	0.01
CAVITY17344	1.08	1.15	0.05	21.12	0.05	0.17	0.21	0.03
CAVITY17370	1.06	1.13	0.05	21.07	0.06	0.14	0.14	0.0
CAVITY17616	1.16	1.26	0.07	21.08	0.06	0.15	0.16	0.01
CAVITY18857	1.04	1.08	0.02	21.34	0.07	0.26	0.32	0.04
CAVITY18904	1.22	1.49	0.15	20.96	0.19	0.14	0.17	0.02
CAVITY19279	1.06	1.12	0.04	21.12	0.01	0.12	0.12	0.01
CAVITY20424	1.07	1.16	0.05	20.91	0.06	0.15	0.18	0.02
CAVITY20787	1.03	1.06	0.02	21.08	0.03	0.14	0.14	0.0
CAVITY29349	1.03	1.06	0.01	21.25	0.02	0.14	0.17	0.02
CAVITY29867	1.26	1.41	0.09	20.87	0.15	0.15	0.21	0.04
CAVITY30685	1.04	1.06	0.01	20.96	0.08	0.15	0.2	0.03
CAVITY31671	1.06	1.1	0.03	20.93	0.18	0.14	0.14	0.0
CAVITY32250	1.38	1.75	0.18	21.13	0.1	0.16	0.19	0.02
CAVITY32895	1.15	1.25	0.07	21.03	0.06	0.12	0.13	0.01
CAVITY32896	1.31	1.49	0.12	21.28	0.19	0.16	0.23	0.05
CAVITY33005	1.05	1.15	0.05	21.39	0.06	0.16	0.35	0.07
CAVITY33678	1.07	1.15	0.05	21.6	0.04	0.15	0.16	0.01
CAVITY33695	1.01	1.02	0.01	20.36	0.32	0.22	0.31	0.07
CAVITY34100	1.02	1.04	0.01	21.3	0.1	0.36	0.37	0.01
CAVITY34101	1.15	1.26	0.08	21.17	0.07	0.3	0.31	0.02
CAVITY34170	1.06	1.19	0.06	20.93	0.18	0.13	0.15	0.01
CAVITY34234	1.1	1.29	0.1	21.21	0.16	0.19	0.22	0.01
CAVITY3430	1.11	1.17	0.04	21.32	0.02	0.19	0.24	0.04
CAVITY35487	1.04	1.12	0.04	21.17	0.08	0.13	0.18	0.02

Table 6: Continued.

Name	OBS_AIRMASS_MEAN	OBS_AIRMASS_MAX	OBS_AIRMASS_STD	OBS_SKYMAG_MEAN [mag arcsec^{-2}]	OBS_SKYMAG_STD [mag arcsec^{-2}]	OBS_EXT_MEAN [mag]	OBS_EXT_MAX [mag]	OBS_EXT_STD [mag]
CAVITY36541	1.11	1.2	0.06	21.01	0.19	0.14	0.14	0.0
CAVITY3666	1.08	1.19	0.06	21.15	0.13	0.27	0.32	0.04
CAVITY3670	1.08	1.12	0.03	21.18	0.03	0.14	0.14	0.0
CAVITY37527	1.07	1.23	0.08	20.96	0.22	0.13	0.17	0.02
CAVITY37605	1.06	1.12	0.04	20.98	0.07	0.36	0.4	0.03
CAVITY37820	1.16	1.24	0.05	21.06	0.05	0.15	0.16	0.01
CAVITY37926	1.04	1.1	0.03	21.17	0.09	0.17	0.18	0.01
CAVITY37963	1.09	1.27	0.12	21.06	0.14	0.14	0.14	0.0
CAVITY38659	1.24	1.33	0.06	20.63	0.12	0.29	0.39	0.06
CAVITY40288	1.02	1.06	0.02	21.48	0.09	0.14	0.14	0.0
CAVITY40821	1.29	1.46	0.12	20.9	0.06	0.16	0.18	0.01
CAVITY40822	1.26	1.41	0.1	21.2	0.05	0.16	0.17	0.01
CAVITY40825	1.21	1.36	0.12	20.88	0.09	0.23	0.3	0.05
CAVITY41235	1.3	1.39	0.08	20.81	0.24	0.33	0.44	0.14
CAVITY41398	1.21	1.34	0.09	21.17	0.05	0.16	0.18	0.01
CAVITY41400	1.03	1.06	0.02	21.63	0.02	0.17	0.19	0.02
CAVITY41401	1.18	1.34	0.11	20.76	0.13	0.3	0.39	0.07
CAVITY41448	1.24	1.39	0.11	20.91	0.05	0.26	0.3	0.03
CAVITY41455	1.24	1.42	0.11	20.84	0.2	0.35	0.44	0.1
CAVITY41495	1.16	1.26	0.08	21.18	0.05	0.15	0.16	0.01
CAVITY43294	1.17	1.28	0.08	21.22	0.04	0.14	0.14	0.0
CAVITY43490	1.15	1.25	0.07	21.29	0.13	0.22	0.24	0.02
CAVITY46819	1.1	1.3	0.1	20.9	0.17	0.34	0.38	0.02
CAVITY47120	1.03	1.12	0.04	21.03	0.14	0.14	0.16	0.02
CAVITY48125	1.07	1.19	0.08	21.06	0.09	0.27	0.37	0.05
CAVITY48399	1.12	1.36	0.12	20.8	0.1	0.17	0.28	0.05
CAVITY48997	1.25	1.41	0.11	20.81	0.08	0.12	0.13	0.01
CAVITY49137	1.09	1.18	0.06	20.79	0.19	0.13	0.14	0.01
CAVITY49935	1.16	1.38	0.1	20.29	0.51	0.25	0.31	0.05
CAVITY50031	1.17	1.29	0.08	21.0	0.13	0.2	0.21	0.01
CAVITY50117	1.07	1.16	0.04	20.93	0.06	0.13	0.14	0.01
CAVITY50532	1.1	1.31	0.1	21.04	0.07	0.14	0.14	0.0
CAVITY50943	1.21	1.29	0.07	20.48	0.26	0.25	0.31	0.09

Table 6: Continued.

