

The SAMI Galaxy Survey: the third and final data release

Scott M. Croom^{1,2}★, Matt S. Owers^{3,4}, Nicholas Scott^{1,2}, Henry Poetrodjojo^{1,2}, Brent Groves^{5,6,2}, Jesse van de Sande^{1,2}, Tania M. Barone^{6,1,2}, Luca Cortese^{5,2}, Francesco D'Eugenio⁷, Joss Bland-Hawthorn^{1,2}, Julia Bryant^{1,2}, Sree Oh^{6,2}, Sarah Brough^{8,2}, James Agostino⁹, Sarah Casura¹⁰, Barbara Catinella^{5,2}, Matthew Colless^{6,2}, Gerald Cecil¹¹, Roger L. Davies¹², Michael J. Drinkwater¹³, Simon P. Driver⁵, Ignacio Ferreras^{14,15,16}, Caroline Foster^{1,2}, Amelia Fraser-McKelvie^{5,2}, Jon Lawrence¹⁷, Sarah K. Leslie^{18,2}, Jochen Liske¹⁰, Ángel R. López-Sánchez^{3,4,2}, Nuria P. F. Lorente¹⁷, Rebecca McElroy^{1,2}, Anne M. Medling^{9,2}, Danail Obreschkow^{5,2}, Samuel N. Richards¹⁹, Rob Sharp⁶, Sarah M. Sweet^{12,2}, Dan S. Taranu²⁰, Edward N. Taylor²¹, Edoardo Tescari²², Adam D. Thomas^{6,2}, James Tocknell¹⁷, Sam P. Vaughan^{1,2}

¹Sydney Institute for Astronomy (SIfA), School of Physics, The University of Sydney, NSW, 2006, Australia

²ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)

³Department of Physics and Astronomy, Macquarie University, NSW, 2109, Australia

⁴Astronomy, Astrophysics and Astrophotonics Research Centre, Macquarie University, Sydney, NSW 2109, Australia

⁵ICRAR, The University of Western Australia, Crawley WA 6009, Australia

⁶Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia

⁷Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium

⁸School of Physics, University of New South Wales, NSW 2052, Australia

⁹Ritter Astrophysical Research Center, University of Toledo, Mail Stop 111, Toledo, OH, 43606, United States

¹⁰Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany

¹¹Dept. Physics and Astronomy University of North Carolina Chapel Hill, NC 27599 USA

¹²Astrophysics, Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Rd., Oxford, OX1 3RH, UK.

¹³School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia

¹⁴Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

¹⁵Instituto de Astrofísica de Canarias, Calle Via Lactea s/n, E38205 La Laguna, Tenerife, Spain

¹⁶Departamento de Astrofísica, Universidad de La Laguna (ULL), La Laguna, E-38206 Tenerife, Spain

¹⁷Australian Astronomical Optics - Macquarie University, NSW 2109, Australia

¹⁸Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

¹⁹SOFIA Science Center, USRA, NASA Ames Research Center, Building N232, M/S 232-12, P.O. Box 1, Moffett Field, CA 94035-0001, USA

²⁰Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA

²¹Centre for Astrophysics and Supercomputing, Swinburne University of Technology, John Street, Hawthorn, VIC 3122, Australia

²²Melbourne Data Analytics Platform (MDAP), The University of Melbourne, Parkville, VIC 3010, Australia

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We have entered a new era where integral-field spectroscopic surveys of galaxies are sufficiently large to adequately sample large-scale structure over a cosmologically significant volume. This was the primary design goal of the SAMI Galaxy Survey. Here, in Data Release 3 (DR3), we release data for the full sample of 3068 unique galaxies observed. This includes the SAMI cluster sample of 888 unique galaxies for the first time. For each galaxy, there are two primary spectral cubes covering the blue (370–570 nm) and red (630–740 nm) optical wavelength ranges at spectral resolving power of $R = 1808$ and 4304 respectively. For each primary cube, we also provide three spatially binned spectral cubes and a set of standardized aperture spectra. For each galaxy, we include complete 2D maps from parameterized fitting to the emission-line and absorption-line spectral data. These maps provide information on the gas ionization and kinematics, stellar kinematics and populations, and more. All data are available online through Australian Astronomical Optics (AAO) Data Central.

Key words: galaxies: general – galaxies: kinematics and dynamics – galaxies: star formation – galaxies: stellar content – galaxies: clusters: general – astronomical data bases: surveys.

1 INTRODUCTION

Our understanding of how galaxies form and evolve has taken great strides in the last few decades, but we are far from a complete picture. No two galaxies are the same, as illustrated by many attempts to find close analogues of the Milky Way or M31 (e.g. [Bland-Hawthorn & Gerhard 2016](#); [Boardman et al. 2020](#)). This complexity is apparent in the range of distinct components within galaxies, e.g. truncated or extended dark matter haloes, one or more disks, long bars and rings, short bars/bulges, smooth or structured stellar bulges and haloes, central star clusters and nuclear disks, or a massive black hole, as well as different gas/dust (molecular, atomic, ionized) phases. To further complicate matters, the different components or phases interact in a variety of ways: gas cooling to form stars, stellar or supernovae feedback, feedback from a central super-massive black hole, dynamical mixing from bars, etc. It is a daunting task to form a coherent narrative for all of these processes over cosmic time, and to demonstrate the robustness of this narrative with a consistent set of cosmological simulations. Yet only when this is achieved can we begin to claim a solid understanding of the primary processes involved in a galaxy’s life cycle.

In recent years, questions have been raised about the limitations of finite resolution in cosmological simulations, as well as chaotic-like behaviour — the ‘butterfly effect’ — particularly in relation to the inherent complexity of so many competing processes and whether we will ever be able to track these interactions meaningfully ([Genel et al. 2019](#); [Keller et al. 2019](#); [Keller & Kruijssen 2020](#)). To make robust tests of these simulations requires us to acquire observations with sufficient information on each galaxy. Coupled with this, we need to sample galaxies over a cosmologically representative volume.

We know that mass is a fundamental parameter in controlling galaxy properties. For example, star formation, mean stellar age and morphology are strongly linked to stellar mass (e.g. [Baldry et al. 2006](#)). We also now understand that a galaxy’s life is shaped by its surroundings. Environment is known to modify galaxy morphology (e.g. [Dressler 1980](#)), current star formation (e.g. [Lewis et al. 2002](#)) and star formation history (e.g. [Blanton & Moustakas 2009](#)). In some cases it appears that the trends in mass and environment may be separable (e.g. [Peng et al. 2010](#)).

Other properties of galaxies also have important evolutionary roles, such as angular momentum, binding energy, gas and dark matter fractions. A challenge for us is to separate the trends in the driving parameters from stochastic effects. In the fullness of time, it may even be possible to reveal the drivers behind apparently stochastic evolution, for example, uncovering the detailed merger history of a galaxy. However, for now we often have to average over the stochasticity, and this is one of the requirements that drives us to large samples of galaxies. To complicate matters further, we observe galaxies over a randomized distribution of viewing angles. This also pushes us to larger samples in order to determine internal properties from projected properties.

The requirement for larger samples has driven the large-scale multi-fibre spectroscopic surveys of the last two decades ([York et al. 2000](#); [Colless et al. 2001](#); [Driver et al. 2011](#)), that have characterized the local galaxy distribution very effectively. However, the need to understand the internal structure of galaxies has driven the development of integral-field spectroscopic surveys. This was pioneered by the SAURON project ([Bacon et al. 2001](#)) and followed by ATLAS3D ([Cappellari et al. 2011](#)), both focussed on early-type galaxies. The CALIFA survey was the first to cover a large number (~ 600) of galaxies of all types ([Sánchez et al. 2012](#)). The com-

plex multi-parameter nature of galaxy formation drives us to much larger samples, which motivated the step to multi-object integral field spectrographs. The SAMI Galaxy Survey ([Croom et al. 2012](#); [Bryant et al. 2015](#)) was the first of these large-scale multiplexed projects, followed by MaNGA ([Bundy et al. 2015](#)). Future projects such as Hector ([Bryant et al. 2016](#)) will further extend the reach of integral-field spectroscopic surveys.

The SAMI Galaxy Survey ([Croom et al. 2012](#); [Bryant et al. 2015](#)) aimed to span the plane of mass and environment with a large sample of galaxies, each with spatially-resolved structural and kinematic measurements. The specific science goals for the survey were to answer the questions i) what is the physical role of environment in galaxy evolution? ii) What is the interplay between gas flows and galaxy evolution? iii) How are mass and angular momentum built up in galaxies? As part of this we aimed to compare and contrast our 3D integral field data cubes for each galaxy with synthetic galaxies emerging from cosmological simulations sampled in the same fashion. Our first detailed comparisons reveal that all simulators are able to match a subset of the galaxy parameters, but often at the expense of other parameters ([van de Sande et al. 2019](#)). The inconsistencies are due largely to our limited understanding of the many complex baryonic processes that work together over billions of years.

The SAMI Galaxy Survey observations took place between 2013 and 2018, obtaining data for over 3000 galaxies. These data have been used in a wide variety of scientific analyses, including studies of galactic winds ([Fogarty et al. 2012](#); [Ho et al. 2014, 2016b](#)), the relationship of angular momentum and spin to galaxy properties and environment ([Fogarty et al. 2014](#); [Cortese et al. 2016](#); [van de Sande et al. 2017b](#); [Brough et al. 2017](#); [Foster et al. 2017](#); [Welker et al. 2020](#)), stellar populations ([Scott et al. 2017](#); [Barone et al. 2018](#); [Ferreras et al. 2019](#); [Santucci et al. 2020](#)), star formation and quenching ([Medling et al. 2018](#); [Schaefer et al. 2019](#); [Owers et al. 2019](#); [Cortese et al. 2019](#); [Varidel et al. 2020](#)), gas-phase metallicity, ([Poetrodjojo et al. 2018](#); [Sánchez et al. 2019](#)) kinematic and structural scaling relations ([Cortese et al. 2014](#); [Bloom et al. 2017](#); [Barat et al. 2019](#); [Oh et al. 2020](#)), detailed comparison to simulations ([van de Sande et al. 2019](#); [Khim et al. 2020](#)), and much more. The SAMI Galaxy Survey team have provided regular data releases ([Allen et al. 2015](#); [Green et al. 2018](#); [Scott et al. 2018](#)). In this current paper we present the third and final data release (DR3) of all SAMI observations, together with value-added products such as stellar kinematics, stellar populations and emission line fits. This is the first SAMI data release to contain data from the eight massive galaxy clusters observed as part of the SAMI Galaxy Survey. We also include for the first time environmental metrics for the entire SAMI sample.

In Section 2 we outline the survey input catalogues, the instrument and the observations. In Section 3 we discuss improvements in data reduction and assessments of data quality. Section 4 describes the primary data products from the survey, including cubes (with various binning schemes) and aperture spectra. Catalogues based on photometric data of SAMI targets are discussed in Section 5. The stellar kinematics and stellar population products are discussed in Sections 6 and 7 respectively. Emission line products are presented in Section 8. In Section 9 we discuss environmental metrics within the SAMI Galaxy Survey. Discussion of data access is provided in Section 10 and we summarize the paper in Section 11. Throughout the paper we assume a cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. All magnitudes are in the AB system (Oke & Gunn 1983) and stellar masses and star-formation rates assume a Chabrier IMF ([Chabrier 2003](#)).

Table 1. The coordinates and number of objects in the GAMA regions. For each region we list the right ascension (RA) and declination (Dec.). For both the primary and secondary samples we list the number of observed galaxies N_{obs} , the number of good targets N_{good} (i.e. those not flagged as bad for photometric reasons such as a bright star in the field-of-view) and the number of all targets N_{all} (including objects with bad flags). The listed completeness for the primary sample is $N_{\text{obs}}/N_{\text{good}}$.

Region	RA (J2000) deg	Dec. (J2000) deg	Primary $N_{\text{obs}}/N_{\text{good}}/N_{\text{all}}$	Secondary $N_{\text{obs}}/N_{\text{good}}/N_{\text{all}}$	Primary Completeness
GAMA 09h	129.0 to 141.0	−1 to +3	575/683/806	82/699/820	84.2%
GAMA 12h	174.0 to 186.0	−2 to +2	637/728/805	65/900/982	87.5%
GAMA 15h	211.5 to 223.5	−2 to +2	728/995/1127	13/915/996	73.2%
Total			1940/2406/2738	160/2514/2798	80.6%

2 THE SAMI GALAXY SURVEY

2.1 The Input Catalogues

The input catalogues used for the SAMI Galaxy Survey are drawn from the three equatorial regions of the Galaxy And Mass Assembly (GAMA) Survey (Driver et al. 2011), as described in Bryant et al. (2015), and eight cluster regions described in Owers et al. (2017). In addition, a small number of observations were made of filler targets when not all the integral field units could be allocated to main survey targets. We describe the different input catalogues below.

2.1.1 GAMA regions

The majority of galaxies in the SAMI Galaxy Survey are within regions observed as part of the GAMA survey. This provides the substantial advantage of deep and complete spectroscopic coverage to select targets and define environment. The GAMA regions also contain excellent photometric data across 21 bands from ultraviolet to far-infrared (Driver et al. 2016). Bryant et al. (2015) describe the selection of targets in the GAMA regions in detail, but for completeness we outline the selection here.

The SAMI Galaxy Survey targets in the GAMA regions were selected within three 4×12 degree regions along the equator (declination ≈ 0 degrees), centred at approximate right ascension (RA) of 9, 12 and 15 hours (see Table 1). Galaxies were targeted based on cuts in the redshift-stellar mass plane, with a stellar mass proxy that used i -band magnitude and $g - i$ colour (Taylor et al. 2011; Bryant et al. 2015). Galaxies were selected within a series of four stepped volumes, with higher stellar mass limits at higher redshift (see Fig. 4 of Bryant et al. 2015). The primary sample is limited to redshift $z < 0.095$. Observations aimed to have high and uniform completeness for this primary sample. A secondary sample was also defined that included high mass galaxies to higher redshift ($z < 0.115$) and used fainter stellar mass limits. As the secondary targets were observed at lower priority, these are less complete (see Table 1). Fig. 1 shows the distribution of primary (observed: red, unobserved: blue) and secondary targets (observed: magenta, unobserved: cyan) in redshift and RA. A set of filler galaxies, at lower priority than the secondary targets, was also defined. These are discussed in Section 2.1.3.

Each potential target was visually checked to identify problems, such as bright nearby stars or sources being a sub-component of a larger galaxy. These were flagged within the input catalogue and their priority set so that they were not observed. The column named BAD_CLASS in the input catalogue (named InputCatGAMADR3) contains flags for different types of problem sources. A full description of the flags is given by Bryant et al. (2015). A small subset of the sources required adjustment of the coordinates for targeting

(0.6 percent of the input catalogue). This was typically in order to place an asymmetric galaxy or a close pair fully into the integral field unit (IFU) and these objects have BAD_CLASS=5. The input catalogue that is publicly available as part of DR3 contains both object coordinates and IFU pointing coordinates. 88.8 percent of the input catalogue remained as a target after flagging for problems. These good objects to be observed had either BAD_CLASS=0 or 5. A further case of BAD_CLASS=8 is also included in observations, but only the cluster regions contain galaxies with these values. BAD_CLASS=8 indicates a galaxy that is the smaller component of a close pair, where the more massive galaxy is outside of the field of view of the IFU. The full GAMA region input sample is contained within the InputCatGAMADR3 catalogue.

A calibration star was observed in one IFU during every galaxy observation. This star allowed improved flux calibration and a good estimate of the spatial point spread function (PSF). These calibration stars were selected to be F-stars, based on their SDSS photometric properties. Bryant et al. (2015) describes the selection of calibration stars in detail. The catalogue of calibration stars is released as part of DR3 and called FstarCatGAMA.

2.1.2 Cluster regions

The cluster targets were drawn from eight regions centred at the positions listed in Table 2. Photometry was based either on SDSS DR9 (Ahn et al. 2012) or the VLT Survey Telescope ATLAS Survey (Shanks et al. 2015), with details of the selection described by Owers et al. (2017). Targets were selected to be within a redshift range defined by the relative velocity of the galaxy with respect to the cluster redshift, z_{clus} , such that $|v_{\text{pec}}|/\sigma_{200} < 3.5$. v_{pec} is the peculiar velocity of the galaxy with respect to the cluster redshift and σ_{200} is the velocity dispersion of the cluster of interest measured within the over-density radius R_{200} [see Section 9.2 and Owers et al. (2017) for further details]. We note that this cut in redshift is less conservative than the criterion used to allocate cluster membership in Owers et al. (2017), and so the input catalogue contains galaxies that are close to the cluster in redshift space, but may not be bona-fide members.

In addition to meeting the aforementioned criteria, targets were further characterised into primary and secondary targets. Primary targets are defined as those that meet the following:

- Cluster-centric distance $R < R_{200}$,
- for $0.045 < z_{\text{clus}} < 0.06$, stellar mass $M^* > 10^{10} M_{\odot}$,
- for $z_{\text{clus}} < 0.045$, stellar mass $M^* > 10^{9.5} M_{\odot}$.

Similar to the GAMA regions, lower priority secondary targets were also included, and were selected based on the following criteria:

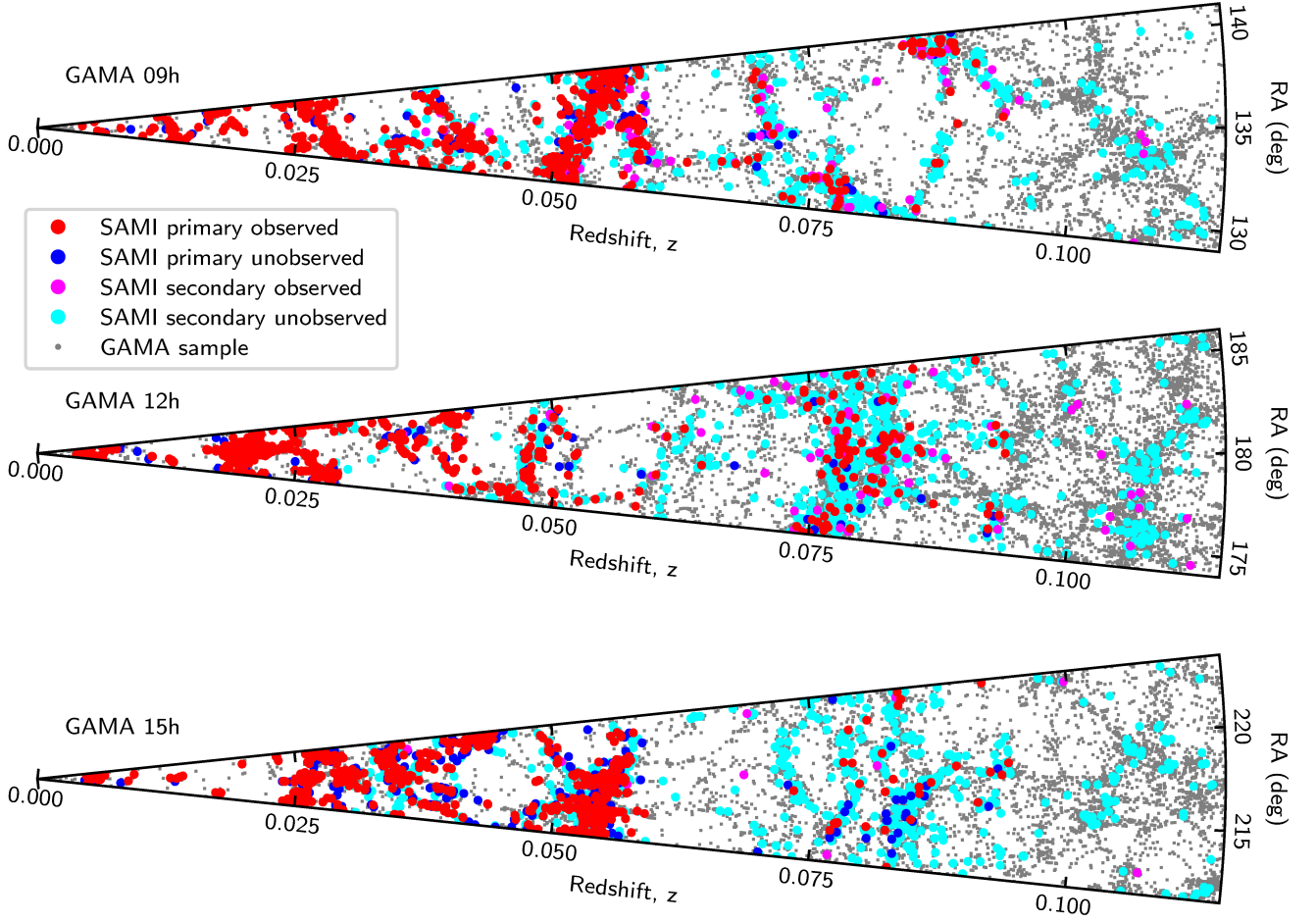


Figure 1. The distribution of SAMI galaxies within the three GAMA regions. We show the original GAMA redshift survey (small gray points), observed (red) and unobserved (blue) primary SAMI galaxies, observed (magenta) and unobserved (cyan) secondary SAMI galaxies.