Name	OBS_AIRMASS_MEAN	OBS_AIRMASS_MAX	OBS_AIRMASS_STD	OBS_SKYMAG_MEAN [mag arcsec^{-2}]	OBS_SKYMAG_STD [mag arcsec^{-2}]	OBS_EXT_MEAN [mag]	OBS_EXT_MAX [mag]	OBS_EXT_STD [mag]
CAVITY51102	1.14	1.29	0.11	20.96	0.05	0.13	0.15	0.01
CAVITY52730	1.22	1.38	0.1	20.75	0.12	0.13	0.13	0.0
CAVITY52954	1.04	1.06	0.01	21.06	0.05	0.14	0.14	0.01
CAVITY53259	1.14	1.24	0.07	21.12	0.06	0.15	0.17	0.02
CAVITY54598	1.25	1.42	0.11	21.34	0.1	0.16	0.19	0.02
CAVITY54706	1.26	1.42	0.11	21.02	0.07	0.25	0.27	0.01
CAVITY55180	1.27	1.42	0.09	20.96	0.05	0.16	0.18	0.01
CAVITY56289	1.22	1.61	0.2	20.94	0.24	0.15	0.15	0.01
CAVITY56627	1.23	1.38	0.1	20.64	0.18	0.17	0.19	0.03
CAVITY57404	1.06	1.15	0.04	20.91	0.04	0.14	0.15	0.0
CAVITY58740	1.17	1.28	0.08	20.94	0.03	0.17	0.17	0.0
CAVITY59465	1.08	1.12	0.03	20.87	0.08	0.14	0.14	0.0
CAVITY59764	1.34	1.56	0.15	20.4	0.09	0.19	0.23	0.03
CAVITY59902	1.15	1.38	0.12	20.89	0.13	0.12	0.17	0.02
CAVITY59906	1.23	1.39	0.1	20.54	0.2	0.14	0.14	0.0
CAVITY59983	1.09	1.16	0.05	20.86	0.04	0.12	0.14	0.01
CAVITY60044	1.08	1.13	0.04	20.93	0.02	0.12	0.15	0.02
CAVITY60187	1.04	1.07	0.01	20.93	0.04	0.13	0.14	0.01
CAVITY60224	1.26	1.52	0.14	20.9	0.11	0.13	0.14	0.01
CAVITY62108	1.2	1.33	0.09	20.96	0.06	0.14	0.14	0.0
CAVITY62262	1.14	1.34	0.1	20.71	0.21	0.14	0.14	0.0
CAVITY62480	1.14	1.26	0.09	21.37	0.09	0.18	0.19	0.01
CAVITY63083	1.13	1.32	0.09	21.1	0.12	0.18	0.22	0.02
CAVITY65288	1.26	1.4	0.1	20.7	0.08	0.43	0.45	0.02
CAVITY65303	1.08	1.09	0.01	21.38	0.03	0.14	0.17	0.02
CAVITY65716	1.29	1.43	0.1	20.59	0.09	0.38	0.4	0.01
CAVITY65887	1.12	1.18	0.04	21.12	0.03	0.17	0.2	0.02
CAVITY66239	1.12	1.18	0.04	20.98	0.02	0.15	0.18	0.02
CAVITY66400	1.16	1.27	0.07	20.52	0.06	0.14	0.14	0.0
CAVITY66803	1.12	1.2	0.05	21.03	0.17	0.29	0.39	0.06
CAVITY67290	1.16	1.26	0.07	20.62	0.04	0.14	0.18	0.03
CAVITY7906	1.06	1.1	0.02	21.14	0.03	0.14	0.17	0.01
CAVITY7926	1.04	1.05	0.01	20.77	0.1	0.25	0.25	0.01

Table 6: Continued.

Name	OBS_AIRMASS_MEAN	OBS_AIRMASS_MAX	OBS_AIRMASS_STD	OBS_SKYMAG_MEAN [mag arcsec^{-2}]	OBS_SKYMAG_STD [mag arcsec^{-2}]	OBS_EXT_MEAN [mag]	OBS_EXT_MAX [mag]	OBS_EXT_STD [mag]
CAVITY8556	1.21	1.33	0.08	21.03	0.08	0.14	0.14	0.0
CAVITY8595	1.13	1.22	0.06	21.07	0.07	0.15	0.18	0.03
CAVITY8646	1.07	1.12	0.03	20.89	0.09	0.37	0.41	0.06

Table 7: Values of the first six quality control flags on the instrumental/-data reduction performance defined for the CAVITY DR1 sample based on the automatic checks described in Sect. 6.2.

Name	RED_MEANSTRAYLIGHT_MAX [counts]	RED_MAXSTRAYLIGHT_MAX [counts]	RED_STDSTRAYLIGHT_MAX [counts]	RED_DISP_MEAN [Å]	RED_DISP_MAX [Å]	RED_DISP_STD [Å]
CAVITY10595	7.53	14.39	2.36	4.57	8.26	0.32
CAVITY11248	8.07	15.02	2.41	4.84	7.84	0.39
CAVITY12190	8.25	13.48	2.04	4.55	10.04	0.31
CAVITY16768	8.01	14.83	2.34	4.77	6.39	0.35
CAVITY16769	10.48	18.36	2.74	4.56	10.28	0.39
CAVITY16881	7.17	13.3	2.14	4.64	6.59	0.3
CAVITY17302	11.84	24.96	3.47	5.08	7.26	0.46
CAVITY17344	9.55	20.58	3.68	4.91	6.8	0.34
CAVITY17370	10.71	18.74	2.51	4.89	6.8	0.41
CAVITY17616	10.59	18.18	2.52	4.88	7.24	0.4
CAVITY18857	8.4	16.42	2.26	4.61	7.26	0.29
CAVITY18904	10.42	20.12	3.16	4.79	7.25	0.37
CAVITY19279	8.47	15.86	2.22	4.9	8.1	0.38
CAVITY20424	9.24	17.42	2.85	4.99	8.09	0.46
CAVITY20787	9.51	19.84	3.34	5.51	8.66	0.63
CAVITY29349	12.08	27.13	4.7	5.13	7.45	0.41
CAVITY29867	8.09	17.33	2.8	5.11	7.81	0.55
CAVITY30685	8.22	15.01	2.51	4.7	7.33	0.34
CAVITY31671	8.53	16.34	2.74	4.76	7.12	0.38
CAVITY32250	11.31	17.29	2.15	4.86	7.79	0.39
CAVITY32895	10.34	17.8	2.72	4.79	6.73	0.39
CAVITY32896	7.17	13.87	2.28	4.53	8.48	0.32
CAVITY33005	8.15	13.65	2.14	4.65	10.74	0.36
CAVITY33678	8.14	16.4	1.7	4.88	7.15	0.4
CAVITY33695	11.31	21.11	3.84	4.94	8.35	0.44
CAVITY34100	8.4	15.66	2.27	4.57	10.24	0.29
CAVITY34101	9.75	14.52	1.83	5.0	7.75	0.43
CAVITY34170	8.86	15.91	2.95	4.74	7.77	0.42
CAVITY34234	9.75	16.12	2.09	4.66	6.73	0.37
CAVITY3430	7.56	13.22	2.12	5.27	7.69	0.53
CAVITY35487	8.99	14.45	2.12	4.58	6.0	0.32

Table 7: Continued.

Name	RED_MEANSTRAYLIGHT_MAX [counts]	RED_MAXSTRAYLIGHT_MAX [counts]	RED_STDSTRAYLIGHT_MAX [counts]	RED_DISP_MEAN [Å]	RED_DISP_MAX [Å]	RED_DISP_STD [Å]
CAVITY36541	8.16	15.45	2.48	4.78	6.37	0.37
CAVITY3666	8.31	15.54	2.19	4.76	7.78	0.33
CAVITY3670	12.05	26.19	4.41	5.11	7.27	0.41
CAVITY37527	15.32	26.4	4.18	4.82	7.25	0.5
CAVITY37605	8.59	13.96	2.05	5.06	8.25	0.44
CAVITY37820	35.72	54.24	10.47	4.87	8.3	0.42
CAVITY37926	10.8	18.41	2.89	4.87	8.13	0.42
CAVITY37963	8.68	13.86	2.43	4.9	8.18	0.43
CAVITY38659	9.14	15.93	2.59	4.78	7.48	0.41
CAVITY40288	12.37	16.88	1.76	5.04	7.06	0.46
CAVITY40821	9.26	17.05	2.57	5.24	8.99	0.51
CAVITY40822	8.8	14.85	1.98	5.08	8.09	0.45
CAVITY40825	11.94	19.18	3.02	4.89	9.27	0.45
CAVITY41235	9.53	16.6	2.49	5.08	9.57	0.52
CAVITY41398	8.52	13.5	2.11	4.95	7.73	0.42
CAVITY41400	8.4	13.73	1.82	4.87	7.28	0.46
CAVITY41401	9.07	17.64	2.53	4.98	8.4	0.43
CAVITY41448	9.59	17.4	2.85	4.97	7.98	0.44
CAVITY41455	10.07	16.9	2.28	5.19	8.9	0.49
CAVITY41495	8.83	13.68	2.13	4.92	7.72	0.4
CAVITY43294	8.12	13.91	1.96	4.65	6.35	0.36
CAVITY43490	16.8	44.94	4.21	4.86	8.09	0.42
CAVITY46819	10.23	18.66	3.02	4.79	6.66	0.36
CAVITY47120	12.16	18.34	2.21	4.73	6.7	0.42
CAVITY48125	8.57	16.91	2.57	4.55	9.03	0.31
CAVITY48399	12.8	22.55	3.71	4.81	7.01	0.41
CAVITY48997	10.72	19.89	3.28	5.07	7.16	0.45
CAVITY49137	11.51	18.22	2.96	4.73	6.62	0.41
CAVITY49935	13.09	22.34	3.83	4.94	8.02	0.42
CAVITY50031	12.22	19.18	2.4	4.81	7.47	0.41
CAVITY50117	12.94	20.27	2.42	4.75	6.91	0.43
CAVITY50532	9.57	17.8	2.58	4.74	8.21	0.39
CAVITY50943	12.45	22.09	3.36	4.87	8.14	0.42

Table 7: Continued.

Name	RED_MEANSTRAYLIGHT_MAX [counts]	RED_MAXSTRAYLIGHT_MAX [counts]	RED_STDSTRAYLIGHT_MAX [counts]	RED_DISP_MEAN [Å]	RED_DISP_MAX [Å]	RED_DISP_STD [Å]
CAVITY51102	9.01	17.73	2.74	5.05	7.13	0.45
CAVITY52730	9.49	19.05	3.05	4.75	7.13	0.43
CAVITY52954	9.18	18.57	2.51	4.64	6.47	0.36
CAVITY53259	9.08	15.66	2.46	4.83	7.76	0.39
CAVITY54598	7.58	14.49	2.11	4.48	6.19	0.25
CAVITY54706	8.13	14.0	2.15	5.12	8.79	0.47
CAVITY55180	9.1	16.39	2.62	5.11	8.87	0.47
CAVITY56289	7.91	18.5	3.36	4.71	6.75	0.43
CAVITY56627	9.23	17.62	2.82	4.71	6.5	0.38
CAVITY57404	9.38	15.9	2.46	4.59	6.09	0.33
CAVITY58740	10.92	23.97	2.83	4.8	7.39	0.39
CAVITY59465	10.96	20.51	2.58	4.7	6.8	0.4
CAVITY59764	12.76	20.62	3.14	4.77	6.98	0.44
CAVITY59902	14.17	23.12	3.43	4.8	6.71	0.33
CAVITY59906	9.63	18.58	3.52	4.73	6.35	0.37
CAVITY59983	9.49	18.27	2.72	5.0	7.18	0.42
CAVITY60044	8.76	16.63	2.54	4.89	7.11	0.37
CAVITY60187	10.79	31.39	2.52	4.68	6.68	0.35
CAVITY60224	10.15	16.2	2.5	4.69	6.6	0.41
CAVITY62108	10.41	15.81	2.46	4.7	6.6	0.41
CAVITY62262	8.7	19.11	3.33	4.71	6.91	0.39
CAVITY62480	8.89	15.6	2.08	4.82	7.1	0.38
CAVITY63083	9.9	15.29	2.4	4.77	7.17	0.37
CAVITY65288	11.79	19.22	2.32	4.79	8.76	0.44
CAVITY65303	9.09	16.14	2.27	4.49	8.32	0.28
CAVITY65716	8.36	15.52	2.54	5.09	8.2	0.45
CAVITY65887	9.04	14.7	1.97	4.77	6.8	0.43
CAVITY66239	13.12	35.11	6.04	5.46	10.07	0.55
CAVITY66400	9.62	17.49	3.0	4.72	6.52	0.38
CAVITY66803	7.86	14.54	2.51	4.57	7.12	0.33
CAVITY67290	8.66	16.52	2.77	4.83	6.91	0.47
CAVITY7906	9.89	14.64	2.09	4.67	6.26	0.33
CAVITY7926	8.06	15.36	2.78	4.51	8.98	0.32

Table 7: Continued.

Name	RED_MEANSTRAYLIGHT_MAX [counts]	RED_MAXSTRAYLIGHT_MAX [counts]	RED_STDSTRAYLIGHT_MAX [counts]	RED_DISP_MEAN [Å]	RED_DISP_MAX [Å]	RED_DISP_STD [Å]
CAVITY8556	9.46	22.72	3.8	5.52	8.41	0.62
CAVITY8595	8.03	14.19	2.12	4.49	8.56	0.28
CAVITY8646	8.97	15.8	2.3	4.77	10.51	0.44

Table 8: Values of the remaining quality control flags on the instrumental/data reduction performance and calibrated data products defined for the CAVITY DR1 sample based on the automatic checks described in Sect. 6.2.

Name	RED_CDISP_MEAN [pixels]	RED_CDISP_MAX [pixels]	RED_CDISP_STD [pixels]	RED_RES5577_MIN [counts]	RED_RES5577_MAX [counts]	RED_STDRES5577_MAX [counts]	CAL_SPECPHOT_FLUXRATIO
CAVITY10595	2.31	3.51	0.17	0.08	0.21	0.77	1.37
CAVITY11248	2.38	3.13	0.15	-0.02	0.02	0.87	0.84
CAVITY12190	2.27	3.39	0.13	0.0	0.16	0.55	1.4
CAVITY16768	2.36	2.86	0.15	0.05	0.09	0.48	0.88
CAVITY16769	2.28	3.3	0.14	0.1	0.15	0.46	1.32
CAVITY16881	2.29	3.15	0.15	0.12	0.16	0.67	1.43
CAVITY17302	2.63	3.39	0.21	0.0	0.04	0.41	0.92
CAVITY17344	2.43	3.32	0.13	-0.0	0.02	0.18	1.2
CAVITY17370	2.46	3.23	0.18	0.02	0.06	0.6	1.02
CAVITY17616	2.48	3.33	0.18	0.01	0.06	0.5	0.97
CAVITY18857	2.34	3.61	0.17	0.06	0.13	0.42	1.39
CAVITY18904	2.3	2.86	0.1	0.0	0.05	0.6	0.9
CAVITY19279	2.45	3.14	0.15	-0.0	0.06	0.47	0.83
CAVITY20424	2.52	3.22	0.21	0.01	0.03	0.42	0.85
CAVITY20787	2.89	4.0	0.29	-0.0	0.02	0.38	0.98
CAVITY29349	2.64	3.53	0.17	-0.01	0.01	0.38	1.04
CAVITY29867	2.55	3.55	0.28	0.02	0.14	0.84	0.93
CAVITY30685	2.26	3.26	0.13	0.01	0.07	0.54	0.95
CAVITY31671	2.29	3.05	0.14	0.02	0.04	0.64	0.92
CAVITY32250	2.4	3.05	0.17	-0.02	0.02	0.43	0.82
CAVITY32895	2.29	2.87	0.1	0.01	0.05	0.66	0.9
CAVITY32896	2.26	3.22	0.14	0.09	0.15	0.46	1.39
CAVITY33005	2.44	3.47	0.2	0.09	0.19	0.68	1.22
CAVITY33678	2.52	3.29	0.22	0.01	0.04	0.28	1.09
CAVITY33695	2.42	3.19	0.16	-0.02	0.02	0.56	0.78
CAVITY34100	2.28	3.51	0.15	0.07	0.21	0.74	1.45
CAVITY34101	2.53	3.37	0.22	-0.03	0.01	0.72	0.94
CAVITY34170	2.21	2.92	0.13	-0.01	0.05	0.78	0.92
CAVITY34234	2.25	3.1	0.14	-0.15	0.01	1.57	0.91
CAVITY3430	2.71	3.55	0.28	0.01	0.03	0.32	1.07

Table 8: Continued.