Table 2. Properties of the eight clusters targeted during the SAMI Galaxy Survey. For each cluster, we list the name, RA and Dec. of the cluster centre, the cluster redshift, velocity dispersion, virial radius estimate, and virial mass as determined in [Owers et al. \(2017\)](#). For the Primary and Secondary targets, we list the number of galaxies observed during the SAMI Galaxy Survey (N_{obs}), the number of galaxies in the input catalogue (N_{all}) and the number of good objects in the input catalogue (N_{good}). The completeness of the Primary sample is determined as $N_{\text{obs}}/N_{\text{good}}$.

Region	RA (J2000) deg	Dec. (J2000) deg	z_{clus}	σ_{200} (km/s)	R_{200} (Mpc)	M_{200} ($10^{14} M_{\odot}$)	Primary $N_{\text{obs}}/N_{\text{good}}/N_{\text{all}}$	Secondary $N_{\text{obs}}/N_{\text{good}}/N_{\text{all}}$	Primary Comp.
APMCC 917	355.397880	-29.236351	0.0509	492	1.19	2.1	25/28/29	8/17/18	89%
Abell 168	18.815777	0.213486	0.0449	546	1.32	3.0	95/126/130	1/40/43	75%
Abell 168 (N)	18.739974	0.430807							
Abell 4038	356.937810	-28.140661	0.0293	597	1.46	2.9	100/104/111	19/85/87	96%
EDCC 442	6.380680	-33.046570	0.0498	583	1.41	3.6	41/50/50	6/47/47	82%
Abell 3880	336.977050	-30.575371	0.0578	660	1.59	4.6	50/51/56	38/57/60	98%
Abell 2399	329.372605	-7.795692	0.0580	690	1.66	6.1	91/92/94	37/48/49	99%
Abell 119	14.067150	-1.255370	0.0442	840	2.04	9.7	202/255/260	0/152/157	79%
Abell 85	10.460211	-9.303184	0.0549	1002	2.42	17.0	152/167/171	23/70/71	91%
Total	–	–	–	–	–	–	756/873/901	132/516/532	87%

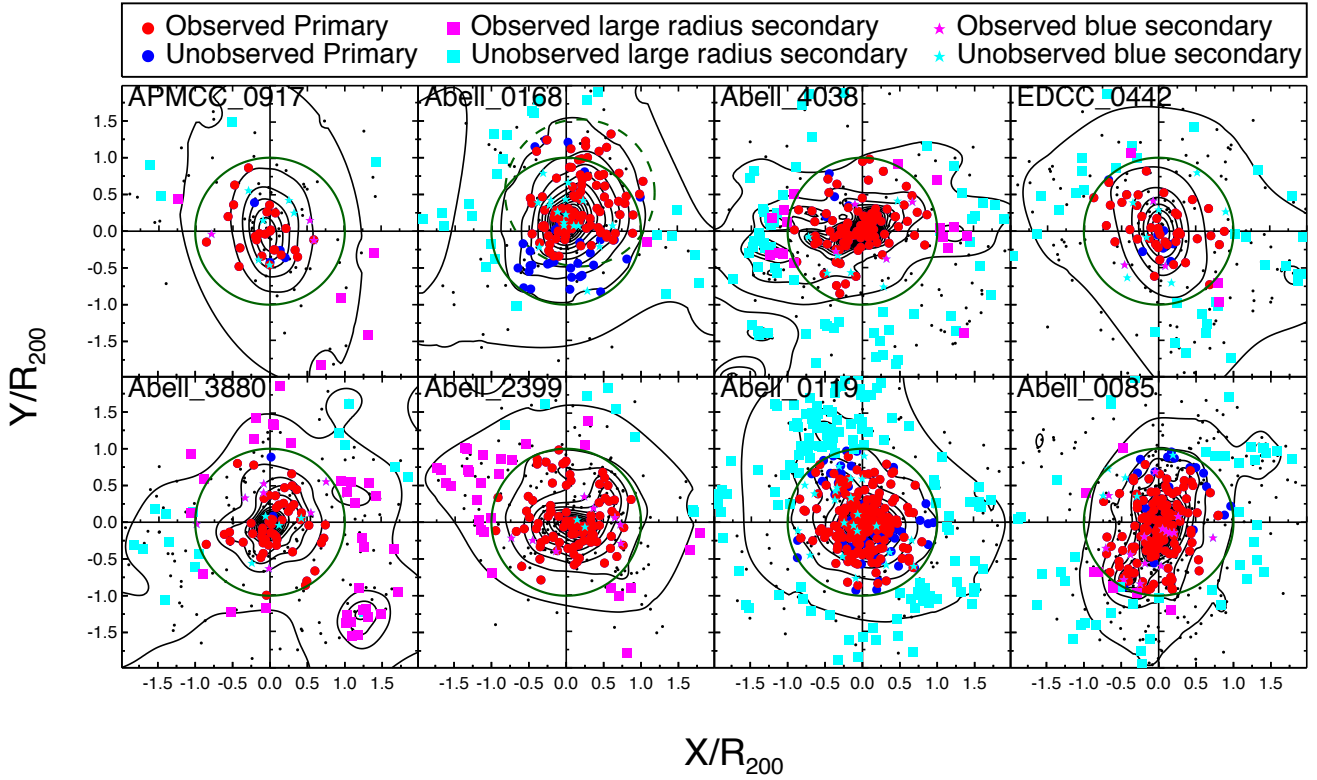


Figure 2. The distribution of SAMI targets in the cluster regions. The small gray points show cluster members selected from the SAMI Cluster Redshift Survey, which are used to define the galaxy surface density isopleths shown as black contours. Large circles show observed (red) and unobserved (blue) primary targets. The magenta (cyan) points show observed (unobserved) secondary targets. These secondary targets are indicated by either squares (for galaxies at large radius) or stars (for blue colour-selected galaxies). The green circles show R_{200} . For Abell 168, the dashed green circle shows the off-centre region used to include targets within R_{200} of the northern BCG, which was used to define the cluster centre during the early parts of the survey.

- Blue secondary targets selected as those galaxies with rest-frame $(g-i)_{\text{corr}} < 0.9$, $R < R_{200}$, and stellar mass 0.5 dex smaller than the stellar mass limit for the primary targets in the cluster of interest.
- large-radius secondary targets were selected to have $R_{200} < R < 2R_{200}$, and with stellar mass above the limit used for the selection of primary targets.

The total and observed number of primary and secondary targets for each cluster is listed in Table 2. In Figure 2, the distribution of the primary and secondary targets is shown for each cluster, with positions measured relative to the centres defined in Table 2 and normalised by R_{200} (green circle). The black contours show galaxy surface density isopleths generated using the cluster members defined in the SAMI Cluster Redshift Survey (black points; Owers et al. 2017). The primary targets are shown as filled circles (observed in red, unobserved in blue), the colour-selected blue secondary galaxies as filled stars (observed in magenta, unobserved in cyan), and the large-radius secondaries as filled squares (observed in magenta, unobserved in cyan).

We note that there are two centres listed for Abell 168 in Table 2 because the ongoing merger in Abell 168 means that there are two sub-clusters. Both the northern and southern sub-clusters contain bright cluster galaxies (Hallman & Markevitch 2004; Fogarty et al. 2014), which leads to some ambiguity in defining the cluster centre. Initially, the cluster centre was defined at the position of the more massive bright cluster galaxy associated with the northern substructure (second listing in Table 2; dashed green circle

Table 3. Table of different filler targets defined by their FILLFLAG parameter that defines which filler sample they are from (see Section 2.1.3 for details). The number of filler targets N_{all} and the number observed N_{obs} are listed.

FILLFLAG	N_{all}	N_{obs}
20	22	1
30	1800	36
40	141	1
50	996	13
90	21	21

in Figure 2). Later, the centre was redefined to the southern bright cluster galaxy due to its proximity to the peak in the galaxy surface density (Owers et al. 2017). For consistency, we define the clustercentric distances of the targets in Abell 168 with respect to the southernmost coordinates listed in Table 2. However, the galaxies initially allocated as primary targets using the northernmost coordinates maintain their status as primary targets, being within R_{200} of the northernmost substructure.

Full information for the cluster region input sample is listed in the InputCatClustersDR3 catalogue. Calibration stars in the cluster regions are listed in the FstarCatClusters catalogue.

2.1.3 Filler targets

In some fields there were not sufficient primary and secondary targets to fill all the IFUs. This was particularly the case for fields observed towards the end of the survey. In some cases spare IFUs were used to target primary and secondary targets that had already been observed. This was either because of low data quality in previous observations, or in order to obtain repeat observations for assessment of data quality. However, extra filler targets were also included for particular science cases in three categories. The filler targets are listed in the catalogue InputCatFiller. The FILLFLAG column in that catalogue identifies the particular class of objects used as fillers. These are broken down into the following categories:

- Galaxies with 21cm H_I detections from the Arecibo Legacy Fast ALFA (ALFALFA) Survey (Giovanelli et al. 2005; Haynes et al. 2018), but that are lower mass than the SAMI selection limits (FILLFLAG=20). We include both high signal-to-noise ratio (S/N) and marginal H_I detections (i.e., ALFALFA detcode=1 and 2).
- Close pairs of galaxies identified from the GAMA Survey (Robotham et al. 2014) that fall outside of the SAMI selection limits. These are in two classes: either both galaxies in the pair are outside of the selection boundaries (FILLFLAG=30), or one of the pair is within the main SAMI sample (FILLFLAG=40).
- Typical star forming disk galaxies at $0.12 < z < 0.15$ (i.e., slightly beyond the SAMI redshift limit), to explore the potential for using velocity information as a means to precision weak lensing experiments (de Burgh-Day et al. 2015; Gurri et al. 2020) (FILLFLAG=50).

The input catalogue for these filler objects (InputCatFiller) forms part of DR3, but is simplified compared to the main SAMI targets, containing only positions, redshifts and FILLFLAG. A small number of cluster galaxies (21) observed early in the survey, but that did not meet the final selection limits are also contained within the filler catalogue (with FILLFLAG=90).

2.2 The SAMI Instrument

The SAMI instrument is a multi-object integral field spectrograph comprising of 13 optical fibre integral field units feeding the AAOmega spectrograph (Croom et al. 2012; Bryant et al. 2015). It is installed on the 1 degree diameter prime focus of the Anglo-Australian Telescope (AAT) in NSW, Australia. The IFUs are *hexabundles* - optical fibre imaging bundles with >75% of the light collected in fibre cores (Bland-Hawthorn et al. 2011; Bryant et al. 2011). These unique bundles are the product of a fusing technique that allows tight packing but without any extra loss of light through focal ratio degradation (Bryant et al. 2014). The hexabundles each contain 61 fibres with core diameter of 105 μm or 1.6 arcsec, subtending 15 arcsec diameter across the hexabundle.

Each field has galaxy and star positions drilled into a plug plate. Typically two different fields are drilled into the same physical plate. The 13 hexabundles and 26 sky fibres, along with 3 guide bundles are plugged by hand. SAMI makes use of the AAOmega spectrograph which is a flexible dual-arm workhorse spectrograph at the AAT (Sharp et al. 2006). For the SAMI survey we used the 580V and 1000R gratings, delivering a wavelength range of 3750–5750 Å and 6300–7400 Å for the blue and red arms respectively. The spectral resolutions are $R = 1808$ and $R = 4304$ for the blue and red arms, equivalent to an effective velocity dispersion of σ of 70.4 and 29.6 km s^{-1} respectively (van de Sande et al. 2017b; Scott et al. 2018).

2.3 Observations

The SAMI Galaxy Survey observations took place over 250 nights from March 2013 to May 2018. Each field, containing 12 galaxy targets and one calibration star, would typically be observed for 7 exposures, each of 1800s. Between exposures the field centre would be offset in a hexagonal pattern to provide uniform coverage of the target in the bundle, allowing for the small gaps between fibres (Sharp et al. 2015). At least one arc frame and one dome flat field frame were taken for each field. Where possible twilight flats were also taken to aid various aspects of calibration. Primary flux standards were observed at the start and end of the night when conditions were photometric.

During observations data quality was checked both in terms of the spatial point spread function (PSF) and transmission. Any exposure where the PSF had FWHM > 3 arcsec, or the transmission was less than 70 percent of the nominal system transmission was flagged for re-observation. Due to scheduling constraints not all flagged data could be re-observed. The threshold in data quality for generating cubes in data reduction was somewhat more relaxed than the above observational constraints (to allow all useful data to be included). Exposures were added to cubes if they had: i) exposure time > 600 sec; ii) a relative transmission of > 33 percent; iii) seeing FWHM of < 4 arcsec. Only targets that had at least 6 frames that passed these quality controls were made into reduced cubes.

Repeat observations of SAMI galaxies fall into two classes. The first is where galaxies are observed within the same plate configuration, but on different observing runs (typically separated by a month or more in time). In this case the individual exposures are combined into a single cube across all the observations. However, we also make a cube from the individual observing runs if there are sufficient exposures that meet our quality control limits. In this case there will be different cubes that share some of the same individual exposures. The main reason for such combinations was insufficient high quality exposures within a single observing run.

The second class of repeat observations was where the same galaxy was observed in two different plates (and therefore usually different hexabundles). This could be during the same or different observing runs. For this class of repeat the data were not combined across the different plates, and hence the galaxy will have two completely independent cubes. There are 215 galaxies with multiple cubes from the same plate, that share some of their data. There are 70 galaxies that have repeated observations from different plates that are fully independent. The best cube, based on seeing FWHM and S/N ratio is identified with the ISBEST flag within the CubeObs catalogue that describes observations of all sources (see Section 4.3 below).

2.4 Survey completeness

Estimates of survey completeness are based on the number of galaxies for which we could successfully construct data cubes (N_{obs}) compared to the number of potential targets that have good quality flags (N_{good} , that is, BAD_CLASS = 0, 5 or 8). The number of unique galaxies successfully observed in each of the GAMA regions along with the completeness is listed in Table 1. In total 1940 unique primary galaxies from the GAMA regions were observed, with a completeness of 80.6 percent. There is some variation in completeness between GAMA regions with the 12h field being 87.5 percent complete, while the 15h field is 73.2 percent complete. This variation is largely driven by the larger number of targets in the 15h region, due to denser large-scale structure (see Fig. 1). As would be

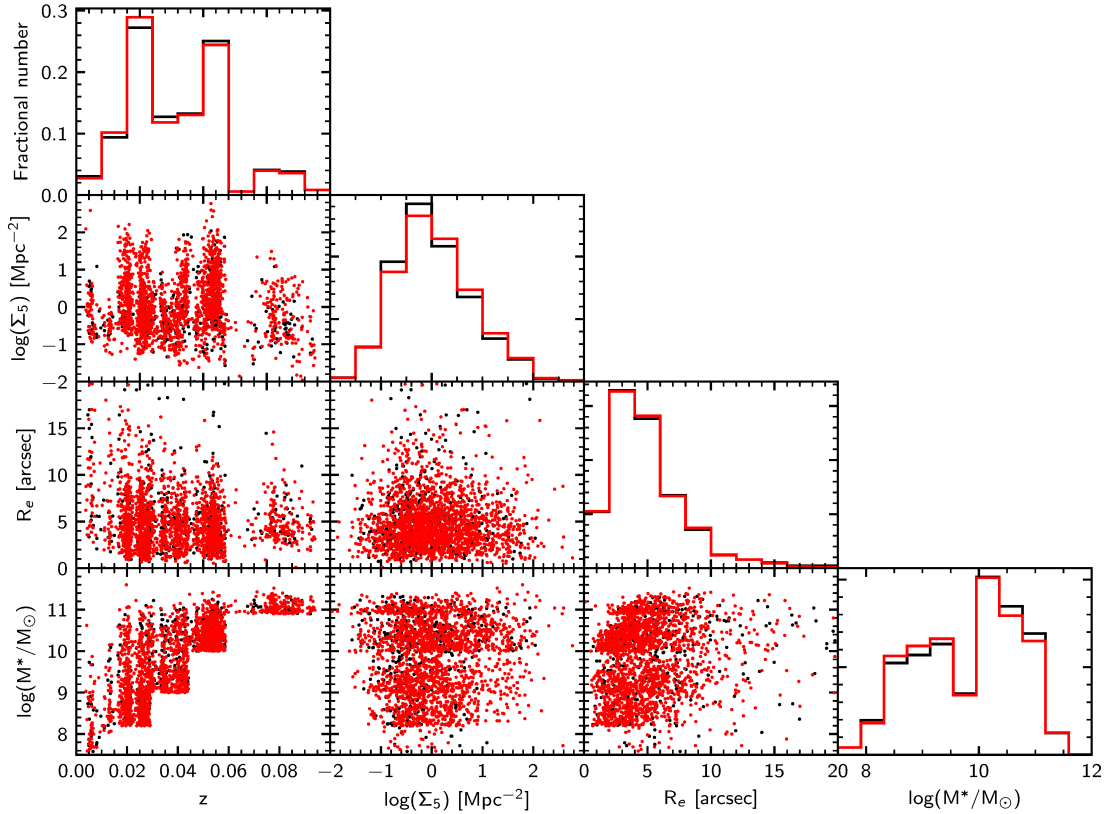


Figure 3. The distribution of good primary SAMI targets in the GAMA regions (black points) compared to the observed targets (red points) as a function of redshift, $\log(M^*/M_\odot)$, R_e (major axis in arcsec) and 5th nearest neighbour density $\log(\Sigma_5)$. Histograms are normalized separately for each sample.

expected, the completeness of secondary targets in the GAMA field is much lower, only 160 of 2514 good targets are observed.

The distribution of good primary targets and observed primary targets in the GAMA regions are shown in Fig. 3 as a function of redshift, stellar mass, major-axis half-light radius (in arcsec) and 5th nearest neighbour density, Σ_5 (see Section 9 for details). The histograms in Fig. 3 are normalized to the total numbers in each population, removing the overall difference in numbers and allowing relative differences to be more visible. In all the displayed parameters the observed distributions are representative of the underlying target distribution. The median stellar masses for the primary target and observed galaxies are $\log(M^*/M_\odot) = 10.04 \pm 0.02$ and 10.02 ± 0.03 respectively. There is a marginally significant difference in the median $\log(\Sigma_5)$, with values of 0.01 ± 0.02 and 0.06 ± 0.02 for primary and observed galaxies. However, this difference is very small compared to the dynamic range of $\log(\Sigma_5)$, which is ~ 4 dex. The distributions of $\log(\Sigma_5)$ for secondary targets are also found to be slightly different, but not significantly so, with median values of 0.04 ± 0.02 and 0.09 ± 0.07 for all and observed objects respectively. We also note that, while not explicitly chosen to fall in regions of low galaxy density, the regularly-spaced redshift cuts used to select the sample were checked to make sure that they did not cut across specific large-scale structures. As a result, the regions with a high density of galaxies in Fig. 3 tend to lie within a single $\log(M^*)$ –redshift region (Bryant et al. 2015).

On average, the completeness of the primary targets in the cluster regions, defined as $N_{\text{obs}}/N_{\text{good}}$ is very good at 87% (Table 2). There are two clusters for which the completeness of observed

primary targets is substantially lower than the average: A168 and A119 with 75 and 79 percent respectively. While A119 has one of the lower completeness values for primaries, it also contains the largest number of primary targets that have been observed.

For the cluster A168, the completeness was affected by a change in the coordinates used to define the cluster centre as outlined in Section 2.1.2. This redefinition of the cluster centre led to the addition of 17 previously-defined secondary targets to the primary sample. Many of these redefined galaxies were ultimately not observed during the survey, leading to an excess of unobserved primaries in the southern part of the cluster (Figure 2).

The distribution of primary targets (black points) and those that have been observed (red points) in the clusters is shown in Fig. 4. As would be expected, these sit at higher 5th nearest neighbour density, $\log(\Sigma_5)$, than the GAMA region galaxies. We find no significant differences between the primary targets and the observed galaxies in any of the parameters shown in Fig. 4. The secondary targets that were observed are significantly different from their parent population in terms of redshift. This is driven by the varying number of secondary targets observed in different clusters. There are no other significant differences between the observed secondaries and their parent population.

The number of objects observed from filler samples can be seen in Table 3. The completeness for filler samples is generally low, as is expected given they were used only when galaxies from the main sample were not available.