Name	RED_CDISP_MEAN [pixels]	RED_CDISP_MAX [pixels]	RED_CDISP_STD [pixels]	RED_RES5577_MIN [counts]	RED_RES5577_MAX [counts]	RED_STDRES5577_MAX [counts]	CAL_SPECPHOT_FLUXRATIO
CAVITY35487	2.22	3.01	0.12	-0.01	0.05	0.72	0.91
CAVITY36541	2.31	2.87	0.14	0.01	0.04	0.29	0.95
CAVITY3666	2.29	3.05	0.12	-0.02	0.06	0.46	0.78
CAVITY3670	2.6	3.52	0.17	-0.01	0.02	0.35	1.05
CAVITY37527	2.21	2.83	0.11	-0.05	0.02	0.52	0.9
CAVITY37605	2.56	3.32	0.21	-0.02	0.03	0.75	0.86
CAVITY37820	2.39	3.2	0.16	-0.01	0.03	0.53	0.87
CAVITY37926	2.4	3.18	0.15	-0.01	0.02	0.59	0.85
CAVITY37963	2.42	3.22	0.16	0.01	0.03	0.72	0.84
CAVITY38659	2.32	3.11	0.14	0.01	0.05	1.05	0.79
CAVITY40288	2.59	3.37	0.24	0.01	0.04	0.23	1.24
CAVITY40821	2.64	4.0	0.26	-0.01	0.52	0.84	0.85
CAVITY40822	2.56	3.42	0.22	-0.03	0.02	0.39	0.88
CAVITY40825	2.42	3.15	0.17	-0.06	0.01	0.76	0.89
CAVITY41235	2.51	3.74	0.26	-0.05	0.03	0.72	0.9
CAVITY41398	2.47	3.21	0.19	-0.02	0.03	0.55	0.86
CAVITY41400	2.48	3.42	0.26	-0.0	0.04	0.27	1.08
CAVITY41401	2.5	3.33	0.2	-0.05	0.03	0.84	0.91
CAVITY41448	2.42	3.34	0.18	-0.01	0.03	0.53	0.9
CAVITY41455	2.65	3.59	0.24	-0.15	0.06	0.89	0.89
CAVITY41495	2.41	3.35	0.18	-0.06	0.02	0.53	0.88
CAVITY43294	2.28	2.99	0.14	-0.03	0.05	0.64	0.92
CAVITY43490	2.38	3.22	0.17	0.01	0.03	0.64	0.73
CAVITY46819	2.28	3.09	0.11	0.0	0.11	0.69	0.91
CAVITY47120	2.21	2.89	0.12	-0.01	0.04	0.62	0.94
CAVITY48125	2.27	3.73	0.15	0.08	0.14	0.51	1.51
CAVITY48399	2.37	2.95	0.17	0.01	0.08	0.61	0.8
CAVITY48997	2.56	3.41	0.2	-0.0	0.03	0.89	0.9
CAVITY49137	2.21	2.85	0.12	0.01	0.03	0.43	0.96
CAVITY49935	2.44	3.22	0.19	-0.03	0.05	0.69	0.85
CAVITY50031	2.36	3.14	0.15	-0.04	0.01	0.93	0.85
CAVITY50117	2.22	2.88	0.13	-0.08	0.06	0.91	0.89
CAVITY50532	2.24	2.96	0.12	0.0	0.07	0.57	0.85

Table 8: Continued.

Name	RED_CDISP_MEAN [pixels]	RED_CDISP_MAX [pixels]	RED_CDISP_STD [pixels]	RED_RES5577_MIN [counts]	RED_RES5577_MAX [counts]	RED_STDRES5577_MAX [counts]	CAL_SPECPHOT_FLUXRATIO
CAVITY50943	2.39	3.2	0.15	-0.01	0.02	0.99	0.87
CAVITY51102	2.55	3.4	0.2	0.0	0.05	0.56	0.91
CAVITY52730	2.25	2.75	0.12	0.03	0.08	1.05	0.86
CAVITY52954	2.21	2.88	0.12	-0.08	0.06	0.9	0.86
CAVITY53259	2.38	3.19	0.15	-0.08	0.03	0.65	0.8
CAVITY54598	2.21	3.27	0.13	0.13	0.2	1.02	1.46
CAVITY54706	2.59	3.44	0.22	-0.0	0.04	0.44	0.84
CAVITY55180	2.53	3.46	0.21	-0.02	0.03	0.6	0.86
CAVITY56289	2.22	2.94	0.12	-0.04	0.05	0.42	0.9
CAVITY56627	2.28	2.97	0.13	-0.0	0.03	0.71	0.96
CAVITY57404	2.2	2.84	0.11	-0.02	0.06	0.9	0.89
CAVITY58740	2.35	3.17	0.14	-0.01	0.03	0.54	0.82
CAVITY59465	2.22	2.88	0.12	-0.0	0.06	0.69	0.89
CAVITY59764	2.27	3.12	0.12	-0.05	-0.0	0.58	0.79
CAVITY59902	2.35	3.09	0.15	-0.0	0.04	0.52	0.93
CAVITY59906	2.25	2.86	0.12	0.01	0.05	0.72	1.03
CAVITY59983	2.51	3.25	0.19	0.0	0.06	0.66	0.93
CAVITY60044	2.43	3.1	0.14	0.03	0.07	0.46	0.85
CAVITY60187	2.32	3.05	0.15	0.02	0.06	0.49	0.86
CAVITY60224	2.24	3.07	0.12	-0.01	0.03	0.68	0.85
CAVITY62108	2.24	3.07	0.12	-0.01	0.01	0.37	0.82
CAVITY62262	2.21	2.83	0.1	-0.03	0.08	1.18	1.0
CAVITY62480	2.45	3.12	0.19	0.02	0.06	0.36	1.1
CAVITY63083	2.29	3.28	0.13	-0.0	0.06	0.56	0.87
CAVITY65288	2.29	3.07	0.14	-0.04	0.08	0.63	0.91
CAVITY65303	2.21	3.13	0.13	0.07	0.14	0.7	1.33
CAVITY65716	2.52	3.66	0.21	-0.01	0.05	0.67	0.93
CAVITY65887	2.26	3.03	0.12	-0.03	0.01	0.51	0.87
CAVITY66239	2.82	4.0	0.24	0.02	0.05	0.28	1.04
CAVITY66400	2.28	2.94	0.13	-0.03	0.02	0.57	1.12
CAVITY66803	2.27	3.53	0.15	0.1	0.17	0.52	1.21
CAVITY67290	2.24	2.85	0.14	-0.12	0.03	0.6	0.86
CAVITY7906	2.31	2.9	0.15	0.02	0.08	0.44	0.89

Table 8: Continued.