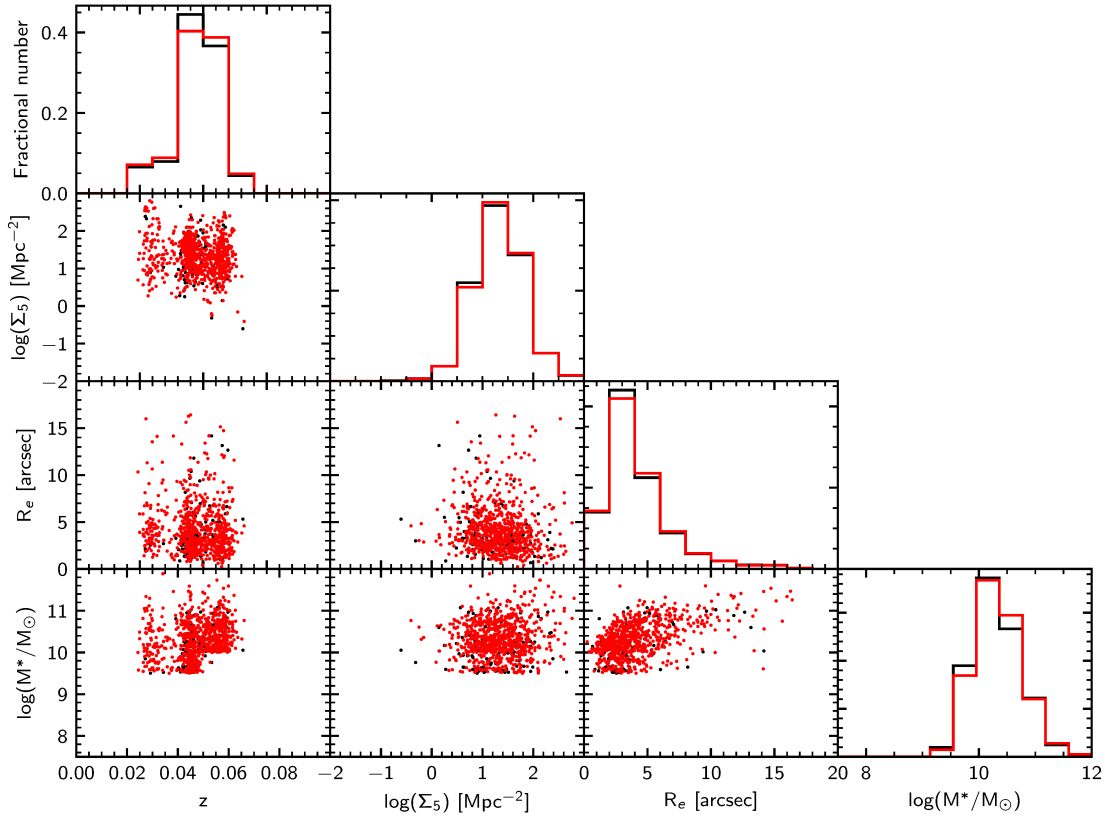


Figure 4. The distribution of good primary SAMI targets in the cluster regions (black points) compared to the observed targets (red points) as a function of redshift, $\log(M^*/M_\odot)$, R_e (major axis in arcsec) and 5th nearest neighbour density $\log(\Sigma_5)$. Histograms are normalized separately for each sample. Axes are on the same scale as Fig. 3.

3 DATA REDUCTION AND DATA QUALITY

The reduction of SAMI data for DR3 follows the procedures described in Allen et al. (2015) and Sharp et al. (2015), including the modifications and additional steps described in Green et al. (2018) and Scott et al. (2018). Here we summarise the process of reducing SAMI data for DR3 and in the following subsections describe in detail significant changes from previous versions of the data. SAMI DR3 data are equivalent to the internal release v0.12. Note that the previous public release, SAMI DR2, used internal release v0.10.1 data.

SAMI data reduction can be neatly divided into two main phases: i) the extraction of row stacked spectra (RSS) from raw observations, and ii) the combining of the RSS frames into 3-dimensional data cubes. The first phase is largely carried out by the 2DFDR data reduction package¹ (AAO software team 2015), with the second phase and the overall process managed by the SAMI PYTHON package (Allen et al. 2014) via the SAMI ‘Manager’.

2DFDR applies the standard steps of overscan subtraction, spectral extraction, flat-fielding, fibre throughput correction, wavelength calibration and sky subtraction. The end result of these steps is a single RSS frame per observation, consisting of 819 1-dimensional spectra corresponding to the 819 fibres of the SAMI instrument. Each RSS frame contains spectra from twelve galaxies (61 spectra

each), one secondary flux standard star (61 spectra) and twenty six sky spectra.

Telluric correction and relative and absolute flux calibration are applied to the individual RSS frames by the SAMI PYTHON package. The flux calibrated, telluric corrected RSS frames are then combined into 3-dimensional datacubes, one per galaxy or secondary standard star, with each cube using spectra from between 6 and 14 separate RSS frames. Fibre spectra from each frame are registered onto a regular grid, including a differential atmospheric refraction correction to align each wavelength slice. Aperture spectra and a set of pre-binned data cubes are also constructed at this stage. The PSF of each datacube is determined from the secondary standard star using a Moffat profile fit to the star datacube at 5000 Å. A dust correction vector, calculated using the dust correction law of Cardelli et al. (1989) and the Planck Milky Way foreground thermal dust map of Planck Collaboration et al. (2014) is also added to each cube (although not applied directly to the data).

For DR3 we have improved various aspects of the SAMI data reduction. This includes modelling of scattered light during spectral extraction, wavelength calibration, sky subtraction, telluric absorption correction, flux calibration, world coordinate system (WCS) calculation and bad pixel rejection. Details of each of these are given below.

¹ <https://www.aao.gov.au/science/software/2dfdr>

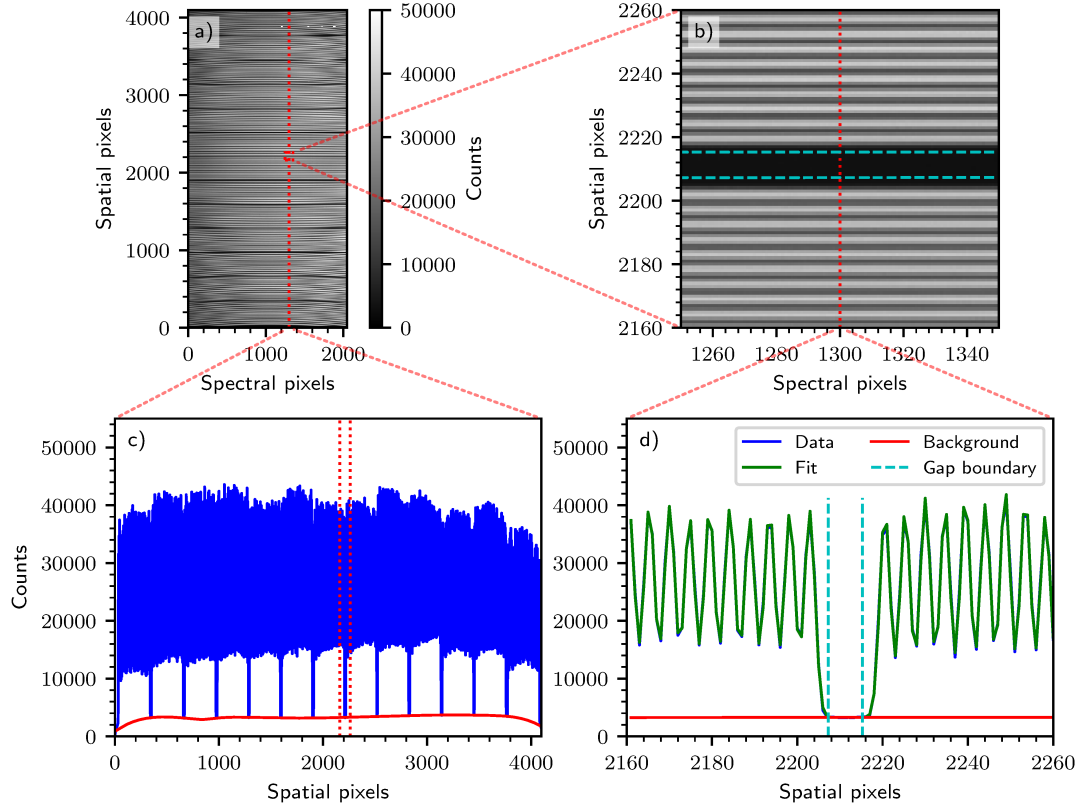


Figure 5. Fibre and scattered light background fits for a SAMI flat field frame. a) A full SAMI fibre flat field. Wavelength varies in the x-direction (spectral pixels) and fibres are arranged in the y-direction (spatial pixels). The red dotted line and box show sub-regions and slices that will be shown in subsequent plots (the aliasing seen in the image of the full detector is purely a display affect). b) A zoom in of the fibre flat field near a gap between slitlets (region indicated by red box in a). Dispersed light from individual fibres can be seen running horizontally. The cyan dashed lines indicate the region defined as the slitlet gap ($\pm 4\sigma_{\text{spat}}$ of the nearest fibre). c) A vertical slice of the fibre flat field at column 1300 showing the observed counts as a function of position (blue) and the best fit scattered light background model (solid red line). The red dashed lines show the region plotted in d. d) A vertical slice in a narrow range around a slitlet gap for column 1300 with the observed counts (blue, mostly hidden by the green line), scattered light background (red) and fitted fibre models (green). The cyan dotted lines show the region defined as the gap.

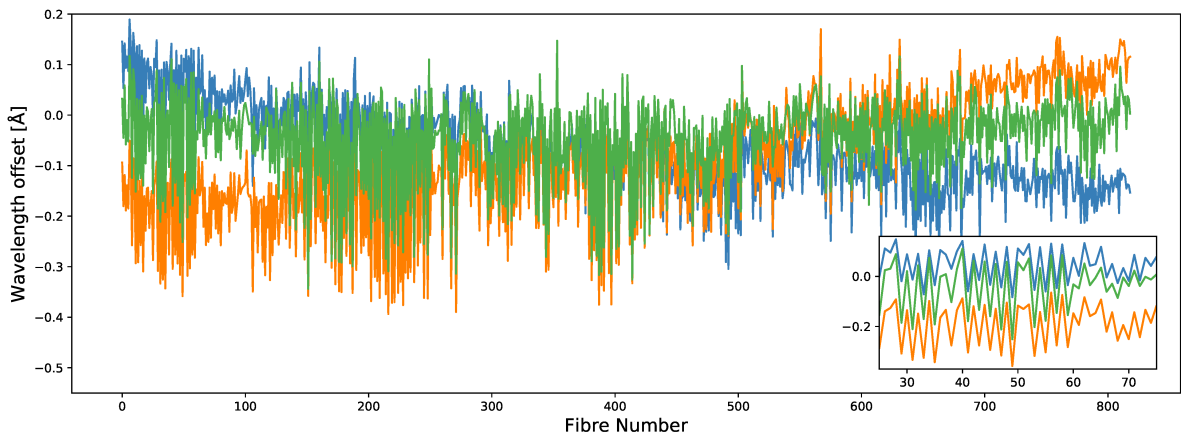


Figure 6. The measured wavelength offset with respect to a high-resolution solar spectrum in Å as a function of fibre number for three twilight sky observations (indicated by the three different colours). The inset shows a zoom in over a small range of fibres to better illustrate the fibre-to-fibre variation.

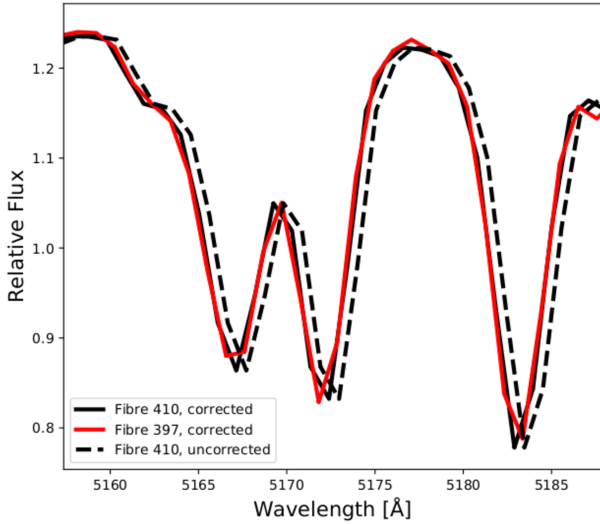


Figure 7. Illustration of the improvement in blue arm wavelength calibration around the Mgb 5177 Å lines. The black solid and dashed lines show a corrected and uncorrected twilight sky spectrum from fibre 410. The solid red line shows a wavelength corrected spectrum from fibre 397. After wavelength correction the agreement between different fibres is significantly improved.

3.1 Extraction and scattered light removal

Extraction of spectra from the 2D CCD image is a fundamental step in data reduction of all fibre-fed spectrographs. The accuracy of this extraction impacts data quality in several ways. The standard approach used within 2DR is to use flat field or twilight exposures to model the width of the fibre profiles in 1D (as a Gaussian of width σ_{spat}) and then simultaneously fit the amplitude of all fibres in a given column on the CCD (see Fig. 5). Each column is fit independently. Previous versions of SAMI data reduction (e.g. [Sharp et al. 2015](#)) also fit a cubic spline simultaneously with the fibre amplitudes to model the scattered light. The spline typically used 12 to 16 uniformly spaced knots across the entire CCD column of 4096 spatial pixels.

The previously used approach suffered problems in two senses. First, the small number of pixels in gaps between blocks of fibres did not have sufficient weight or S/N to place strong constraints on the scattered light model. Second, some data exhibited excess scattered light that was more localized and could not be adequately modelled. The excess in scattered light was particularly significant after the CCD in the blue arm of AAOmega was replaced in the first half of 2014. Contamination of the CCD dewar during the replacement led to low-level condensation on the field-flattening lens in front of the detector within the dewar. This was slowly reduced by repeated pumping over a number of months. The contamination led to extra scattering that was particularly visible around the bright 5577 Å night-sky emission line as a circular Lorentzian halo around each line.

To address the above issues we revised the scattered light modelling so that the new procedure was as follows:

- (i) Identify the location of gaps between slitlets on the CCD image (see Fig. 5) by selecting pixels that are $> 4\sigma_{\text{spat}}$ from the centre of fibres either side of the gap. The gaps are located between groups of 63 fibres (61 for the hexabundle, 2 for sky fibres).
- (ii) For each gap calculate the average flux in bins 30 pixels

wide (10 pixels for bright frames such as flat fields and twilights) along the spectral direction, clipping outliers. The averages are not variance weighted as the count rates in the gaps are low, which can lead to biases in the estimated variance.

(iii) Fit a cubic spline to the average flux along each gap, with 8 (48 for flat fields and twilights) uniformly spaced knots along the spectral direction. For the blue arm of AAOmega, we ignore the ± 60 pixels either side of the 5577 Å night-sky line when fitting the spline.

(iv) The smooth model for scattered light along the gaps is then fit using a cubic spline across the gaps (in the spatial direction) with 6 evenly spaced knots. This provides a full 2D model of the smooth component of the scattered light (red solid line in Fig. 5c, d).

(v) For data from the blue arm of AAOmega the spline is then subtracted from the average flux in the fibre gaps and the residual compared to a 2D Lorentzian model with $\Gamma = 17.6$ pixels. The value of $\Gamma = 17.6$ pixels was chosen based on investigation of the typical width of the scattered light profile across different frames. The model consists of 819 2D Lorentzians, each one centred on the 5577 Å line for a fibre. The normalization of the model is allowed to vary smoothly in the spatial direction on the CCD, and this variation is parameterized by a 4th order polynomial fit to all the fibre gaps simultaneously. The model for scattering around the 5577 Å line is then added to the smooth 2D scattered light model generated in the previous step.

The combined model is then subtracted from the 2D image prior to fitting the fibre profiles to extract the flux in each spectrum. The key outcome from this modified approach is a substantial improvement in sky subtraction that will be further discussed below in Section 3.3. In one particular case an individual field (Y15SAR3_P006_12T097) contained a bright extended source (galaxy 273336) that caused excess scattered light. In this case we fit a high-order (order 16 and 60 in the red and blue arms respectively) polynomial simultaneously with the fit to the fibre profiles using the old scattered light approach.

3.2 Wavelength calibration

When performing a combined analysis of data from the red and blue arms we identified a small (≤ 0.5 Å) offset in the wavelength solution between the two arms. Further investigation identified residual wavelength calibration errors in the blue arm spectra of ~ 0.3 Å. While the blue arm wavelength calibration relies solely upon the arc frame solution, a further refinement is applied to the red arm from a fit to night sky emission lines. Therefore an offset between the two is not unexpected. This offset varies from fibre to fibre but is constant with wavelength and shows no variation in the instrumental dispersion, consistent with the detailed analysis presented in [Scott et al. \(2018\)](#).

In the DR3 data release we implemented an additional wavelength calibration refinement step based on twilight sky observations for the blue arm only. We convolve the extremely high resolution solar spectrum of [Neckel \(1999, R > 300,000\)](#) to the resolution of the SAMI blue arm. We then normalise both the solar spectrum and the SAMI twilight sky spectrum using a 20th order polynomial to remove the continuum shape. We interpolate the SAMI twilight spectrum onto the high spectral sampling wavelength scale of the solar spectrum (0.01 Å pixel⁻¹), then cross-correlate each individual fibre twilight spectra with the solar spectrum to determine the wavelength offset between the two, repeating this process for all 819 fibres.

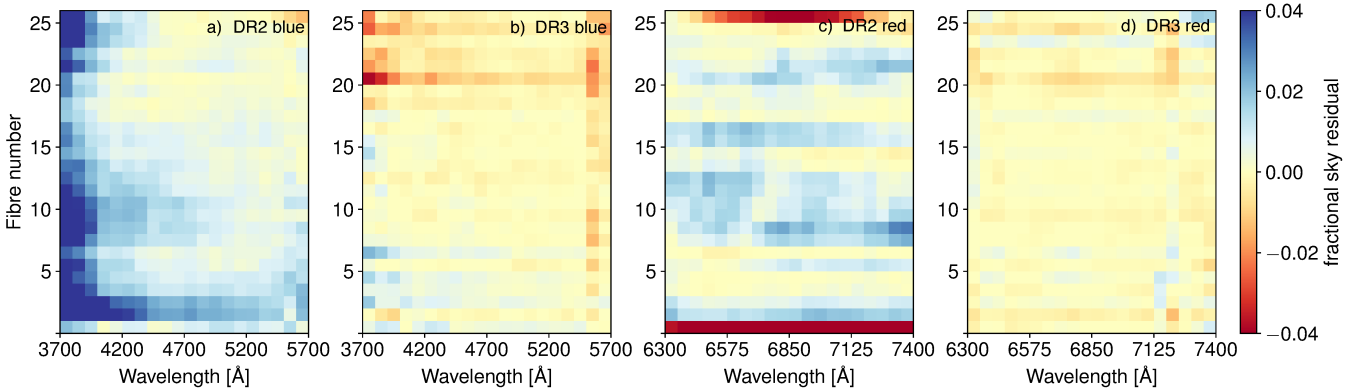


Figure 8. Median fractional sky subtraction residuals as a function of wavelength and fibre number (equivalent to slit location) for sky fibres. On the left we show the blue arm data for DR2 (a) and DR3 (b). On the right we show the red arm data for DR2 (c) and DR3 (d). The median is calculated over all data frames with exposures > 900 s from the entire set of survey observations.

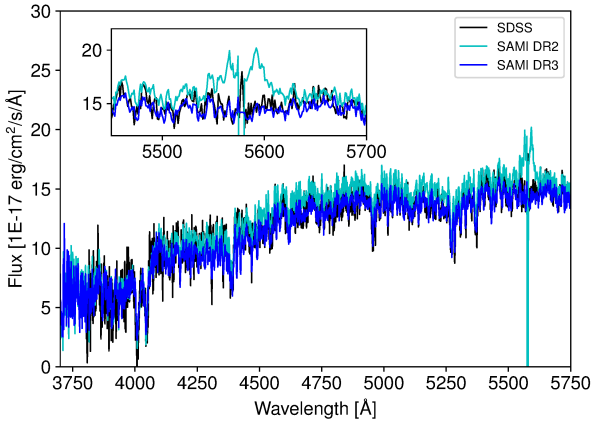


Figure 9. A comparison of SAMI blue arm spectra (3 arcsec diameter aperture) from DR2 (cyan) and DR3 (blue) for object 289198 that showed particularly strong residual scattered light around the $5577 \text{ \AA}$ night-sky line in DR2. We also compare the spectra to an SDSS spectrum of the same object (black). Inset is a close-up of the region around the $5577 \text{ \AA}$ line.