Name	RED_CDISP_MEAN [pixels]	RED_CDISP_MAX [pixels]	RED_CDISP_STD [pixels]	RED_RES5577_MIN [counts]	RED_RES5577_MAX [counts]	RED_STDRES5577_MAX [counts]	CAL_SPECPHOT_FLUXRATIO
CAVITY7926	2.22	3.18	0.12	0.15	0.35	1.08	1.66
CAVITY8556	2.91	3.97	0.28	0.0	0.03	0.54	0.84
CAVITY8595	2.23	3.28	0.13	0.1	0.2	0.94	1.39
CAVITY8646	2.4	3.07	0.2	0.1	0.16	0.66	1.38

B Database tables

In this section, we review the details of all the catalogs provided through CAVITY DR1. These catalogs are accessible from the CAVITY database, as explained in Section 7, and are available in both FITS and CSV formats. Below, we briefly review the main properties of each:

- **PPAK_cubes:** This table contains the FITS files of observed CAVITY galaxies provided by the PPAK data reduction pipeline.
- **PPAK_images:** This table contains the footprint and radar imagery of the CAVITY galaxies. In all the images, the effective radius of the galaxy is marked in white, and the PPak hexagonal footprint is marked in green.
- **Master:** This table contains the main physical properties of CAVITY galaxies and is a summary of all the catalogues described here. The majority of parameters presented in this table are taken from publicly available catalogues, as outlined in Tab. 5. A few parameters are derived by the team members, for which we will discuss their details here. The "effr-frac" is the effective radius fraction of the location of galaxies within voids, assuming them as spheres with similar volume as the actual void. "d25" represents the galaxy's major axis, which is three times the value of its "petroR90-r." The "petroR90-r" is the radius encompassing 90% of the Petrosian flux in the R-band. These values are taken from Alam et al. (2015). "incl" is galaxy's inclination defined as arccos(b/a) where b and a are minor and major axis, respectively. "A_v" is Galactic r band extinction along line-of-sight and is derived following the steps outlined for table AG in the following.
- **Voids:** This table describes the main physical properties of CAVITY voids. All the properties are directly taken from Pan et al. (2012) public void catalogue. All x,y,z coordinates reported in this table are calculated as distances in Mpc/h, using the following equations: x = rcos(ra)cos(dec); y = rsin(ra)cos(dec); z = rsin(dec). Their assumed cosmology is $\Omega_l = 0.73$ and $\Omega_m = 0.27$.
- **AG:** This table contains the Galactic extinction data for CAVITY galaxies. Optical Galactic extinctions in different SDSS filters are calculated based on the SFD Dust survey (Schlegel et al. 1998) and SDSS filters (Schlafly & Finkbeiner 2011) publicly available gdpyc python package[8].
- **ALFALFA_SDSS:** ALFALFA-SDSS Galaxy Catalogue for CAVITY galaxies.
- **IFU:** This table contains information on IFU data availability from CALIFA[9] and MaNGA[10] surveys for CAVITY galaxies. Columns are set to 1 if the corresponding data is available and 0 otherwise.
- **PAN_MAGLIM:** This table contains properties of CAVITY galaxies directly taken from the publicly available magnitude limited catalogue of Pan et al. (2012).
- **SDSS:** This table contains photometric properties of CAVITY galaxies, directly taken from the 12th data release of the Sloan Digital Sky Survey (Alam et al. 2015).
- **Morpho:** The morphological classification catalogue of CAVITY galaxies, constructed from four public catalogues and studies (Lintott et al. 2008; Willett et al. 2013; Huertas-Company et al. 2011; Domínguez Sánchez et al. 2018).
- **MPA:** This table contains properties of CAVITY galaxies, directly taken from the MPA-JHU Catalogue[11] (Brinchmann et al. 2004; Kauffmann et al. 2003; Tremonti et al. 2004).
- **LEDA:** This table contains properties of CAVITY galaxies directly taken from the Hyperleda Catalogue[12].

[8] https://gdpyc.readthedocs.io/en/latest/
[9] https://califa.caha.es/
[10] https://www.sdss4.org/surveys/manga/
[11] https://live-sdss4org-dr12.pantheonsite.io/spectro/galaxy_mpajhu/
[12] http://leda.univ-lyon1.fr/leda/meandata.html

Table 9: PPAK_cubes

column	description	unit	UCD	datatype
galaxy	CAVITY ID	nan	meta.id;meta.main;	int
name	CAVITY name	nan	meta.id;	char
date	Observation date	nan	time.epoch;	date
version	Pipelide version	nan	nan	char
file	Fits file of the galaxy	nan	meta.ref;meta.file;meta.fits;	char

Table 10: PPAK_images

column	description	unit	UCD	datatype
galaxy	CAVITY ID	nan	meta.id;meta.main;	int
img_footprint	PPak footprint on the SDSS image of the galaxy.	nan	meta.ref;meta.image;	char
img_moment	Moment 0 map (from 450 to 700 nm) of the galaxy.	nan	meta.ref;meta.image;	char
img_spectrum	Spectrum of the center of the galaxy.	nan	meta.ref;meta.image;	char

Table 11: Voids

column	description	unit	UCD	datatype
void	CAVITY Void ID	nan	meta.id;meta.main;	int
ra	Right ascension of the center of the maximal sphere	[degree]	pos.eq.ra;	float
dec	Declination of the center of the maximal sphere	[degree]	pos.eq.dec;	float
x	Equatorial x coordinate of the center of the maximal sphere	[Mpc/h]	pos.cartesian.x;	float
y	Equatorial y coordinate of the center of the maximal sphere	[Mpc/h]	pos.cartesian.y;	float
z	Equatorial z coordinate of the center of the maximal sphere	[Mpc/h]	pos.cartesian.z;	float
r	Radius of the maximal sphere	[Mpc/h]	pos.spherical.r;	float
volx	Equatorial x coordinate of the center of the void region	[Mpc/h]	pos.cartesian.x;	float
voly	Equatorial y coordinate of the center of the void region	[Mpc/h]	pos.cartesian.y;	float
volz	Equatorial z coordinate of the center of the void region	[Mpc/h]	pos.cartesian.z;	float
effr	Radius of the sphere which has the same volume as the void region	[Mpc/h]	pos.spherical.r;	float
numgals	Number of volume limited galaxies that reside in the entire void region	nan	meta.number;	int

Table 12: AG

column	description	unit	UCD	datatype
galaxy	CAVITY ID	nan	meta.id;meta.main;	int
LEDA	Galactic extinction in B-band (ag from Hyperleda catalogue)	[mag]	phys.absorption.gal;em.opt.B;	float
SDSS_u	Galactic extintion using SDSS U-Band filter	[mag]	phys.absorption.gal;em.opt.U;	float
SDSS_g	Galactic extintion using SDSS G-Band filter	[mag]	phys.absorption.gal;em.opt.B;	float
SDSS_r	Galactic extintion using SDSS R-Band filter	[mag]	phys.absorption.gal;em.opt.R;	float
SDSS_i	Galactic extintion using SDSS I-Band filter	[mag]	phys.absorption.gal;em.opt.I;	float
SDSS_z	Galactic extintion using SDSS Z-Band filter	[mag]	phys.absorption.gal;em.opt.I;	float