Inspection of the fibre wavelength offsets showed two distinct variations: a pattern of fibre-to-fibre wavelength offsets that was consistent between different observations, and an overall shape variation between observations taken on different nights (or between evening and morning twilight on the same night). This overall shape variation is well approximated by a simple linear fit. In Fig. 6 we illustrate these two trends of the wavelength offset as a function of fibre number for a small subset of all the twilight sky observations taken as part of the SAMI Galaxy Survey. The constant, small-scale, fibre-to-fibre variations are caused by small misalignments between fibres in the slit, combined with the arc lamps feeding the fibres at an f-ratio that is different to the sky (due to the location of the arc lamps in front of the corrector in the AAT top end). The larger-scale effect that changes with time is caused by small shifts in the relative position of camera and slit through the night (Sharp et al. 2015).

Given these two observed patterns we implemented a three-step correction for the blue arm wavelength solution. For each observing run, we first perform and subtract a linear fit to the wavelength offset vs. fibre number relation for each twilight sky observation. We

then average these offsets across all observations for each fibre, resulting in an accurate measurement of the fibre-to-fibre wavelength offsets for each observing run. For each object frame we identify the closest-in-time twilight sky observation and combine the linear fit to the wavelength offset vs. fibre measurements for that twilight sky observation with the averaged fibre-to-fibre wavelength offset described above. As a final correction to account for any linear shift between the nearest twilight and the data frame in question we use a robust linear fit to the shift of the $5577 \text{ \AA}$ sky line as a function of fibre number. This combined wavelength shift is then applied to the wavelength solution for each blue object frame as an additional refinement to the original arc-derived wavelength calibration.

The wavelength correction is illustrated in Fig. 7, which shows a wavelength corrected spectrum from fibre 397 (red line), and a wavelength corrected and uncorrected spectrum from fibre 410 (black solid and dashed lines respectively) for a typical twilight sky observation. As well as improving the wavelength solution in the blue arm, we note that the improved fibre-to-fibre wavelength calibration results in a reduction of the residuals around the $5577 \text{ \AA}$ sky line. See the following section for further details.

3.3 Sky subtraction

A number of improvements have been made to the sky subtraction for SAMI DR3. The first of these are changes discussed above concerning scattered light lead to improved continuum sky subtraction accuracy.

A second modification is related to the way the sky fibre spectra are combined before subtraction. Previous versions of 2dFDR used inverse variance weighting to combine sky spectra. However, at very low count rates in the far blue ($\sim 4000 \text{ \AA}$ and below) this leads to a small but significant bias in the combined sky (typically ~ 0.5 counts). The cause of the bias is that pixels that scatter low will have estimates of their variance that are smaller, and so have greater weight. The DR3 version of 2dFDR uses an unweighted combination of the sky spectra that leads to lower systematic sky-subtraction residuals below $4000 \text{ \AA}$.

To quantify these improvements we carry out a similar analysis to previous SAMI data releases, using residuals in the sky fibres to estimate sky subtraction residuals. In Fig. 8 we show the median sky-subtraction residuals as a function of sky fibre number (equiva-

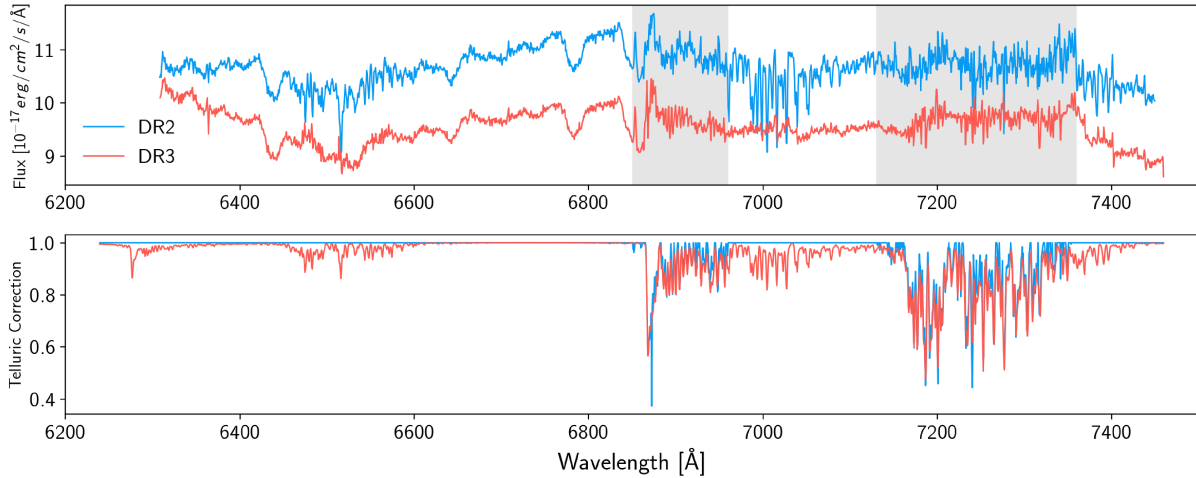


Figure 10. A comparison of the red arm 3 arcsec aperture spectrum for galaxy 136602 from DR2 (blue) and DR3 (orange). The top panel shows the flux, while the bottom panel shows the telluric correction function applied to each spectrum. The grey shaded regions indicate the two bands that telluric correction was applied to in DR2, between 6850 and 6960 Å, and between 7130 and 7360 Å. Comparatively, in DR3 we allow for telluric correction to the entire red spectrum. The offset between the DR2 and DR3 fluxes is due to differences in the flux calibration.

lent to distance along the slit) and wavelength. We compare the new sky subtraction residuals (Figs. 8b and d) to those from DR2 [Figs. 8a and c; see also Fig. 3 of Scott et al. (2018)]. In both the blue and red arms of AAOmega the systematic sky residuals are lower in the DR3 reductions. For example, the increased residuals near the edge of the CCD seen in DR2 data are substantially reduced. The systematic residuals below 4000 Å are also reduced, largely because of the revised sky combination.

A weak residual gradient from top to bottom of the CCD remains in the sky subtraction. This is at the level of $\approx \pm 0.5$ percent. We suspect that this is caused by a small ($\ll 0.1$ pixels from top to bottom) relative stretch of the fibre locations as the slit moves through the night. We previously noted (Sharp et al. 2015) that due to the boiling off of liquid nitrogen the AAOmega cameras systematically shift through the night by ~ 0.03 pixels per hour. To account for this we apply a shift between the fibre ‘tramlines’ measured on a flat field and those observed in the on-sky data. However, for each 0.1 pixel shift, the optics of the spectrograph also cause a relative ≈ 0.03 pixel stretch in the fibre locations from top to bottom of the CCD. This small misalignment leads to slightly different fluxes extracted. Future updates to 2DFDR should incorporate a correction for this effect.

To further quantify our sky subtraction accuracy we estimate the median fractional continuum residual in each sky fibre (this minimizes the impact of shot-noise on the measurement). We then calculate the median of the absolute value of all these residuals across all sky fibres in the entire SAMI DR3 data set. In the blue arm we find a median continuum sky residual of 0.76 percent (and a 90th percentile of 2.9 percent), compared to the DR2 value of 1.2 percent (90th percentile of 4.6 percent). In the red arm the median continuum sky residual is 0.72 percent (90th percentile 2.8 percent) compared to 0.90 percent (90th percentile of 3.1 percent) in DR2.

A third improvement focused on the 5577 Å night sky line. As discussed in Section 3.1, contamination in the dewar of the blue arm of AAOmega causes enhanced scattered light. The scattered light modelling approach described in Section 3.1 reduced the residual scattered light, typically by a factor of 2 or more. However, for the more severely affected frames a significant residual was still visible. To remove this we applied a two-stage principal component analysis

(PCA) in the reduced wavelength calibrated and sky-subtracted data that makes use of the algorithm built into 2DFDR (Sharp & Parkinson 2010).

In the first stage of the PCA correction we generate principal components of the 5577 Å line residual in a narrow 6 Å window around the line. A smooth (median filtered) continuum is subtracted from the spectra and then the faintest 10 per cent of fibres (which will be dominated by sky emission) are used to generate principal components. Only the first three principal components are used at this point. Weights for each component are calculated to minimize the sky line residual in the core of the line. The aim of this first stage is to minimize the residual in the core of the line so that it does not dominate the PCA in the next stage.

The second stage of the PCA correction generates components over a 100 Å window centred on the 5577 Å line. Again only the 10 percent faintest fibres are used to generate the principal components. In this second stage the first ten principal components are used to minimize the sky residuals. Once the model is generated for the sky residuals it is smoothed using a median filter in the fibre direction (of width 101 fibres). This is because in a small number of cases very bright galaxies can have real small-scale structure in their spectra removed by the PCA. However, the scattered light residual around the 5577 Å line varies slowly across the slit. The median filtering maintains excellent correction of scattered light, and removes any negative impact on bright galaxies. An example is shown in Fig. 9 for object 289198 that compares SAMI 3 arcsec diameter aperture spectra from DR2 and DR3, as well as an SDSS spectrum of the same object. The DR2 data shows a significant residual around the 5577 Å line (cyan line) that is completely removed in DR3 (blue line). It is also worth noting that the overall shape of the spectrum from DR3 more closely matches the SDSS spectrum due to improvements in flux calibration that will be detailed below.

3.4 Telluric Correction

The handling of telluric correction has changed substantially since previous data releases, going from a comparatively straightforward polynomial interpolation approach, to using the dedicated telluric feature-fitting software *moIecfit* (Smette et al. 2015; Kausch et al.

2015). While the two methods both rely on the secondary standard star observations, *molecfit* uses a physically motivated model to fit and then correct for atmospheric absorption, therefore requiring additional atmospheric and meteorological data. We fit for two molecules, H_2O and O_2 , using the *equ.atm* reference atmospheric profile supplied in the software installation. The improvement in the telluric correction is illustrated in Fig. 10 for galaxy ID 136602. Previously only two regions were corrected for telluric features, between 6850 and 6960 Å, and between 7130 and 7360 Å (illustrated by the shaded grey regions in Fig. 10). In contrast, for DR3 we allow for correction to the whole red arm spectrum. Therefore, in addition to improving the telluric correction applied to these regions by using a physically motivated model, we now correct for additional features outside these bands that were not accounted for in previous data releases. Most notably, residual telluric absorption around 7000 Å present in DR2 (blue line in Fig. 10), are now modelled and removed in DR3 (orange line).

3.5 Flux calibration

Flux calibration for SAMI makes use of ‘primary’ standards² observed separately to the galaxies and ‘secondary’ standards observed at the same time as galaxies using a hexabundle. The secondary standards are colour-selected to be early F-stars.

The flux calibration of DR2 and earlier was based on the primary flux standards that were typically observed at the start and end of a night. As a result there could be several hours’ difference between observations of galaxy data and standard stars. In some cases (where nights were not photometric) no standard observations were made and the nearest in time and date was chosen to use as calibration. This approach provided relative flux calibration, but to obtain the correct normalization the secondary standard stars observed at the same time as the galaxies were compared to their photometric magnitudes to obtain a scaling for each observed frame. Once cubes were made this scaling was re-applied so that all SAMI cubes should in principle be normalized to the input imaging data (either SDSS or VST).

For DR3 a number of modifications were made to flux calibration procedures. In the processing of the primary calibrators, three changes were made. First, the original reference calibration spectra, that are relatively low resolution (often 50 Å bins) and sometimes had residual telluric absorption still present, were replaced with higher resolution telluric-corrected versions from the Supernovae Factory project (Aldering et al. 2002)³. Second, *molecfit* was applied to the primary standards before determining the spectrophotometric transfer function, allowing us to better sample the transfer function in the red arm. Third, the fitting of the standard star flux within the individual RSS frames was modified to use differential atmospheric refraction (DAR) parameters based on atmospheric conditions and zenith distance. Previous versions had directly fit the positional shift of the star with wavelength, but the under-sampling of the PSF by the fibres meant that in some cases the direct fit was not robust.

The processing of the secondary flux calibration stars was substantially revised, so that a full transfer function could be estimated from them. This followed a similar approach to that used by SDSS (Abazajian et al. 2004). In detail the procedure was as follows:

(i) Individual RSS frames were corrected for atmospheric extinction, approximately flux calibrated using the primary standards and then telluric absorption corrected.

(ii) For each reduced RSS frame the flux of the secondary standard was extracted fitting a Moffat profile allowing for DAR.

(iii) The extracted secondary spectra were fit to model template spectra based on Kurucz (1992) model atmospheres. We used Penalized Pixel-Fitting (pPXF; Cappellari & Emsellem 2004) and this was done in two stages. The first stage fitted individual templates in a grid of effective temperature (T_{eff}), metallicity ($[\text{Fe}/\text{H}]$) and surface gravity. For the best fit surface gravity the nearest 4 templates in T_{eff} and $[\text{Fe}/\text{H}]$ were then refit allowing a linear combination of templates. The fitting was only done on the blue arm data (3700–5700 Å) and included a multiplicative polynomial of 8th order to take out residual transfer function errors. The weights of the templates were saved to the RSS frame.

(iv) The weights of the templates are averaged across all the observations of a field, as these typically contain 7 (or more) observations of the same star.

(v) From the average weights a best fit template spectrum was derived, including Galactic extinction from Planck (Planck Collaboration et al. 2014), as the secondary flux calibration stars are sufficiently distant that they are all in the Galactic halo. The best fit template was normalized by comparing to the observed *g*- and *r*-band photometry for the star in question and applying the average normalization from the two bands.

(vi) Transfer functions for each RSS frame were derived by comparing the observed spectrum to the best fit template. While in principle this could be applied to the individual frames, we only have a single secondary calibration star observed per field, and this can lead to some scatter in the transfer function. Instead we average the transfer functions across all the observations of a field in a given night.

(vii) The average transfer function was applied to the data, but allowing for an individual normalization for each frame to account for variations in transmission.

Once the transfer function was applied, no further scaling of the data for transmission was carried out.

Various tests were carried out to examine the robustness of this approach. We investigated whether variations in the assumed atmospheric extinction curve caused residuals, but there was no correlation found between spectrophotometric residuals and airmass. Residual flux calibration differences (see below) also did not appear to have the same shape as known occasional changes due to excess dust and smoke in the atmosphere (Mike Bessell, private communication). We also found that the scatter in spectrophotometric calibration increased when we applied transfer functions derived from individual frames.

To test the improvement of the spectrophotometric calibration we compare 3 arcsec aperture spectra derived from SAMI cubes (see Section 4) in DR2 and DR3 to single fibre SDSS spectra of the same object. This is not an absolute test, as varying seeing and aperture effects (including differences in pointing) can influence the aperture spectra, particularly for single fibre spectra. However, differences between SDSS and SAMI spectra provide a lower limit on the accuracy of our spectrophotometric calibration.

As a first test we compare the median SAMI flux (across both SAMI arms) to the median SDSS flux in the same spectral windows. As SDSS 3-arcsec diameter spectra are renormalized to the PSF magnitudes we first correct them (on average) back to fibre fluxes with a scaling of 0.35 mag. The distribution of SAMI to SDSS

² These were selected as A or F stars from the ESO spectrophotometric standards list: <https://www.eso.org/sci/observing/tools/standards/spectra.html>

³ <https://snfactory.lbl.gov/snf/snf-specstars.html>

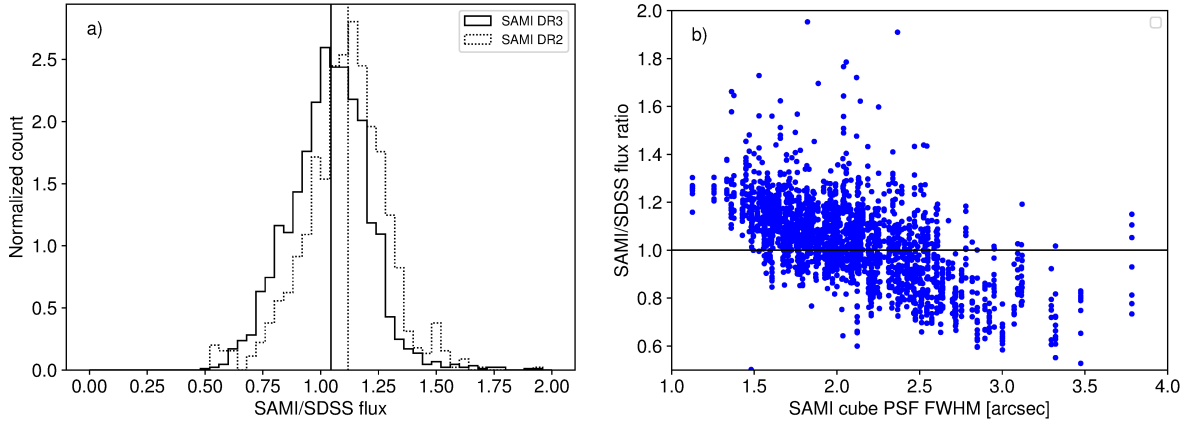


Figure 11. a) the distribution of SAMI to SDSS flux ratio from 3 arcsec diameter aperture spectra for SAMI DR3 (solid line) and DR2 (dotted line). The vertical lines show the median for each histogram. b) The SAMI to SDSS flux ratio for DR3 as a function of the FWHM of the SAMI PSF in each cube.

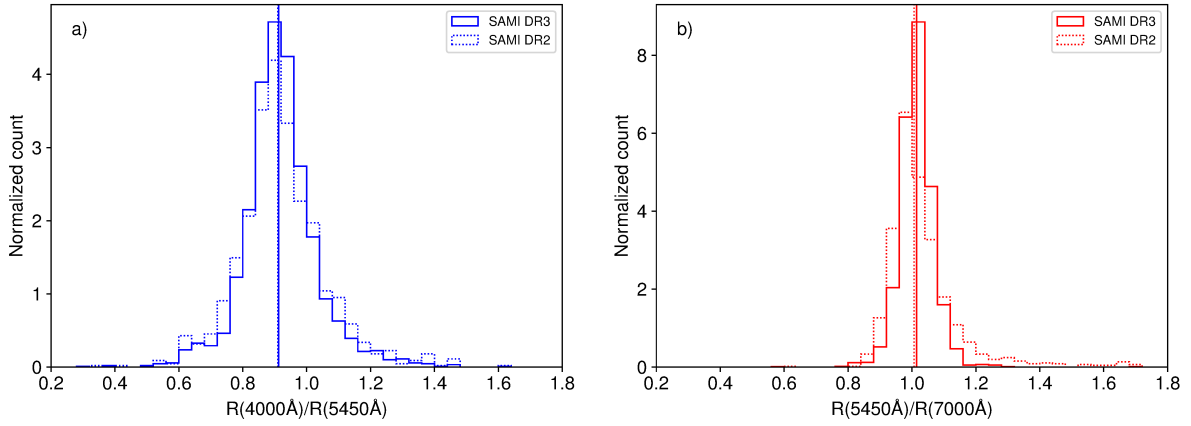


Figure 12. The ratio of SAMI to SDSS spectral flux ratios at different wavelengths, showing difference in spectral shape between 4000 Å and 5450 Å (a) and 5450 Å and 7000 Å (b). The distributions are shown for both DR2 (dotted lines) and DR3 (solid lines). Vertical lines denote the medians of each distribution.

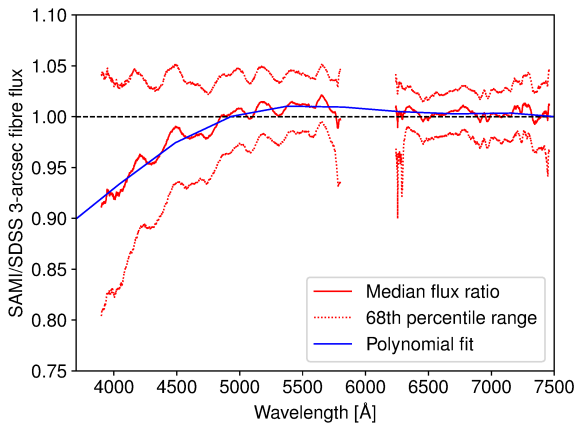


Figure 13. The median ratio of SAMI to SDSS 3-arcsec aperture spectra in the AAOmega blue and red arms. The dotted lines show the 68th percentile range. Spectra are smoothed using a 51 pixel median filter prior to combining, to remove the impact of individual bad pixels. The blue line is a 5th order polynomial fit to the median ratio.