Table 13: ALFALFA_SDSS

column	description	unit	UCD	datatype
galaxy	CAVITY ID	nan	meta.id;meta.main;	int
AGC	Entry number in the AGC catalogue	nan	meta.id.cross;	int
Vhelio	Heliocentric velocity of the HI profile mid-point	[km·s^{-1}]	spect.dopplerVeloc; em.line.HI;	int
W50	Width at 50% level	[km·s^{-1}]	nan	int
W20	Width at 20% level	[km·s^{-1}]	nan	int
HIflux	HI line flux density	[Jy·km·s^{-1}]	phot.flux.density; em.line.HI; spect.line;	float
sigflux	Uncertainty in HIflux	[Jy·km·s^{-1}]	stat.error;phot.flux.density;	float
Dist	Distance used in MHI, Mstar, SFR	[Mpc]	nan	float
sigDist	Uncertainty of Dist	[Mpc]	nan	float
logMH	HI mass, where distance is the one adopted in	[log(M$_\odot$)]	phys.mass;	float
logMH_err	Uncertainty in logMH	nan	stat.error;phys.mass;	float
Hicode	HI source code. Good S/N or tentative detection (1 or 2, respectively)	nan	nan	int
logMstarTaylor	Stellar mass from SDSS optical photometry using the Taylor method	[log(M$_\odot$)]	phys.mass;	float
logMstarTaylor_err	Uncertainty in logMstarTaylor	nan	stat.error;phys.mass;	float
logMstarMcGaugh	Stellar mass from infrared unWISE photometry using the McGaugh method	[log(M$_\odot$)]	phys.mass;	float
logMstarMcGaugh_err	Uncertainty in logMstarMcGaugh	nan	stat.error;phys.mass;	float
logMstarGSWLC	Stellar mass from GSWLC-2	[log(M$_\odot$)]	phys.mass;	float
logMstarGSWLC_err	Uncertainty in logMstarGSWLC	nan	stat.error;phys.mass;	float
logSFR22	Star Formation Rate (SFR) from unWISE infrared photometry using the flux at the frequency of 22 m	[log(M$_\odot yr^{-1}$)]	phys.SFR;	float
logSFR22_err	Uncertainty in SFR22	nan	stat.error;phys.SFR;	float
logSFRNUVIR	Corrected near-ultraviolet SFR from GALEX NUV photometry SFRNUVcor for galaxies with NUV fluxes available in NASA-Sloan Atlas	[log(M$_\odot yr^{-1}$)]	phys.SFR;	float
logSFRNUVIR_err	Uncertainty in logSFRNUVIR	nan	stat.error;phys.SFR;	float
logSFRGSWLC	SFR from GSWLC-2	[log(M$_\odot yr^{-1}$)]	phys.SFR;	float
logSFRGSWLC_err	Uncertainty in logSFRGSWLC	nan	stat.error;phys.SFR;	float

Table 14: IFU

column	description	unit	UCD	datatype
galaxy	CAVITY ID [3]	nan	meta.id;meta.main;	int
CALIFA	Flag for CALIFA galaxies in CAVITY	nan	meta.code;	int
MaNGA	Flag for MaNGA galaxies in CAVITY	nan	meta.code;	int

Table 15: PAN_MAGLIM

column	description	unit	UCD	datatype
galaxy	CAVITY ID	nan	meta.id;meta.main;	int
ra	Right Ascension of the galaxy	[degree]	pos.eq.ra;	float
dec	Declination of the galaxy	[degree]	pos.eq.dec;	float
redshift	Redshift of the void galaxy	nan	src.redshift;	float
dist	Comoving distance to the void galaxy	[Mpc/h]	pos.distance;	float
rabsmag	r-band absolute magnitude of the void galaxy	[mag]	phys.magAbs;em.opt.R;	float
g-r	g-r band color of the void galaxy	[mag]	phot.color;em.opt.B;em.opt.R;	float
x	Equatorial x coordinate of the void galaxy	[Mpc/h]	pos.cartesian.x;	float
y	Equatorial y coordinate of the void galaxy	[Mpc/h]	pos.cartesian.y;	float
z	Equatorial z coordinate of the void galaxy	[Mpc/h]	pos.cartesian.z;	float
plate	SDSS plate number	nan	meta.number;instr.plate;	int
fiber	SDSS fiber number	nan	meta.number;	int
mjd	Modified Julian date of the observation	[day]	time.epoch;obs;	int

Table 16: SDSS

column	description	unit	UCD	datatype
galaxy	CAVITY ID	nan	meta.id;meta.main;	int
objID	Unique SDSS identifier composed from [skyVersion,rerun,run,camcol,field,obj]	nan	meta.id;meta.corr;	char
petroR50_r	Radius containing 50% of Petrosian flux in the R-band	[arcsec]	phys.size.radius;em.opt.R;	float
petroR50Err_r	Error in petroR50_r	[arcsec]	stat.error; phys.size.radius;em.opt.R;	float
petroR90_r	Radius containing 90% of Petrosian flux in the R-band	[arcsec]	phys.size.radius;em.opt.R;	float
petroR90Err_r	Error in petroR90_r	[arcsec]	stat.error; phys.size.radius;em.opt.R;	float
expAB_r	Exponential fit b/a in the R-band	nan	phys.size.axisRatio;em.opt.R;	float
expABErr_r	Error in expAB_r	nan	stat.error; phys.size.axisRatio;em.opt.R;	float
modelMag_u	Model magnitude in u filter	[mag]	phot.mag;em.opt.U;	float
modelMag_g	Model magnitude in g filter	[mag]	phot.mag;em.opt.B;	float
modelMag_r	Model magnitude in r filter	[mag]	phot.mag;em.opt.R;	float
modelMag_i	Model magnitude in i filter	[mag]	phot.mag;em.opt.I;	float
modelMag_z	Model magnitude in z filter	[mag]	phot.mag;em.opt.I;	float
modelMagErr_u	Error in modelMag_u	[mag]	stat.error; phot.mag;em.opt.U;	float
modelMagErr_g	Error in modelMag_g	[mag]	stat.error; phot.mag;em.opt.B;	float
modelMagErr_r	Error in modelMag_r	[mag]	stat.error; phot.mag;em.opt.R;	float
modelMagErr_i	Error in modelMag_i	[mag]	stat.error; phot.mag;em.opt.I;	float
modelMagErr_z	Error in modelMag_z	[mag]	stat.error; phot.mag;em.opt.I;	float

Table 17: Morpho

column	description	unit	UCD	datatype
galaxy	CAVITY ID	nan	meta.id;meta.main;	int
GZ1Ell	GZ1 catalogue of ELLIPTICAL classified galaxy. Ellipticals are marked as 1, else as 0	nan	meta.code;src.morph.type;	float
GZ1S	GZ1 catalogue of SPIRAL classified galaxy. Spirals are marked as 1, else as 0	nan	meta.code;src.morph.type;	float
GZ1U	GZ1 catalogue of UNCERTAIN classified galaxy. Galaxies with high uncertainty in classification are marked as 1, else as 0	nan	meta.code;src.morph.type;	float
GZ2class	The most common morphology for galaxies among different classifications	nan	meta.code.class;src.morph.type;	char
GZ2formula	GZ2 morphology classification	nan	src.morph.type;	float
pdominant	Weighted fraction of GZ2 votes for “dominant bulge”. Galaxies with dominant bulge have higher values	nan	stat.value;	float
pnobulge	Weighted fraction of GZ2 votes for “no bulge”. Galaxies with no bulge have higher values	nan	stat.value;	float
pobvious	Weighted fraction of GZ2 votes for “obvious bulge”. Galaxies with obvious bulge have higher values	nan	stat.value;	float
probaEarly	Probability of a galaxy being a early-type galaxy (’probaEarly’ = ’probaE’ + ’probaS0’)	nan	stat.probability;	float
probaEll	Probability of a galaxy being elliptical	nan	stat.probability;	float
probaLate	Probability of a galaxy being a late-type galaxy (’probaLate’ = ’probaSab’ + ’probaScd’)	nan	stat.probability;	float
probaS0	Probability of a galaxy being lenticular	nan	stat.probability;	float
probaSab	Probability of a galaxy being spiral Sa or Sb type	nan	stat.probability;	float
probaScd	Probability of a galaxy being spiral Sc or Sd type	nan	stat.probability;	float
TType_DS	T-type, early–type galaxies are marked with values 0 and late–type–galaxies with values ¿ 0	nan	meta.code.class;src.morph.type;	float

Table 18: MPA

column	description	unit	UCD	datatype
galaxy	CAVITY ID	nan	meta.id;meta.main;	int
bestObjID	Unique ID in SDSS database	nan	meta.id;meta.corr;	char
specObjID	Spectral ID in SDSS database	nan	meta.id;meta.corr;	char
sn_median	Median S/N per pixel of the whole spectrum [from galSpecInfo]	nan	stat.snr;stat.median;	float
oii_3726_flux	[OII]3726 Flux from Gaussian fit to continuum subtracted data.	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	phot.flux;	float
oii_3726_flux_err	Uncertainty in oii_3726_flux	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	stat.error;phot.flux;	float
oii_3729_flux	[OII]3729 Flux from Gaussian fit to continuum subtracted data.	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	phot.flux;	float
oii_3729_flux_err	Uncertainty in oii_3729_flux	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	stat.error;phot.flux;	float
h_beta_flux	H Flux from Gaussian fit to continuum subtracted data.	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	phot.flux;	float
h_beta_flux_err	Uncertainty in the h_beta_flux	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	stat.error;phot.flux;	float
oiii_5007_flux	[OIII]5007 Flux from Gaussian fit to continuum subtracted data.	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	phot.flux;	float
oiii_5007_flux_err	Uncertainty in oiii_5007_flux	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	stat.error;phot.flux;	float
h_alpha_cont	Continuum at line center from 200 pixel median smoothing of the emission-line subtracted continuum	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}\cdot\mathrm{Å}^{-1}]$	phot.flux;spect.continuum;	float
h_alpha_cont_err	Uncertainty in h_alpha_cont	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}\cdot\mathrm{Å}^{-1}]$	stat.error;phot.flux;	float
h_alpha_reqw	The equivalent width of the continuum-subtracted emission line computed from straight integration over the bandpasses listed in [galSpecLine] (emission is negative). In this case the continuum comes from a running 200 pixel median and does not properly account for stellar absorption. In general the EQW measurements provide a better measure of the true equivalent width. The purpose of this measurement is to help characterize the stellar absorption effecting the line. (EW_stellar = REQW – EQW)	Å	spect.line.eqWidth;spect.continuum;	float
h_alpha_reqw_err	Uncertainty in h_alpha_reqw	Å	stat.error;spect.line.eqWidth;	float
h_alpha_eqw	The equivalent width of the CB08 continuum-subtracted emission line computed from straight integration over the bandpasses listed in [galSpecLine] (emission is negative). This properly takes into account stellar absorption. Note that it will not be correct in the case of blended lines (i.e., NII & H)	Å	spect.line.eqWidth;spect.continuum;	float
h_alpha_eqw_err	Uncertainty in h_alpha_eqw	Å	stat.error;spect.line.eqWidth;	float
h_alpha_flux	H Flux from Gaussian fit to continuum subtracted data.	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	phot.flux;	float
h_alpha_flux_err	Uncertainty in h_alpha_flux	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	stat.error;phot.flux;	float
nii_6584_flux	[NII]6584 Flux from Gaussian fit to continuum subtracted data.	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	phot.flux;	float
nii_6584_flux_err	Uncertainty in nii_6584_flux	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	stat.error;phot.flux;	float
sii_6717_flux	[SII]6717 Flux from Gaussian fit to continuum subtracted data.	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	phot.flux;	float
sii_6717_flux_err	Uncertainty in sii_6717_flux	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	stat.error;phot.flux;	float
sii_6731_flux	[SII]6731 Flux from Gaussian fit to continuum subtracted data.	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	phot.flux;	float
sii_6731_flux_err	Uncertainty in sii_6731_flux	$[10^{-17}\cdot\mathrm{erg}\cdot\mathrm{cm}^{-2}\cdot\mathrm{s}^{-1}]$	stat.error;phot.flux;	float

Table 18 MPA (Continued)

column	description	unit	UCD	datatype
oh_p16	The 16th percentile of the Oxygen abundance derived using Charlot & Longhetti models. The values are reported as 12 + Log O/H. See Tremonti et al (2004) and Brinchmann et al (2004) for details. [from galSpecExtra]	nan	phys.abund;stat.median;	float
oh_p50	The median estimate of the Oxygen abundance derived using Charlot & Longhetti models. The values are reported as 12 + Log O/H. See Tremonti et al (2004) and Brinchmann et al (2004) for details. [from galSpecExtra]	nan	phys.abund;stat.median;	float
oh_p84	The 84th percentile of the Oxygen abundance derived using Charlot & Longhetti models. The values are reported as 12 + Log O/H. See Tremonti et al (2004) and Brinchmann et al (2004) for details. [from galSpecExtra]	nan	phys.abund;stat.median;	float
lgm_tot_p16	The 16th percentile of the Log total stellar mass PDF using model photometry. [from galSpecExtra]	[log($M_\odot$)]	phys.mass;stat.median;	float
lgm_tot_p50	The median estimate of the Log total stellar mass PDF using model photometry. [from galSpecExtra]	[log($M_\odot$)]	phys.mass;stat.median;	float
lgm_tot_p84	The 84th percentile of the Log total stellar mass PDF using model photometry. [from galSpecExtra]	[log($M_\odot$)]	phys.mass;stat.median;	float
sfr_tot_p16	The 16th percentile of the Log total SFR PDF. This is derived by combining emission line measurements from within the fibre where possible and aperture corrections are done by fitting models ala Gallazzi et al (2005), Salim et al (2007) to the photometry outside the fibre. For those objects where the emission lines within the fibre do not provide an estimate of the SFR, model fits were made to the integrated photometry. [from galSpecExtra]	[log($M_\odot yr^{-1}$)]	phys.SFR;stat.median;	float
sfr_tot_p50	The median estimate of the Log total SFR PDF. This is derived by combining emission line measurements from within the fibre where possible and aperture corrections are done by fitting models ala Gallazzi et al (2005), Salim et al (2007) to the photometry outside the fibre. For those objects where the emission lines within the fibre do not provide an estimate of the SFR, model fits were made to the integrated photometry. [from galSpecExtra]	[log($M_\odot yr^{-1}$)]	phys.SFR;stat.median;	float
sfr_tot_p84	The 84th percentile of the Log total SFR PDF. This is derived by combining emission line measurements from within the fibre where possible and aperture corrections are done by fitting models ala Gallazzi et al (2005), Salim et al (2007) to the photometry outside the fibre. For those objects where the emission lines within the fibre do not provide an estimate of the SFR, model fits were made to the integrated photometry. [from galSpecExtra]	[log($M_\odot yr^{-1}$)]	phys.SFR;stat.median;	float

Table 18 MPA (Continued)

column	description	unit	UCD	datatype
specsfr_tot_p16	The 16th percentile of the Log total SPECSFR PDF. This is derived by combining emission line measurements from within the fibre where possible and aperture corrections are done by fitting models ala Gallazzi et al (2005), Salim et al (2007) to the photometry outside the fibre. For those objects where the emission lines within the fibre do not provide an estimate of the SFR, model fits were made to the integrated photometry. [from galSpecExtra]	[log(yr^{-1})]	phys.SFR;stat.median;	float
specsfr_tot_p50	The median estimate of the Log total SPECSFR PDF. This is derived by combining emission line measurements from within the fibre where possible and aperture corrections are done by fitting models ala Gallazzi et al (2005), Salim et al (2007) to the photometry outside the fibre. For those objects where the emission lines within the fibre do not provide an estimate of the SFR, model fits were made to the integrated photometry. [from galSpecExtra]	[log(yr^{-1})]	phys.SFR;stat.median;	float
specsfr_tot_p84	The 84th percentile of the Log total SPECSFR PDF. This is derived by combining emission line measurements from within the fibre where possible and aperture corrections are done by fitting models ala Gallazzi et al (2005), Salim et al (2007) to the photometry outside the fibre. For those objects where the emission lines within the fibre do not provide an estimate of the SFR, model fits were made to the integrated photometry. [from galSpecExtra]	[log(yr^{-1})]	phys.SFR;stat.median;	float
d4000_n	4000AA break, Balogh et al (1999) definition. [from galSpecIndx]	nan	phot.flux.density;arith.ratio;	float
d4000_n_err	Uncertainty in d4000_n	nan	stat.error;phot.flux.density;arith.ratio;	float