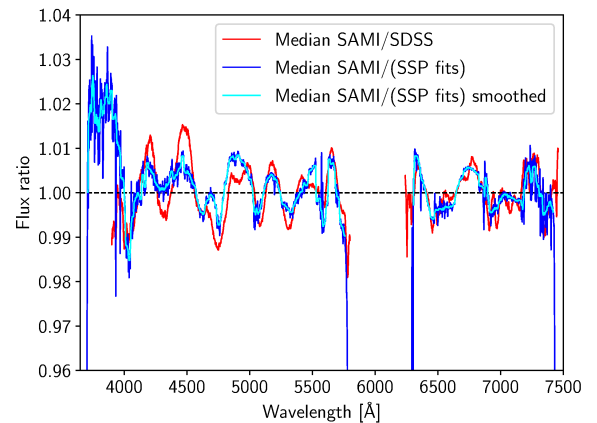


Figure 14. The median ratio of SAMI to SDSS 3-arcsec aperture spectra (red line), compared to the ratio of SAMI to stellar population fits (blue line). Also shown is a median filtered version of the SAMI to stellar population ratio (cyan line).

flux ratios is shown in Fig. 11a. The median offset (solid vertical line) is 1.04 with a 68th percentile width of 0.16. For DR2 the median offset was larger at 1.12, but with the same width of 0.16. The width of the distribution is dominated by variations in seeing, as can be seen in Fig. 11b. The SAMI/SDSS flux ratio declines with increased SAMI seeing as less flux falls within the 3-arcsec diameter aperture defined within the SAMI cubes. There is not the same trend with SDSS seeing, as the SDSS spectra are normalized to PSF magnitudes. Given that the median seeing for the SDSS and SAMI data in this comparison are very similar, 1.96 and 2.04 arcsec respectively, it is not surprising that we have good agreement between the flux scales of the two data sets.

In Fig. 12 we make a comparison of the spectral shapes of SDSS and SAMI aperture spectra. We calculate the median flux in 100 Å bands centred at 4000 Å, 5450 Å and 7000 Å. We then define the ratio

$$R(\lambda) = F(\lambda)_{\text{SAMI}}/F(\lambda)_{\text{SDSS}} \quad (1)$$

for each wavelength, where $F(\lambda)$ is the median flux in a 100 Å window centred at a wavelength of λ . The distribution of the ratio of $R(\lambda)$ at different wavelengths is then shown in Fig. 12. The ratio $R(4000)/R(5450)$ has a median value of 0.91, and the values for both DR2 and DR3 are similar. This implies that on average SAMI spectra have less blue flux than SDSS spectra. The 68th percentile width of the $R(4000)/R(5450)$ distribution is 0.09 for DR3, slightly narrower than the value for DR2 (0.12). The ratio $R(5450)/R(7000)$ (Fig. 12b) is very close to 1, with a median of 1.01 for DR3. The scatter is significantly reduced from DR2 to DR3, with the 68th percentile width reducing from 0.073 to 0.045. The new flux calibration in DR3 has in particular removed the tail to high $R(5450)/R(7000)$ seen in DR2. This shows that the blue and red arms are now better matched spectrophotometrically.

Fig. 13 shows the median ratio of SAMI to SDSS flux, once each spectrum is scaled by the median flux ratio (to take out normalization differences due to seeing and other aperture effects). The trend seen in Fig. 13 is consistent with that in Fig. 12a. At ~ 4000 Å SAMI spectra have $\simeq 7$ percent less blue flux. The 68th percentile range of flux ratios (dotted lines in Fig. 13) is seen to increase towards the blue. To enable a direct mapping to the SDSS spectral flux scale we fit a 5th order polynomial to the flux ratio (blue line in Fig. 13). The best fit function is

$$R(x) = 8.25975 - 41.8455x + 90.374x^2 - 93.2674x^3 + 46.5418x^4 - 9.0518x^5, \quad (2)$$

where $x = \lambda/5500$ and λ is the wavelength in Å. We have investigated various explanations of the difference between SDSS fibre spectra and SAMI aperture spectra that could be caused by the flux calibration of either or both data sets. Small known differences in the pipeline approaches, such as SAMI using Planck dust extinction estimates compared to SDSS using estimates from Schlegel et al. (1998), can only account for ~ 1 percent differences. A small number of individual SDSS plates show differences in the relative flux calibration that are factors of 2–3 larger than the median relation in Fig. 13, but removing these does not noticeably modify the median trend. We do not find correlations of the difference with airmass or seeing. We do not apply Eq. 2 directly to the SAMI DR3 data products, as the source of the difference between SAMI and SDSS flux calibration is unclear. However, we provide it here to allow users to correct (on average) between the two data sets.

Once the SAMI/SDSS flux ratio is corrected using Eq. 2 there are weak (less than 1 percent) residual small-scale features remaining. To examine whether these are real we create a flux ratio between

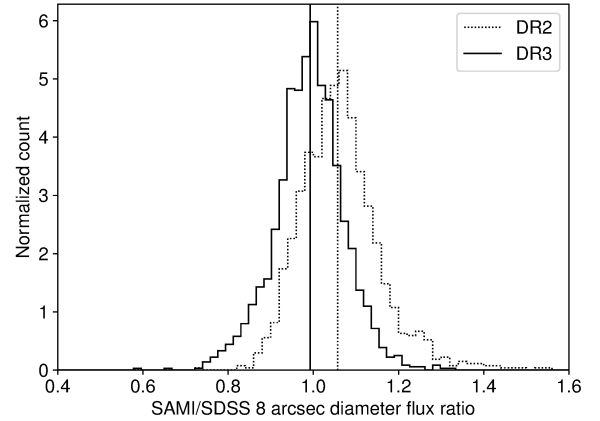


Figure 15. The distribution of flux ratios of SAMI blue cubes to SDSS g -band imaging for DR3 (solid lines) and DR2 (dotted lines). The vertical lines show the median values in each case.

SAMI and model stellar population fits made to 69428 SAMI spectra (each binned to a minimum S/N ratio of 20 Å^{-1}). These fits are performed using the penalised pixel fitting code pPXF (Cappellari & Emsellem 2004; Cappellari 2017). We use templates from the simple stellar population (SSP) models of Vazdekis et al. (2015), as well as including templates representing a number of common optical emission lines in the (rest-frame) wavelength range of 3500 Å to 7000 Å. Full details of these stellar population fits will be presented in an upcoming paper (Vaughan et al. in prep).

A comparison between the SAMI/SDSS flux ratio (red line) and the SAMI/(SSP fits) flux ratio (blue line) is shown in Fig. 14. The small-scale variations in the flux ratios are in relatively good agreement, suggesting that most of the ripples seen are low-level systematic features in the SAMI flux calibration. In most cases these flux calibration residuals will only impact spectral fits of very high S/N spectra. However, as part of DR3 we provide a correction vector to remove these residuals based on the SAMI/(SSP fits) flux ratio. This is median filtered using a 21 pixel window and only specified within the wavelength ranges 3710 – 5760 Å and 6305 – 7420 Å where the correction is well defined. The correction vector is shown as the cyan line in Fig. 14 and SAMI data should be divided by this vector to apply the correction.

We perform an independent fit of the ~ 3000 SAMI aperture spectra (see Section 4.2) to test the above correction using the STARLIGHT code (Cid Fernandes et al. 2005) and templates from Bruzual & Charlot (2003). The correction leads to very slightly reduced residuals, but these are not significant for individual spectra. For example, using the correction the mean fractional residuals are reduced from -0.029 ± 0.187 to -0.028 ± 0.177 in the window 4000–4500 Å. At 4500–5000 Å the change is slightly larger from 0.017 ± 0.084 to 0.014 ± 0.082 . The errors quoted are the RMS scatter about the mean. Tests using other code and templates lead to similar results. There is also no significant difference in the estimated ages and metallicities. In summary, correcting the low-level ripples in flux calibration does not significantly impact our spectral fitting.

Because of the impact of aperture effects when comparing to SDSS 3-arcsec fibre spectra, we also make a comparison to the SDSS g -band imaging. We follow the same procedure described by Scott et al. (2018), taking an 8 arcsec diameter aperture to measure the flux in the SAMI cubes and SDSS g -band imaging data. The flux

from the SAMI cube is convolved with the *g*-band filter response prior to making the comparison. The ratio of SAMI to SDSS flux is shown in Fig. 15. Consistent with the comparison to SDSS spectra, SAMI DR3 data is a closer match to SDSS photometry. The median ratio for DR3 is 0.991 ± 0.002 , compared to 1.051 ± 0.003 for DR2. The 68th and 95th percentile ranges of the flux ratio are ± 0.074 and ± 0.135 for DR3 compared to ± 0.081 and ± 0.144 for DR2. Hence the scatter is also slightly reduced in DR3 relative to DR2.

3.6 Centring and WCS

The observed flux of each galaxy in each individual exposure was fit with a two-dimensional Gaussian to identify the centre of the source. This approach sometimes generates an inaccurate centroid when there is a bright star or a secondary object within the field-of-view of the hexabundle. In such a case, we masked out the flux from the secondary object before fitting a Gaussian, which reduces the mis-identification of the galaxy of interest. After this, only 0.5% of the entire survey (12 galaxies) is assessed to have inaccurate centroids due to a bright secondary object/star (catalogue IDs 218717, 228278, 273256, 380682, 509892, 549182, 610816, 760733, 91579, 93384, 9008500120, and 9016800113). The dithers are combined after aligning the centroids. The galaxy centre is located at cube spaxel coordinates (25.5, 25.5) where we assign the catalogue coordinate of the galaxy. Then, we defined a world-coordinate system (WCS) for each cube using the relative position to the spatial coordinate of the galaxy centre.

We characterise the accuracy of the SAMI WCS by cross-correlating the reconstructed SAMI images with *g*-band images from the SDSS. We applied the SDSS *g*-band filter response to each cube to reconstruct a two-dimensional image. The SDSS images have been re-sampled to have the same pixel scale as the SAMI cubes. Fig. 16 presents the offset in RA and Dec. between the SAMI and SDSS WCS. The median offset is -0.030 ± 0.004 arcsec in RA and -0.017 ± 0.005 arcsec in Dec. We visually inspected 51 outliers whose offset is greater than 0.5 arcsec (one SAMI pixel) either in RA or Dec. We found 8 galaxies whose catalogue coordinates do not correspond to the object centre in the SDSS images though they are well centred on the SAMI cubes. The circular Gaussian distribution used for centroiding may not represent the centre of highly disturbed or edge-on galaxies very well. We found an offset greater than 0.5 arcsec from 31 disturbed and 12 edge-on galaxies. We corrected the cube WCS of 12 mis-identifications (listed above) and the 51 outliers to match that estimated from the SDSS imaging.

3.7 Cubing and bad pixel rejection

The procedure for cubing is based on a Drizzle-like algorithm, described by Sharp et al. (2015). For DR3 the only modification to the cubing was to implement a more robust bad pixel rejection algorithm. In DR2 and prior, the bad-pixel rejection was a simple clipping of pixels across multiple frames that are more than 5σ from the median value in a given output pixel. However, in some cases this was not stable, particularly for very bright objects, where small seeing variations between frames might lead to spurious rejection. The result of this was that a small number of individual spaxels in the output cubes had sudden steps in their data as a function of wavelength.

The bad pixel rejection as part of cubing is particularly important for the red arm of AAOmega. This is because the thick detector in the red camera causes some cosmic ray events to be spread out,

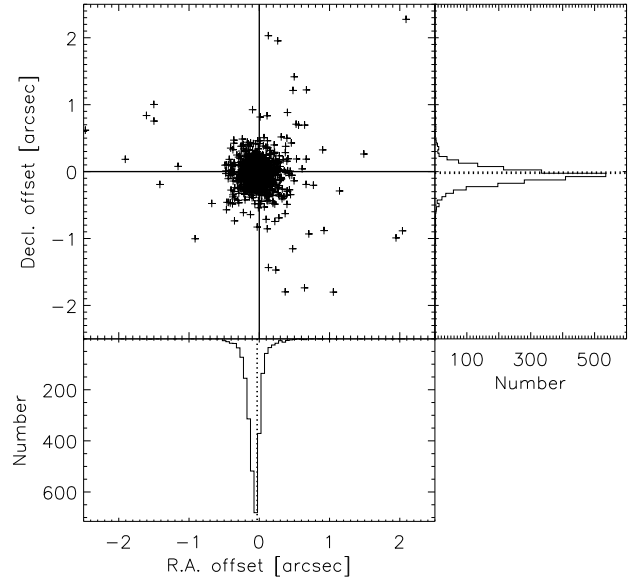


Figure 16. The distribution of R.A. and Decl. offset between SAMI and SDSS with histograms of the offset in R.A. and Decl. along the axes. Dotted lines in the histograms show the median offset in R.A. (-0.030 arcsec) and Decl. (-0.017 arcsec). Note that 12 mis-identified galaxies are not presented.

to be as broad as the instrumental PSF. As a result, regular edge-detection algorithms (van Dokkum 2001; Husemann et al. 2012) are not effective.

To improve the bad-pixel rejection we implemented a more robust algorithm that compares fibre spectra prior to projecting them into a cube. For each fibre spectrum we identify the six fibres spatially nearest to the fibre in question (using all RSS frames that will contribute to the cube). Each fibre has the median filtered continuum subtracted and they are then combined to generate a median spectrum (from seven fibres including the fibre being tested for bad pixels). The median absolute deviation is estimated from the seven continuum subtracted spectra. The median spectrum is then compared to the individual spectra, allowing for a scaling between the two, and pixels that were more than 5 times the median absolute deviation away from the median spectrum were flagged to be rejected. Application of this new method resolved the problems that the previous algorithm had caused.

4 PRIMARY DATA PRODUCTS

4.1 Cubes

The main data products from the SAMI Galaxy Survey are spectral cubes for each observed galaxy. The structure of these is unchanged from DR2 (Scott et al. 2018), but we give a brief description here for completeness. There is a separate cube for the blue and red arms of the spectrograph. The cubes have 50×50 spatial pixels (spaxels), each one 0.5×0.5 arcsec in area, and 2048 wavelength slices. Each cube contains the measured flux, a variance cube, a weight map and the compressed covariance between spaxels (Sharp et al. 2015).

In DR3 we release all cubes that pass our quality control limits (see Section 2.3). The release includes repeat observations of galaxies as separate cubes. There are a total of 3426 pairs of galaxy cubes of 3068 unique galaxies. We also include the calibration star

cubes used for flux calibration and PSF estimation. There are 286 pairs of calibration star cubes of 177 unique stars.

As well as the default cubes with 0.5×0.5 arcsec spaxels we release cubes with different binning schemes. These are described in detail by [Scott et al. \(2018\)](#). In summary these are:

- ‘adaptive’ binning, using the Voronoi method of [Cappellari & Copin \(2003\)](#) to a median blue arm S/N of 10.
- ‘annular’ binning into five linearly spaced elliptical annuli. The ellipses are defined using the `FIND_GALAXY` routine ([Cappellari 2002](#)) applied to images generated by collapsing the cubes in the wavelength direction.
- ‘sector’ binning that azimuthally subdivides each of the annular bins into eight equal area regions.

All the binned cubes are generated using the binning module within the SAMI data reduction pipeline ([Allen et al. 2014](#)).

4.2 Aperture spectra

To allow easy comparison to single aperture data (e.g. single fibre surveys), and to measure quantities in consistent physical or relative apertures, we generate aperture spectra. The data structure for aperture spectra is unchanged compared to DR2 ([Scott et al. 2018](#)). The aperture spectra are calculated using the binning module in the SAMI data reduction pipeline. The apertures used are the following:

- circular 1.4, 2.0, 3.0 and 4.0 arcsec diameter apertures.
- a circular 3 kpc diameter aperture based on the flow corrected redshift, z_{tonry} (for the GAMA regions), or the cluster redshift (for cluster galaxies).
- an elliptical $1R_e$ aperture, based on Sersic fits to SDSS or VST imaging data, as described by [Kelvin et al. \(2012\)](#) and [Owers et al. \(2019\)](#).
- an elliptical $1R_e$ aperture based on multi-Gaussian expansion (MGE) fitting of SDSS or VST photometry. The MGE fitting is described in Section 5 and is new to DR3.

For a small number of galaxies aperture spectra are not available in all apertures. For example MGE measurements have not been made for galaxies in the filler samples.

4.3 Input and observational catalogues

Input catalogues for all sources are provided as part of DR3. This is a total of 5 catalogues listed in Table 4. The input catalogue for the GAMA regions (named InputCatGAMADR3) is identical to that presented by [Bryant et al. \(2015\)](#) with two exceptions. The first is that two sets of coordinates are now listed. One set, named RA_OBJ and DEC_OBJ are the object coordinates taken from the input photometric catalogues. A second set of coordinates is also provided, named RA_IFU and DEC_IFU. These give the nominal IFU pointing for each target. In most cases these two sets of coordinates are identical, but in a small number of cases the IFU pointing was modified (see Section 2.1). The second change for the GAMA input catalogue is that the source position angle (PA) has been corrected to the standard on-sky North through East definition. Previous versions of the input catalogue had PA defined in image coordinates, taken directly from the GAMA Sersic fit catalogues ([Kelvin et al. 2012](#)).

The input catalogue for the SAMI cluster regions (named InputCatClustersDR3) is newly presented in this data release. It is based on the SAMI Cluster Redshift Survey ([Owers et al. 2017](#))

carried out using the 2-degree Field (2dF) instrument on the AAT. Selection is described in Section 2.1.2 and [Owers et al. \(2017\)](#). The cluster input catalogue includes similar measurements to the GAMA input catalogue, including redshift, stellar mass, $g-i$ colour. It also includes R_e , ellipticity and PA derived from Sersic photometric fits from [Owers et al. \(2019\)](#).

With DR3 we release the complete catalogues of all potential filler targets for the SAMI Survey (catalogue name InputCatFiller, see also Section 2.1.3). These catalogues do not have the same extensive set of parameters as the main survey targets. The input catalogues for colour selected F-stars is also provided as part of DR3 in both the GAMA and cluster regions (these catalogues are named FstarCatGAMA and FstarCatClusters).

Details of all the observations that passed our quality criteria and that led to cubes being produced are listed in the catalogue CubeObs. The primary role of this catalogue is to provide quality assessments such as seeing and relative transmission for each observed cube. The CubeObs catalogue includes observations of calibration stars. Where repeat observations have led to multiple cubes being made, the CubeObs catalogue separately lists each set of cubes. The flag ISBEST is used to indicate in the case of repeats which of the observations is considered the highest quality. The CubeObs catalogue also contains a number of flags indicating a variety of issues, including missing value-added products, multiple objects in the IFU field-of-view and various calibration and measurements problems. Description of these flags is contained as part of the schema in the online database.

5 PHOTOMETRY BASED CATALOGUES

5.1 Visual Morphology

All galaxies in the survey have been visually classified by several members of our team taking advantage of RGB combined colour images from either the SDSS Data Release 9 or VST ATLAS surveys. The classification scheme is simple, with just 4 types (ellipticals, S0s, early- and late-spirals), and it is based on the prominence/presence of a bulge component and spiral arms. For a small fraction of the sample (148 galaxies, $\sim 5\%$ of the sample), no classification is provided as it was considered too uncertain/subjective. We refer the reader to [Cortese et al. \(2016\)](#) for a more extensive description of our classification procedure. The morphology measurements are available in catalogue VisualMorphologyDR3.

5.2 MGE fits

For each galaxy, we measure the photometric parameters of magnitude, position angle, projected ellipticity and projected half-light radius using the Multi Gaussian Expansion algorithm (MGE, [Emmellem et al. 1994](#)). The measurements are based on r -band SDSS and VST images. To characterise the PSF, we use the source detection software SExtractor ([Bertin & Arnouts 1996](#)) to retrieve all the stars in a 400×400 arcsec cutout image around the galaxy in question. We then use the Python package MGEFIT to obtain the MGE of each star ([Cappellari 2002](#)). To model the galaxy, the adopted PSF is the best-fit MGE of the star that is the best compromise between goodness of the MGE, magnitude and distance to the target. To mask regions contaminated by neighbours and interlopers, we use segmentation maps from SExtractor. The position angle is measured using either the method of moments, or SExtractor, whichever gives the lowest χ^2 . The total magnitude m

Table 4. A summary of the catalogues provided as part of DR3. We list the catalogue name, the number of rows (N_{rows}), a brief description of the catalogue and the primary references for the catalogue.