Table 19: LEDA

column	description	unit	UCD	datatype
galaxy	CAVITY ID	nan	meta.id;meta.main;	int
pgc	LEDA internal number	nan	meta.id.cross;	int
name	LEDA search string	nan	meta.id;	char
objname	Principal designation	nan	meta.id;	char
d25	Major axis	[arcsec]	phys.angSize.smajAxis;	float
objtype	Type of object (G=galaxy; S=Star ...)	nan	meta.code;	char
al1950	RA 1950 (hours decimal value)	nan	pos.eq.ra;	float
de1950	DEC 1950 (hours decimal value)	nan	pos.eq.dec;	float
al2000	RA 2000 (hours decimal value)	nan	pos.eq.ra;	float
de2000	DEC 2000 (hours decimal value)	nan	pos.eq.dec;	float

Table 19 LEDA (Continued)

column	description	unit	UCD	datatype
l2	Galactic longitude	[degree]	pos.galactic.lon;	float
b2	Galactic latitude	[degree]	pos.galactic.lat;	float
sgl	Supergalactic longitude	[degree]	pos.supergalactic.lon;	float
sgb	Supergalactic latitude	[degree]	pos.supergalactic.lat;	float
type	Morphological type	nan	meta.code;src.morph.type;	char
bar	Bar (B or blank)	nan	meta.code;src.morph.param;	char
ring	Ring (R or blank)	nan	meta.code;src.morph.param;	char
multiple	Multiple (M or blank)	nan	meta.code;src.morph.param;	char
compactness	Compactness (C=Compact or D=Diffuse or blank)	nan	meta.code;src.morph.param;	char
t	Morphological type code	nan	src.morph.type;	float
e_t	Actual error on t	nan	stat.error;src.morph.type;	float
logd25	log10 of apparent diameter (d25 in 0.1")	[log(arcsec)]	phys.size.diameter;	float
e_logd25	Actual error on logd25	[log(arcsec)]	stat.error;phys.size.diameter;	float
logr25	log10 of axis ratio (major axis/minor axis)	nan	phys.size.axisRatio;	float
e_logr25	Actual error on logr25	nan	stat.error;phys.size.axisRatio;	float
pa	Major axis position angle (North Eastwards)	[degrees]	pos.posAng;	float
brief	Mean effective surface brightness	[mag·arcsec^{-2}]	phot.mag.sb;stat.mean;	float
e_brief	Actual error on brief	[mag·arcsec^{-2}]	stat.error;phot.mag.sb;stat.mean;	float
bt	Total B-band magnitude	[mag]	phot.mag;em.opt.B;	float
e_bt	Actual error on bt	[mag]	stat.error;phot.mag;em.opt.B;	float
it	Total I-band magnitude	[mag]	phot.mag;em.opt.I;	float
e_it	Actual error on it	[mag]	stat.error;phot.mag;em.opt.I;	float
ut	Total U-band magnitude	[mag]	phot.mag;em.opt.U;	float
vmaxg	Apparent maximum rotation velocity of gas	[km·s^{-1}]	phys.veloc.rotat;stat.max;	float
e_vmaxg	Actual error on vmaxg	[km·s^{-1}]	stat.error;phys.veloc.rotat;	float
m21	21-cm line flux in magnitude	[mag]	phys.flux;em.line.HI;	float
e_m21	Actual error on m21	[mag]	stat.error;phot.flux;em.line.HI;	float
mfir	Far infrared magnitude	[mag]	phot.mag;	float
vrad	Heliocentric radial velocity from radio measurements	[km·s^{-1}]	spect.dopplerVeloc.radio;pos.heliocentric;	float
e_vrad	Actual error on vrad	[km·s^{-1}]	stat.error;spect.dopplerVeloc.radio;	float
vopt	Heliocentric radial velocity from optical measurements	[km·s^{-1}]	spect.dopplerVeloc.opt;pos.heliocentric;	float
e_vopt	Actual error on vopt	[km·s^{-1}]	stat.error;spect.dopplerVeloc.opt;	float
ai	Internal extinction in B magnitude	[mag]	phys.absorption;em.opt.B;	float
incl	Inclination	[degree]	src.orbital.inclination;	float
a21	21-cm self absortion	[mag]	phys.absorption;em.line.HI;	float
logdc	log10 of apparent corrected diameter (dc in 0.1')	[log(arcmin)]	phys.size.diameter;	float
btc	Total apparent corrected B-magnitude	[mag]	phot.mag;em.opt.B;	float

Table 19 LEDA (Continued)

column	description	unit	UCD	datatype
bri25	Mean surface brightness within isophote 25	[mag·arcsec^{-2}]	phot.mag.sb;stat.mean;	float
vrot	Maximum velocity rotation corrected for inclination	[km·s^{-1}]	phys.veloc.rotat;stat.max;	float
e_vrot	Actual error on vrot	[km·s^{-1}]	stat.error;phys.veloc.rotat;	float
m21c	Corrected 21-cm line flux in magnitude	[mag]	phys.flux;phot.mag;em.line.HI;	float
hic	21-cm color index = m21c – btc	[mag]	phot.color;em.line.HI;em.opt.B;	float
vlg	Radial velocity (cz) with respect to the Local Group	[km·s^{-1}]	spect.dopplerVeloc;	float
vgsr	Radial velocity (cz) with respect to the GSR	[km·s^{-1}]	spect.dopplerVeloc;	float
vvir	Radial velocity (cz) corrected for Virgocentric infall	[km·s^{-1}]	spect.dopplerVeloc;	float
v3k	Radial velocity (cz) with respect to the CMB radiation	[km·s^{-1}]	spect.dopplerVeloc;	float
modz	Redshift distance modulus	[mag]	phot.mag.distMod;	float
mod0	True distance modulus (from parameters)	[mag]	phot.mag.distMod;	float
mabs	Absolute B-magnitude	[mag]	phys.magAbs;em.opt.B;	float
f_astrom	Precision flag on celestial position	nan	nan	char
e_mod0	Actual error on mod0	[mag]	stat.error;phot.mag.distMod;	float
e_modz	Actual error on modz	[mag]	stat.error;phot.mag.distMod;	float
e_mabs	Actual error on mabs	[mag]	stat.error;phys.magAbs;	float
modbest	True distance modulus, combining mod0 and modz	[mag]	phot.mag.distMod;	float
e_modbest	Actual error on modbest	[mag]	stat.error;phot.mag.distMod;	float
kt	Total K-band magnitude	[mag]	phot.mag;em.IR.K;	float
e_kt	Actual error on kt	[mag]	stat.error;phot.mag;em.IR.K;	float
e_ut	Actual error on ut	[mag]	stat.error;phot.mag;em.opt.U;	float
hl_names(pgc)	List of all names	nan	meta.id;	char
celposb(pgc)	B1950 position (character string)	nan	pos;	char
celposj(pgc)	J2000 position (character string)	nan	pos;	char