Name	N_{rows}	Description	Reference
InputCatGAMADR3	5536	Targets and their properties in GAMA regions	Bryant et al. (2015)
InputCatClustersDR3	1433	Targets and their properties in cluster regions	Owers et al. (2017)
InputCatFiller	2980	Filler targets	This paper
FstarCatGAMA	2578	Colour selected calibration stars	Bryant et al. (2015)
FstarCatClusters	183	Colour selected calibration stars	Owers et al. (2017)
VisualMorphologyDR3	3068	Visual morphological classification	Cortese et al. (2016)
CubeObs	3712	Observed cubes and quality flags	This paper
MGEPhotomUnregDR3	3150	MGE photometric measurements	D'Eugenio et al. in prep
samiDR3Stelkin	3426	Stellar kinematic measurements	van de Sande et al. (2017b)
samiDR3gaskinPA	3426	Gas kinematics position angle	Scott et al. (2018)
IndexAperturesDR3	3375	Stellar continuum index measurements	Scott et al. (2017)
SSPAperturesDR3	3375	SSP age, [Z/H] and [α /Fe]	Scott et al. (2017)
DensityCatDR3	6969	Local density estimates	This paper
EmissionLine1compDR3	3245	Aperture em. line measurements, 1-component fits	Scott et al. (2018)
EmissionLineRecomcompDR3	3245	Aperture em. line measurements, recommended component fits	Scott et al. (2018)

and the circularised half-light radius R_e are calculated from the best non-regularised MGE model. The projected ellipticity is the ellipticity of the isophote of area πR_e^2 . Our measurements are validated against the corresponding GAMA measurements, that are based on 2-dimensional Sérsic fits (Kelvin et al. 2012). In log-space, the best-fit linear relation between Sérsic- and MGE-based half-light radii has $rms = 0.067$ dex. Based on this result, we estimate the measurement uncertainty of the MGE half-light radii as $0.067/\sqrt{2} = 0.047$ dex, similar to the estimate of ATLAS3D (Cappellari et al. 2013). In the Abell 85 cluster both SDSS and VST photometry is available and so we provide measurements for both sets of data within the MGE catalogue. More information on these measurements will be provided in a separate paper (D'Eugenio et al., in prep.)

6 STELLAR KINEMATICS VALUE ADDED PRODUCTS

6.1 Method

We extract the line-of-sight stellar velocity distribution (LOSVD) from SAMI cubed data using the penalised Pixel Fitting code (pPXF) to fit all spectra (Cappellari & Emsellem 2004; Cappellari 2017). The method is described in detail by van de Sande et al. (2017b) with the various value added products outlined by Scott et al. (2018). We give a brief overview below.

In DR3 we provide two stellar kinematic data products: one where we assume a Gaussian LOSVD (two moments: V_{m2} , σ_{m2}), and an alternative where the LOSVD is parametrized with a truncated Gauss-Hermite series (van der Marel & Franx 1993; Gerhard 1993) with four kinematic moments: V_{m4} , σ_{m4} , h_3 and h_4 . Here, h_3 and h_4 are related to the skewness and kurtosis of the LOSVD. Local optimal templates are derived from annular binned spectra (Section 4) using the MILES stellar library (Sánchez-Blázquez et al. 2006; Falcón-Barroso et al. 2011) that consists of 985 stars covering a large range in stellar atmospheric parameters.

To extract the LOSVD, we fit each spaxel three times with pPXF. Once for determining a precise measure of the noise scaling from the residual of the fit, a second time to clip outliers using the CLEAN parameter in pPXF, and a third time where pPXF is provided with the optimal templates from the annular bin in which the spaxel is located (derived from the previous fits to the annular bins),

as well as the optimal templates from neighbouring annular bins. A 12th order additive Legendre polynomial is used to remove residuals from small errors in the flux calibration. Uncertainties on the LOSVD parameters are estimated using a Monte-Carlo approach. The same method is applied to the binned data (Section 4). For the aperture spectra (Section 4), optimal templates are constructed for each individual aperture and we then use the same procedure as described above to extract the LOSVD.

For SAMI DR3 data we recommend applying the following quality criteria to the stellar kinematic maps. For the Gaussian LOSVD maps: $S/N > 3 \text{ Å}^{-1}$, $\sigma_{\text{obs}} > \text{FWHM}_{\text{instr}}/2 \sim 35 \text{ km s}^{-1}$, $V_{\text{error}} < 30 \text{ km s}^{-1}$ (Q_1 from van de Sande et al. 2017b), and $\sigma_{\text{error}} < \sigma_{\text{obs}} \times 0.1 + 25 \text{ km s}^{-1}$ (Q_2 from van de Sande et al. 2017b). For the four-moment LOSVD fits, a reliable estimate of h_3 and h_4 requires a higher S/N quality cut of $> 20 \text{ Å}^{-1}$ and $\sigma > 70 \text{ km s}^{-1}$ (Q_3 from van de Sande et al. 2017b).

6.2 Products

Maps of the two- and four-moment kinematic fitted parameters are provided as part of DR3 for all the binning schemes we release. Additionally we provide four-moment fits for another binning scheme, adaptively binned to $S/N = 20 \text{ Å}^{-1}$. From the two-moment kinematic maps we derive the following value added products.

(i) The kinematic asymmetry of the galaxy velocity fields. We determine the amplitudes of the Fourier harmonics k_5/k_1 on all velocity data that pass the quality cut Q_1 , measured using the KINEMETRY routine (Krajnović et al. 2006, 2008) following the method outlined in Krajnović et al. (2011) and van de Sande et al. (2017b).

(ii) The position angle (PA) of the stellar rotation. We measure the PA from the stellar velocity maps on all spaxels that pass the quality cut Q_1 using the FIT_KINEMATIC_PA code that is based on the method described in Appendix C of Krajnović et al. (2006).

(iii) The ratio of ordered versus random motions given by $(V/\sigma)_e$ and λ_{R_e} . We use the definitions described by Cappellari et al. (2007) and Emsellem et al. (2007). We calculate the sum of the unbinned flux, velocity, and velocity dispersion maps over all spaxels that pass the kinematic quality cuts Q_1 and Q_2 (van de Sande et al. 2017b) within an ellipse with semi-major axis R_e and axis ratio b/a . In the calculation of λ_R we adopt a radius, R , definition as described by Cortese et al. (see 2016); i.e. R is the semi-major axis

of the ellipse on which each spaxel sits, not the polar radius. A fill factor of 95 per cent of good spaxels within the aperture is required for producing λ_{R_e} and $(V/\sigma)_e$ measurements. When the largest kinematic aperture radius is smaller than the effective radius, we apply an aperture correction to V/σ and λ_R as described in [van de Sande et al. \(2017a\)](#).

For these kinematic data products, we provide a flag derived from a visual inspection where we have identified maps with irregular kinematics due to nearby objects or mergers that influence the stellar kinematics of the main object. The main stellar kinematic catalogue is *samiDR3Stelkin*, however the flags for problem objects are contained within the main observational catalogue *CubeObs*, as they can influence measurements other than the stellar kinematics.

7 STELLAR POPULATION VALUE ADDED PRODUCTS

In DR3 we provide a table of single stellar population (SSP) equivalent measurements of the light-weighted age, metallicity ($[Z/H]$) and alpha-enhancement ($[\alpha/Fe]$) derived from the SAMI DR3 aperture spectra included in this release. We provide SSP measurements for all apertures with sufficient S/N, averaged over the entire blue arm spectrum. We impose a minimum S/N limit of 10 per Å for age and $[Z/H]$ and a minimum S/N of 20 per Å for $[\alpha/Fe]$ measurements – we consider measurements that satisfy this criterion ‘successful’. The value -99 is used to indicate measurements that do not satisfy these S/N criteria, or are otherwise missing. We also provide Lick absorption line index measurements for 20 different indices for all apertures.

We attempt to measure SSP equivalent parameters for all apertures of all galaxies that have 3-kpc aperture spectra in DR3. The Lick absorption line index and SSP measurement tables provided as part of DR3 include 3375 objects, of which 358 are duplicate observations. We successfully measure age and $[Z/H]$ ($[\alpha/Fe]$) for 2989 (2754) unique galaxies in the 4 arcsec circular aperture (the highest average S/N aperture), decreasing to 2772 (2241) in the 1.4 arcsec circular aperture (the lowest average S/N aperture). Median formal uncertainties in the 4 arcsec aperture are 1.4 Gyr in age, 0.20 dex in $[Z/H]$ and 0.18 dex in $[\alpha/Fe]$. Comparison of 70 galaxies with independent duplicate observations within DR3 gives rms scatters in reasonable agreement with the formal uncertainties (0.15, 0.27 and 0.12 dex in log age, $[Z/H]$ and $[\alpha/Fe]$); however, we note that these estimates are less reliable given the relatively small number of duplicate observations of any given galaxy. When comparing DR2 and DR3 population measurements for galaxies in common between the two we find rms scatters of 0.12 dex in log age, 0.21 dex in $[Z/H]$ and 0.14 dex in $[\alpha/Fe]$.

The method used to measure absorption line indices and SSP-equivalent parameters is identical to that outlined in [Scott et al. \(2018\)](#), with details provided in [Scott et al. \(2017\)](#). Briefly, absorption line indices are measured on emission and Milky Way-extinction corrected blue-arm aperture spectra using a PYTHON implementation of the *EZ_AGES* IDL software package ([Graves 2014](#)), which includes a correction for the intrinsic velocity dispersion of each galaxy. We adopt index definitions from [Trager et al. \(1998\)](#). From the Lick index measurements we determine stellar population parameters by comparing to the stellar population synthesis models of [Thomas et al. \(2011\)](#), for metallicity and alpha-enhancement) and [Schiavon \(2007\)](#), for age. For a discussion of why these two different sets of models are used, see [Scott et al. \(2017\)](#). We select the best-fitting SSP-equivalent age, $[Z/H]$ and $[\alpha/Fe]$ using a χ^2 -minimisation approach following [Proctor et al. \(2004\)](#).

8 EMISSION LINE VALUE ADDED PRODUCTS

The emission line value-added products have been one of the key parts of the previous data releases, and thus a systematic process has been built to create these. The final SAMI galaxy survey data release follows a similar fitting process, and full details of the process can be found in the earlier data releases ([Green et al. 2018](#); [Scott et al. 2018](#)). We briefly summarise that process here and emphasize key differences with the previous data releases.

For each of the data products we described in Section 4 (the cubes, the various binned cubes and the aperture spectra) we fit the strong emission lines in each unique spatial element with one to three Gaussian profiles using the fitting code LZIFU ([Ho et al. 2016a](#)). As in DR2, we first continuum subtract each spatial element using pPXF with the MILES simple stellar population (SSP) spectral library ([Vazdekis et al. 2010](#)) supplemented with younger SSP templates drawn from [González Delgado et al. \(2005\)](#). However, for the full resolution cubes there is not sufficient S/N in the continuum in each spatial element for a good fit, so we first fit the continuum and lines simultaneously in the Voronoi-binned cubes and then use this fit to refit the full-resolution cube using a limited set of templates with priors on the weights, as described in [Owers et al. \(2019\)](#). However, even with this approach, there are spaxels (and even Voronoi-binned elements) where the spectra has a zero or even negative (due to sky-subtraction residuals) signal-to noise. In these cases we set the continuum to 0, but still attempt to fit any faint emission line that may exist. Due to differential atmospheric refraction some spaxels will not have complete spectral coverage at the edge of the SAMI field-of-view. For emission line fitting we flag and mask (set to NaN) any spaxel that has more than 500 pixels without spectral data in either the blue or red cubes. This is because the missing data can prevent a good continuum fit, and at times prevents lines from being fit as well.

Once the continuum is subtracted, LZIFU then fits the emission lines in each spatial element in the data products with 1, 2, and 3 Gaussian components. The velocity centroids of the components are measured relative to the redshift of the galaxy (as defined by the input catalogue), and components are ordered by increasing velocity dispersion. As with previous data releases, we fit all emission lines simultaneously, tying the velocities and velocity dispersions together, but allowing the relative strengths of these components to vary. The emission lines we fit are $[O II] \lambda\lambda 3727, 3729$, $H\beta$, $[O III] \lambda 5007$, $H\alpha$, $[N II] \lambda 6583$, $[S II] \lambda 6716$, and $[S II] \lambda 6731$. For the aperture spectra we also fit for the higher order Balmer lines, $H\gamma$, $H\delta$ and $H\epsilon$, as well as $[Ne III] \lambda 3869$, due to the generally higher S/N in these spectra.

The resulting spectral fits are then run through a trained Neural Network LZCOMP ([Hampton et al. 2017](#)) that determines whether 1, 2, or 3 Gaussian components are necessary to describe the observed emission line structure. This is returned as our "recommended" fit where each spatial element is described by the necessary number of components. As with previous data releases, we only provide the different fluxes in each component for the $H\alpha$ line as it is generally the strongest. The $H\alpha$ line also lies in the higher spectral resolution red spectrograph, providing the most information on the spectral decomposition of the emission lines. For all other lines we only provide the total flux. The central velocity and velocity dispersion for each necessary component are given for each spatial element.

An exception to this is the annular-binned data and the aperture spectra. In this instance for many objects the annular binning leads to double-peaked (or horned) profiles due to the sampling of the

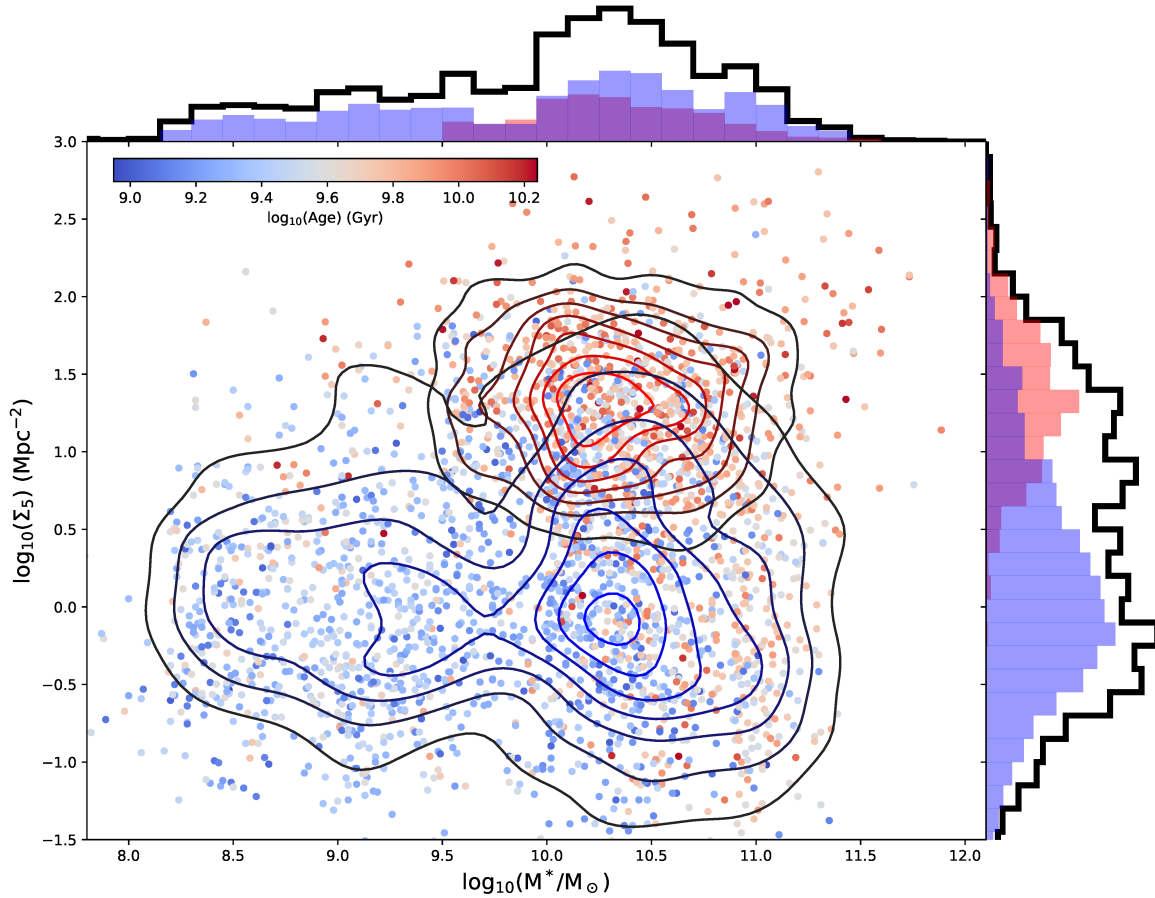


Figure 17. The filled circles show the distribution of SAMI DR3 galaxies in fifth-nearest-neighbour surface density (Σ_5 , Mpc^{-2}) versus stellar mass ($M^*/M_\odot$). For each galaxy, the filled circles are colour-coded based on the light-weighted age determined from an aperture spectrum containing light from within R_e in Section 7. The colourbar inset at the top left-hand corner shows the mapping from age to colour. The blue and red contours show the kernel-density estimates for the separate distributions of the GAMA and cluster samples, respectively, highlighting the contribution of the cluster regions, which allows the SAMI survey to probe the full range in environmental density. Histograms show the marginal distributions for the full DR3 sample (black line), the GAMA galaxies (blue shaded), and the cluster galaxies (red shaded). Note that only cluster galaxies with $\log_{10}(M^*/M_\odot) > 9.5$ were observed (see Section 2.1.2).

rotational profile which LZCOMP is unable to process. Thus for all annular binned data we only present our 2-component Gaussian fits with LZIFU and return the flux in all emission lines for these data. Similarly for the aperture spectra we only report the 1 component fits within the tables. Given the generally higher S/N in these larger bins, most emission lines have sufficient flux for significant decomposition.

The final DR3 released emission line products for full-resolution and binned SAMI data cubes are: total line flux maps for [O II] $\lambda\lambda 3727, 3729$, $\text{H}\beta$, [O III] $\lambda 5007$, [N II] $\lambda 6583$, $\text{H}\alpha$, [S II] $\lambda 6716$, and [S II] $\lambda 6731$ using 1-component and multi-component Gaussian fits, velocity and velocity dispersion based on the 1-component fits and for each recommended component of the multi-Gaussian fits, and for the multi-Gaussian fits the decomposed flux of each component for $\text{H}\alpha$.

8.1 Emission line derived value-added products

In addition to the emission line properties extracted from the SAMI data, we also derive standard value added products from this emission line data, namely: Balmer decrement derived extinction maps, classification maps based on standard emission line diagnostic di-

agrams, and star-formation maps. Full details of the derivations of these can be found in the DR2 paper and in [Medling et al. \(2018\)](#).

The extinction maps are given as an $\text{H}\alpha$ correction factor and assume an intrinsic Balmer decrement of $\text{H}\alpha/\text{H}\beta = 2.86$ and a [Cardelli et al. \(1989\)](#) extinction curve with $R_V = 4.1$. We note, as described in [Green et al. \(2018\)](#) and in [Medling et al. \(2018\)](#), that the Balmer decrement is sensitive to the aliasing introduced by DAR and the sampling of the PSF of the SAMI hexabundles. For this reason, our extinction maps have been smoothed by a Gaussian kernel with a FWHM of 1.6 spaxels (0.8 arcsec) which corrects this issue (with full details of the Kernel and the results shown in [Medling et al. \(2018\)](#)).

The classification maps distinguish star-formation dominated spaxels from spaxels where another ionising mechanism, such as AGN, supernova or diffuse ionised gas contribute significantly to the overall emission using emission-line ratio based diagnostics. As emphasized in previous papers, these classification maps are a conservative mask, capturing all spaxels where the SFR can be confidently determined. While star-formation can, and is likely to, contribute to the emission in masked spaxels, the other contributing sources such as AGN need to be accounted for in a non-straightforward process (see, e.g., [Davies et al. 2016](#); [D’Agostino et al. 2019](#), for examples of

this). A full description of the classification algorithm is in [Medling et al. \(2018\)](#).

The star-formation maps are based on the extinction-corrected $H\alpha$ flux for all star-forming dominated spaxels and use the [Kennicutt et al. \(1994\)](#) relation, corrected to a [Chabrier \(2003\)](#) stellar initial mass function, to convert to a SFR;

$$\text{SFR} [\text{M}_\odot \text{yr}^{-1}] = 5.16 \times 10^{-42} F_{H\alpha} [\text{erg s}^{-1}]. \quad (3)$$

For full SFR maps where the impact of AGN etc. are neglected, the same equation can be applied directly to the extinction-corrected $H\alpha$ maps.

A final emission line related value-added product is the measurement of the gas kinematic PA. This is measured in an identical way to the stellar kinematic PA discussed in Section 6.2, but applied to the 1-component gas kinematic maps.

9 ENVIRONMENTAL METRICS

As outlined in Section 2.1, DR3 includes SAMI data for the cluster regions for the first time. Therefore, when combined with the SAMI data drawn from the GAMA regions, the DR3 data span the full range of environmental densities, from the low-density field to the high-density regions found in the cores of rich clusters. In this Section, we outline the environment metrics included as value-added data products in DR3.

9.1 Fifth-nearest neighbour surface density estimates

Fifth-nearest neighbour surface density estimates are obtained for targets in both the GAMA and cluster regions. The surface density estimates are determined in a similar manner to that described in [Brough et al. \(2013\)](#) and [Brough et al. \(2017\)](#). Briefly, we use the redshift information provided by the GAMA survey ([Driver et al. 2011](#); [Liske et al. 2015](#)) and the SAMI Cluster Redshift Survey ([Owers et al. 2017](#)) to specify a density-defining population consisting of galaxies that have absolute r -band magnitudes $M_r < -18.6$ or $M_r < -19.0$ mag, with the latter limit included to allow densities to be estimated for the secondary targets with $z > 0.1$ in the GAMA regions. For the GAMA sample, the magnitudes of the density-defining population are k -corrected as outlined in [Loveday et al. \(2012\)](#), while for the cluster sample we use the CALC_KCORR software ([Chilingarian et al. 2010](#)).

The surface density for each SAMI target is determined as $\Sigma_5 = 5/\pi d_n^2$, where d_n is the projected comoving distance to the 5th nearest galaxy in the density-defining population that is within ± 1000 km/s of the SAMI target redshift. The Σ_5 values are corrected to account for the effect of spectroscopic incompleteness. In 95% of the cases that require a correction, the Σ_5 values increase by a factor that is smaller than 1.15. Uncertainties on the Σ_5 values are determined as $\Sigma_{5, \text{err}} = \max(|\Sigma_5 - \Sigma_4|, |\Sigma_5 - \Sigma_6|)$, where Σ_4 and Σ_6 are the surface densities determined to the 4th and 6th nearest neighbours, respectively. For the GAMA Σ_5 measurements, quality flags are assigned based on proximity to survey edges and bright star exclusion regions. For the cluster Σ_5 measurements, flags are assigned based on proximity to survey edges, although in practice this only occurs for a small number of cluster targets ($< 1\%$).

Fig. 17 shows the distribution of observed SAMI targets on the $\log_{10}(M^*/\text{M}_\odot)$ versus $\log_{10}(\Sigma_5)$ plane, where the points have been colour-coded based on the light-weighted SSP ages measured within $1 R_e$ apertures as described in Section 7. Contours show the kernel-density estimates separately for the GAMA and cluster

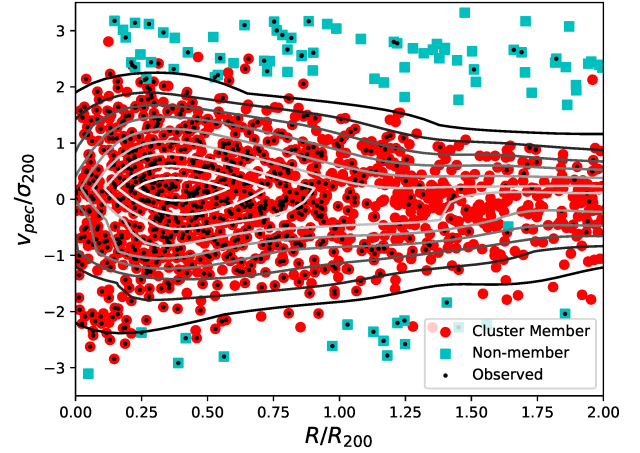


Figure 18. The combined projected-phase-space diagram for all galaxies in the input catalogue described in Section 2.1.2. Galaxies assigned cluster membership in [Owers et al. \(2017\)](#) are shown as red circles, while those non-members with $|v_{\text{pec}}/\sigma_{200}| < 3.5$ are shown as cyan squares. Targets that were observed during the survey are highlighted by black points, while the contours show the kernel density estimate derived from all spectroscopically confirmed cluster members in the SAMI Cluster Redshift Survey.

targets, while the histograms show the marginal distributions of the full sample (black) and the GAMA (blue shaded) and cluster (red shaded) samples. Fig. 17 demonstrates the breadth of environmental density that the combined DR3 sample probes.

9.2 Cluster-specific measurements

To complement the newly released cluster targets in DR3, within the cluster input catalogues we include three cluster-specific environment metrics: a member/non-member flag, normalised clustercentric distances (R/R_{200}), and normalised peculiar velocities ($v_{\text{pec}}/\sigma_{200}$). The description of the measurements of these metrics can be found in [Owers et al. \(2017\)](#), and we provide a brief overview here.

The R/R_{200} measurements are determined as the projected angular distance of the target to the cluster centre listed in Table 2, where the conversion from angular to a physical separation is determined using the angular diameter distance to the cluster. The normalisation by the virial radius estimator, R_{200} , allows for more physically meaningful comparisons between the clustercentric distance of galaxies in clusters of disparate halo mass. Here, $R_{200} = 0.17\sigma_{200}/H(z)$ Mpc is determined iteratively using the cluster velocity dispersion, σ_{200} , which is derived from the cluster members within R_{200} using the biweight estimator ([Beers et al. 1990](#)). The peculiar velocities for the SAMI targets are determined as $v_{\text{pec}} = c[(1+z_{\text{pec}})^2 - 1]/[(1+z_{\text{pec}})^2 + 1]$ where $z_{\text{pec}} = (z - z_{\text{clus}})/(1 + z_{\text{clus}})$, z is the redshift of the SAMI target determined from the SAMI Cluster Redshift Survey, and z_{clus} is the cluster redshift determined using the biweight mean of the cluster members (z_{clus}). The values of R_{200} , z_{clus} , and σ_{200} for each cluster are listed in Table 2.

As noted in Section 2.1.2, the selection of SAMI targets used a more relaxed cut in v_{pec} when compared with that used to define cluster members in [Owers et al. \(2017\)](#). This less-conservative cut

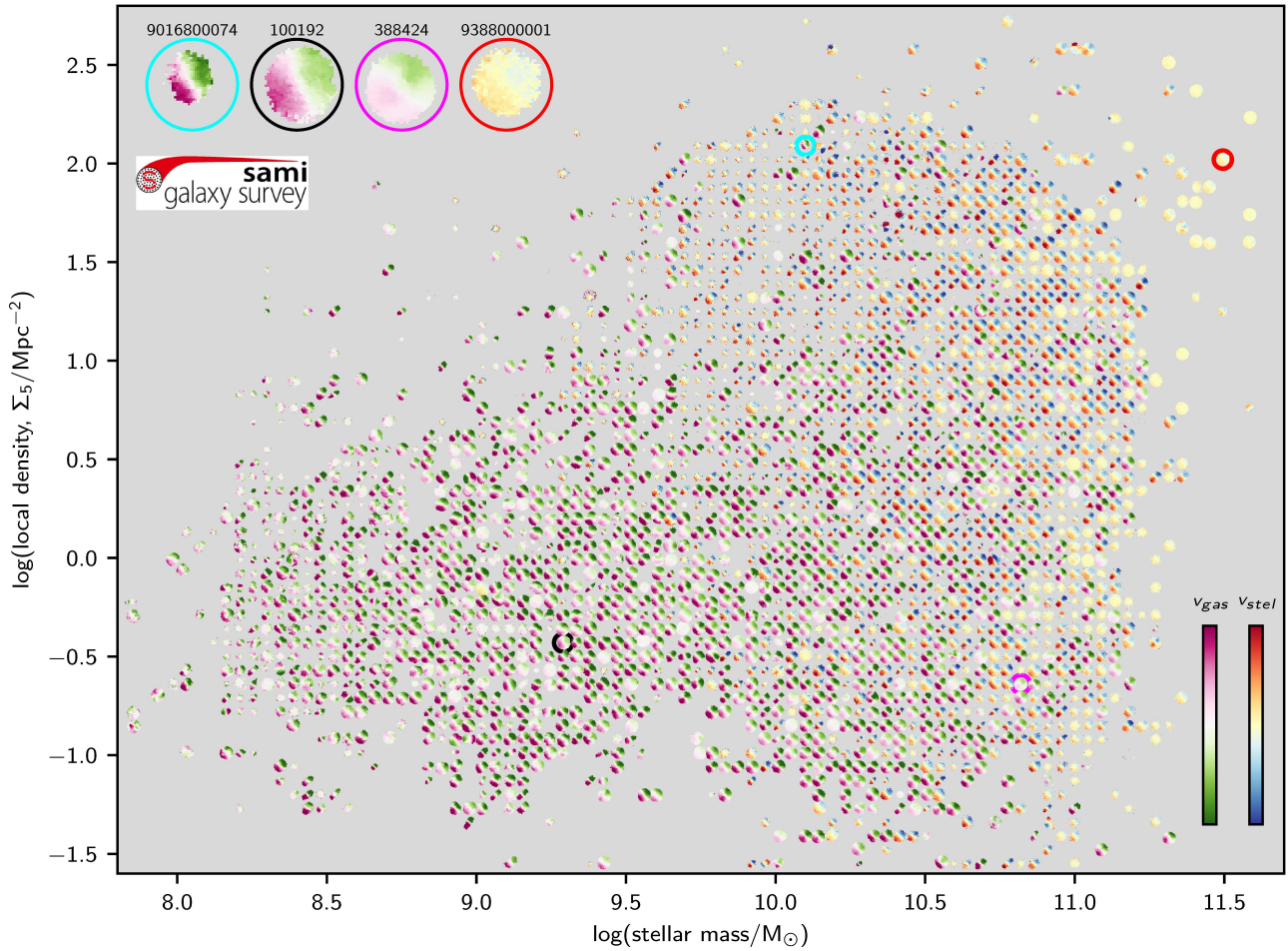


Figure 19. Velocity fields for SAMI galaxies in the plane of fifth-nearest-neighbour surface density (Σ_5) vs. stellar mass. The velocity scale for each galaxy is normalized to a maximum of $0.7\times$ the velocity inferred from the stellar mass Tully-Fisher relation (hence there is no scale on the velocity colour bars). Stellar velocities are shown for early-type galaxies, while gas kinematics are shown for late-type galaxies, as defined by the visual morphology catalogue. Each velocity map has the same image scale in arcseconds. Top left we show zoom-ins of four example galaxies, where their locations in Σ_5 and stellar mass are indicated by the coloured circles. Zoom-ins of the same galaxies are shown in Figs. 20 and 21.

was designed to allow for the inherent uncertainty in assigning cluster membership, which may exclude important high-velocity galaxies infalling onto the cluster for the first time. We therefore include a flag in our catalogue that indicates whether a SAMI target was allocated as a cluster member in [Owers et al. \(2017\)](#). In Fig. 18 we show the projected-phase-space (PPS) diagram for the SAMI cluster targets. The cluster environmental metrics are contained within the input catalogue for the cluster sample, InputCatClustersDR3.

Also listed in Table 2 are estimates for the virial mass of the clusters, which are determined using the spectroscopically confirmed members as outlined in [Owers et al. \(2017\)](#). Halo masses for groups in the GAMA regions will be made available in a forthcoming release ([Robotham et al. in prep.](#)), and the SAMI DR3 galaxies in the GAMA regions can be linked to these groups via their CATIDs. It is worth reiterating that the virial mass determinations for the clusters in [Owers et al. \(2017\)](#) differ from those determined by [Robotham et al. \(2011\)](#) for the groups in the GAMA survey both in terms of the assumed cosmology and the methodology. As outlined in [Owers et al. \(2017\)](#), care must therefore be taken in combining the catalogues. After rescaling masses to the same assumed cosmology, [Owers et al. \(2017\)](#) recommend that the cluster virial mass

estimates should be scaled by a factor of 1.25 in order to match the virial-like mass estimates determined for the GAMA groups in [Robotham et al. \(2011\)](#).

9.3 SAMI galaxies across the mass-environment plane

A fundamental aim of the SAMI Galaxy Survey was to measure the internal structure of galaxies across mass and environment. To visualise this Figs. 19, 20 and 21 present velocity maps, $H\alpha$ flux maps and $\log([N\text{ II}]/H\alpha)$ maps in the plane of stellar mass vs. Σ_5 . Below we provide a qualitative description of some of the trends seen when visualizing the SAMI data in this way.

In Fig. 19 we show gas velocity maps for late-type galaxies (green to purple colour scale) and stellar velocity maps for early-type galaxies (blue to red colour scale). For each plotted velocity map the maximum and minimum velocity is set to be $0.7\times$ the velocity inferred from the stellar mass of the galaxy, using the Tully-Fisher ([Tully & Fisher 1977](#)) relation as measured by the SAMI Galaxy Survey ([Bloom et al. 2017](#)). The position angle of the velocity maps is set by the stellar kinematic PA, or the gas kinematic PA if the error

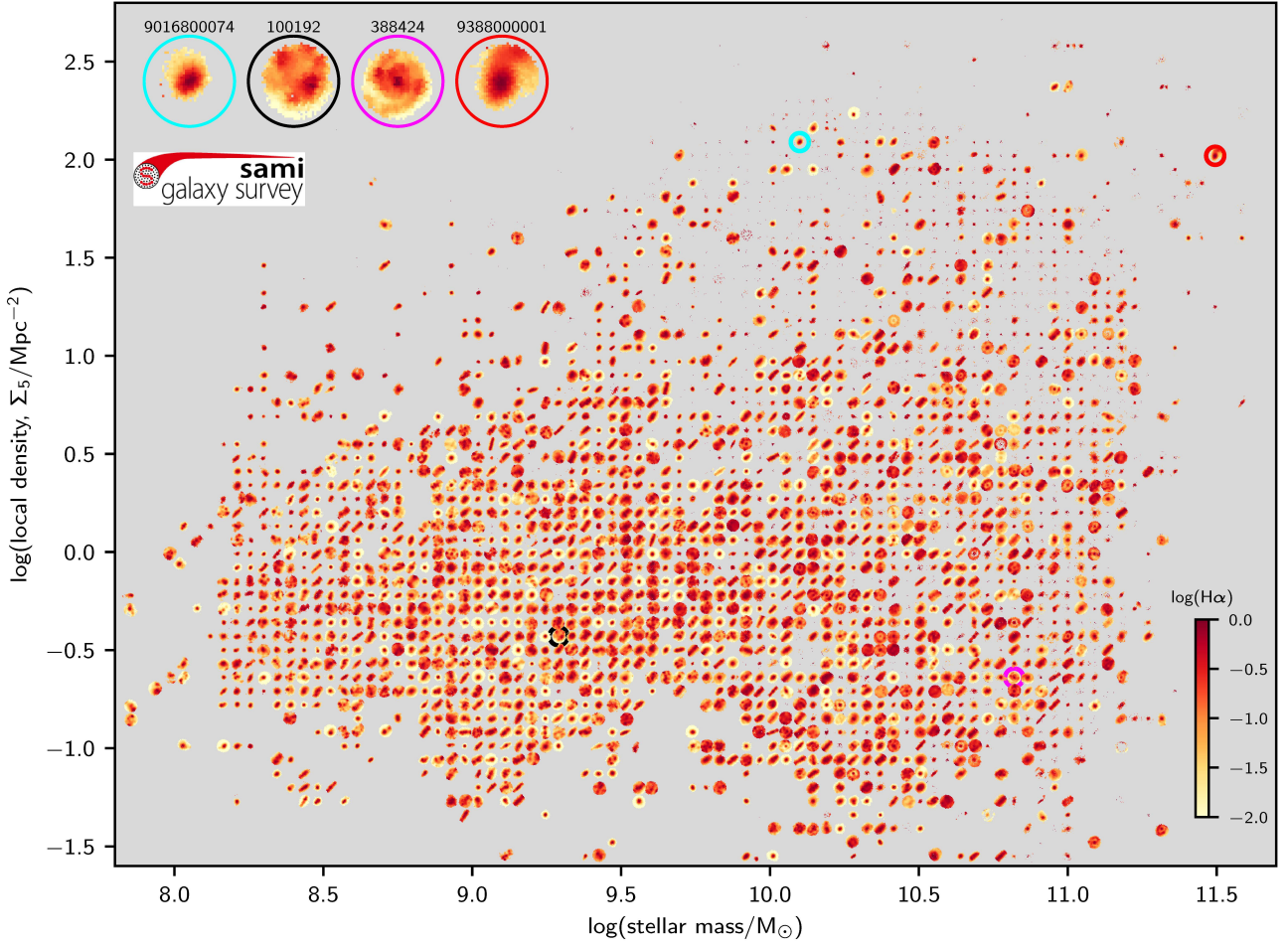


Figure 20. $H\alpha$ maps for SAMI galaxies in the plane of fifth-nearest-neighbour surface density (Σ_5) vs. stellar mass. The $H\alpha$ flux is on a log scale and normalized such that $\log(\text{flux}) = 0$ for the maximum $H\alpha$ in each galaxy. The zoom-ins in the top left are the same galaxies as in Figs. 19 and 21.

in the stellar kinematic PA is larger than 20 degrees. The location of the maps approximately coincides with the value of stellar mass and Σ_5 for each galaxy, but to reduce overlapping of the maps each galaxy was positioned in the nearest available location in a grid.

Above $\Sigma_5 \approx 10 \text{ Mpc}^{-2}$ the majority of SAMI galaxies are from the cluster fields. The increased fraction of early-type galaxies above this value can be clearly seen, as can the increased number of early types towards high stellar mass. At all masses and environments there are some galaxies with low velocity gradients relative to their expected Tully-Fisher rotational velocity. These low gradients are largely due to inclination effects. However, at the highest masses [above $\log(M_*/M_\odot) = 11$] there is an increased number of galaxies with low velocity gradients, consistent with the increased fraction of slow rotators in this mass range (e.g. Brough et al. 2017). It is less clear from Fig. 19 whether the galaxies with the highest Σ_5 have lower rotation. In fact, we see suggestions of an increased fraction of slow rotators with mass at all values of Σ_5 . This will be investigated fully by van de Sande et al. (in preparation). At the lowest masses ($\log(M_*/M_\odot) < 8.5$), more galaxies also tend to have lower rotation compared to their expected values from the Tully-Fisher relation, as pointed out by Bloom et al. (2017).

Fig. 20 shows $H\alpha$ maps across the stellar mass vs. Σ_5 plane. The $H\alpha$ maps are log-scaled and normalized relative to the bright-

est $H\alpha$ flux in each galaxy. As Σ_5 increases, fewer galaxies show evidence of extended $H\alpha$ emission, with either no emission at all, or the emission appearing more compact. The $H\alpha$ morphology also varies with stellar mass. As mass increases, galaxies often have a bright central peak, that is sometimes surrounded by a ring, providing a ‘target-like’ morphology.

The ionization state of the gas can be visualized using $\log([\text{N II}]/H\alpha)$ flux ratio maps, as in Fig. 21. The most significant trend visible here is the change from low $\log([\text{N II}]/H\alpha)$ (blue colours) at low mass, to high $\log([\text{N II}]/H\alpha)$ (green colours) at high mass. This is driven by increasing gas-phase metallicity with increasing mass. $\log([\text{N II}]/H\alpha)$ values above ~ -0.2 cannot be caused by star formation, but must be driven by other sources of ionization, such as AGN or shocks. This non-star-formation emission becomes more dominant at high stellar mass. The $[\text{N II}]/H\alpha$ ratio maps show a diversity of structure, including examples of galaxies with enhanced $[\text{N II}]/H\alpha$ at large radius, indicative of shocks in galactic-scale outflows or diffuse ionized gas. At high stellar mass many of the galaxies with central $H\alpha$ flux peaks (sometimes surrounded by rings) have high central $[\text{N II}]/H\alpha$ ratios, indicating LINER or AGN-like activity. This is likely to be driven by increased bulge fraction in high mass disks that have central regions that are partially or completely quenched.

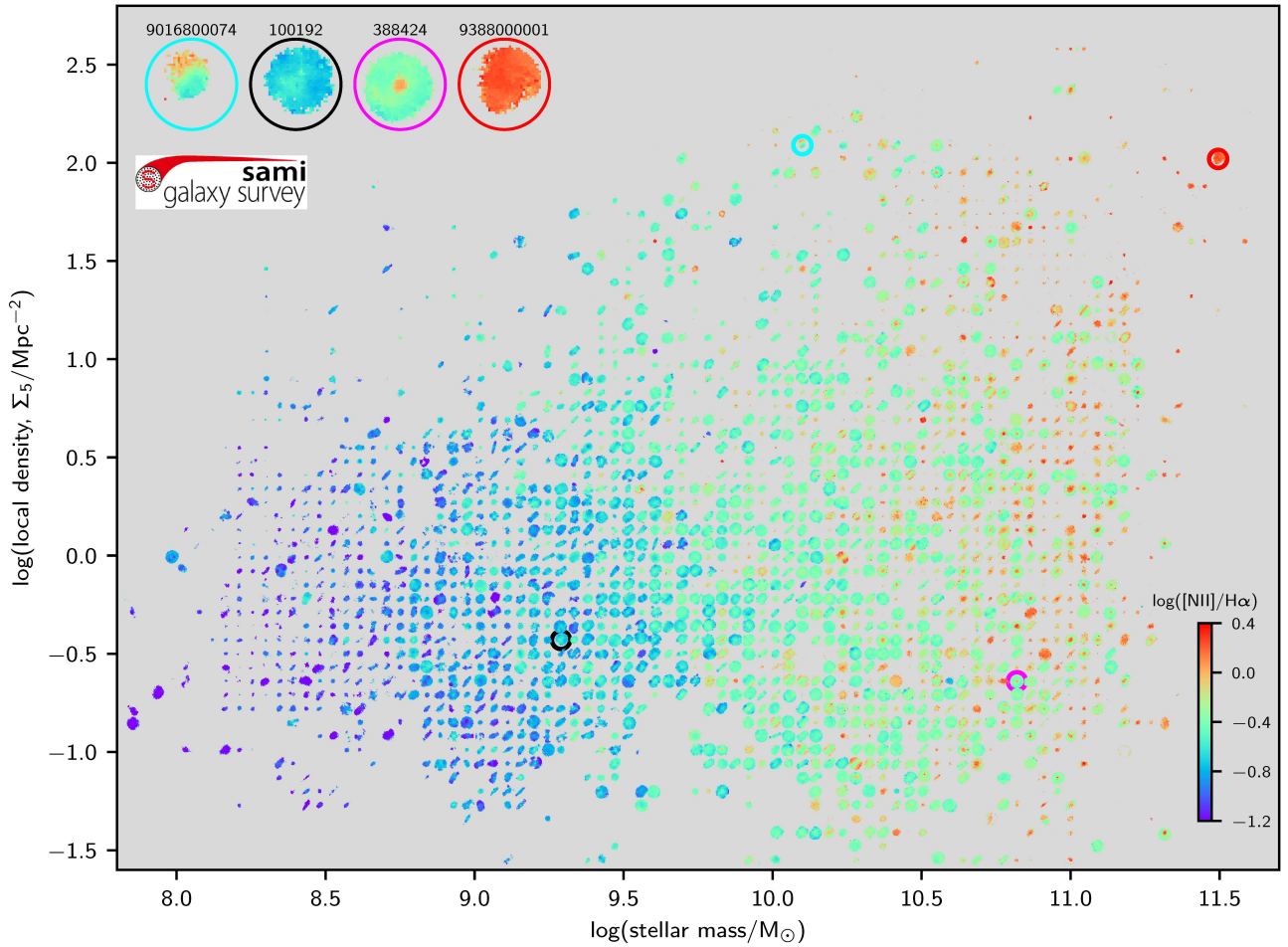


Figure 21. Maps of $\log([\text{N II}]/\text{H}\alpha)$ flux ratio for SAMI galaxies in the plane of fifth-nearest-neighbour surface density (Σ_5) vs. stellar mass. The zoom-ins in the top left are the same galaxies as in Figs. 19 and 20.

In Figs. 19, 20 and 21 we show zoom-ins of four example galaxies (top left), using coloured circles to indicate their location in Σ_5 and stellar mass. The first example (galaxy 9016800074, cyan circle) is a late-type galaxy in cluster Abell 168 with regular rotation, but assymmetric $\text{H}\alpha$ emission on one side of the galaxy. The assymmetric gas has a high $[\text{N II}]/\text{H}\alpha$ ratio, suggestive of shocks, possibly caused by ram-pressure stripping (Owers et al. 2019). Our second example (galaxy 100192, black circle) is a lower-mass late-type galaxy in a low density environment. 100192 has a regular rotating disk, including clumpy star formation and a radial $[\text{N II}]/\text{H}\alpha$ gradient, indicating a metallicity gradient. The third example (galaxy 288424, magenta circle) is a higher mass disk galaxy that is close to face-on. The central $[\text{N II}]/\text{H}\alpha$ ratio for 288424 is high, suggestive of AGN or LINER-like emission. Our final example (galaxy 9388000001, red circle) is the brightest cluster galaxy in Abell 3880. Its stellar kinematics show it to be a slow rotator, but the galaxy has strong $\text{H}\alpha$ emission. The $\text{H}\alpha$ emission is centrally concentrated, but with an extended plume that was also noted by Hamer et al. (2016). All the ionized gas emission in 9388000001 has a high $[\text{N II}]/\text{H}\alpha$ ratio, consistent with shocks or AGN emission. This galaxy is a known bright radio source (PKS 2225-308; Shimmins & Bolton 1974), and the ionization is likely related to this.

The maps presented here give just a few different examples of the richness of large-scale IFS data sets such as SAMI.

10 SAMI DR3 PRODUCTS AND ACCESS

All SAMI DR3 data and products are available through Australian Astronomical Optics’ Data Central service at <https://datacentral.org.au/>. Data Central provides query tools based on a SQL framework to search and connect data from different SAMI catalogues and also federate it with data from other surveys such as GAMA (Driver et al. 2011). A simple query to select observed galaxies in a particular stellar mass range is given below:

```
SELECT t1.CATID, t1.RA_OBJ, t1.DEC_OBJ, t1.Mstar,
       t2.CUBEID, t2.CUBEIDPUB
FROM sami_dr3.InputCatGAMADR3 as t1
INNER JOIN sami_dr3.CubeObs as t2 on
t2.CATID = t1.CATID
WHERE (t1.Mstar BETWEEN 10.5 and 11.0) and
(t2.ISBEST = True)
```

Data products related to the output of a given SQL search can then be downloaded.

Cubes, images and spectra can be downloaded as FITS files. Tables can be downloaded in a range of formats including FITS, CSV and VOTABLE files. Data Central contains a full schema browser that describes the data products as well as detailed documentation.

11 SUMMARY

With this paper we have released data for all the observations made as part of the SAMI Galaxy Survey, as well as a range of value-added data products, including gas and stellar measurements. This includes data for 3068 unique galaxies. For the first time, the SAMI Galaxy Survey is releasing data within the eight clusters targeted as part of the survey. These data provide a unique view of the role of dense environments in influencing galaxy evolution, as seen in the local Universe. The SAMI Galaxy Survey has already enabled a diverse array of investigations into galaxy evolution, and we hope that the public release of this data set will enable many more.

There are a number of important next steps beyond the SAMI Galaxy Survey that can further elucidate the complexities of galaxy formation. In the local Universe, larger samples and higher resolution (both spatial and spectral) are key directions. Higher spectral resolution is particularly important to understand the kinematic signatures in low-mass galaxies, as well as high-mass disks and in cases with complex kinematics such as outflows. The Hector instrument, to be commissioned on the Anglo-Australian Telescope in 2021 will provide this higher spectral resolution for over 10000 galaxies (Bryant et al. 2016). Larger telescopes with greater sensitivity are also starting to make IFS measurements outside of the local Universe, although the number that have included stellar measurements as well as gas is limited. New projects such as the Middle-Ages Galaxy Properties with Integral Field Spectroscopy Survey (MAGPI; Foster et al. 2020) are using the Multi-Unit Spectroscopic Explorer (MUSE) instrument (Bacon et al. 2010) to push stellar kinematics measurements to higher redshift. In the future larger telescopes still will be required to overcome the surface brightness limitations of current facilities. The prospect of multiplexing extremely large telescopes, for example using the MANIFEST facility on the Giant Magellan Telescope (Saunders et al. 2010), opens up the possibility of carrying out SAMI-like surveys in the distant Universe.

DATA AVAILABILITY

The data presented in this paper are available from Astronomical Optics' Data Central service at <https://datacentral.org.au/>.

ACKNOWLEDGEMENTS

The SAMI Galaxy Survey is based on observations made at the Anglo-Australian Telescope. The Sydney-AAO Multi-object Integral field spectrograph (SAMI) was developed jointly by the University of Sydney and the Australian Astronomical Observatory. The SAMI input catalogue is based on data taken from the Sloan Digital Sky Survey, the GAMA Survey and the VST ATLAS Survey. The SAMI Galaxy Survey website is <http://sami-survey.org/>. The SAMI Galaxy Survey is supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013, the Australian Research Council Centre of Excellence for All-sky

Astrophysics (CAASTRO), through project number CE110001020, and other participating institutions.

Based on data acquired at the Anglo-Australian Telescope under programs A/2013B/012 and A/2016B/16. We acknowledge the traditional owners of the land on which the AAT stands, the Gamilaraay people, and pay our respects to elders past and present.

GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programmes including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is <http://www.gama-survey.org/>.

SMC acknowledges the hospitality of the Astronomy, Astrophysics and Astrophotonics Research Centre at Macquarie University while preparing parts of this work. MSO acknowledges the funding support from the Australian Research Council through a Future Fellowship (FT140100255). JJB acknowledges support of an Australian Research Council Future Fellowship (FT180100231). FDE acknowledges funding through the H2020 ERC Consolidator Grant 683184. JvdS acknowledges support of an Australian Research Council Discovery Early Career Research Award (project number DE200100461) funded by the Australian Government. NS acknowledges support of an Australian Research Council Discovery Early Career Research Award (project number DE190100375) funded by the Australian Government. TMB is supported by an Australian Government Research Training Program Scholarship. DO is a recipient of an Australian Research Council Future Fellowship (FT190100083) funded by the Australian Government. LC is the recipient of an Australian Research Council Future Fellowship (FT180100066) funded by the Australian Government. SB acknowledges funding support from the Australian Research Council through a Future Fellowship (FT140101166). This work was supported by the UK Science and Technology Facilities Council through the 'Astrophysics at Oxford' grant ST/K00106X/1. RLD acknowledges travel and computer grants from Christ Church, Oxford and support from the Oxford Hintze Centre for Astrophysical Surveys which is funded by the Hintze Family Charitable Foundation. The work was undertaken in collaboration with the Melbourne Data Analytics Platform (MDAP) at the University of Melbourne. This material is based in part upon work supported by the National Science Foundation under Grant No. 2009416. We thank Mike Bessell for discussions regarding extinction at Siding Spring Observatory.

This paper made use of the *astropy* PYTHON package (Astropy Collaboration et al. 2013), as well as the *matplotlib* plotting software (Hunter 2007) and the scientific libraries *numpy* (Harris et al. 2020), *IPython* (Perez & Granger 2007) and *scipy* (Virtanen et al. 2020).

REFERENCES

- AAO software team 2015, 2dfdr: Data reduction software, Astrophysics Source Code Library (ascl:1505.015)
- Abazajian K., et al., 2004, *AJ*, **128**, 502
- Ahn C. P., et al., 2012, *ApJS*, **203**, 21
- Aldering G., et al., 2002, in Tyson J. A., Wolff S., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4836, Proc. SPIE. pp 61–72, doi:[10.1117/12.458107](https://doi.org/10.1117/12.458107)

- Allen J. T., et al., 2014, SAMI: Sydney-AAO Multi-object Integral field spectrograph pipeline, Astrophysics Source Code Library (ascl:1407.006)
- Allen J. T., et al., 2015, *MNRAS*, **446**, 1567
- Astropy Collaboration et al., 2013, *A&A*, **558**, A33
- Bacon R., et al., 2001, *MNRAS*, **326**, 23
- Bacon R., et al., 2010, in Ground-based and Airborne Instrumentation for Astronomy III. p. 773508, doi:10.1117/12.856027
- Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, *MNRAS*, **373**, 469
- Barat D., et al., 2019, *MNRAS*, **487**, 2924
- Barone T. M., et al., 2018, *ApJ*, **856**, 64
- Beers T. C., Flynn K., Gebhardt K., 1990, *AJ*, **100**, 32
- Bertin E., Arnouts S., 1996, *A&AS*, **117**, 393
- Bland-Hawthorn J., Gerhard O., 2016, *ARA&A*, **54**, 529
- Bland-Hawthorn J., et al., 2011, *Optics Express*, **19**, 2649
- Blanton M. R., Moustakas J., 2009, *ARA&A*, **47**, 159
- Bloom J. V., et al., 2017, *MNRAS*, **472**, 1809
- Boardman N., et al., 2020, *MNRAS*, **498**, 4943
- Brough S., et al., 2013, *MNRAS*, **435**, 2903
- Brough S., et al., 2017, *ApJ*, **844**, 59
- Bruzual G., Charlot S., 2003, *MNRAS*, **344**, 1000
- Bryant J. J., O'Byrne J. W., Bland-Hawthorn J., Leon-Saval S. G., 2011, *MNRAS*, **415**, 2173
- Bryant J. J., Bland-Hawthorn J., Fogarty L. M. R., Lawrence J. S., Croom S. M., 2014, *MNRAS*, **438**, 869
- Bryant J. J., et al., 2015, *MNRAS*, **447**, 2857
- Bryant J. J., et al., 2016, in Proc. SPIE. p. 99081F (arXiv:1608.03921), doi:10.1117/12.2230740
- Bundy K., et al., 2015, *ApJ*, **798**, 7
- Cappellari M., 2002, *MNRAS*, **333**, 400
- Cappellari M., 2017, *MNRAS*, **466**, 798
- Cappellari M., Copin Y., 2003, *MNRAS*, **342**, 345
- Cappellari M., Emsellem E., 2004, *PASP*, **116**, 138
- Cappellari M., et al., 2007, *MNRAS*, **379**, 418
- Cappellari M., et al., 2011, *MNRAS*, **413**, 813
- Cappellari M., et al., 2013, *MNRAS*, **432**, 1709
- Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, **345**, 245
- Chabrier G., 2003, *PASP*, **115**, 763
- Chilingarian I. V., Melchior A.-L., Zolotukhin I. Y., 2010, *MNRAS*, **405**, 1409
- Cid Fernandes R., Mateus A., Sodré L., Stasińska G., Gomes J. M., 2005, *MNRAS*, **358**, 363
- Colless M., et al., 2001, *MNRAS*, **328**, 1039
- Cortese L., et al., 2014, *ApJ*, **795**, L37
- Cortese L., et al., 2016, *MNRAS*, **463**, 170
- Cortese L., et al., 2019, *MNRAS*, **485**, 2656
- Croom S. M., et al., 2012, *MNRAS*, **421**, 872
- D'Agostino J. J., Kewley L. J., Groves B. A., Medling A., Dopita M. A., Thomas A. D., 2019, *MNRAS*, **485**, L38
- Davies R. L., et al., 2016, *MNRAS*, **462**, 1616
- Dressler A., 1980, *ApJ*, **236**, 351
- Driver S. P., et al., 2011, *MNRAS*, **413**, 971
- Driver S. P., et al., 2016, *MNRAS*, **455**, 3911
- Emsellem E., Monnet G., Bacon R., 1994, *A&A*, **285**, 723
- Emsellem E., et al., 2007, *MNRAS*, **379**, 401
- Falcón-Barroso J., Sánchez-Blázquez P., Vazdekis A., Ricciardelli E., Cardiel N., Cenarro A. J., Gorgas J., Peletier R. F., 2011, *A&A*, **532**, A95
- Ferreras I., et al., 2019, *MNRAS*, **489**, 608
- Fogarty L. M. R., et al., 2012, *ApJ*, **761**, 169
- Fogarty L. M. R., et al., 2014, *MNRAS*, **443**, 485
- Foster C., et al., 2017, *MNRAS*, **472**, 966
- Foster C., et al., 2020, arXiv e-prints, p. arXiv:2011.13567
- Genel S., et al., 2019, *ApJ*, **871**, 21
- Gerhard O. E., 1993, *MNRAS*, **265**, 213
- Giovannelli R., et al., 2005, *AJ*, **130**, 2598
- González Delgado R. M., Cerviño M., Martins L. P., Leitherer C., Hauschildt P. H., 2005, *MNRAS*, **357**, 945
- Graves G. J., 2014, EZ_Ages: Stellar population age calculator (ascl:1407.019)
- Green A. W., et al., 2018, *MNRAS*, **475**, 716
- Gurri P., Taylor E. N., Fluke C. J., 2020, *MNRAS*
- Hallman E. J., Markevitch M., 2004, *ApJ*, **610**, L81
- Hamer S. L., et al., 2016, *MNRAS*, **460**, 1758
- Hampton E. J., et al., 2017, *MNRAS*, **470**, 3395
- Harris C. R., et al., 2020, *Nature*, **585**, 357–362
- Haynes M. P., et al., 2018, *ApJ*, **861**, 49
- Ho I. T., et al., 2014, *MNRAS*, **444**, 3894
- Ho I.-T., et al., 2016a, *Ap&SS*, **361**, 280
- Ho I.-T., et al., 2016b, *MNRAS*, **457**, 1257
- Hunter J. D., 2007, *Computing in Science and Engineering*, **9**, 90
- Husemann B., Kamann S., Sandin C., Sánchez S. F., García-Benito R., Mast D., 2012, *A&A*, **545**, A137
- Kausch W., et al., 2015, *A&A*, **576**, A78
- Keller B. W., Kruijssen J. M. D., 2020, arXiv e-prints, p. arXiv:2004.03608
- Keller B. W., Wadsley J. W., Wang L., Kruijssen J. M. D., 2019, *MNRAS*, **482**, 2244
- Kelvin L. S., et al., 2012, *MNRAS*, **421**, 1007
- Kennicutt Robert C. J., Tamblyn P., Congdon C. E., 1994, *ApJ*, **435**, 22
- Khim D. J., et al., 2020, *ApJ*, **894**, 106
- Krajinović D., Cappellari M., de Zeeuw P. T., Copin Y., 2006, *MNRAS*, **366**, 787
- Krajinović D., et al., 2008, *MNRAS*, **390**, 93
- Krajinović D., et al., 2011, *MNRAS*, **414**, 2923
- Kurucz R. L., 1992, in Barbuy B., Renzini A., eds, IAU Symposium Vol. 149, The Stellar Populations of Galaxies. p. 225
- Lewis I., et al., 2002, *MNRAS*, **334**, 673
- Liske J., et al., 2015, *MNRAS*, **452**, 2087
- Loveday J., et al., 2012, *MNRAS*, **420**, 1239
- Medling A. M., et al., 2018, *MNRAS*, **475**, 5194
- Neckel H., 1999, *Sol. Phys.*, **184**, 421
- Oh S., et al., 2020, *MNRAS*, **495**, 4638
- Owers M. S., et al., 2017, *MNRAS*, **468**, 1824
- Owers M. S., et al., 2019, *ApJ*, **873**, 52
- Peng Y.-j., et al., 2010, *ApJ*, **721**, 193
- Perez F., Granger B. E., 2007, *Computing in Science and Engineering*, **9**, 21
- Planck Collaboration et al., 2014, *A&A*, **571**, A11
- Poetrodjojo H., et al., 2018, *MNRAS*, **479**, 5235
- Proctor R. N., Forbes D. A., Beasley M. A., 2004, *MNRAS*, **355**, 1327
- Robotham A. S. G., et al., 2011, *MNRAS*, **416**, 2640
- Robotham A. S. G., et al., 2014, *MNRAS*, **444**, 3986
- Sánchez-Blázquez P., et al., 2006, *MNRAS*, **371**, 703
- Sánchez S. F., et al., 2012, *A&A*, **538**, A8
- Sánchez S. F., et al., 2019, *MNRAS*, **484**, 3042
- Santucci G., et al., 2020, *ApJ*, **896**, 75
- Saunders W., Colless M., Saunders I., Hopkins A., Goodwin M., Heijmans J., Brzeski J., Farrell T., 2010, in Ground-based and Airborne Instrumentation for Astronomy III. p. 773568, doi:10.1117/12.856353
- Schaefer A. L., et al., 2019, *MNRAS*, **483**, 2851
- Schiavon R. P., 2007, *ApJS*, **171**, 146
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, **500**, 525
- Scott N., et al., 2017, *MNRAS*, **472**, 2833
- Scott N., et al., 2018, *MNRAS*, **481**, 2299
- Shanks T., et al., 2015, *MNRAS*, **451**, 4238
- Sharp R., Parkinson H., 2010, *MNRAS*, **408**, 2495
- Sharp R., et al., 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. p. 62690G (arXiv:astro-ph/0606137), doi:10.1117/12.671022
- Sharp R., et al., 2015, *MNRAS*, **446**, 1551
- Shimmins A. J., Bolton J. G., 1974, Australian Journal of Physics Astrophysical Supplement, **32**, 1
- Smette A., et al., 2015, *A&A*, **576**, A77
- Taylor E. N., et al., 2011, *MNRAS*, **418**, 1587
- Thomas D., Maraston C., Johansson J., 2011, *MNRAS*, **412**, 2183
- Trager S. C., Worthey G., Faber S. M., Burstein D., González J. J., 1998, *ApJS*, **116**, 1

- Tully R. B., Fisher J. R., 1977, *A&A*, **500**, 105
Varidel M. R., et al., 2020, *MNRAS*, **495**, 2265
Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., 2010, *MNRAS*, **404**, 1639
Vazdekis A., et al., 2015, *MNRAS*, **449**, 1177
Virtanen P., et al., 2020, *Nature Methods*, **17**, 261
Welker C., et al., 2020, *MNRAS*, **491**, 2864
York D. G., et al., 2000, *AJ*, **120**, 1579
de Burgh-Day C. O., Taylor E. N., Webster R. L., Hopkins A. M., 2015, *MNRAS*, **451**, 2161
van Dokkum P. G., 2001, *PASP*, **113**, 1420
van de Sande J., et al., 2017a, *MNRAS*, **472**, 1272
van de Sande J., et al., 2017b, *ApJ*, **835**, 104
van de Sande J., et al., 2019, *MNRAS*, **484**, 869
van der Marel R. P., Franx M., 1993, *ApJ*, **407**, 525

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.